



Universidad Autónoma de Madrid
Departamento de Química



Université Evry Val d'Essonne
LAMBE

Fragmentation dynamics of biomolecules in gas phase and water environment

A thesis submitted for the degree of
Doctor in Chemistry

Candidate :

Pablo López Tarifa

Supervisors :

Manuel Alcamí Pertejo

Marie-Pierre Gaignot

Madrid, January 2011

Agradecimientos

Esta tesis se ha sido realizada entre el Departamento de Química de la Universidad Autónoma de Madrid y el Laboratoire Analyse et Modélisation pour la Biologie et l'Environnement (LAMBE) de la Université Evry Val d'Essonne durante los años 2008-2010 bajo la co-dirección de los profesores Manuel Alcamí Pertejo y Marie-Pierre Gageot, a quienes me gustaría agradecer todo su apoyo, tiempo, esfuerzo y dedicación para que este trabajo saliera adelante.

También quisiera agradecer a los demás profesores responsables no sólo del trabajo aquí recogido, sino también de mi formación tanto científica como personal recibida durante estos años:

-) A los profesores Marie-Françoise Politis y Fernando Martín por su apoyo incondicional, dedicación y consejos; por haber llevado cuidado de cuanto he hecho.

-) A los profesores Marie-Anne Hervé du Penhoat y Rodolphe Villeumier por su inestimable ayuda, desde el principio hasta el final, tanto *in situ* como a la distancia. Agradezcoles su cuidado y su cálida acogida en mi visita a sus laboratorios.

-) Al profesor Ivano Tavernelli por su inestimable ayuda y su excelente acogida en Lausanne; por haberme brindado la oportunidad de aprender otras aplicaciones de la TD-DFT.

A todos ellos en definitiva agradecerles su apuesta en este proyecto conjunto y su confianza depositada tanto en mí como en mi trabajo.

Me gustaría también agradecer al profesor Patrick Moretto-Capelle y Lamri Adoui por su amabilidad y atención prestada.

A todos los demás profesores y compañeros del departamento C-IX por su incansable apoyo y sabios consejos. Gracias por haberme hecho más livianos y llevaderos los momentos más duros.

También quisiera agradecer a los miembros del tribunal de esta tesis, Manuel Yañez, Marie-Pierre Gageot, Alberto Castro, Ivano Tavernelli y Paola Bolognesi por aceptar gentilmente formar parte del mismo.

Fuera del mundo de la ciencia, me gustaría agradecer a mis amigos y familiares por haber soportado los momentos más difíciles. Gracias Ele por tu paciencia, apoyo infinito y serenidad. Gracias a mis padres y hermanos por vuestro cariño, esfuerzo, humildad e ilusión.

Resumen

Hoy en día, el empleo de iones muy energéticos y altamente cargados se ha convertido en una exitosa alternativa a los tradicionales rayos-X en el tratamiento del cáncer. Cuando son incididos sobre tejidos vivos, dichos iones presentan una región en la cual su depósito de energía es máximo (región de Bragg); permitiendo así una mayor localización de la dosis a una determinada profundidad de penetración. Desafortunadamente la formación de las especies químicas resultantes de dicha interacción (y que pueden llegar a inducir la muerte celular) no pueden ser establecidos experimentalmente ya que ocurren a una escala de tiempo muy reducida ($\approx 1 - 100$ fs).

Las simulaciones basadas en Dinámica Molecular (MD) han demostrado ser la herramienta idónea para el estudio de tales mecanismos, sin embargo, hasta ahora muy pocos trabajos se han centrado en estos procesos primarios. En la presente tesis hemos aplicado una combinación de MD basada en la Teoría del Funcional de la Densidad Tiempo-Dependiente (TD-DFT) y en la aproximación de Born-Oppenheimer (BO), con el fin de investigar la fragmentación de especies doblemente cargadas tales como uracilo, pirimidina y pequeños agregados de agua, generados mediante la ionizaciones de capas internas.

El estado inicial doblemente cargado es preparado mediante la sustracción de dos electrones de un orbital específico de Kohn-Sham o Wannier de la especie neutra. La densidad electrónica evoluciona según la llamada aproximación Ehrenfest MD donde una superficie electrónica promedio, calculada al nivel TD-DFT (mediante la propagación tiempo-dependiente de las ecuaciones de Kohn-Sham) conduce el movimiento nuclear. Las fuerzas que actúan sobre cada átomo son calculadas mediante el teorema de Hellmann-Feynman. Cuando la dinámica es esencialmente adiabática, la superficie de energía potencial es la correspondiente al estado fundamental y en consecuencia, la dinámica es extendida mediante una BO MD hasta un tiempo total de ≈ 100 fs.

Los resultados obtenidos para la fragmentación de uracilo²⁺ en fase gas están en buen acuerdo con las coincidencias ion-ion medidas en los experimentos de colisión de uracilo con protones de 100 keV. Orbitales de misma energía y/o localizados en los mismos enlaces pueden llevar a muy diferentes fragmentaciones, mostrando así la importancia del medio químico intramolecular. Los fragmentos observados casi nunca corresponden con los caminos de disociación más favorables, lo que pone de manifiesto la importancia de los efectos dinámicos que ocurren en los primeros fs después de la ionización. Cuando la biomolécula se encuentra hidratada, la ionización de los mismos orbitales de Kohn-Sham que los propios de la molécula aislada lleva a diferentes patrones de fragmentación. El disolvente representa un papel activo en el proceso ya que no sólo evita la movilidad de la carga en el anillo durante los primeros femtosegundos, sino que también actúa tomando carga positiva de la biomolécula y llevándola a capas de hidratación más externas. Para evaluar los efectos indirectos de la radiación, hemos supuesto también la doble ionización de una molécula de agua vecina del uracilo. Así, ionizaciones tanto de una molécula de agua como del propio uracilo llevan a la formación de oxígeno atómico.

El mismo método ha sido empleado en el estudio de la disociación de agregados de agua por impacto de iones altamente cargados. Nuestros resultados respaldan la alta movilidad

de carga sugerida por los espectros de masa en coincidencia. El modelo basado en TD-DFT MD también ha sido utilizado para estudiar la fragmentación del ion pirmimidina²⁺. Particularmente hemos estudiado dos casos en los que el ion ha sido generado mediante la sustracción de dos electrones pertenecientes a los dos orbitales de valencia más profundos. Ambas dinámicas han dado unos patrones de fragmentación muy diferentes.

Abstract

Nowadays, swift heavy ion beams have become a successfully alternative to traditional X-Ray radiation for cancer treatment. Their high energy loss region (Bragg peak) allows to target the maximum deposited energy at a selected body depth. Unfortunately, formation of the chemical species resulting from the irradiation, which might lead to the cell death, cannot be established by experiments since these processes occur in a very short time scale ($\approx 1 - 100$ fs).

Molecular Dynamics (MD) simulations have demonstrated to be a well suited tool to unravel such mechanisms, however, so far only a few studies have focused on these primary processes. In the present thesis we apply a combination of Time-Dependent Density Functional Theory (TD-DFT) MD and Born-Oppenheimer (BO) MD methods to investigate the fragmentation of doubly-charged uracil, pyrimidine and small water clusters arisen from inner shell ionizations.

The initial doubly charged state is prepared by removing two electrons from a specific Kohn-Sham or Wannier orbital of the neutral species. The electronic density evolves within the so-called Ehrenfest MD approximation where the mean field potential energy surface driving the nuclear dynamics is computed at the TD-DFT level (by propagation of the time-dependent Kohn-Sham equations), and nuclear forces are computed using the Hellmann-Feynman theorem. When the dynamics becomes essentially adiabatic, the system is switched to a ground state BO MD and extended until total times next to ≈ 100 fs.

Results for uracil²⁺ fragmentation in gas phase are in good agreement with ion/ion coincidence measurements performed in uracil collisions with protons of 100 keV. Orbitals of similar energy and/or localized in similar bonds can lead to very different fragmentation patterns, thus showing the importance of the intramolecular chemical environment. The observed fragments almost never correspond to the energetically most favorable dissociation paths, which is due to dynamical effects occurring in the first few fs after electron removal. When the biomolecule is immersed in water, ionization from equivalent Kohn-Sham orbitals than in gas phase lead to different fragmentation patterns. The solvent is revealed as an active participant in the dissociation process since it not only prevents the charge mobility of the biomolecule during the first femtoseconds of dynamics, but also drives electron density from outer hydration shells to the uracil. To evaluate the indirect effect of the radiation, double ionizations of a surrounding water molecule of uracil are also considered. Ionizations either from one water molecule or the uracil lead to the formation of an atomic oxygen.

Same techniques have been also applied to the study of dissociative ionization of water clusters by heavy ion impact. Our results support the high charge mobility suggested by the mass spectrum coincidences. TD-DFT molecular dynamics have been also applied to investigate fragmentation of doubly-charged gas-phase pyrimidine. Dynamics starting from the two deepest inner valence orbitals predict the formation of very different fragmentation patterns.

Contents

I	Introduction	7
1	Introduction	9
1.1	Radiobiological context	9
1.2	Time scales of the biological damage	10
1.3	Motivation	11
1.4	Structure of this thesis	13
II	Methodology	15
2	<i>Ab-initio</i> Molecular Dynamics	17
2.1	Introduction	17
2.2	Schrödinger Equation	17
2.3	Hellmann-Feynman theorem	19
2.3.1	Hellmann-Feynman for time-independent wavefunction	20
2.3.2	Hellmann-Feynman for time-dependent wavefunction	22
2.4	The Born-Oppenheimer approximation	23
2.5	Born-Oppenheimer ansatz and electronic couplings	25
2.6	Born-Oppenheimer molecular dynamics	26
2.6.1	Classical nuclei approximation	27
2.7	Car-Parrinello molecular dynamics	28
2.8	Non-adiabatic dynamics in the Ehrenfest formalism	29
2.8.1	Time-dependent adiabatic couplings, and nuclear forces	31
3	Density Functional Theory	33
3.1	Electronic density and Energy functional	33
3.2	Hohenberg-Kohn theorems	34
3.3	The Kohn-Sham method	35
3.3.1	The exchange-correlation functional	36
3.4	Solutions of the Kohn-Sham equations	39
3.4.1	Plane waves basis sets	43
3.4.2	Pseudopotential Approximation	44
3.4.3	Localized Molecular Orbitals	45
3.5	Nuclear forces in Density Functional Theory	46
3.6	Outline of molecular dynamics simulation based on the DFT formalism	47
3.7	Verlet algorithm	47

4	Basics on Time Dependent Density Functional Theory	49
4.1	Runge-Gross theorem	49
4.2	The Time-Dependent Kohn-Sham Equations	50
4.3	Adiabatic Approximation	51
4.4	Outline of molecular dynamics simulation based on time propagation	52
5	Modelling double ionization	55
5.1	Computational details	55
5.2	Choosing the time-step	57
5.3	Typical CPU times	58
6	Observables	61
6.1	Ensembles	61
6.2	Time averages	62
6.3	Temperature in Molecular Dynamics	63
6.4	Charge analysis on DFT calculations	64
6.4.1	Finding the zero flux surfaces	65
6.4.2	Henkelman's grid-based algorithm	66
6.4.3	Fixing the lattice bias	67
III	Results	69
	Introduction	71
7	Fragmentation of Uracil²⁺ in the gas phase	73
7.1	Introduction	73
7.2	Dynamics of fragmentation	75
7.3	Comparison with experiment	81
8	Double ionization of Uracil in the liquid phase	85
8.1	Introduction	85
8.2	MD simulations subsequent to ionization of uracil	87
8.2.1	Ionization of KS1 orbital of uracil	88
8.2.2	Ionization of KS2 orbital of uracil	89
8.3	MD simulations subsequent to ionization of water	92
8.4	Liquid vs gas phase	95
9	Double ionization of small water clusters	99
9.1	Introduction	99
9.2	Experimental results	100
9.3	TD-DFT MD simulation of (H ₂ O) ₃ ²⁺ and (H ₂ O) ₁₁ ²⁺	103
10	Fragmentation of pyrimidine²⁺ and Auger effect	107
10.1	Introduction	107
10.2	Theoretical calculation of the Auger spectra	108
10.3	TD-DFT MD simulation of pyrimidine ²⁺	112
IV	Conclusions	115
	Conclusiones	117
	Conclusions	119

V	Appendixes	121
A	Appendix	123
A.1	Ehrenfest theorem	123
A.2	The BO ansatz	124
A.3	Semiclassical equations	126
A.4	Hartree-Fock Approximation	127
A.4.1	Variational theorem	127
A.4.2	Slater determinant	127
A.4.3	Hartree-Fock method	128
B	List of Publications	129
	List of figures	130
	List of tables	132
	Bibliography	133

Abbreviations

ALDA Adiabatic Local Density ApproximationMolecular

BO Born-Oppenheimer

DFT Density Functional Theory

KS Kohn-Sham

MD Molecular Dynamics

MO Molecular Orbital

LET Linear Energy Transfer

PBC Periodic Boundary Conditions

PES Potential Energy Surface

ROS Reactive Oxygen Species

TD-DFT Time-Dependent Density Functional Theory

TDSE Time-Dependent Schrödinger Equation

TISE Time-Independent Schrödinger Equation

Part I

Introduction

Chapter 1

Introduction

1.1 Radiobiological context

The aim of this thesis is to use dynamical calculations to interpret, at a molecular level, the different physical and physico-chemical phenomena following the interaction of ionizing radiation (electromagnetic radiation, beams of electrons, protons, or heavy ions) with model bio-molecular systems consisting of small RNA building blocks, either isolated or solvated by water molecules. As a consequence of the interaction, the biomolecule may end up in a highly charged state in which electrons have been extracted from inner orbitals. Our work is focused on the theoretical study of the fragmentation of these multiply charged and excited molecules.

The general interest of this study is to better understand the process of radiation damage. Ionizing radiation can produce damage in biological tissues; these damages are nowadays considered the results of attacks of DNA molecules and are usually classified as either direct or indirect effects [1]. Direct effects are considered when the radiation directly interacts with the DNA molecules; while indirect effects refer to the interaction with the surrounding media, generally water molecules. Radiolysis of water produces reactive oxygen species (ROS) that may further react with DNA. The frontier between direct and indirect effects is however a puzzling question since hydration shells are part of DNA molecules, and as such also govern their conformations [2]. Numerous studies have suggested that ionization of these tightly bound water molecules results in damage to the DNA due to charge transfer, a phenomenon called the quasi-direct effect [3,4]. The fact that no hydroxyl radicals and trapped electrons were detected below about 10 and 21 water molecules per nucleotide, respectively, has been presented as evidence to this effect [5,6]. Moreover, most of the commonly used radiations, like high-energy photons or swift light ions, generate tracks of ionizations where a significant amount of the close-neighbor water molecules are likely to be ionized simultaneously. The extraordinary complexity of a living cell makes the elementary process, ionization of water molecules and direct DNA ionization, difficult to decouple from each other.

Ionizing radiations are nowadays used in cancer therapy to irradiate “risk zones” or tumours. Among them proton- and hadron-therapies have emerged in the past years as promising methods. Swift charged ions (as H^+ or C^{6+}) deposit most of their energy at a specific depth, called the Bragg peak, that depends on the velocities of the ions Fig.(1.1) shows such effect in comparison with the energy deposition of the traditional X-rays beams. Indeed, the energy released by a charged ion in the medium depends to a large extent on the velocity and the charge state of the ions. In the so-called high velocity regime (energy range > 1 MeV) the ion velocity is much larger than the velocity of the electrons of the target:

the dominant phenomenon is the ionization of the target and consequently high energy secondary electrons tracks are produced. As a consequence the ion slows down and when its velocity is of the order of magnitude of the velocity of the electrons of the target molecule, more complex phenomena occur [7,8], with the capture of electrons by the incident ion and the formation of an excited quasi-molecule (with subsequent charge transfers between ion and target). Thus, the energy loss of the ion strongly increases (Bragg peak). For practical purposes, this means that the velocity and the charge of the incident ion can be adjusted to target the adequate energy release for the depth of a tumour, leaving the surrounding healthy tissue less affected.

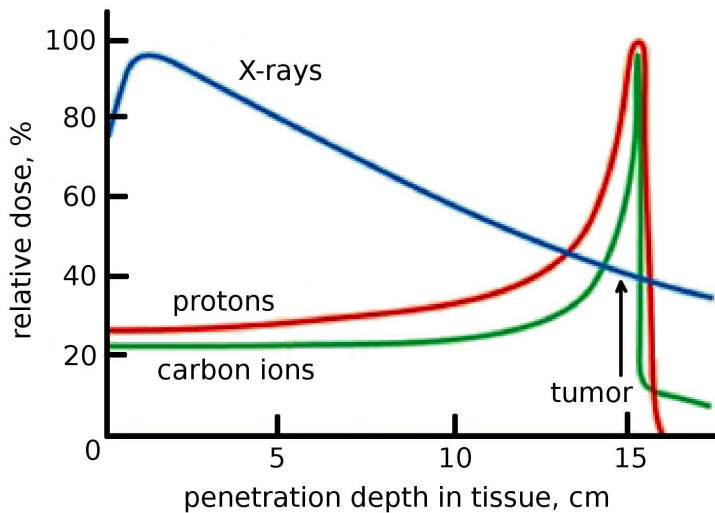


Figure 1.1: X-rays (photons) lose energy rapidly by ionization as they travel through the body. On the other hand, charged particles such as protons and carbon ions deposit most of their energy at a specific depth that depends on their energy (called the Bragg peak). This means that they can deliver a high radiation dose at a tumour site, while sparing the surrounding healthy tissue.

1.2 Time scales of the biological damage

Relating biological damage to primary physical events is difficult because it involves knowledge of the physical, chemical and biological processes that follow cells irradiation. As it is shown in Fig.(1.2), phenomena following irradiation are classified into different time stages: a) The physical stage that lasts only about 10^{-15} s, and during which the ions (and all the secondary electrons that they generate) collides with the biomolecules. In this stage molecules along the ions paths are excited and ionized. b) The physical-chemical stage that gathers the sequence of physical-chemical events that can lead DNA damage. In a time scale from 10^{-15} until 10^{-12} s, the fragmentation of the excited and ionized molecules and the formation of primary radical species occur. These species then diffuse, recombine and react to generate new species during the chemical stage (time scale up to seconds). c) If the cell is damaged very complex enzymatic processes try to fix it. d) If the attempt is unsuccessful, the cell either dies or is left with mutations which may lead to cancer in a major time scale.

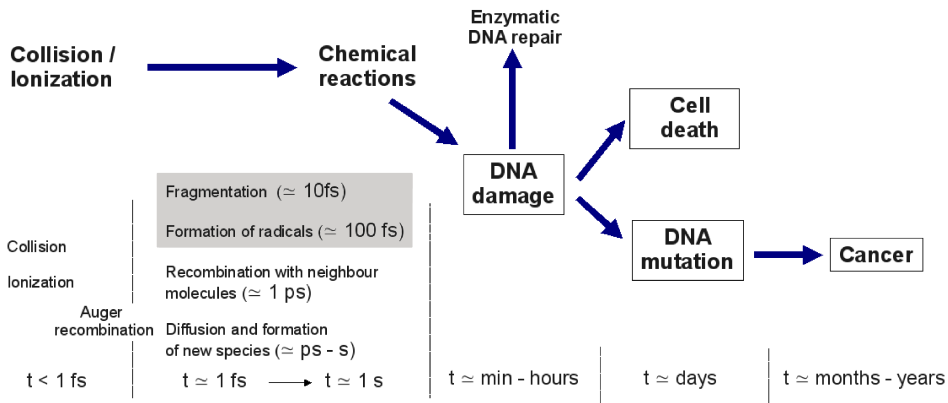


Figure 1.2: Schematic time-diagram of molecular damage. Highlighted in grey the time scale studied.

Regarding this scheme, the present work focuses on the physico-chemical stage for double ionized molecules up to ≈ 100 fs. Our objective is to characterize the chemical products that arise from the ionization of the biomolecule (Direct effect, fragmentation of ionized RNA components) and from ionization of the surrounding water hydration shells (indirect effects).

1.3 Motivation

Establishing the origins of radiation induced DNA damage thus seems an important step towards the application of proton- and hadron- therapies and the investigation of their possible undesirable side effects. Radiation damage of biological tissues has been studied for a long time in order to clarify the mechanisms and the phenomena occurring during the interaction of high-energy radiation with living cells [9,10]: a major goal is the identification of the critical physical events which ultimately lead to the main biological endpoints, i.e. cell inactivation and chromosomal aberrations.

Thus, many experiments have been performed on a mesoscopic level, for example irradiating cells or larger segments of DNA in solution and analyzing the results with biological and chemical methods [11–13]. These studies clearly showed that the most severe consequences of irradiation are due to the damage of DNA, more precisely to the production of single and double strand breaks and in particular to the clustering of these lesions [14].

The hypothesis of a possible biological role of core ionizations in DNA constituent atoms was first raised from studies with ion particles. It was showed that the observed maxima for cell-inactivation cross-sections as a function of linear energy transfer (LET) closely parallel those exhibited by the inner-shell ionization cross-sections calculated for the strategic DNA atoms such as carbon, nitrogen and oxygen [15]. The properties of ultrasoft X-rays to interact with light atoms predominantly through inner-shell ionizations was then used to probe the biological consequences of core ionizations. The range of endpoints covered includes cellular inactivation [11, 16, 17], exchange-type chromosomal aberrations [18, 19], plasmid [11, 20] and cellular [21] DNA double strand break induction and repair. This led to the hypothesis that inner-shell ionizations of DNA atoms trigger a whole set of events, hereafter called “core events”, which induce a class of complex double strand break responsible for cellular inactivation. The overall proposed mechanism was the following: electrons are removed from the biomolecules leaving an electron hole in the core shells. Relaxation of this excited

molecule can occur through an Auger decay, in which an electron from an upper shell decays to the core hole. The energy gained in the process is transferred to another bound electron of the molecule, which is also emitted thus leading to a doubly charged molecule. It is worth mentioning that Auger transition probabilities are nearly 100% in DNA constituent atoms. The lifetime of the Auger decay takes typically a few fs, so most of the molecules are in a double charged state before fragmentation starts.

Thus, the Auger effect leads to the doubly-ionization of the molecule and to the emission of two electrons with a kinetic energy of a few hundreds of eV. Both facts are important to fully understand radiation damage: a) On the one hand these electrons can, in turn, locally damage DNA *via* interactions directly within DNA or indirectly through reactions with radicals produced in the surrounding free water molecules. Furthermore, Boudaiffa et al. [22] showed that already at very low electron beam energies, strand breaks occur due to the process of dissociative electron attachment. b) On the other hand the biomolecule is left in a doubly ionized state with the electrons extracted from inner orbitals. These states have a strongly dissociative character. In particular, they may induce ultra-fast dissociation reactions in which the nuclear dynamics takes place on time scales comparable to the lifetime of the core vacancies.

The exact specificity and mechanisms of DNA damage induction by core events however remain elusive. For instance, 760 eV photons were found more efficient at inducing strand breaks [11, 20] and base damage [23] in dry plasmid DNA films than 350 eV photons. This could be ascribed either to differences in the core-ionized atom (C, N, O or P) and its location relative to DNA (sugar backbone, base or DNA hydration shell) or to differences in the energy of the two electrons subsequently emitted. Experiments have aimed at investigating whether specific fragmentation may result from core hole vacancies in thin dry films of DNA [24] or in DNA components [25–29]. Specific DNA damage, due to oxygen core ionization, and different than that due to secondary electrons, was reported in Ref. [24].

The advent of femtosecond time-resolved laser spectroscopy has enabled the investigation of early charge-transfer processes and elementary radical reactions in molecular liquids and solutions [30]. Gauduel et al. have observed ultrafast one-electron reduction of oxidized pyridine nucleotides [31] and cystamine [32] by prehydrated electrons. These findings and more recent real-time observation of dissociative electron transfer reactions of prehydrated electrons with DNA nucleotides in aqueous solutions reported in Ref. [33], challenge the conventional notion that damage to the genome by ionizing radiation is mainly induced by OH radicals.

There are another mechanisms that can lead to doubly charged molecules. Multiple ionization events are also frequent in swift heavy ions tracks (they represents about 10% of primary events) and may also appear in the Bragg peak region. Experimental [11, 15, 20] and theoretical works [34, 35] support the conclusion that though less probable than other events (for instance electronic excitations, electron attachments), multiple ionizations are systematically more damageable, and therefore more lethal to cells.

All these findings enhance the importance of theoretical works to help obtaining a comprehensive view of the physics and chemistry of the primary events. The ionization process by itself and the subsequent chemical processes occurring at the femtosecond time-scale are pure quantum events and thus should be modelled at the *ab-initio* level, i.e. taking into account the electronic wavefunction as governed by the Schrödinger equation (time-dependent or time-independent). Therefore theoretical simulations based on *ab-initio* molecular dynamics should help to follow the dynamics of a core-ionized biomolecule.

So far, very few theoretical works have focused on these primary processes leading to DNA molecular fragmentation. Most of the quantum chemistry calculations exploring the mechanisms of DNA damage have mainly focused on singly ionized radicals of the DNA bases

in the gas phase. These calculations have characterized geometries and energetics, mostly for the purpose of assessing ionization energies or to identify the most stable fragmentation pathways. Specific processes like the chemical reactivity of OH species on DNA bases in the gas phase [36,37] or hydrogen abstraction have been also studied. Reviews on these calculations can be found in Refs. [38–43].

The present thesis focuses on the study of the fragmentation following double ionization events which result from the relaxation of core hole vacancies after Auger effect or from direct double ionizations.

It is important to note that we do not intend to describe the collisional process (swift ion - biomolecule interaction) or the dynamics involved in the Auger decay. Our aim is to predict the fragmentation pathways once the molecule is in a double ionized and excited state. To describe this process we chose Time-Dependent Density Functional Theory molecular dynamics simulations (TD-DFT MD) in which effective molecular orbitals are propagated in time [44,45]. With that methodology we can define whether the two-electron hole is located in the biomolecule (a RNA component) or in any of the water molecules from the surrounding. We can also control the molecular orbital from which this double ionization proceeds. In order to reach bigger scale of time we have also used Born-Oppenheimer (BO) MD. All our calculations are performed with the electronic representation at the Density Functional Theory (DFT) level as a good compromise between accuracy, size of the simulated system and affordable calculation time (CPU time).

This methodology has been previously checked in the dissociation of doubly ionized water molecule embedded in water. That study allowed to verify that the double ionization of one target water molecule leads to the formation of an atomic oxygen as a direct consequence of the molecule coulomb explosion into its three atomic fragments in less than 4 fs, independently of molecular orbital which is ionized [45]. This study has completed the simulation of the production of HO₂ in liquid water in swift ion track, proving unambiguously that this radical is a product of multiple ionization in water [35].

The molecules selected for our study are uracil, pyrimidine and small water cluster. The first one is a basic components of RNA and pyrimidine is a good model for DNA bases. Both molecules have been widely used as a model system. From the experimental point they have also been extensively studied due to their high stability in gas phase (which prevents, e.g., its thermal degradation). Gas phase ionization by different radiation (photon, electron, ions) has been studied and many data were available in the literature [46–55]. On the other hand, ionization of water clusters is relevant to radiation damage studies because water is the natural environment of biomolecules.

In order to study the effect of the solvent and to have a deeper insight of the importance of direct and indirect effects, uracil fragmentation has been studied in both gas and liquid phase. For the latter case water molecules have been explicitly included in the simulations. Uracil in the gas phase and water clusters have been studied in close collaboration with experimental groups, in these cases we will also provide comparisons with these experiments.

1.4 Structure of this thesis

This dissertation is organized as follows: The second part of the manuscript presents the foundations of the methodology organized in five chapters. The first one, Chapter 2, is devoted to expose the main approaches of the Molecular Dynamics (MD) performed in this thesis, namely: Born-Oppenhemier MD, Car-Parrinello MD and Ehrenfest MD. In the next two ones the implementation of such approximations in DFT and TD-DFT are explained. Chapter 5 will explain the computational details and hypothesis assumed in our simulations. Finally the last chapter of this part (Chapter 6) will explain how the

macroscopical information can be obtained from the theoretical simulations, especially the importance of calculating charges from electronic densities.

The results will be presented in the third part of the thesis. The chapters are based on articles published (or submitted) in international journals of chemistry and physics, and are classified according to the system studied. Thus, Chapters 7 and 8 deal with double ionizations of uracil in the gas and liquid phase respectively; Chapter 9 with double ionization of small water clusters and, finally, Chapter 10 shows the results for double ionizations of pyrimidine. We have extracted the most relevant conclusions in the fourth part of the thesis. Some appendices are also included at the end to give details on mathematical deductions.

Part II

Methodology

Chapter 2

Ab-initio Molecular Dynamics

2.1 Introduction

Molecular Dynamics (MD) simulations has had a long history [56–60] and has evolved into an important and widely used theoretical tool that allow to model different types of systems, including gases, liquids, solids, surfaces and clusters [61–64]. It aims at simulating both nuclei and electrons all simultaneously, although, due to the complex description of their interactions, different approaches have to be considered. The methods can be then classified in several groups: Classical MD simulations based on Newtons equations of movement and model potentials [65]. Amongst them one of the most used is the Force Fields MD [66] in which the nuclei evolve according to an empirical potential designed as close as possible to the real system. Another important sets of simulations are the *Ab-initio* MD [67–69] for which quantum description is used for electrons while the nuclei are kept as classical particles, and Quantum MD simulations that aim to a more complex quantum description of electrons and nuclei [70,71]. The choice of one method over the others mainly concerns the information that we want to extract from the simulation, though in a practical sense computational resources are the main limiting factor.

In this Chapter we will summarize the theoretical concepts behind the MD techniques used in this thesis. All of them belong to the group of *Ab-initio* MD (nuclei treated classically/electrons using quantum mechanics). There are many of such methods that can be classified attending to the approximations made in the description of the electronic wavefunction [72]. The ones used here are included in the so-called *on-the-fly* methods for which the total calculation of the Potential Energy Surface (PES) is avoided. Instead, the surface is reduced only to those nuclear configurations given by the classical trajectory. Namely, we will describe the essentials of Born-Oppenheimer, Ehrenfest and Car-Parrinello approximations first using a wavefunction-based formulation [73,74], and afterwards, we will go through their implementation in the Density Functional Theory (DFT) and Time-Dependent DFT (TD-DFT) approach.

2.2 Schrödinger Equation

We are interested in describing systems formed by electrons and nuclei. According to wave mechanics proposed by E. Schrödinger [75], the physical stage of such systems can be mathematically expressed by a wavefunction $|\Phi(\mathbf{r}, \mathbf{R}, t)\rangle$ where $\mathbf{r}$ and $\mathbf{R}$ represent electronic and nuclear degrees of freedom. This function contains all the information that is possible to

know about the system and its time variation is postulated as follows:

$$i\hbar \frac{\partial}{\partial t} |\Phi(\mathbf{r}, \mathbf{R}, t)\rangle = \hat{H} |\Phi(\mathbf{r}, \mathbf{R}, t)\rangle \quad (2.1)$$

The Eq.(2.1) is known as *Time-dependent Schrödinger Equation* (TDSE), where the non-relativistic Hamiltonian operator $\hat{H}$, given a system of N nuclei and n electrons, is defined by:

$$\begin{aligned} \hat{H} &= \hat{T}_N(\mathbf{R}) + \hat{T}_e(\mathbf{r}) + \hat{V}_{eN}(\mathbf{r}, \mathbf{R}) + \hat{V}_{NN}(\mathbf{R}) + \hat{V}_{ee}(\mathbf{r}) = \\ &= - \sum_{\alpha=1}^N \frac{\hbar^2}{2M_{\alpha}} \nabla_{\alpha}^2 - \sum_{i=1}^n \frac{\hbar^2}{2m_e} \nabla_i^2 - \sum_{\alpha=1}^N \sum_{i=1}^n \frac{Z_{\alpha} q_e^2}{r_{\alpha i}} + \sum_{\alpha=1}^N \sum_{\beta > \alpha}^N \frac{Z_{\alpha} Z_{\beta} q_e^2}{r_{\alpha \beta}} + \sum_{i=1}^n \sum_{j > i}^n \frac{q_e^2}{r_{ij}} \end{aligned} \quad (2.2)$$

where $\nabla^2 = \frac{\partial^2}{\partial x^2} \mathbf{i} + \frac{\partial^2}{\partial y^2} \mathbf{j} + \frac{\partial^2}{\partial z^2} \mathbf{k}$ (cartesian coordinates), subindex α and β denote nuclei, and i and j electrons.

In Eq.(2.2), $\hat{T}_N(\mathbf{R})$ represents the kinetic nuclear operator and $\hat{T}_e(\mathbf{r})$ is the kinetic electronic operator. The third term, $\hat{V}_{eN}(\mathbf{r}, \mathbf{R})$, is the potential nuclear-electron attraction energy operator where $r_{\alpha i}$ is the distance between nucleus α and electron i . The fourth term, $\hat{V}_{NN}(\mathbf{r}, \mathbf{R})$, represents the nuclear-nuclear repulsion operator where $r_{\alpha \beta}$ is the distance between nuclei α and β . Last term $\hat{V}_{ee}(\mathbf{r})$, is the electron-electron repulsion operator where r_{ij} is the distance between electrons i and j .

Equivalent expressions for Eq.(2.2) will be used throughout this text:

$$\hat{H} = \hat{T}_N(\mathbf{R}) + \hat{T}_e(\mathbf{r}) + \hat{V}_{n-e}(\mathbf{r}, \mathbf{R}) = \hat{T}_N(\mathbf{R}) + \hat{H}_e(\mathbf{r}, \mathbf{R}) \quad (2.3)$$

where $\hat{H}_e(\mathbf{r}, \mathbf{R})$ is the electronic Hamiltonian. If there is not any time-dependent external field applied, the molecular Hamiltonian is explicitly time-independent.

A particular solution of the TDSE can be found separating the electronic and nuclear variables from time:

$$|\Phi(\mathbf{r}, \mathbf{R}, t)\rangle = |\psi(\mathbf{r}, \mathbf{R})\rangle |f(t)\rangle \quad (2.4)$$

Replacing Eq.(2.4) in Eq.(2.1) we get:

$$-\frac{\hbar}{i} \frac{1}{|f(t)\rangle} \frac{\partial |f(t)\rangle}{\partial t} = [\hat{T}_N(\mathbf{R}) + \hat{H}_e(\mathbf{r}, \mathbf{R})] |\psi(\mathbf{r}, \mathbf{R})\rangle \quad (2.5)$$

Eq.(2.5) is separable into two independent differential equations. The first one is a time-dependent equation:

$$\frac{d|f(t)\rangle}{|f(t)\rangle} = -\frac{iE}{\hbar} dt \quad (2.6)$$

While the second one is a spatial-dependent eigenvalue equation:

$$[\hat{T}_N(\mathbf{R}) + \hat{H}_e(\mathbf{r}, \mathbf{R})] |\psi(\mathbf{r}, \mathbf{R})\rangle = E |\psi(\mathbf{r}, \mathbf{R})\rangle \quad (2.7)$$

that is called the *Time-independent Schrödinger Equation* (TISE). The constant E is a real number (since $\hat{H}$ is an hermitian operator) that corresponds to the energy of the state represented by the eigenfunction $|\psi(\mathbf{r}, \mathbf{R})\rangle$. Notice that the number of solutions of the TISE

can vary from a discrete number of them for bounded systems to a continuous range in unbounded potentials. Differential equations Eq.(2.6) has a solution:

$$|f(t)\rangle = A \exp\left(\frac{-i}{\hbar}Et\right) \quad (2.8)$$

Then, the total wavefunction solutions of the TDSE can be written as:

$$|\Phi(\mathbf{r}, \mathbf{R}, t)\rangle = |\psi(\mathbf{r}, \mathbf{R})\rangle \exp\left(\frac{-i}{\hbar}Et\right) \quad (2.9)$$

where the amplitude A is included in $|\psi(\mathbf{r}, \mathbf{R})\rangle$. Eq.(2.9) represent the so-called *stationary states* in which the system has a well defined energy and also *probability density* constant on time:

$$\langle\Phi(\mathbf{r}, \mathbf{R}, t)|\Phi(\mathbf{r}, \mathbf{R}, t)\rangle = |\psi(\mathbf{r}, \mathbf{R})|^2 \quad (2.10)$$

It is worth noting that in both TDSE and TISE there is a time-dependence of the total wavefunction. In the later, however, the dependence is given as a wave phase that makes the energy and also the probability density time-independent.

A huge number of chemical problems have been successfully studied relying on these stationary states. Specially, the state that corresponds to the lowest energy in Eq.(2.7) (*ground state*) plays an important role for many chemical and physical processes. On the contrary, if we were interested in cases in which the Hamiltonian depends on time (e.g. a laser source); the TDSE must be solved by using more complex wavefunction that really represents the time dependence. For instance, it can be demonstrated [76] that linear combinations of different stationary solutions:

$$|\Phi(\mathbf{r}, \mathbf{R}, t)\rangle = \sum_{n=0}^{\infty} a_n |\psi_n(\mathbf{r}, \mathbf{R})\rangle \exp\left(\frac{-i}{\hbar}E_n t\right) \quad (2.11)$$

are also solution of the TDSE due to the linearity of the time-dependent equation. If energy values are continuous the summation can be then changed into an integral:

$$|\Phi(\mathbf{r}, \mathbf{R}, t)\rangle = \int_0^{\infty} a(E) |\psi_E(\mathbf{r}, \mathbf{R})\rangle \exp\left(\frac{-i}{\hbar}Et\right) dE \quad (2.12)$$

and we obtain a *wavepacket*, that represents one way to include the nontrivial time dependence of the wavefunction.

The complexity of the total wavefunction, $|\Phi(\mathbf{r}, \mathbf{R}, t)\rangle$, for big molecular systems (far from diatomic molecules) becomes unaffordable for many practical purposes. Mainly, a pure wavefunction involves degrees of freedom of two different natures (nuclear and electrons) and their variation on time is really hard to capture properly with theoretical models. Different approximations have been proposed to simplify the complexity of the wavefunction by imposing simpler expressions.

2.3 Hellmann-Feynman theorem

In *Ab-initio* MD, the nuclear forces are extracted as analytical derivations of the PES generated by the electrons:

$$\mathbf{F}_{\alpha} = -\frac{\partial E}{\partial \mathbf{R}_{\alpha}} \quad (2.13)$$

where $\mathbf{R}_\alpha$ are the nuclear coordinates of a given α -nuclei.

The procedures implemented in the codes nowadays lay on the Hellmann-Feynman theorem that significantly reduces the CPU cost involved in the derivative of the right hand side of Eq.(2.13). Formally, the theorem relates the derivative of the total energy with respect to a parameter, Λ , to the expectation value of the derivative of the Hamiltonian:

$$\frac{\partial E}{\partial \Lambda} = \langle \Phi | \frac{\partial \hat{H}}{\partial \Lambda} | \Phi \rangle \quad (2.14)$$

In the present section we will particularize Eqs.(2.13) and (2.14) to time-independent and time-dependent wavefunctions, following a common route: first, we will consider the quantum definition of the force and after, we will connect this expression to the Λ -derivative of the expectation value of the Hamiltonian via Hellmann-Feynman theorem.

2.3.1 Hellmann-Feynman for time-independent wavefunction

Let's consider the quantum definition of force as the time derivative of the expectation value of the momentum operator [77] applied to a time-independent wavefunction, $|\Phi(\mathbf{r}, \mathbf{R})\rangle$:

$$\begin{aligned} \mathbf{F}_\lambda &= \frac{\hbar}{i} \frac{d}{dt} \langle \frac{\partial}{\partial \lambda} \rangle = \frac{\hbar}{i} \frac{d}{dt} \langle \Phi(\mathbf{r}, \mathbf{R}) | \frac{\partial}{\partial \lambda} | \Phi(\mathbf{r}, \mathbf{R}) \rangle = \\ &= \frac{\hbar}{i} \langle \frac{\partial}{\partial t} \Phi(\mathbf{r}, \mathbf{R}) | \frac{\partial}{\partial \lambda} | \Phi(\mathbf{r}, \mathbf{R}) \rangle + \frac{\hbar}{i} \langle \Phi(\mathbf{r}, \mathbf{R}) | \frac{\partial}{\partial t} \frac{\partial}{\partial \lambda} | \Phi(\mathbf{r}, \mathbf{R}) \rangle + \\ &+ \frac{\hbar}{i} \langle \Phi(\mathbf{r}, \mathbf{R}) | \frac{\partial}{\partial \lambda} | \frac{\partial}{\partial t} \Phi(\mathbf{r}, \mathbf{R}) \rangle \end{aligned} \quad (2.15)$$

where λ can be any electronic or nuclear degree of freedom

Since there is no time-dependency of the momentum operator, Eq.(2.15) is reduced to:

$$\mathbf{F}_\lambda = \frac{\hbar}{i} \langle \frac{\partial}{\partial t} \Phi(\mathbf{r}, \mathbf{R}) | \frac{\partial}{\partial \lambda} | \Phi(\mathbf{r}, \mathbf{R}) \rangle + \frac{\hbar}{i} \langle \Phi(\mathbf{r}, \mathbf{R}) | \frac{\partial}{\partial \lambda} | \frac{\partial}{\partial t} \Phi(\mathbf{r}, \mathbf{R}) \rangle \quad (2.16)$$

If the wavefunction is stationary, i.e. it verifies Eq.(2.9), then:

$$\begin{aligned} \mathbf{F}_\lambda &= \frac{\hbar}{i} \langle \frac{d}{dt} \psi(\mathbf{r}, \mathbf{R}) \exp(\frac{i}{\hbar} Et) | \frac{\partial}{\partial \lambda} | \psi(\mathbf{r}, \mathbf{R}) \exp(\frac{-i}{\hbar} Et) \rangle + \\ &+ \frac{\hbar}{i} \langle \psi(\mathbf{r}, \mathbf{R}) \exp(\frac{i}{\hbar} Et) | \frac{\partial}{\partial \lambda} | \frac{d}{dt} \psi(\mathbf{r}, \mathbf{R}) \exp(\frac{-i}{\hbar} Et) \rangle \\ &= \langle \psi(\mathbf{r}, \mathbf{R}) | E \frac{\partial}{\partial \lambda} | \psi(\mathbf{r}, \mathbf{R}) \rangle - \langle \psi(\mathbf{r}, \mathbf{R}) | \frac{\partial}{\partial \lambda} E | \psi(\mathbf{r}, \mathbf{R}) \rangle \end{aligned} \quad (2.17)$$

Considering the hermiticity of the Hamiltonian:

$$\mathbf{F}_\lambda = \langle \psi(\mathbf{r}, \mathbf{R}) | \hat{H} \frac{\partial}{\partial \lambda} | \psi(\mathbf{r}, \mathbf{R}) \rangle - \langle \psi(\mathbf{r}, \mathbf{R}) | \frac{\partial}{\partial \lambda} \hat{H} | \psi(\mathbf{r}, \mathbf{R}) \rangle = \langle [H, \frac{\partial}{\partial \lambda}] \rangle \quad (2.18)$$

where the term in square brackets corresponds to the commutator of both operators.

By using some useful equivalences from Ref. [78]:

$$\mathbf{F}_\lambda = \langle [\hat{H}, \frac{\partial}{\partial \lambda}] \rangle = -\langle [\frac{\partial}{\partial \lambda}, H] \rangle = -\frac{i}{\hbar} \langle [p, \hat{H}] \rangle = -\langle \Phi(\mathbf{r}, \mathbf{R}) | \frac{\partial}{\partial \lambda} \hat{H} | \Phi(\mathbf{r}, \mathbf{R}) \rangle \quad (2.19)$$

Particularizing Eq.(2.19) to a nuclear position ($\lambda = \alpha$), we get:

$$\mathbf{F}_\alpha = -\langle \Psi(\mathbf{r}, \mathbf{R}) | \nabla_\alpha \hat{H} | \Psi(\mathbf{r}, \mathbf{R}) \rangle \quad (2.20)$$

that represents the force acting on a given nuclei.

Now we want to connect right hand side of Eq.(2.20) to the variation of the energy expected value with any of the nuclear coordinates, $\mathbf{R}_\alpha$. To that end, we need to consider the TISE (2.7) expressed in a mean value form:

$$E = \langle \Psi(\mathbf{r}, \mathbf{R}) | \hat{H} | \Psi(\mathbf{r}, \mathbf{R}) \rangle \quad (2.21)$$

and take its derivative to respect to any α nuclear position:

$$\begin{aligned} \nabla_\alpha E &= \nabla_\alpha \langle \Psi(\mathbf{r}, \mathbf{R}) | \hat{H} | \Psi(\mathbf{r}, \mathbf{R}) \rangle = \\ &= \langle \nabla_\alpha \Psi(\mathbf{r}, \mathbf{R}) | \hat{H} | \Psi(\mathbf{r}, \mathbf{R}) \rangle + \langle \Psi(\mathbf{r}, \mathbf{R}) | \nabla_\alpha \hat{H} | \Psi(\mathbf{r}, \mathbf{R}) \rangle + \langle \Psi(\mathbf{r}, \mathbf{R}) | \hat{H} | \nabla_\alpha \Psi(\mathbf{r}, \mathbf{R}) \rangle = \\ &= E \langle \nabla_\alpha \Psi(\mathbf{r}, \mathbf{R}) | \Psi(\mathbf{r}, \mathbf{R}) \rangle + \langle \Psi(\mathbf{r}, \mathbf{R}) | \nabla_\alpha \hat{H} | \Psi(\mathbf{r}, \mathbf{R}) \rangle + E \langle \Psi(\mathbf{r}, \mathbf{R}) | \nabla_\alpha \Psi(\mathbf{r}, \mathbf{R}) \rangle \end{aligned} \quad (2.22)$$

Since the electronic wavefunction is normalized:

$$\langle \Psi(\mathbf{r}, \mathbf{R}) | \Psi(\mathbf{r}, \mathbf{R}) \rangle = 1 \quad (2.23)$$

the following two expressions are verified:

$$\nabla_\alpha \langle \Psi(\mathbf{r}, \mathbf{R}) | \Psi(\mathbf{r}, \mathbf{R}) \rangle = 0 \quad (2.24)$$

$$\langle \nabla_\alpha \Psi(\mathbf{r}, \mathbf{R}) | \Psi(\mathbf{r}, \mathbf{R}) \rangle = -\langle \Psi(\mathbf{r}, \mathbf{R}) | \nabla_\alpha \Psi(\mathbf{r}, \mathbf{R}) \rangle \quad (2.25)$$

Thus, Eq.(2.22) is reduced to the following expression:

$$\nabla_\alpha E = \langle \Psi(\mathbf{r}, \mathbf{R}) | \nabla_\alpha \hat{H} | \Psi(\mathbf{r}, \mathbf{R}) \rangle \quad (2.26)$$

that represents the Hellmann-Feynman theorem for a time-independent wavefunction. This theorem was proven by many authors (Paul Güttinger [79], Wolfgang Pauli [80], Hans Hellmann [81] and Richard Feynman [82]) independently.

Combining both Eq.(2.20) and Eq.(2.26), finally we obtain the following equality:

$$\mathbf{F}_\alpha = -\langle \Psi(\mathbf{r}, \mathbf{R}) | \nabla_\alpha \hat{H} | \Psi(\mathbf{r}, \mathbf{R}) \rangle = -\nabla_\alpha E \quad (2.27)$$

As we said at the beginning of the section, the practical importance of the Hellmann-Feynman theorem is that energy derivatives are difficult to compute numerically, whereas the expression in Eq.(2.20) can be computed efficiently. For instance, in the molecular problem, only two components of the Hamiltonian contribute to the required derivative—the electron-nucleus and nucleus-nucleus terms:

$$\begin{aligned} \frac{\partial \hat{H}}{\partial \mathbf{R}_\gamma} &= \frac{\partial}{\partial \mathbf{R}_\gamma} \left[-\sum_{\alpha=1}^N \sum_{i=1}^n \frac{Z_\alpha q_e^2}{r_{\alpha i}} + \sum_{\alpha=1}^N \sum_{\beta > \alpha}^N \frac{Z_\alpha Z_\beta q_e^2}{r_{\alpha \beta}} \right] = \\ &= Z_\gamma q_e^2 \sum_{i=1}^n \frac{\mathbf{R}_\gamma - \mathbf{r}_i}{|\mathbf{R}_\gamma - \mathbf{r}_i|^3} - Z_\gamma q_e^2 \sum_{\alpha \neq \gamma}^N Z_\alpha \frac{\mathbf{R}_\gamma - \mathbf{R}_\alpha}{|\mathbf{R}_\gamma - \mathbf{R}_\alpha|^3} \end{aligned} \quad (2.28)$$

where $\mathbf{R}_\gamma$ and $\mathbf{r}_i$ are nuclear and electronic positions. Nevertheless, it was recognized by early computational work using atom-centered basis functions that the Hellmann-Feynman force gave manifestly wrong results [83]. In fact, whenever a variational wavefunction is used:

$$E \langle \Psi(\mathbf{r}, \mathbf{R}) | \Psi(\mathbf{r}, \mathbf{R}) \rangle = \langle \Psi(\mathbf{r}, \mathbf{R}) | \hat{H} | \Psi(\mathbf{r}, \mathbf{R}) \rangle \quad (2.29)$$

where E and $|\Psi(\mathbf{r}, \mathbf{R})\rangle$ now represent the variational energy and the variational (unnormalized) wavefunction respectively. Eq.(2.20) is transformed into the following expression:

$$\begin{aligned} \nabla_\alpha E \langle \Psi(\mathbf{r}, \mathbf{R}) | \Psi(\mathbf{r}, \mathbf{R}) \rangle &= \langle \Psi(\mathbf{r}, \mathbf{R}) | \nabla_\alpha \hat{H} | \Psi(\mathbf{r}, \mathbf{R}) \rangle + \\ &+ \langle \nabla_\alpha \Psi(\mathbf{r}, \mathbf{R}) | \hat{H} - E | \Psi(\mathbf{r}, \mathbf{R}) \rangle + \langle \Psi(\mathbf{r}, \mathbf{R}) | \hat{H} - E | \nabla_\alpha \Psi(\mathbf{r}, \mathbf{R}) \rangle \end{aligned} \quad (2.30)$$

It can be shown that Eq.(2.30) is reduced to the Hellmann-Feynman theorem (2.26) for exact variational wavefunctions (e.g. Hartree-Fock) expanded in a infinite basis set [84]. If finite localized basis sets are used however, second and third term of Eq.(2.30) (known as Pulay forces [83]) have to be evaluated explicitly. On the contrary, using originless basis functions, such as plane waves, Pulay forces vanish exactly since they do not depend on $\mathbf{R}_\alpha$. That applies of course to all *ab-initio* molecular dynamics schemes using that particular basis set scheme.

2.3.2 Hellmann-Feynman for time-dependent wavefunction

We now address the issue of connecting the rigorous force definition to the Hellmann-Feynman force for a normalized time-dependent wavefunction, $|\Phi(\mathbf{r}, \mathbf{R}, t)\rangle$:

$$\mathbf{F}_\lambda = \frac{\hbar}{i} \frac{d}{dt} \left\langle \frac{\partial}{\partial \lambda} \right\rangle = \frac{\hbar}{i} \frac{d}{dt} \langle \Phi(\mathbf{r}, \mathbf{R}, t) | \frac{\partial}{\partial \lambda} | \Phi(\mathbf{r}, \mathbf{R}, t) \rangle \quad (2.31)$$

that satisfies the TDSE (2.1):

$$\begin{aligned} \mathbf{F}_\lambda &= \langle \hat{H} \Phi(\mathbf{r}, \mathbf{R}, t) | \frac{\partial}{\partial \lambda} | \Phi(\mathbf{r}, \mathbf{R}, t) \rangle + \\ &+ \frac{\hbar}{i} \langle \Phi(\mathbf{r}, \mathbf{R}, t) | \frac{\partial}{\partial t} \frac{\partial}{\partial \lambda} | \Phi(\mathbf{r}, \mathbf{R}, t) \rangle - \langle \Phi(\mathbf{r}, \mathbf{R}, t) | \frac{\partial}{\partial \lambda} \hat{H} | \Phi(\mathbf{r}, \mathbf{R}, t) \rangle \end{aligned} \quad (2.32)$$

Eq.(2.31) can be easily related to the Ehrenfest theorem detailed in Appendix A.1:

$$\mathbf{F}_\lambda = \frac{\hbar}{i} \left[\langle \Phi(\mathbf{r}, \mathbf{R}, t) | \frac{\partial}{\partial t} \frac{\partial}{\partial \lambda} | \Phi(\mathbf{r}, \mathbf{R}, t) \rangle + \frac{i}{\hbar} \langle [\hat{H}, \frac{\partial}{\partial \lambda}] \rangle \right] \quad (2.33)$$

in which the first term is equal to zero (the momentum operator does not depend on time). Therefore:

$$\mathbf{F}_\lambda = \langle [\hat{H}, \frac{\partial}{\partial \lambda}] \rangle = -\frac{i}{\hbar} \langle [p, H] \rangle = -\langle \Phi(\mathbf{r}, \mathbf{R}, t) | \frac{\partial}{\partial \lambda} \hat{H} | \Phi(\mathbf{r}, \mathbf{R}, t) \rangle \quad (2.34)$$

that for a particular nuclear position ($\lambda = \alpha$):

$$\mathbf{F}_\alpha = -\langle \Phi(\mathbf{r}, \mathbf{R}, t) | \nabla_\alpha \hat{H} | \Phi(\mathbf{r}, \mathbf{R}, t) \rangle \quad (2.35)$$

Now let's analyze the derivative of the expectation value of the energy to any of the degrees of freedom, λ :

$$\begin{aligned} \frac{\partial}{\partial \lambda} \langle \Phi(\mathbf{r}, \mathbf{R}, t) | \hat{H} | \Phi(\mathbf{r}, \mathbf{R}, t) \rangle &= \\ &= \langle \frac{\partial}{\partial \lambda} \Phi(\mathbf{r}, \mathbf{R}, t) | \hat{H} | \Phi(\mathbf{r}, \mathbf{R}, t) \rangle + \langle \Phi(\mathbf{r}, \mathbf{R}, t) | \frac{\partial}{\partial \lambda} \hat{H} | \Phi(\mathbf{r}, \mathbf{R}, t) \rangle + \\ &+ \langle \Phi(\mathbf{r}, \mathbf{R}, t) | \hat{H} | \frac{\partial}{\partial \lambda} \Phi(\mathbf{r}, \mathbf{R}, t) \rangle \end{aligned} \quad (2.36)$$

Regarding the TISE (2.1) and its conjugate ¹:

$$\begin{aligned}
& \frac{\partial}{\partial \lambda} \langle \Phi(\mathbf{r}, \mathbf{R}, t) | \hat{H} | \Phi(\mathbf{r}, \mathbf{R}, t) \rangle = \\
& = i\hbar \left(\frac{\partial}{\partial \lambda} \Phi(\mathbf{r}, \mathbf{R}, t) \middle| \frac{\partial}{\partial t} \Phi(\mathbf{r}, \mathbf{R}, t) \right) + \langle \Phi(\mathbf{r}, \mathbf{R}, t) | \frac{\partial}{\partial \lambda} \hat{H} | \Phi(\mathbf{r}, \mathbf{R}, t) \rangle - \\
& - i\hbar \left(\frac{\partial}{\partial t} \Phi(\mathbf{r}, \mathbf{R}, t) \middle| \frac{\partial}{\partial \lambda} \Phi(\mathbf{r}, \mathbf{R}, t) \right)
\end{aligned} \tag{2.37}$$

Applying the chain rule to Eq.(2.37):

$$\frac{\partial}{\partial t} = \frac{\partial \lambda}{\partial t} \frac{\partial}{\partial \lambda} \tag{2.38}$$

we obtain:

$$\begin{aligned}
& \frac{\partial}{\partial \lambda} \langle \Phi(\mathbf{r}, \mathbf{R}, t) | \hat{H} | \Phi(\mathbf{r}, \mathbf{R}, t) \rangle = \\
& = i\hbar \left(\frac{\partial}{\partial \lambda} \Phi(\mathbf{r}, \mathbf{R}, t) \middle| \frac{\partial \lambda}{\partial t} \frac{\partial}{\partial \lambda} \Phi(\mathbf{r}, \mathbf{R}, t) \right) + \langle \Phi(\mathbf{r}, \mathbf{R}, t) | \frac{\partial}{\partial \lambda} \hat{H} | \Phi(\mathbf{r}, \mathbf{R}, t) \rangle - \\
& - i\hbar \left(\frac{\partial \lambda}{\partial t} \frac{\partial}{\partial \lambda} \Phi(\mathbf{r}, \mathbf{R}, t) \middle| \frac{\partial}{\partial \lambda} \Phi(\mathbf{r}, \mathbf{R}, t) \right) = \\
& = i\hbar \frac{\partial \lambda}{\partial t} \left[\left\langle \frac{\partial}{\partial \lambda} \Phi(\mathbf{r}, \mathbf{R}, t) \middle| \frac{\partial}{\partial \lambda} \Phi(\mathbf{r}, \mathbf{R}, t) \right\rangle - \left\langle \frac{\partial}{\partial \lambda} \Phi(\mathbf{r}, \mathbf{R}, t) \middle| \frac{\partial}{\partial \lambda} \Phi(\mathbf{r}, \mathbf{R}, t) \right\rangle \right] + \\
& + \langle \Phi(\mathbf{r}, \mathbf{R}, t) | \frac{\partial \hat{H}}{\partial \lambda} | \Phi(\mathbf{r}, \mathbf{R}, t) \rangle
\end{aligned} \tag{2.39}$$

that finally is reduced to the following expression:

$$\frac{\partial}{\partial \lambda} \langle \Phi(\mathbf{r}, \mathbf{R}, t) | \hat{H} | \Phi(\mathbf{r}, \mathbf{R}, t) \rangle = \langle \Phi(\mathbf{r}, \mathbf{R}, t) | \frac{\partial \hat{H}}{\partial \lambda} | \Phi(\mathbf{r}, \mathbf{R}, t) \rangle \tag{2.40}$$

that represents the Hellmann-Feynman theorem for time-dependent wavefunctions.

By combining Eq.(2.35) and Eq.(2.40) we obtain a final expression of the nuclear force:

$$\mathbf{F}_\alpha = -\langle \Phi(\mathbf{r}, \mathbf{R}, t) | \nabla_\alpha \hat{H} | \Phi(\mathbf{r}, \mathbf{R}, t) \rangle = -\nabla_\alpha E \tag{2.41}$$

For a variational wavefunction a more complex expression of forces, similar to Eq.(2.30), can be also derived. The resulting Pulay forces are zero if the basis functions do not depend on the nuclear coordinates (see Ref. [85] for further details).

In conclusion, we have demonstrated that forces acting on a nuclei can be obtained by deriving the mean value of the electronic energy to the selected nuclear coordinate. This statement is valid for both time-dependent and time-independent wavefunctions whenever a delocalized basis set is used.

2.4 The Born-Oppenheimer approximation

Regarding the TISE (2.7), solutions to this equation cannot be expressed as a simple separation of electronic and nuclear variables, $|\Psi_k(\mathbf{r})\rangle|\chi_N(\mathbf{R})\rangle$, since the molecular Hamiltonian operator has a crossed term, $\hat{V}_{eN}(\mathbf{r}, \mathbf{R})$, that depends on the two sets of coordinates. Nevertheless, due to the mass difference between the electron and nuclei, $m/M_\alpha \sim 1/1000$, the

¹ $-i\hbar \partial/\partial t \langle \Phi(\mathbf{r}, \mathbf{R}, t) | = \hat{H} \langle \Phi(\mathbf{r}, \mathbf{R}, t) |$

nuclei move very slowly in comparison with electrons and they can be considered as nearly fixed to respect to electron movement. Therefore, if nuclei are considered fixed, nuclear coordinates act as parameters. Under this approximation the total spatial function might be expressed as a product of a electronic wavefunction and a nuclear wavefunction:

$$|\psi(\mathbf{r}, \mathbf{R})\rangle = |\Psi_k(\mathbf{r}; \mathbf{R})\rangle |\chi_N(\mathbf{R})\rangle \quad (2.42)$$

Where the electronic wavefunction has a parametric dependency on nuclear positions expressed by the semicolon. Wavefunctions like Eq.(2.42) are solution of the TISE (2.7), i.e.:

$$\hat{H}|\Psi_k(\mathbf{r}; \mathbf{R})\rangle |\chi_N(\mathbf{R})\rangle = E_k |\Psi_k(\mathbf{r}; \mathbf{R})\rangle |\chi_N(\mathbf{R})\rangle \quad (2.43)$$

Or equivalently:

$$\begin{aligned} \{\hat{T}_N(\mathbf{R}) + \hat{T}_e(\mathbf{r}) + \hat{V}_{eN}(\mathbf{r}; \mathbf{R}) + \hat{V}_{NN}(\mathbf{R}) + \hat{V}_{ee}(\mathbf{r})\} |\Psi_k(\mathbf{r}; \mathbf{R})\rangle |\chi_N(\mathbf{R})\rangle = \\ = E_k |\Psi_k(\mathbf{r}; \mathbf{R})\rangle |\chi_N(\mathbf{R})\rangle \end{aligned} \quad (2.44)$$

If the kinetic energy of the nuclei is neglected we arrive to the electronic equation:

$$\hat{H}_e(\mathbf{r}; \mathbf{R}) |\Psi_k(\mathbf{r}; \mathbf{R})\rangle = E_k(\mathbf{R}) |\Psi_k(\mathbf{r}; \mathbf{R})\rangle \quad (2.45)$$

If we consider now the nuclear movement, going back to Eq.(2.44) the nuclear kinetic term is given by:

$$\begin{aligned} \hat{T}_N(\mathbf{R}) |\Psi_k(\mathbf{r}; \mathbf{R})\rangle |\chi_N(\mathbf{R})\rangle = \\ = - \sum_{\alpha=1}^N \frac{\hbar^2}{2M_\alpha} |\Psi_k(\mathbf{r}; \mathbf{R})\rangle \nabla_\alpha^2 |\chi_N(\mathbf{R})\rangle - \sum_{\alpha=1}^N \frac{\hbar^2}{M_\alpha} \nabla_\alpha |\Psi_k(\mathbf{r}; \mathbf{R})\rangle \nabla_\alpha |\chi_N(\mathbf{R})\rangle - \\ - \sum_{\alpha=1}^N \frac{\hbar^2}{2M_\alpha} |\chi_N(\mathbf{R})\rangle \nabla_\alpha^2 |\Psi_k(\mathbf{r}; \mathbf{R})\rangle \end{aligned} \quad (2.46)$$

Since $\hat{T}_e(\mathbf{r})$ contains no $\mathbf{R}$ dependence:

$$\hat{T}_e(\mathbf{r}) |\chi_N(\mathbf{R})\rangle |\Psi_k(\mathbf{r}; \mathbf{R})\rangle = |\chi_N(\mathbf{R})\rangle \hat{T}_e(\mathbf{r}) |\Psi_k(\mathbf{r}; \mathbf{R})\rangle \quad (2.47)$$

and considering the parametric dependence:

$$\hat{V}_{eN} |\Psi_k(\mathbf{r}; \mathbf{R})\rangle |\chi_N(\mathbf{R})\rangle = |\chi_N(\mathbf{R})\rangle \hat{V}_{eN} |\Psi_k(\mathbf{r}; \mathbf{R})\rangle \quad (2.48)$$

If we rearrange terms:

$$\begin{aligned} - \sum_{\alpha=1}^N \frac{\hbar^2}{2M_\alpha} |\Psi_k(\mathbf{r}; \mathbf{R})\rangle \nabla_\alpha^2 |\chi_N(\mathbf{R})\rangle - \sum_{\alpha=1}^N \frac{\hbar^2}{M_\alpha} \nabla_\alpha |\Psi_k(\mathbf{r}; \mathbf{R})\rangle \nabla_\alpha |\chi_N(\mathbf{R})\rangle - \\ - \sum_{\alpha=1}^N \frac{\hbar^2}{2M_\alpha} |\chi_N(\mathbf{R})\rangle \nabla_\alpha^2 |\Psi_k(\mathbf{r}; \mathbf{R})\rangle + |\chi_N(\mathbf{R})\rangle \hat{T}_e |\Psi_k(\mathbf{r}; \mathbf{R})\rangle + \\ + |\chi_N(\mathbf{R})\rangle \hat{V}_{eN}(\mathbf{R}) |\Psi_k(\mathbf{r}; \mathbf{R})\rangle + |\Psi_k(\mathbf{r}; \mathbf{R})\rangle \hat{V}_{NN}(\mathbf{R}) |\chi_N(\mathbf{R})\rangle + \\ + |\chi_N(\mathbf{R})\rangle \hat{V}_{ee}(\mathbf{R}) |\Psi_k(\mathbf{r}; \mathbf{R})\rangle = E_k |\Psi_k(\mathbf{r}; \mathbf{R})\rangle |\chi_N(\mathbf{R})\rangle \end{aligned} \quad (2.49)$$

We can see that a typical contribution of the third term, in atomic units, has the form $1/(2M_\alpha)\nabla_\alpha^2|\Psi_k(\mathbf{r};\mathbf{R})\rangle$ [86]. The value of $\nabla_\alpha|\Psi_k(\mathbf{r};\mathbf{R})\rangle$ is of the same order as $\nabla_i|\Psi_k(\mathbf{r};\mathbf{R})\rangle$ since the derivatives operate over approximately the same dimensions. The latter is $|\Psi_k(\mathbf{r};\mathbf{R})\rangle p_e$, where p_e is the momentum of an electron. Therefore $1/(2M_\alpha)\nabla_\alpha^2|\Psi_k(\mathbf{r};\mathbf{R})\rangle \approx p_e^2/(2M_\alpha) = (m_e/M_\alpha)T_e$. Since $m/M_\alpha \sim 1/1000$, the third term can be dropped, and also the second since it is even much smaller. Ignoring both, we get:

$$|\Psi_k(\mathbf{r};\mathbf{R})\rangle\hat{T}_N(\mathbf{R})|\chi_N(\mathbf{R})\rangle + |\chi_N(\mathbf{R})\rangle\hat{T}_e(\mathbf{r})|\Psi_k(\mathbf{r};\mathbf{R})\rangle + |\Psi_k(\mathbf{r};\mathbf{R})\rangle\hat{V}_{eN}(\mathbf{r},\mathbf{R})|\chi_N(\mathbf{R})\rangle + |\Psi_k(\mathbf{r};\mathbf{R})\rangle\hat{V}_{NN}(\mathbf{R})|\chi_N(\mathbf{R})\rangle + |\chi_N(\mathbf{R})\rangle\hat{V}_{ee}(\mathbf{r})|\Psi_k(\mathbf{r};\mathbf{R})\rangle \approx E|\Psi_k(\mathbf{r};\mathbf{R})\rangle|\chi_N(\mathbf{R})\rangle \quad (2.50)$$

Thus, regarding Eq.(2.45), Eq.(2.50) can be re-written as:

$$\left(\hat{T}_N + \langle\Psi_k(\mathbf{r};\mathbf{R})|\hat{H}_e(\mathbf{r};\mathbf{R})|\Psi_k(\mathbf{r};\mathbf{R})\rangle\right)|\chi_N(\mathbf{R})\rangle \approx E_k|\chi_N(\mathbf{R})\rangle \quad (2.51)$$

Due to the mass difference the molecular problem is translated to a new picture where the nuclei are moving on a PES generated by the electrons. These electrons instantaneously adjust their wavefunction according to the nuclear one. The electronic level to which corresponds the electronic wavefunction is expressed by the subindex “ k ”. The number of electronic states is not unique for every molecule, in fact k can run from a discrete number to infinite. Linear combinations of these states can be also proposed as a solution of the TISE leading to coupling between states. If we constrain the electrons to be in only one of them (*adiabatic approximation*) the problem of coupling between them is eliminated since all coupling terms go to zero. In this respect, M. Born and R. Oppenheimer in 1927 [87] studied, by a perturbative analysis of a time-independent Hamiltonian, the influence of the nuclei on the electronic wavefunction in the mass ratio $k = (m_e/M)^{1/4}$, where M is the average mass weight of the molecule. They concluded that electrons do not undergo transitions between states under *appropriate* conditions.

2.5 Born-Oppenheimer ansatz and electronic couplings

Let us suppose the solutions of Eq.(2.45) are known and assume that the set of eigenvalues of $\hat{H}_e(\mathbf{r},\mathbf{R})$ is a finite number. If the eigenfunctions are orthonormalized:

$$\langle\Psi_k(\mathbf{r};\mathbf{R})|\Psi_l(\mathbf{r};\mathbf{R})\rangle = \delta_{k,l} \quad (2.52)$$

the total time-dependent wavefunction $|\Phi(\mathbf{r},\mathbf{R},t)\rangle$ can be expanded in terms of the eigenfunctions of $\hat{H}_e$ since these form a complete set:

$$|\Phi(\mathbf{r},\mathbf{R},t)\rangle = \sum_{l=0}^{\infty} |\Psi_l(\mathbf{r};\mathbf{R})\rangle|\chi_l(\mathbf{R},t)\rangle \quad (2.53)$$

This product constitutes the so-called BornOppenheimer ansatz [78,88,89] and represents the simplest way to include the time variable into the total wavefunction. Insertion of this ansatz into the time-dependent Schrödinger Eq.(2.1) followed by multiplication from the left by $\langle\Psi_k(\mathbf{r};\mathbf{R})|$ leads to a set of coupled differential equations (see Appendix A.2 for details):

$$i\hbar\frac{\partial|\chi_k(\mathbf{R},t)\rangle}{\partial t} = [\hat{T}_N(\mathbf{R}) + E_k(\mathbf{R})]|\chi_k(\mathbf{R},t)\rangle + \sum_{l=0}^{\infty} C_{kl}|\chi_l(\mathbf{R},t)\rangle \quad k = 0, 1, 2, \dots, \infty \quad (2.54)$$

where the exact coupling operator C_{kl} is defined as:

$$C_{kl} = \langle\Psi_k(\mathbf{r};\mathbf{R})|\hat{T}_N(\mathbf{R})|\Psi_l(\mathbf{r};\mathbf{R})\rangle - \sum_{\alpha=1}^N \frac{\hbar^2}{M_\alpha} \{ \langle\Psi_k(\mathbf{r};\mathbf{R})|\nabla_\alpha|\Psi_l(\mathbf{r};\mathbf{R})\rangle \} \nabla_\alpha \quad (2.55)$$

The first term is a matrix element of the kinetic energy operator of the nuclei, whereas the second term depends on their momenta and is zero when the electronic wavefunction is real. An alternative notation can be introduced:

$$\mathbf{d}_{kl}^\alpha(\mathbf{R}) = \langle \Psi_k(\mathbf{r}; \mathbf{R}) | \nabla_\alpha | \Psi_l(\mathbf{r}; \mathbf{R}) \rangle \quad (2.56)$$

$$\mathbf{D}_{kl}^\alpha(\mathbf{R}) = -\langle \Psi_k(\mathbf{r}; \mathbf{R}) | \nabla_\alpha^2 | \Psi_l(\mathbf{r}; \mathbf{R}) \rangle \quad (2.57)$$

where the terms $\mathbf{d}_{kl}(\mathbf{R})$ and $\mathbf{D}_{kl}(\mathbf{R})$ are called the *first* and *second derivative couplings* respectively. We also refer $\mathbf{d}_{kl}(\mathbf{R})$ as the *nonadiabatic coupling* since, as it can be seen in Eq.(2.55), the diagonal contribution of C_{kk} depends only on a single adiabatic wavefunction $|\Psi_k(\mathbf{r}; \mathbf{R})\rangle$:

$$C_{kk} = \langle \Psi_k(\mathbf{r}; \mathbf{R}) | \hat{T}_N(\mathbf{R}) | \Psi_k(\mathbf{r}; \mathbf{R}) \rangle \quad (2.58)$$

and corresponds to a correction to the (adiabatic) eigenvalue E_k of the Schrödinger equation (2.45). The *adiabatic approximation* consists in taking into account only this diagonal elements, reducing the set of Eq.(2.54) to:

$$[\hat{T}_N(\mathbf{R}) + E_k(\mathbf{R}) + C_{kk}(\mathbf{R})]|\chi_k(\mathbf{R}, t)\rangle = i\hbar \frac{\partial |\chi_k(\mathbf{R}, t)\rangle}{\partial t} \quad k = 0, 1, 2, \dots, \infty \quad (2.59)$$

Neglecting the off-diagonal terms we are assuming that the motion of the nuclei proceeds without changing the quantum state “ k ” of the electronic subsystem during time evolution. Consequently, the wavefunction Eq.(2.53) is reduced to a single product term:

$$|\Phi(\mathbf{r}, \mathbf{R}, t)\rangle \approx |\Psi_k(\mathbf{r}; \mathbf{R})\rangle |\chi_k(\mathbf{R}, t)\rangle \quad (2.60)$$

The final simplification consists in neglecting also the coupling term:

$$[\hat{T}_N(\mathbf{R}) + E_k(\mathbf{R})]|\chi_k(\mathbf{R}, t)\rangle = i\hbar \frac{\partial |\chi_k(\mathbf{R}, t)\rangle}{\partial t} \quad (2.61)$$

which leads back to the Born-Oppenheimer approximation.

2.6 Born-Oppenheimer molecular dynamics

In this section we will derive the semiclassical (or quantum-classical) molecular dynamics from Quantum Mechanics in the particular case of the Born-Oppenheimer approximation. The idea of semiclassical dynamics is to keep the electrons as a quantum objects while treating the atoms as classical bodies. The formal deduction will be done following the route of Tully [74], which starts by rewriting the corresponding nuclear wavefunction in the following manner [78, 90]:

$$|\chi_k(\mathbf{R}, t)\rangle = A_k(\mathbf{R}, t) \exp\left(\frac{i}{\hbar} S_k(\mathbf{R}, t)\right) \quad (2.62)$$

which can always be done for any complex function depending on $\mathbf{R}$ and t . $A_k(\mathbf{R}, t)$ and $S_k(\mathbf{R}, t)$ are the amplitude factor and phase, both considered to be real and $A_k(\mathbf{R}, t) > 0$, noticing that $A_k(\mathbf{R}, t)^2$ corresponds to the particle probability density $|\chi(\mathbf{R}, t)|^2 \equiv \rho_k(\mathbf{R}, t)$.

Some algebra is needed in order to introduce Eq.(2.62) into the time-dependent Eq.(2.61) (see Appendix A.3 for further details). Once the real and imaginary part are separated, we

obtain the following two expressions:

$$\frac{\partial S_k(\mathbf{R}, t)}{\partial t} + \sum_{\alpha=1}^N \frac{1}{2M_\alpha} (\nabla_\alpha S_k(\mathbf{R}, t))^2 + E_k(\mathbf{R}) = \sum_{\alpha=1}^N \frac{\hbar^2}{2M_\alpha} \frac{\nabla_\alpha^2 A_k(\mathbf{R}, t)}{A_k(\mathbf{R}, t)} \quad (2.63)$$

$$\frac{\partial A_k(\mathbf{R}, t)}{\partial t} + \sum_{\alpha=1}^N \frac{1}{M_\alpha} (\nabla_\alpha A_k(\mathbf{R}, t)) (\nabla_\alpha S_k(\mathbf{R}, t)) + \sum_{\alpha=1}^N \frac{1}{2M_\alpha} A_k(\mathbf{R}, t) (\nabla_\alpha^2 S_k(\mathbf{R}, t)) = 0 \quad (2.64)$$

that are equivalent to the TISE and closely related with Bohmian (fluid) dynamics [91].

Multiplying Eq.(2.64) by $2A_k(\mathbf{R}, t)$ we can obtain an expression related with the probability flux, $\mathbf{J}_{k,\alpha}(\mathbf{R}, t)$:

$$\frac{\partial \rho_k(\mathbf{R}, t)}{\partial t} + \sum_{\alpha=1}^N \nabla_\alpha \mathbf{J}_{k,\alpha}(\mathbf{R}, t) = 0 \quad (2.65)$$

where:

$$\mathbf{J}_{k,\alpha}(\mathbf{R}, t) = \frac{\rho_k(\mathbf{R}, t) (\nabla_\alpha^2 S_k(\mathbf{R}, t))}{M_\alpha} \quad (2.66)$$

We see that the spatial variation of the phase of the wavefunction characterizes the probability flux; the stronger the phase variation, the more intense the flux.

On the other hand, we can also define a velocity field given by:

$$\mathbf{v}_\alpha = \frac{\mathbf{J}_{k,\alpha}(\mathbf{R}, t)}{\mathbf{p}_\alpha} = \frac{S_k(\mathbf{R}, t)}{M_\alpha} \quad (2.67)$$

where $\mathbf{p}_\alpha$ is the linear momentum of particle α .

2.6.1 Classical nuclei approximation

Heisenberg's Uncertainty Principle establishes that it is not possible to know at the same time two conjugate magnitudes with the same level of accuracy. The precision in one of them runs in favor of the imprecision in the other and vice versa. The error of both measurements is related with the value of $\hbar$. This natural impossibility spoils the idea of classical (deterministic) trajectories and is the underlying origin of the probability character of Quantum Mechanics. However, Classical Mechanics can be derived from Quantum Mechanics imposing the limit $\hbar \rightarrow 0$, known as the *classical limit* [92].

In this limit, the right hand side term of Eq.(2.63) can be neglected since it explicitly depends on $\hbar$. Doing so, we get a similar equation to the Hamilton-Jacobi equation in Classical Mechanics [93]:

$$\frac{\partial S_k(\mathbf{R}, t)}{\partial t} + \sum_{\alpha=1}^N \frac{1}{2M_\alpha} (\nabla_\alpha S_k(\mathbf{R}, t))^2 + E_k(\mathbf{R}) = 0 \quad (2.68)$$

We could consider that, in the classical approximation, $|\chi_k(\mathbf{R}, t)\rangle$ describes a fluid of non-interacting classical particles of mass m subject to the potential $E_k(\mathbf{R})$. We can transform Eq.(2.68) using the equivalences of Eq.(2.67) into:

$$\frac{\partial S_k(\mathbf{R}, t)}{\partial t} + \sum_{\alpha=1}^N \frac{1}{2} M_\alpha \mathbf{v}_\alpha^2 + E_k(\mathbf{R}) = 0 \quad (2.69)$$

Taking the gradient in both sides:

$$\nabla \frac{\partial S_k(\mathbf{R}, t)}{\partial t} + \sum_{\alpha=1}^N \frac{1}{2} \nabla_{\alpha} M_{\alpha} \mathbf{v}_{\alpha}^2 = -\nabla E_k(\mathbf{R}) \quad (2.70)$$

Plus some algebra:

$$\frac{\partial M_{\alpha} \mathbf{v}_{\alpha}}{\partial t} + \sum_{\alpha=1}^N \mathbf{v}_{k,\alpha} M_{\alpha} \nabla_{\alpha} \mathbf{v}_{\alpha} = -\nabla E_k(\mathbf{R}) \quad (2.71)$$

Lastly, if we assume an incompressible flow:

$$\nabla_{\alpha} \cdot \mathbf{v}_{\alpha} = 0 \quad (2.72)$$

the second term on the left hand side of Eq.(2.71) vanish, and then we obtain a set of Newton equations:

$$M_{\alpha} \ddot{\mathbf{R}}_{\alpha} = -\nabla_{\alpha} E_k^{BO}(\mathbf{R}) \quad \alpha = 1, 2, 3, \dots, N \quad (2.73)$$

separately for each decoupled electronic state k .

Reading Eq.(2.73), the nuclei move according to classical mechanics in an effective potential, E_k^{BO} . Thus, the force acting on the nuclei α is obtained by the derivation of the PES that comes from solving the simultaneous TISE (2.45) of clamped nuclei for the “kth” state at the given nuclear configuration $\{\mathbf{R}(t)\}$. Applying the Hellmann-Feynman theorem (2.26) we get a final expression of the force acting on each nuclei:

$$M_{\alpha} \ddot{\mathbf{R}}_{\alpha} = -\langle \Psi_k(\mathbf{r}; \mathbf{R}) | \nabla_{\alpha} \hat{H}_e(\mathbf{r}, \mathbf{R}) | \Psi_k(\mathbf{r}; \mathbf{R}) \rangle \quad \alpha = 1, 2, 3, \dots, N \quad (2.74)$$

2.7 Car-Parrinello molecular dynamics

Aiming to cut down the computational expenses of molecular dynamics, Car and Parrinello [94] introduced a non-obvious approach in which the electronic wavefunctions are accounted as dynamical variables. It can be considered to combine the advantages of both Ehrenfest and Born-Oppenheimer molecular dynamics in an optimal way.

Briefly, in Ehrenfest dynamics the time-step to integrate the nuclear and electronic movement (given by Eq.(2.88) and Eq.(2.89) in next section) is dictated by the later, i.e., by the intrinsic dynamics of the electrons (much faster than nuclear motion). Contrary to that, there is no electron dynamics involved in solving the BO equations of motion, Eq.(2.45) and Eq.(2.74), because the electronic problem is treated within the TISE. This implies that these equations of motion can be integrated on the time scale given by nuclear motion, thus allows us to use a larger molecular dynamics time-step. However, this means that the electronic structure problem has to be solved self-consistently at each molecular dynamics step, whereas this is avoided in Ehrenfest dynamics due to the possibility of propagating the wave function simply by applying the Hamiltonian to an initial wave function (obtained by a single self-consistent optimization at the very beginning of such a simulation).

In this sense, Car-Parrinello (CP) approach integrates the equations of motion on the (long) time scale set by the nuclear motion, and takes advantage of the smooth time evolution of the dynamically evolving electronic subsystem (avoiding the minimization each time-step of dynamic). A Lagrangian is defined for the electronic system as follows:

$$\mathcal{L} = \sum_i \mu \langle \dot{\psi}_i | \dot{\psi}_i \rangle - E[\{\psi_i\}, \{\mathbf{R}_I\}, \{\alpha_n\}] \quad (2.75)$$

where μ is a fictitious mass associated with the electronic wavefunctions, E is the Kohn-Sham energy functional, $\mathbf{R}_I$ is the position of the ion I , and α_n define the size and shape of the unit cell. The kinetic-energy term in the Lagrangian is due to the fictitious dynamics of the electronic degrees of freedom. The Kohn-Sham energy functional takes the place of the potential energy in a conventional Lagrangian formulation. The electronic wavefunctions are subject to the constraints of orthonormality:

$$\langle \psi_i(\mathbf{r}) | \psi_j(\mathbf{r}) \rangle = \delta_{ij} \quad (2.76)$$

The corresponding Newtonian equations of motion are obtained from the associated Euler-Lagrange equations [93] but in terms of nuclear positions and orbitals:

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{R}}_I} = \frac{\partial \mathcal{L}}{\partial \mathbf{R}_I} \quad (2.77)$$

$$\frac{d}{dt} \frac{\delta \mathcal{L}}{\delta \dot{\psi}_i^*} = \frac{\delta \mathcal{L}}{\delta \psi_i^*} \quad (2.78)$$

The second equation involves functional derivatives because the orbitals are continuous scalar fields. In practice, however, these fields are represented in a basis (e.g. on a discrete real-space mesh or in terms of plane wave components), and the concept of functional derivation reduces to the partial derivation with respect to the components of the field in the basis. By performing the operations indicated above, we arrive at the Car-Parrinello equations of motion:

$$M_I \ddot{\mathbf{R}}_I(t) = \frac{\partial}{\partial \mathbf{R}_I} \langle \Psi_0(\mathbf{r}) | \hat{H}_e(\mathbf{r}, \mathbf{R}) | \Psi_0(\mathbf{r}) \rangle + \frac{\partial}{\partial \mathbf{R}_I} \{constrains\} \quad (2.79)$$

$$\mu \ddot{\psi}_i(t) = \frac{\delta}{\delta \psi_i^*} \langle \Psi_0(\mathbf{r}) | \hat{H}_e(\mathbf{r}, \mathbf{R}) | \Psi_0(\mathbf{r}) \rangle + \frac{\delta}{\delta \psi_i^*} \{constrains\} \quad (2.80)$$

These equations can be particularized to one particle Hamiltonians such as those resulting from Kohn-Sham theory (described in next section). The proper orbital orthonormality Eq.(2.76) must be imposed by Lagrange multipliers Λ_{ij} :

$$M_I \ddot{\mathbf{R}}_I(t) = -\nabla_I \langle \Psi_0(\mathbf{r}) | \hat{H}_e^{KS}(\mathbf{r}, \mathbf{R}) | \Psi_0(\mathbf{r}) \rangle \quad (2.81)$$

$$\mu \ddot{\psi}_i(t) = -\hat{H}_e^{KS}(\mathbf{r}, \mathbf{R}) |\psi_i(\mathbf{r})\rangle + \sum_j \Lambda_{ij} |\psi_j(\mathbf{r})\rangle \quad (2.82)$$

2.8 Non-adiabatic dynamics in the Ehrenfest formalism

The Born-Oppenheimer approximation can be applied for a big number of systems however the non-adiabaticity is presented in many others (charge transfer, photoisomerization reactions, etc.) in which the change in the nuclear configurations might vary the adiabatic states populations. Such a distortion of the electron cloud will, in turn, influence the nuclear trajectory. To introduce the electronic-nuclear interaction the quantum-mechanical time evolution of the electrons must be maintained. The simplest way of doing this is given by the so-called *Mean field method* or *Ehrenfest method*. It starts by using the simplest possible product ansatz that regards the separation between the nuclear and electronic contributions to the total wavefunction [74, 95]:

$$|\Phi(\mathbf{r}, \mathbf{R}, t)\rangle \approx |\Psi(\mathbf{r}, t)\rangle |\chi(\mathbf{R}, t)\rangle \exp\left[-\frac{i}{\hbar} \int E_k(\mathbf{R}) dt\right] \quad (2.83)$$

where the electronic and nuclear wavefunctions are separately normalized to unity at every instant of time, and the exponent is a convenient phase factor chosen in such way that the final expression is easy to handle [73].

Substituting Eq.(2.83) above into TDSE (2.1), we obtain two expressions:

$$i\hbar \frac{\partial |\Psi(\mathbf{r}, t)\rangle}{\partial t} = - \sum_{i=1}^n \frac{\hbar^2}{2m_e} \nabla_i^2 |\Psi(\mathbf{r}, t)\rangle + \{ \langle \chi(\mathbf{R}, t) | \hat{V}_{n-e}(\mathbf{r}, \mathbf{R}) | \chi(\mathbf{R}, t) \rangle \} |\Psi(\mathbf{r}, t)\rangle \quad (2.84)$$

$$i\hbar \frac{\partial |\chi(\mathbf{R}, t)\rangle}{\partial t} = - \sum_{\alpha=1}^N \frac{\hbar^2}{2M_\alpha} \nabla_\alpha^2 |\chi(\mathbf{R}, t)\rangle + \{ \langle \Psi(\mathbf{r}, t) | \hat{H}_e(\mathbf{r}, \mathbf{R}) | \Psi(\mathbf{r}, t) \rangle \} |\chi(\mathbf{R}, t)\rangle \quad (2.85)$$

that are the basics of the Time-dependent self-consistent field (TDSCF) introduced by Dirac [96] in 1930. First applications to atomic and molecular system came, however, much later in the 70-80's [97–103]. Noticing that in this scheme both electrons and nuclei move quantum-mechanically in time-dependent effective potentials (given by the expressions in the braces). The assumption of a single-determinant ansatz like Eq.(2.83) produces an anticipated mean-field description of the coupled nuclear-electronic dynamics.

Using the same equation for the nuclear wavefunction (2.62) (and following almost the same algebra as in Appendix A.3), we can obtain one expression in terms of phase from Eq.(2.85):

$$\begin{aligned} \frac{\partial S_k(\mathbf{R}, t)}{\partial t} + \sum_{\alpha=1}^N \frac{1}{2M_\alpha} (\nabla_\alpha S_k(\mathbf{R}, t))^2 + \langle \Psi(\mathbf{r}, t) | \hat{H}_e(\mathbf{r}, \mathbf{R}) | \Psi(\mathbf{r}, t) \rangle = \\ = \sum_{\alpha=1}^N \frac{\hbar^2}{2M_\alpha} \frac{\nabla_\alpha^2 A_k(\mathbf{R}, t)}{A_k(\mathbf{R}, t)} \end{aligned} \quad (2.86)$$

The classical limit can be also introduced here as we did in Sec.(2.6.1), obtaining a final expression:

$$M_\alpha \ddot{\mathbf{R}}_\alpha(t) = -\nabla_\alpha V^E(\mathbf{R}(t)) = -\nabla_\alpha \langle \Psi(\mathbf{r}, \mathbf{R}) | \hat{H}_e(\mathbf{r}, \mathbf{R}) | \Psi(\mathbf{r}, \mathbf{R}) \rangle \quad \alpha = 1, 2, 3, \dots, N \quad (2.87)$$

where V^E is often called the Ehrenfest potential. Noticing that $\mathbf{R}(t)$ represents the classical-mechanical nuclear positions and not the quantum ones (given by the full nuclear wavefunction $|\chi(\mathbf{R}, t)\rangle$). If we use the HF theorem for time-dependent wavefunctions introduced in Sec.(2.3.2):

$$M_\alpha \ddot{\mathbf{R}}_\alpha(t) = -\langle \Psi(\mathbf{r}, \mathbf{R}) | \nabla_\alpha \hat{H}_e(\mathbf{r}, \mathbf{R}) | \Psi(\mathbf{r}, \mathbf{R}) \rangle \quad \alpha = 1, 2, 3, \dots, N \quad (2.88)$$

we get an easier expression to implement for the nuclear force acting on nuclei α .

In the classical limit, TDSCF Eq.(2.84) can be reduced to:

$$i\hbar \frac{\partial}{\partial t} |\Psi_k(\mathbf{r}, \mathbf{R})\rangle = \hat{H}_e(\mathbf{r}, \mathbf{R}) |\Psi_k(\mathbf{r}, \mathbf{R})\rangle \quad (2.89)$$

if the quantum nuclear positions are replaced by the classical positions (see Ref. [104]).

The Ehrenfest molecular dynamics that relies on solving Newtons equation for the nuclei, Eq.(2.88), simultaneously with Schrödinger equations for the electrons, Eq.(2.89), is often called Ehrenfest MD in honor of Paul Ehrenfest who was the first to address the essential question of how Newtonian classical dynamics of point particles can be derived from Schrödinger time-dependent wave equation [105].

2.8.1 Time-dependent adiabatic couplings, and nuclear forces

The total wavefunction given by Eq.(2.83) can be written as a linear combination of adiabatic eigenfunctions:

$$|\Psi(\mathbf{r}, t)\rangle = \sum_{l=0}^{\infty} a_l |\Psi_l(\mathbf{r}; \mathbf{R})\rangle \exp\left[-\frac{i}{\hbar} \int E_k(\mathbf{R}) dt\right] \quad (2.90)$$

that are solutions of the TISE. Insertion of this ansatz into the time-dependent electronic Eq. (2.89) followed by multiplication from the left by $\langle \Psi_l(\mathbf{r}; \mathbf{R}) |$ and integration over the electronic coordinates leads to a set of coupled differential equations:

$$\frac{da_k}{dt} = \sum_{l=0}^{\infty} a_l C_{kl} \exp\left[-\frac{i}{\hbar} \int (E_l(\mathbf{R}) - E_k(\mathbf{R})) dt\right] \quad (2.91)$$

where:

$$C_{kl} = \langle \Psi_l(\mathbf{r}; \mathbf{R}) | \frac{\partial}{\partial t} | \Psi_l(\mathbf{r}; \mathbf{R}) \rangle \quad (2.92)$$

are the nonadiabatic coupling elements, that can be related with the nonadiabatic coupling vector by using the chain rule:

$$C_{kl} = \frac{d\mathbf{R}}{dt} \langle \Psi_k(\mathbf{r}; \mathbf{R}) | \nabla | \Psi_l(\mathbf{r}; \mathbf{R}) \rangle = \frac{d\mathbf{R}}{dt} \mathbf{d}_{kl} \quad (2.93)$$

Integration of Eq.(2.91) yields the expansion coefficients $a_k(t)$ whose square modulus, $|a_k(t)|^2$, can be interpreted as the probability of finding the system in the adiabatic state k at time t .

Using the definition of the Ehrenfest potential and inserting Eq.(2.90) is quite straightforward to deduce that:

$$V^E(\mathbf{R}(t)) = \langle \Psi(\mathbf{r}, \mathbf{R}) | \hat{H}_e(\mathbf{r}, \mathbf{R}) | \Psi(\mathbf{r}, \mathbf{R}) \rangle = |a_k(t)|^2 \sum_{k=0}^{\infty} E_k(\mathbf{R}) \quad (2.94)$$

meaning that the atoms evolve on an effective potential representing an average over the adiabatic states weighted by their state populations $|a_k(t)|^2$.

Following the Ref. [106], nuclear forces of Eq.(2.88) can be also expressed in terms of probabilities:

$$\mathbf{F}_\alpha = - \sum_{k=0}^{\infty} |a_k(t)|^2 \nabla_\alpha E_k(\mathbf{R}) + \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} a_k(t)^* a_l(t) (E_l(\mathbf{R}) - E_k(\mathbf{R})) d_{lk}^\alpha \quad (2.95)$$

The Ehrenfest method has been applied with great success to a number of chemical problems including energy transfer at metal surfaces, the study of excited state lifetimes of organic molecules and currents in nanowires [107–109]. The classical description of the heavy ions provides an intuitive way to understand their motion and is readily handled using molecular dynamics. Another advantage is that the electrons can be treated accurately using a single-particle approach, such as TD-DFT theory [44, 110–113]. However, due to its mean-field character the method has some serious limitations. A system that was initially prepared in a pure adiabatic state will be in a mixed state when leaving the region of strong nonadiabatic coupling [106].

Chapter 3

Density Functional Theory

3.1 Electronic density and Energy functional

Density functional theory (DFT) is based on the electron probability density function (or electron density function) $n(\mathbf{r})$. Given a normalized electronic wavefunction $|\Psi(\mathbf{r})\rangle$, the probability of finding an electron in the space-spin volume element $d\mathbf{r}_1$ located in $\mathbf{r}_1$ while simultaneously another electron is in $d\mathbf{r}_2$ at $\mathbf{r}_2$ and so on, is given by the product $\Psi^*(\mathbf{r}_1 \dots \mathbf{r}_n) \Psi(\mathbf{r}_1 \dots \mathbf{r}_n) d\mathbf{r}_1 \dots d\mathbf{r}_n$. If we are interested only in probability of finding one electron in $d\mathbf{r}_1$ at $\mathbf{r}_1$, independent of where the other electrons are located, then we must average over all space-spin coordinates of the other electrons. Thus the electron density for a single electron $n(\mathbf{r})$ is defined as the following multiple integral¹:

$$n(\mathbf{r}) = N \int \dots \int |\Psi(\mathbf{r}_1, \mathbf{r}_2 \dots \mathbf{r}_N)|^2 d\mathbf{r}_2 d\mathbf{r}_3 \dots d\mathbf{r}_N \quad (3.1)$$

As the electrons are indistinguishable, the integral of $n(\mathbf{r})$ over the volume element $d\mathbf{r}$ determines the probability of finding any of the electrons within that volume. If the wavefunction is normalized, the integral of the density function over the whole space is:

$$\int n(\mathbf{r}) d\tau = N \quad (3.2)$$

where N is the total number of electrons.

Unlike the wavefunction, the electronic density is measurable (i.e. by X-ray diffraction or electron diffraction) and also is a non-negative function of only three spatial variables which vanishes at infinity. No matter how big the molecule may be, the electron density remains a function of three variables while the complexity of the wavefunction is increased with the number of electrons (exactly it depends on $6n$ spin and spatial variables). Another advantage is that in about the same time needed for a Hartree-Fock calculation, DFT can often obtain results about the same quality as MP2 calculations [114]. As we will see later, the disadvantages of using DFT is the unknown form of the *exchange-correlation functional*, E_{XC} [115]. Approximations are necessary to build up this term, but unlike the wavefunction methods where the solutions can be improved by going to bigger basis sets or higher correlation levels, in DFT there is no way to systematically improve the functional. The choice of the functional really holds the success or the failure of the calculation.

¹In this chapter the bracket notation will be exchanged for the traditional integral notation.

The average electronic energy can be expressed as a functional of the ground state electron density:

$$E[n] = T[n] + V_{Ne}[n] + V_{ee}[n] \quad (3.3)$$

where it has been separated into three functionals: One associated with the average kinetic electron energy, one with the average nuclear-electron attraction energy and another with the average electron-electron repulsion energy.

For any closed-shell system:

$$\hat{V}_{Ne} = - \sum_{\alpha=1}^N \sum_{i=1}^n \frac{Z_{\alpha}}{r_{\alpha i}} = \sum_{i=1}^n \nu(r_i) \quad (3.4)$$

where $\nu(r_i)$ is the *external potential* created by all the nuclei over the i^{th} -electron. It can be shown [116] that the functional density form of the nuclear-electron potential is known:

$$V_{Ne}[n] = \int n(\mathbf{r})\nu(\mathbf{r})d\mathbf{r} \quad (3.5)$$

The Hohenberg-Kohn functional $F_{HK}[n]$ is defined as:

$$F_{HK}[n] = T[n] + V_{ee}[n] \quad (3.6)$$

Thus substituting equations (3.5) and (3.6) into equation (3.3) we get:

$$E[n] = \int n(\mathbf{r})\nu(\mathbf{r})d\mathbf{r} + F_{HK}[n] \quad (3.7)$$

3.2 Hohenberg-Kohn theorems

The Hohenberg-Kohn (HK) theorems [117,118] are the base of the density functional theory and can be summarized as follows:

1) For an external potential $\nu(\mathbf{r})$ there is a unique ground-state non-degenerate wavefunction $|\Psi_0(\mathbf{r})\rangle$ that gives rise to a unique ground state density $n_0(\mathbf{r})$. The first HK theorem states that this mapping can be done backwards too, that is, for any v-representable electronic density² $n(\mathbf{r})$, the external potential which gives rise to it, is unique. This can be expressed by the following bijective map:

$$\begin{array}{ccc} \nu(\mathbf{r}) & \rightarrow & |\Psi_0(\mathbf{r})\rangle \rightarrow n_0(\mathbf{r}) \\ & \longleftarrow & \\ & & HK \text{ theorem} \end{array} \quad (3.8)$$

All the observable properties of a molecule in a ground electronic state are then determined by the ground-state electron density function $n_0(\mathbf{r})$. For instance, the energy E_0 can be represented by:

$$E_0 = E_0[n_0] \quad (3.9)$$

2) The second Hohenberg-Kohn theorem says that any trial electron density function $n(\mathbf{r}_{prob})$ will give an energy higher than or equal to (if it were the exact electron density function) the true ground state energy:

$$E[n_{prob}] \geq E_0[n_0] \quad (3.10)$$

²A electron density is v-representable only if it is the ground state of some external potential.

Similar to the variational theorem for wavefunction-based methods (Appendix A.4.1), the second Hohenberg-Kohn theorem assures that we can keep choosing different densities and those that provide lower energies will be closer to the correct.

3.3 The Kohn-Sham method

Following the first Hohenberg-Kohn theorem we can rewrite Eq.(3.7) as follows:

$$E[n_0] = \int n_0(\mathbf{r})\nu(\mathbf{r})d\mathbf{r} + F_{HK}[n_0] \quad (3.11)$$

showing the relation between the energy and the ground-state electron density. As it is suggested by the second HK theorem, finding the ground-state can be done by minimizing this expression. Unfortunately, the process is not simple because we do not know the relationship between $F_{HK}[n_0]$ and the density. The key point of the Kohn-Sham method is to consider an auxiliary system of N non-interacting electrons for estimating the kinetic energy of the real (interacting) system [118–121].

The Hamiltonian (in atomic units) for n non-interacting electrons moving under an external potential $\nu_s(\mathbf{r})$ is given by:

$$\hat{H}_s = \sum_{i=1}^n \hat{h}_s(\mathbf{i}) = -\frac{1}{2} \sum_{i=1}^n \nabla^2(\mathbf{i}) + \sum_{i=1}^n \nu_s(\mathbf{i}) \quad (3.12)$$

And the wavefunction is a single Slater determinant (see appendix A.4.2):

$$|\Psi_s(\mathbf{r})\rangle = (n!)^{-1/2} |\psi_1(\mathbf{r})\bar{\psi}_1(\mathbf{r})\psi_2(\mathbf{r})\bar{\psi}_2(\mathbf{r})\dots\psi_N(\mathbf{r})\bar{\psi}_N(\mathbf{r})\rangle \quad (3.13)$$

Each $|\psi_i(\mathbf{r})\rangle$ represents a Kohn-Sham spin-orbital. If we suppose that the system is closed-shell, then the density will be:

$$n_s(\mathbf{r}) = \sum_{i=1}^{n/2} |\psi_i(\mathbf{r})|^2 \quad (3.14)$$

where the subindex i runs only over the occupied orbitals. At the same time, the kinetic energy functional of the system is given by:

$$T_s[n_s] = - \sum_{i=1}^{n/2} \langle \psi_i(\mathbf{r}) | \frac{1}{2} \nabla^2 | \psi_i(\mathbf{r}) \rangle \quad (3.15)$$

We can impose the condition that Kohn-Sham orbitals will be chosen in such a way that its ground state electron density $n_s(\mathbf{r})$ is exactly the same as the real ground state of the non-interacting system: $n_s(\mathbf{r}) = n_0(\mathbf{r})$ (hereafter simply $n(\mathbf{r})$). For the real system, we can rewrite the expression of the energy (3.3) in terms of quantities of the reference non-interacting system:

$$E[n] = T_s[n] + V_{Ne}[n] + V_{ee}[n] + \Delta T[n] + \Delta V_{ee}[n] \quad (3.16)$$

The first and the second term represent the kinetic energy and the nuclear-electron interaction for a non-interacting system, given by equations (3.15) and (3.5) respectively. The third term is the classical electrostatic repulsion energy term, also called Hartree term:

$$V_{ee}[n] = \frac{1}{2} \int \int \frac{n(\mathbf{r}_1)n(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_1 d\mathbf{r}_2 \quad (3.17)$$

If we knew the set of Kohn-Sham spin-orbitals $\{|\psi_i\rangle\}$ that determinate the ground state density $n(\mathbf{r})$, then the value of these three functionals are readily calculated. By contrast, the fourth and fifth terms of the equation (3.16) involve unknown quantities: $\Delta T[n]$ includes all the deviations of the real kinetic functional from the reference system, while $\Delta V_{ee}[n]$ includes the corrections to electron classical self-interaction (for instance, in Eq.(3.17) any of the electrons is forced to repel itself).

Both corrections are grouped within the exchange-correlation functional, E_{XC} :

$$E_{XC}[n] = \Delta T[n] + \Delta V_{ee}[n] \quad (3.18)$$

Substituting the equations (3.5), (3.17) and (3.18) into the equation (3.16) leads to the *Kohn-Sham energy equation*:

$$E[n] = T_s[n] + \int n(\mathbf{r})\nu(\mathbf{r})d\mathbf{r} + \frac{1}{2} \iint \frac{n(\mathbf{r}_1)n(\mathbf{r}_2)}{r_{12}}d\mathbf{r}_1d\mathbf{r}_2 + E_{XC}[n] \quad (3.19)$$

that is exact if the exchange-correlation energy functional $E_{XC}[n]$ is known.

3.3.1 The exchange-correlation functional

As it has been previously mentioned, the “main problem of DFT” is the unknown exchange-correlation energy functional $E_{XC}[n]$. The dependence of this functional on the electron density is normally expressed as an interaction between the electron density $n(\mathbf{r})$ and “energy density” functional $\epsilon_{XC}[n]$:

$$E_{XC}[n] = \int n(\mathbf{r})\epsilon_{XC}[n]d\mathbf{r} \quad (3.20)$$

Taking the expression (3.20) and deriving it with respect to electron density, we obtain the exchange correlation potential $\nu_{XC}(\mathbf{r})$:

$$\nu_{XC}(\mathbf{r}) = \frac{\partial E_{XC}[n]}{\partial n} = \epsilon_{XC}[n] + n(\mathbf{r})\frac{\partial \epsilon_{XC}[n]}{\partial n} \quad (3.21)$$

Many approximations for $E_{XC}[n(\mathbf{r})]$ have been proposed. In the following sections, we will revise a few of them focusing the attention on BLYP functional, which is the one used in this thesis.

Local density approximation (LDA)

The local density approximation [122] represents the simplest approximation to $E_{XC}[n]$. It derives from analysis of the uniform electron gas assuming that the local exchange functional can be derived from the exact solution of this model. In this approximation the energy density functional $\epsilon_{XC}[n(\mathbf{r})]$ is always treated as a sum of individual exchange and correlation contributions:

$$\epsilon_{XC}[n] = \epsilon_X[n] + \epsilon_C[n] \quad (3.22)$$

The first term of equation (3.22) is:

$$\epsilon_X[n] = -\frac{3}{4} \left(\frac{3}{\pi} \right)^{\frac{1}{3}} (n(\mathbf{r}))^{\frac{1}{3}} \quad (3.23)$$

If the previous expression is employed without any correlation functional it is known as “ X_α method” and was proposed by Slater [123]. Vosko, Wilk and Nusair (VWN) obtained an expression for the correlation energy $\epsilon_C[n(\mathbf{r})]$ of a uniform electron gas using Monte Carlo calculations [124]. Already simple approximations such as the local density approximation (LDA) have turned out to be quite successful in the calculation of structures and total energies of molecules and solids [125, 126].

Local spin density approximation (LSDA)

Local spin density approximation is suited to treat the open shell systems or molecules next to dissociation. While in the LDA two paired electrons with opposite spin have the same spatial Kohn-Sham orbital $|\psi(\mathbf{r})\rangle$, in the LSDA they occupy different spatial orbitals: $|\psi_\alpha(\mathbf{r})\rangle$ and $|\psi_\beta(\mathbf{r})\rangle$ for spin up and spin down respectively. Thus α and β densities are separately minimized:

$$n_r(\mathbf{r}) = n_\alpha^r(\mathbf{r}) + n_\beta^r(\mathbf{r}) = \sum_{i=1}^{N_\alpha} |\psi_{\alpha i}(\mathbf{r})|^2 + \sum_{i=1}^{N_\beta} |\psi_{\beta i}(\mathbf{r})|^2 \quad (3.24)$$

Thus we define two different exchange-correlation functionals:

$$E_{XC}^{LSDA}[n_\alpha] = E_X^{LSDA}[n_\alpha] + E_C^{LSDA}[n_\alpha, n_\beta] \quad (3.25)$$

$$E_{XC}^{LSDA}[n_\beta] = E_X^{LSDA}[n_\beta] + E_C^{LSDA}[n_\alpha, n_\beta] \quad (3.26)$$

Like in LDA the exchange-correlation functional $E_{XC}^{LDA}[n]$ and its derivate $\nu_{XC}^{LDA}(\mathbf{r})$, $E_{XC}^{LSDA}[n_\alpha]$ and $E_{XC}^{LSDA}[n_\beta]$ and their functional derivatives can be accurately calculated [120, 127].

Generalized gradient approximation (GGA)

In the local density approximation (LDA) the functional used to obtain the exchange correlation energy only depends on the electron density. An improvement of this method is to consider exchange correlation functionals that not only involve the electron density $n(\mathbf{r})$ but also its local gradients (first derivative with respect to position).

The introduction of gradient correction terms to the LDA functional by Perdew and Yang [128] lead to the *Generalized Gradient Approximation* (GGA) and the *gradient-corrected* functionals. These functionals are also known as *nonlocal* functionals, in contrast to the local LDA and LSDA functionals, since the gradient is the change over an infinitesimal distance beyond the “local” point of the coordinates $\mathbf{r}$. One of the earliest and most popular GGA exchange functionals was proposed by A. D. Becke (B or B88) as a correction to the LSDA exchange energy [129]:

$$\epsilon_X^{B88}[n] = \epsilon_X^{LDA}[n] + \Delta\epsilon_X^{B88}[n] \quad (3.27)$$

where:

$$\Delta\epsilon_X^{B88}[n] = -\beta n(\mathbf{r})^{1/3} \frac{x^2}{1 + 6\beta x \sinh^{-1}x} \quad (3.28)$$

and:

$$x = \frac{|\nabla n(\mathbf{r})|}{n(\mathbf{r})^{4/3}} \quad (3.29)$$

Among other GGA exchange functionals it is also worth to mention the G96 [130] and PBE [131, 132] functionals.

There have been several GGA functionals proposed for the correlation energy. One popular functional is due to Lee, Yang and Parr [133] (LYP) composed by a density gradient expansion, based on the orbital functional for the correlation energy of Colle and Salvetti [134] (closed-shell expression):

$$\begin{aligned} \epsilon_C^{LYP}[n] = & -a(F'n(\mathbf{r}) + F) - abC_F n(\mathbf{r})^{5/3}(G'n(\mathbf{r}) + \frac{8}{3}G) - \\ & - \frac{ab}{4}[G''n(\mathbf{r})|\nabla n(\mathbf{r})|^2 + G'(3|\nabla n(\mathbf{r})|^2 + 2n(\mathbf{r})\nabla^2 n(\mathbf{r})) + 4G\nabla^2 n(\mathbf{r})] - \\ & - \frac{ab}{72}[3G''n(\mathbf{r})|\nabla n(\mathbf{r})|^2 + G'(5|\nabla n(\mathbf{r})|^2 + 6n(\mathbf{r})\nabla^2 n(\mathbf{r})) + 4G\nabla^2 n(\mathbf{r})] \end{aligned} \quad (3.30)$$

where:

$$G[n] = F[n]n(\mathbf{r})^{-5/3}\exp\left(-cn(\mathbf{r})^{-1/3}\right) \quad \text{and} \quad F[n] = \frac{1}{1 + dn(\mathbf{r})^{-1/3}} \quad (3.31)$$

G' and G'' refer to first and second derivatives of the $G[n]$ functional to respect to $n(\mathbf{r})$. The same stands for F' and F'' for $F[n]$ functional. The a , b , c and d parameters are determined by fitting to data for the helium atom. Natural constant C_F is given by:

$$C_F = \frac{3}{10}(3\pi^2)^{2/3} \quad (3.32)$$

and related with the Tomas-Fermi kinetic energy density.

The combination of the exchange correction of Becke in Eq.(3.27) and the LYP correlation potential given by Eq.(3.30) leads to the BLYP functional used in the MD simulation of the present work.

The GGA functionals represent a noticeable improvement upon LDA especially for molecular systems but also work very well for atoms and solids [135]. They can in particular reproduce well hydrogen bonded system like liquid water [136].

Hybrid functionals: B3LYP

The hybrid functionals [137] incorporate a fraction of Hartree-Fock exchange (see Appendix A.4.3):

$$\epsilon_X^{HF} = - \sum_{i=1}^n \sum_{j=1}^n \langle \psi_i(\mathbf{1})\psi_j(\mathbf{2}) | \frac{1}{r_{ij}} | \psi_j(\mathbf{1})\psi_i(\mathbf{2}) \rangle \quad (3.33)$$

where the summations run over the n occupied KS orbitals.

For an infinite expansion basis set, ϵ_X^{HF} is the exact exchange energy for a system of noninteracting electrons with electron density equals to the real system. Including the Hartree-Fock exchange energy (3.33) in to a LSDA gradient-corrected DFT expression gives a HF/DFT exchange-correlation functional, commonly called a *hybrid* DFT functional.

The most popular hybrid functional is based on an exchange-energy functional developed by Becke [138], and modified by Stevens et al. [139] by introduction of the LYP [133] correlation-energy functional. This exchange-correlation functional, called the Becke3LYP or B3LYP functional is:

$$\epsilon_{XC}^{B3LYP} = (1 - a - b)\epsilon_X^{LSDA} + a\epsilon_X^{HF} + b\epsilon_X^{B88} + (1 - c)\epsilon_C^{VWN} + c\epsilon_C^{LYP} \quad (3.34)$$

where a , b and c (usually 0.20, 0.72 and 0.81) are those that give the best fit of the calculated energy to molecular atomization energies, ionization potentials, electron affinities and total

atomic energies. ϵ_X^{LSDA} is the LSDA exchange energy functional and ϵ_X^{HF} is the corresponding one to the Hartree-Fock the given by equation (3.33). Likewise, ϵ_X^{B88} is the B88 exchange functional mentioned above and either E_C^{VWN} as E_C^{LYP} are correlation energy functionals. The former is the correlation energy from the Vosko, Wilk and Nusair functional [124], and the later is the Lee, Yang and Parr correlation functional.

3.4 Solutions of the Kohn-Sham equations

After discussing the foundations of DFT, we will shortly describe methods for solving the Kohn-Sham problem. There are basically two methods to find the ground-state density within the Kohn-Sham framework for a given external potential [69]. The first one focuses on the self-consistent resolution of the equation obtained from the derivation of the energy functional (3.19) with respect to a small change in the electron density. The second one aims at directly minimizing Eq.(3.19) while imposing some constraints.

Diagonalization technique

From the second HK theorem we can extract a condition for an electronic density to minimize the energy functional:

$$\delta E_v[n] = \delta T_s[n] + \int d\mathbf{r} \left(\nu(\mathbf{r}) + \frac{\partial V_{ee}[n]}{\partial n(\mathbf{r})} + \nu_{XC}(\mathbf{r}) \right) \delta n(\mathbf{r}) = 0 \quad (3.35)$$

Where:

$$\frac{\partial V_{ee}[n]}{\partial n(\mathbf{r})} = \int \frac{n(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_2 = \nu_J(\mathbf{r}) \quad (3.36)$$

is the Hartree potential. The variation of the kinetic functional, $\delta T_s[n]$, can be expressed in terms of the external potential of the non-interacting system:

$$\delta T_s[n] = - \int d\mathbf{r} \nu_s(\mathbf{r}) \delta n(\mathbf{r}) \quad (3.37)$$

Introducing Eq.(3.37) into Eq.(3.35) we get:

$$\delta E_v[n] = - \int d\mathbf{r} \nu_s(\mathbf{r}) \delta n(\mathbf{r}) + \int d\mathbf{r} (\nu(\mathbf{r}) + \nu_J(\mathbf{r}) + \nu_{XC}(\mathbf{r})) \delta n(\mathbf{r}) = 0 \quad (3.38)$$

The condition for $n(\mathbf{r})$ to be the minimum of $E_v[n]$ is that:

$$\int d\mathbf{r} \nu_s(\mathbf{r}) \delta n(\mathbf{r}) = \int d\mathbf{r} (\nu(\mathbf{r}) + \nu_J(\mathbf{r}) + \nu_{XC}(\mathbf{r})) \delta n(\mathbf{r}) \quad (3.39)$$

or equally, we can choose $\nu_s(\mathbf{r}) = \nu_{KS}(\mathbf{r})$ in such a way that:

$$\nu_{KS}(\mathbf{r}) = \nu(\mathbf{r}) + \nu_J(\mathbf{r}) + \nu_{XC}(\mathbf{r}) \quad (3.40)$$

where $\nu_{KS}(\mathbf{r})$ is the so-called *Kohn-Sham potential* or *effective potential*. Thus Schrödinger equation for the set of non-interacting electrons moving under this external potential is given by:

$$-\frac{1}{2} \nabla_i^2 |\psi_i(\mathbf{r})\rangle + \nu_{KS}(\mathbf{r}) |\psi_i(\mathbf{r})\rangle = \epsilon_i |\psi_i(\mathbf{r})\rangle \quad i = 1, 2 \dots n \quad (3.41)$$

where i runs to the number of electrons. This set of integro-differential equations known as *Kohn-Sham equations*.

Thus, the problem of finding a ground-state electronic density $n(\mathbf{r})$ that minimizes the energy functional (3.19) has been changed into finding the external potential of a non-interacting system that satisfies Eq.(3.40). These are the basis of the diagonalization technique that resolves the problem recursively:

1. The procedure starts from choosing a basis set, an initial molecular geometry and an exchange-correlation functional. After, a guess density $n(\mathbf{r}) = n_{old}(\mathbf{r})$ and the associated Kohn-Sham potential is constructed:

$$\nu_{KS}^{old}(\mathbf{r}) = \nu(\mathbf{r}) + \int \frac{n^{old}(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_2 + \nu_{XC}(\mathbf{r}; n^{old}) \quad (3.42)$$

The Kohn-Sham equations for the non-interacting electrons is then given by:

$$-\frac{1}{2}\nabla_i^2|\psi_i^{new}(\mathbf{r})\rangle + \nu_{KS}^{old}(\mathbf{r})|\psi_i^{new}(\mathbf{r})\rangle = \epsilon_i^{new}|\psi_i^{new}(\mathbf{r})\rangle \quad i = 1, 2...n \quad (3.43)$$

If we define the the Kohn-Sham operator $\hat{H}_{KS}$ as:

$$\hat{H}_{KS}(\mathbf{i}) = -\frac{1}{2}\nabla^2(\mathbf{i}) + \nu_{KS}(\mathbf{r}) \quad (3.44)$$

then Eq.(3.43) can be rewritten as:

$$\hat{H}_{KS}^{old}(\mathbf{i})|\psi_i^{new}(\mathbf{r})\rangle = \epsilon_i^{new}|\psi_i^{new}(\mathbf{r})\rangle \quad i = 1, 2...n \quad (3.45)$$

2. The following step is to expand the electronic orbitals into a basis set:

$$|\psi_i(\mathbf{r})\rangle = \sum_{n=1}^m c_{ni}|\phi_n(\mathbf{r})\rangle \quad (3.46)$$

that allow us to write the set of equations (3.45) as the following matrix equation:

$$\mathbf{H}^{KS}\mathbf{C} = \mathbf{S}\mathbf{C}\epsilon \quad (3.47)$$

where F_{ij}^{KS} and S_{ij} are the elements of the Kohn-Sham matrix and overlap matrix respectively defined as:

$$H_{ij}^{KS} = \int \phi_n^*(\mathbf{1})\hat{H}_{KS}(\mathbf{1})\phi_n(\mathbf{1})d\mathbf{r} \quad (3.48)$$

$$S_{ij} = \int \phi_n^*(\mathbf{1})\phi_n(\mathbf{1})d\mathbf{r} \quad (3.49)$$

3. Eq.(3.47) can be solved by diagonalization when the basis set is not too large. The process is performed by constructing an explicit expression of the Kohn-Sham operator $\hat{f}_{KS}$ and calculating the overlap one-electron and two-electron integrals to obtain the matrix elements F_{ij}^{KS} and S_{ij} . The Kohn-Sham matrix is then orthogonalized and diagonalized giving a new set of the c 's in the basis set expansion of Eq.(3.46).

Using these new coefficients we construct the new ground state density of the electric system simply as:

$$n^{new}(\mathbf{r}) = \sum_{i=1}^N |\psi_i^{new}(\mathbf{r})|^2 \quad (3.50)$$

The whole process is repeated until the convergence is achieved, that is until $n_{new} = n_{old}$ (Self Consistent Approach). If this condition is not satisfied an improved $\hat{H}_{KS}^{new}$ operator is calculated and a new set of F_{ij}^{KS} and S_{ij} matrix elements are updated. Then the HK matrix is diagonalized and so on and so forth.

For the case of large systems and/or large basis sets, other iterative techniques are used. The Lanczos method [140–143], for example, avoids the explicit construction of the Kohn-Sham matrix: it is sufficient in these methods to have a procedure to apply (successively) the Kohn-Sham matrix on vectors $\{c_{ni}\}$. Only these vectors need then to be stored.

It is important to notice that in the DFT approximation only the electronic density, the energy of the Kohn-Sham HOMO orbital [144] and the *chemical potential*, μ :

$$\mu = \delta T_s[n] + \nu_{KS}(\mathbf{r}) \quad (3.51)$$

have physical sense. On the contrary, the rest of the orbital energies and the Kohn-Sham wavefunction do not have a strict physical meaning. In fact, the Kohn-Sham wavefunction, $|\Psi(\mathbf{r})\rangle$, is nothing but a wavefunction that belongs to the set of functions whose integral is $n(\mathbf{r})$.

Constrained minimization technique

In the constrained minimization technique we look for an electronic density $n(\mathbf{r})$:

$$n(\mathbf{r}) = \sum_{i=1}^N |\psi_i(\mathbf{r})|^2$$

that minimizes the functional for the Kohn-Sham energy (3.19):

$$E[\{\psi_i\}] = - \sum_{i=1}^n \langle \psi_i(\mathbf{r}) | \frac{1}{2} \nabla^2 | \psi_i(\mathbf{r}) \rangle + \int n(\mathbf{r}) \nu(\mathbf{r}) d\mathbf{r} + \frac{1}{2} \iint \frac{n(\mathbf{r}_1)n(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_1 d\mathbf{r}_2 + E_{XC}[n]$$

with respect to variations of $\{\psi_i\}$, that is:

$$\delta E[\{\psi_i\}] = 0 \quad (3.52)$$

This minimization process is carried out with the constraint that the spin-orbitals remain orthonormal:

$$\int \psi_i^*(\mathbf{r}) \psi_j(\mathbf{r}) d\mathbf{r} = \delta_{ij} \quad (3.53)$$

using the Lagrange multiplier technique:

$$\mathcal{L}[\{\psi_i\}] = E[\{\psi_i\}] - \sum_{i=1}^n \sum_{j=1}^n \Lambda_{ij} \left(\int \psi_i^*(\mathbf{r}) \psi_j(\mathbf{r}) d\mathbf{r} - \delta_{ij} \right) \quad (3.54)$$

where Λ_{ij} are the Lagrangian multipliers.

So that, the problem of finding a minimum energy of Eq.(3.52) is changed into solving the equation:

$$\delta \mathcal{L}[\{\psi_i\}] = 0 \quad (3.55)$$

At the minimum:

$$\frac{\partial E}{\partial \psi_i^*(\mathbf{r})} - \frac{\partial}{\partial \psi_i^*(\mathbf{r})} \sum_{i=1}^n \sum_{j=1}^n \Lambda_{ij} \left(\int \psi_i^*(\mathbf{r}) \psi_j(\mathbf{r}) d\mathbf{r} - \delta_{ij} \right) = 0 \quad (3.56)$$

therefore:

$$\frac{\partial E}{\partial \psi_i^*(\mathbf{r})} = \sum_{j=1}^n \Lambda_{ij} |\psi_j(\mathbf{r})\rangle \quad (3.57)$$

Evaluating the left hand side term of Eq.(3.57):

$$\frac{\partial E}{\partial \psi_i^*(\mathbf{r})} = -\frac{1}{2} \nabla^2 |\psi_i(\mathbf{r})\rangle + \int d\mathbf{r} (\nu(\mathbf{r}) + \nu_J(\mathbf{r}) + \nu_{XC}(\mathbf{r})) \frac{\partial n(\mathbf{r})}{\partial \psi_i^*(\mathbf{r})} \quad (3.58)$$

From Ref. [121]:

$$\frac{\partial n(\mathbf{r})}{\partial \psi_i^*(\mathbf{r})} = |\psi_i(\mathbf{r})\rangle \quad (3.59)$$

Thus Eq.(3.58) is reduced to:

$$\frac{\partial E}{\partial \psi_i^*(\mathbf{r})} = \left[-\frac{1}{2} \nabla^2 + \nu_{KS}(\mathbf{r}) \right] |\psi_i(\mathbf{r})\rangle \equiv \hat{H}_{KS} |\psi_i(\mathbf{r})\rangle \quad (3.60)$$

where $\nu_{KS}(\mathbf{r})$ and $\hat{H}_{KS}$ are the Kohn-Sham potential and Kohn-Sham operators that previously appeared in Eqs.(3.40) and (3.44).

In order to minimize $E[\{\psi_i\}]$, the orbitals are once more expanded on a basis set:

$$|\psi_i(\mathbf{r})\rangle = \sum_{n=1}^m c_{ni} |\phi_n(\mathbf{r})\rangle$$

and the energy functional is written as a function of the vectors $\{c_{ni}\}$. The minimization then amounts to minimizing a function in a high dimension space which can be performed by standard iterative methods like the conjugate gradient method [68, 145] or Direct Inversion in Iterative Space method [146–148] (DIIS). These methods rely on the calculation of the gradient $\partial E / \partial \psi_i^*(\mathbf{r})$; avoiding in this way an explicit construction of the Kohn-Sham matrix elements H_{ij}^{KS} given by Eq.(3.48).

Using Eq.(3.57) and Eq.(3.60), at the minimum it is satisfied that:

$$\hat{H}_{KS} |\psi_i(\mathbf{r})\rangle = \Lambda_{ij} |\psi_j(\mathbf{r})\rangle \quad (3.61)$$

The symmetric Λ_{ij} matrix can be diagonalized by an unitary transformation:

$$(U^T \Lambda U)_{ij} = \epsilon_i \delta_{ij} \quad (3.62)$$

where $\{\epsilon_i\}$ is the set of the one-electron orbitals energies of the auxiliary non-interacting electron system that diagonalizes the Kohn-Sham Hamiltonian. It can also be shown that the set of orbitals $\{\psi_i(\mathbf{r})\}$ are related with those of the non-interacting system $\{\tilde{\psi}(\mathbf{r})\}$ by:

$$|\tilde{\psi}_i(\mathbf{r})\rangle = U_{ij}^T |\psi_j(\mathbf{r})\rangle \quad (3.63)$$

This unitary transformation leaves both the density $n(\mathbf{r})$ and the kinetic terms $\langle \psi_i(\mathbf{r}) | \frac{1}{2} \nabla^2 | \psi_i(\mathbf{r}) \rangle$ invariant, so that the total energy E is also invariant by this unitary transformation. Finding the minimum of $E[\{\psi_i\}]$ with respect to variations of $\{\psi_i\}$ is thus strictly equivalent to solving the Kohn-Sham problem. The one-electron orbitals $\{\tilde{\psi}_i\}$ which diagonalize the Kohn-Sham Hamiltonian are called the canonical Kohn-Sham orbitals of the system, while any unitary transformation $\{\psi_i\}$ of these which also minimize $E[\{\psi_i\}]$ are called minimal orbitals.

3.4.1 Plane waves basis sets

In this section we will quickly discuss the basis expansion of Eq.(3.46) in the special case of plane wave basis set. Bloch's theorem [149] states that in a periodic solid each electronic wavefunction can be written as the product of a cell-periodic part and a wavelike part:

$$|\psi_i(\mathbf{r})\rangle = \exp(i\mathbf{k} \cdot \mathbf{r})|f_i(\mathbf{r})\rangle \quad (3.64)$$

The cell-periodic part of the wave function can be expanded using a basis set made by a discrete set of plane waves whose wave vectors are reciprocal lattice vectors of the crystal:

$$|f_i(\mathbf{r})\rangle = \sum_{\mathbf{G}} c_{i,\mathbf{G}} \exp(i\mathbf{G} \cdot \mathbf{r}) \quad (3.65)$$

The reciprocal lattice vectors $\mathbf{G}$ are defined by $\mathbf{G} \cdot \mathbf{l} = 2\pi m$ where $\mathbf{l}$ is a lattice vector of the crystal and m is an integer. For example, if we consider an orthorhombic box with lengths L_x , L_y and L_z , the wavevectors $\mathbf{G}$ are:

$$\mathbf{G} = i\frac{2\pi}{L_x}\mathbf{x} + j\frac{2\pi}{L_y}\mathbf{y} + k\frac{2\pi}{L_z}\mathbf{z} \quad \text{with } i, j, k \in \mathbb{Z} \quad (3.66)$$

A final expression of the electronic wavefunction as a sum of plane waves is obtained if we insert Eq.(3.65) into Eq.(3.64):

$$|\psi_i(\mathbf{r})\rangle = \sum_{\mathbf{G}} c_{i,\mathbf{k}+\mathbf{G}} \exp(i(\mathbf{k} + \mathbf{G}) \cdot \mathbf{r}) \quad (3.67)$$

Instead of having to solve an infinite number of wavefunctions over whole infinite space, the Bloch theorem changes the problem to one of solving a finite number of wavefunctions at an infinite number of possible values for $\mathbf{k}$. In order to simplify the problem to manageable proportions, it is necessary to impose some boundary conditions on the wavefunction, which restrict the allowed values of $\mathbf{k}$. Thus, the system is contained within a supercell which is then replicated periodically throughout space. The supercell must be large enough so that the systems contained within each one, which in reality are isolated, do not interact significantly.

In principle, even if the boundary conditions are applied, an infinite plane wave basis set is required to expand the electronic wavefunctions. In a practical way, since the basis set elements are eigenvectors of the kinetic energy operator:

$$-\frac{1}{2}\nabla^2|f_i(\mathbf{r})\rangle = \frac{1}{2}||\mathbf{G}||^2|f_i(\mathbf{r})\rangle \quad (3.68)$$

the size of the basis set is determined by inserting a cut-off:

$$E_{kin} = \frac{1}{2}||\mathbf{G}||^2 < E_{cut} \quad (3.69)$$

that includes all plane waves whose kinetic energy is below a defined limit E_{cut} [68].

Kohn-Sham equations (3.41) can be written in terms of plane-waves using Eq.(3.67):

$$\sum_{\mathbf{G}'} [|\mathbf{k} + \mathbf{G}|^2 \delta_{\mathbf{G}\mathbf{G}'} + \nu(\mathbf{G}-\mathbf{G}') + \nu_J(\mathbf{G}-\mathbf{G}') + \nu_{XC}(\mathbf{G}-\mathbf{G}')] c_{i,\mathbf{k}+\mathbf{G}'} = \epsilon_i c_{i,\mathbf{k}+\mathbf{G}} \quad (3.70)$$

where the kinetic energy is diagonal and the various potentials are described in terms of their Fourier transforms [68].

For MD, the main advantage of a plane wave basis set is the independence of the basis set elements with respect to the ionic positions. As it was described in Sec.(2.3), the Hellmann-Feynman theorem can be applied without additional Pulay terms arising from a basis set that would be dependent on the nuclei positions. Therefore, the forces on the ions will be calculated at no extra-cost. There is also no Basis Set Superposition Error for the same reasons. Another advantage of plane wave basis sets is that their quality depends only on the cut-off considered; it is thus easier both to compare results and to make convergence studies with only one number defining the quality of the basis set. Finally, concerning the CPU time, plane wave basis sets have the advantage of being orthonormal and they can not become over-complete (one element being exactly or even approximately a linear combination of other elements, a fact that leads to divergences).

However, plane wave basis sets also have disadvantages. The first one is probably the very large number of basis set elements required, specially to cover more localized inner electrons. Indeed calculations might become intractable and it is absolutely necessary to employ pseudopotentials: only valence electrons are thus considered, not core electrons; in consequence, the electron-ion interactions are not simply the fundamental coulomb attraction but there are also pseudo-potential interactions. Another interesting point is that, not being atom centered orbitals, plane wave basis sets do not easily lead to chemical insight on the electronic structure of the system studied: it is hard to describe the result of a plane wave calculation in a Linear Combination of Atomic Orbital framework.

Other characteristics can be both an advantage or disadvantage, depending on the application considered. For instance, plane wave basis set elements fill equally the whole simulation box. Thus, regions of high electronic densities near the atomic cores are described at the same level as regions of lower density like the valence region. When an isolated molecule is considered, this turns out to be a further disadvantage as the vacuum region around the molecule is still described at the same level of accuracy. On the contrary, it plays in favor of condensed systems since no region of very low electronic density is present. Another particular situation that makes this feature an advantage is when an electron is localized in a cavity instead of around an atom for example. This is the case for a solvated electron [150] or a solvated silver atom in aqueous solution [151,152].

3.4.2 Pseudopotential Approximation

As it has been already said, the plane wave basis set represents a appropriate way to expand electronic wavefunctions on periodic external potentials like crystals or surfaces. However, the very large number of them needed to expand the tightly bound core orbitals and to represent the rapid oscillations of the valence electron wavefunction in the core region³, make this choice useless when an all-electron calculation is considered.

The Pseudopotential Approximation [153–155] solves both problems by replacing the bare coulomb attraction between electrons and nuclei by a pseudopotential that acts on a set of pseudo wavefunctions (rather than the true valence wavefunctions) within the core region. The most common pseudopotential approach consists of not allowing the relaxation of the core states according to the environment (*frozen core approximation*), although some polarizable core approaches have been proposed [156].

The pseudo-functions proposed are combination of the true valence $|\psi_v(\mathbf{r})\rangle$ and core $|\psi_c(\mathbf{r})\rangle$ wavefunctions:

$$|\tilde{\psi}_v(\mathbf{r})\rangle = |\psi_v(\mathbf{r})\rangle + \sum_c \alpha_{cv} |\psi_c(\mathbf{r})\rangle \quad (3.71)$$

³These oscillations are a consequence of maintaining the orthogonality of core wavefunctions with the valence wavefunctions.

and (3.71) satisfies the modified Schrödinger equation:

$$\left[\hat{H} + \sum_c (\epsilon_v - \epsilon_c) |\psi_c(\mathbf{r})\rangle \langle \psi_c(\mathbf{r})| \right] |\tilde{\psi}_v(\mathbf{r})\rangle = \epsilon_v |\tilde{\psi}_v(\mathbf{r})\rangle \quad (3.72)$$

where the potential nuclear-electron attraction term of the Hamiltonian operator is replaced by $\hat{V}_{eN} = (Z_c/r)\hat{I}$, where $\hat{I}$ is the identity operator. It is possible to construct a pseudo-Hamiltonian:

$$\hat{H}_{PS} = \hat{H} + \sum_c (\epsilon_v - \epsilon_c) |\psi_c(\mathbf{r})\rangle \langle \psi_c(\mathbf{r})| \quad (3.73)$$

with the same eigenvalues of the original Hamiltonian but smoother, nodeless wavefunction. The associated potential:

$$\hat{V}_{PS} = \frac{Z_c}{r} \hat{I} + \sum_c (\epsilon_v - \epsilon_c) |\psi_c(\mathbf{r})\rangle \langle \psi_c(\mathbf{r})| \quad (3.74)$$

is called a pseudopotential.

It is required however that these pseudopotential functions be transferable, i.e., the pseudopotential for one element can be valid in different calculations for many different chemical environments of this element [157]. A major step forward in the theory of pseudopotentials was the introduction of a “norm-conservation” condition with the proposition by Hamann, Schlüter and Chiang [158] of a set of conditions to ensure transferability of pseudopotentials. Different recipes for constructing pseudopotentials satisfying these conditions have been proposed. They can be either analytic [158, 159] or numeric [160] and also differ by the way valence wave functions are made smooth in the core region.

For these norm-conserving pseudopotentials, a different potential needs to be applied on each orbital depending on its angular component l . These pseudopotentials then have a semi-local form:

$$\hat{V}_{PS}(\mathbf{r}) = \sum_{l=0}^{\infty} \sum_{m=-l}^l \nu_{PS}^l(\mathbf{r}) |lm\rangle \langle lm| \quad (3.75)$$

where $\nu_{PS}^l(\mathbf{r})$ is the pseudopotential corresponding to the angular component l .

3.4.3 Localized Molecular Orbitals

For computational purposes, it is convenient to work with Kohn-Sham molecular orbitals, i.e. those that make the Kohn-Sham matrix F^{KS} diagonal; and that are eigenfunctions of the Kohn-Sham $\hat{H}_{KS}$ operator at convergence. This corresponds to a specific choice of a unitary (orthogonal) transformation of the occupied molecular orbitals. Once the self-consistent procedure has converged, however, other sets of orbitals may be chosen by forming linear combinations of the canonical molecular orbitals. The total wavefunction, and thus all observable properties, is independent of such a rotation of the molecular orbitals:

$$|\psi'(\mathbf{r})\rangle = U|\psi(\mathbf{r})\rangle \quad \text{where} \quad |\psi'_i(\mathbf{r})\rangle = \sum_{j=1}^N u_{ij} |\psi_j(\mathbf{r})\rangle \quad (3.76)$$

The traditional view of molecular bonds is that they are due to an increased probability of finding electrons between two nuclei. The Kohn-Sham orbitals are delocalized over the whole molecule and do not readily reflect this, since the density between two nuclei is the result

of many small contributions from many (all) molecular orbitals. The canonical molecular orbitals therefore do not readily allow identification of the bonding properties of the system.

The goal of Localized Molecular Orbitals is to define molecular orbitals that are spatially confined to a relatively small volume, and therefore clearly display which atoms are bonded and furthermore have the property of being approximately constant between structurally similar units in different molecules. A set of localized orbitals may be defined by optimizing the expectation value of a two-electron operator Ω [161, 162]:

$$\langle \Omega \rangle = \sum_{j=1}^N \langle \psi'_i(\mathbf{1}) \psi'_i(\mathbf{2}) | \Omega | \psi'_i(\mathbf{1}) \psi'_i(\mathbf{2}) \rangle \quad (3.77)$$

In Eq.(3.77), the expectation value of Ω depends on the u_{ij} parameters in Eq.(3.76), that represents again a function optimization problem. The unitary transformation of the orbitals preserves the orthogonality, i.e. the resulting localized molecular orbitals are also orthogonal. Since all observable properties depend only on the total electron density, and not the individual molecular, there is no unique choice for Ω . For example Boys-Foster localization scheme [163] uses the square of the distance between two electrons as the operator, and minimizes the expectation value:

$$\langle \Omega \rangle_{Boys} = \sum_{j=1}^N \langle \psi'_i(\mathbf{1}) \psi'_i(\mathbf{2}) | (r_1 - r_2)^2 | \psi'_i(\mathbf{1}) \psi'_i(\mathbf{2}) \rangle \quad (3.78)$$

This corresponds to determining a set of localized molecular orbitals that minimizes the spatial extent, that is, they are as compact as possible. For systems described by plane wave basis functions the equivalent of the Boys localized orbitals is called Wannier orbitals [164], that will be use in the present work.

3.5 Nuclear forces in Density Functional Theory

In this section we will explicitly write the expressions of the Hellmann-Feynman forces in DFT. For this purpose, we need to recall the total energy of a many-electron system with charge density $n(\mathbf{r})$ in a external potential $\nu(\mathbf{r})$ written as a functional of the density:

$$E[n] = T_s[n] + \int n(\mathbf{r})\nu(\mathbf{r})d\mathbf{r} + \frac{1}{2} \iint \frac{n(\mathbf{r}_1)n(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_1 d\mathbf{r}_2 + E_{XC}[n]$$

According to Eq.(2.27), the electronic contribution to the force acting on a nuclei α is given by:

$$\mathbf{F}_\alpha = \frac{\partial E[n]}{\partial \mathbf{R}_\alpha} = - \int \frac{\partial \nu(\mathbf{r})}{\partial \mathbf{R}_\alpha} n(\mathbf{r}) d\mathbf{r} - \int \frac{\partial E(\mathbf{r})}{\partial n(\mathbf{r})} \frac{n(\mathbf{r})}{\partial \mathbf{R}_\alpha} d\mathbf{r} \quad (3.79)$$

The first term in Eq.(3.79) is the Hellmann-Feynman force, while the second is the Pulay-like term that vanishes if a plane wave basis sets used. We emphasize that it disappears not because plane waves are complete, but because they do not depend on atomic positions. Thus the expression of the force is reduced to:

$$\mathbf{F}_\alpha = Z_\alpha \int \frac{n(\mathbf{r})}{|\mathbf{r} - \mathbf{R}_\alpha|^3} (\mathbf{r} - \mathbf{R}_\alpha) \quad (3.80)$$

3.6 Outline of molecular dynamics simulation based on the DFT formalism

In the previous sections we have seen that finding a ground-state density is equal to solve the self-consistent electronic Kohn-Sham equations:

$$\left[-\frac{1}{2}\nabla_i^2 + \nu(\mathbf{r}) + \nu_J(\mathbf{r}) + \nu_{XC}(\mathbf{r}) \right] |\psi_i(\mathbf{r})\rangle = \epsilon_i |\psi_i(\mathbf{r})\rangle$$

where:

$$\nu(\mathbf{r}) = -\sum_{\alpha=1}^N \frac{Z_{\alpha}}{r_{\alpha i}} \quad ; \quad \nu_J(\mathbf{r}) = \int \frac{n(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_2 \quad ; \quad \nu_{XC}(\mathbf{r}) = \frac{\partial E_{XC}[n]}{\partial n(\mathbf{r})}$$

are the external, Hartree and exchange-correlation potential respectively.

In a Born-Oppenheimer MD simulation this process is combined with the nuclear motion following the general outline:

1. A set of atoms and initial positions and velocities $\{\mathbf{r}_0\}$, $\{\mathbf{v}_0\}$ are needed. We also need the basis set cut-off, time-step, and the proposed exchange-correlation potential, $E_{XC}[n]$.
2. Given the positions of the atoms, the external potential $\nu(\mathbf{r})$ is set up.
3. A trial electron density $n(\mathbf{r})$ is constructed.
4. $\nu_J(\mathbf{r})$ and $\nu_{XC}(\mathbf{r})$ potentials are calculated.
5. Kohn-Sham equations are solved by diagonalization. This is the highest CPU time consuming process.
6. With the set of the new Kohn-Sham orbitals, a new election density is constructed.
7. The solution is checked to be consistent, if it is not, go back to step 4.
8. Once a ground electronic density is achieved, the following step is to calculate the atomic forces, $\{\mathbf{f}_0\}$, following the Hellmann-Feynman theorem frameworked in DFT as Eq.(3.80) indicates.
9. From $\{\mathbf{r}_0\}$, $\{\mathbf{v}_0\}$, $\{\mathbf{f}_0\}$, a new set of positions and velocities $\{\mathbf{r}\}$, $\{\mathbf{v}\}$ are obtained by using the Verlet algorithm (detailed in next section) with the time-step read from the input.
10. Positions and velocities are updated and step 2 reinitiated.

3.7 Verlet algorithm

The forces are used to calculate the velocities and positions at the next time-step. In molecular dynamics, the most commonly used time integration algorithm is probably the so-called Verlet algorithm [165, 166]. The basic idea is to write two third-order Taylor expansions for the positions $\mathbf{r}(t)$, one forward and one backward in time. Calling $\mathbf{v}$ the velocities, $\mathbf{a}$ the accelerations, and $\mathbf{b}$ the third derivatives of $\mathbf{r}$ with respect to t , one has:

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \mathbf{v}(t)\Delta t + \dots \mathbf{a}(t)\Delta t^2 + (1/6)\mathbf{b}(t)\Delta t^3 + O(\Delta t^4) \quad (3.81)$$

$$\mathbf{r}(t - \Delta t) = \mathbf{r}(t) - \mathbf{v}(t)\Delta t + \dots \mathbf{a}(t)\Delta t^2 - (1/6)\mathbf{b}(t)\Delta t^3 + O(\Delta t^4) \quad (3.82)$$

Adding both expressions gives:

$$\mathbf{r}(t + \Delta t) = 2\mathbf{r}(t) - \mathbf{r}(t - \Delta t) + \mathbf{a}(t)\Delta t^2 + O(\Delta t^4) \quad (3.83)$$

This is the basic form of the Verlet algorithm. Since we are integrating Newton's equations, $\mathbf{a}(t)$ is just the force divided by the mass, and the force is in turn a function of the positions $\mathbf{r}(t)$:

$$\mathbf{a}(t) = -(1/m)\nabla V(\mathbf{r}(t)) \quad (3.84)$$

As one can immediately see, the truncation error of the algorithm when evolving the system by Δt is of the order of Δt^4 , even if third derivatives do not appear explicitly. This algorithm is at the same time simple to implement, accurate and stable, explaining its large popularity among molecular dynamics simulators.

A problem with this version of the Verlet algorithm is that velocities are not directly generated. While they are not needed for the time evolution, their knowledge is necessary. Moreover, they are required to compute the kinetic energy K , whose evaluation is necessary to test the conservation of the total energy $E=K+V$ in a Microcanonical MD. This is one of the most important tests to verify that a MD simulation is proceeding correctly. One could compute the velocities from the positions by using

$$\mathbf{v}(t) = \frac{\mathbf{r}(t + \Delta t) - \mathbf{r}(t - \Delta t)}{2\Delta t}. \quad (3.85)$$

However, the error associated to this expression is of order Δt^2 rather than Δt^4 . To overcome this difficulty, some variants of the Verlet algorithm have been developed. They give rise to exactly the same trajectory, and differ in what variables are stored in memory and at what times. The leap-frog algorithm, not reported here, is one of such variants [167] where velocities are handled somewhat better. An even better implementation of the same basic algorithm is the so-called velocity-Verlet scheme (used in the CPMD code [168,169]), where positions, velocities and accelerations at time $t + \Delta t$ are obtained from the same quantities at time t in the following way:

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \mathbf{v}(t)\Delta t + (1/2)\mathbf{a}(t)\Delta t^2 \quad (3.86)$$

$$\mathbf{v}(t + \Delta t/2) = \mathbf{v}(t) + (1/2)\mathbf{a}(t)\Delta t \quad (3.87)$$

$$\mathbf{a}(t + \Delta t) = -(1/m)\nabla V(\mathbf{r}(t + \Delta t)) \quad (3.88)$$

$$\mathbf{v}(t + \Delta t) = \mathbf{v}(t + \Delta t/2) + (1/2)\mathbf{a}(t + \Delta t)\Delta t \quad (3.89)$$

Note how we need $9N$ memory locations to save the $3N$ positions, $3N$ velocities and $3N$ accelerations, but we never need to have simultaneously stored the values at two different times for any one of these quantities.

Chapter 4

Basics on Time Dependent Density Functional Theory

Up to now we have seen that ground-state DFT is an alternative to wavefunction calculations for electronic structure and properties. DFT is in principle exact, built on the Hohenberg-Kohn theorems: the ground-state density of an interacting N -electron system in a static external potential completely determines all of its properties (e.g., energy, bond lengths ...) [118, 120]. The practical power of DFT lies in mapping the system to a fictitious non-interacting system of the same ground-state density, the Kohn-Sham (KS) system, with a one-body potential. Solution of these self-consistent orbital equations is faster than solution of the Schrödinger equation, allowing calculations on large systems of interest in materials science and chemistry.

Time Dependent Density Functional Theory (TD-DFT) is based on the Runge-Gross theorem [170] that represents an extension of HK theorems for interacting electrons in time-dependent external potentials. A set of equivalent Time-Dependent Kohn-Sham equations (TD-KS) for fictitious noninteracting electrons are also derived. The theory aims for example to treat atoms and molecules in intense laser fields, find electronic excitation spectra, calculate polarizabilities, etc. Following Ref. [171], it is convenient to distinguish two regimes in TD-DFT: (a) If the external time-dependent potential is “small”, the complete numerical solution of the time-dependent Kohn-Sham equations can be avoided by the use of linear response theory. This is the case, e.g., for the calculation of photoabsorption spectra. (b) For a “strong” external potential, a full solution of the time-dependent Kohn-Sham equations is needed. This situation is encountered, for instance, when matter interacts with intense laser fields.

In this Chapter we will introduce some of the basic concepts on TD-DFT focusing our attention on the time propagation of the TD-KS equations in the Ehrenfest MD used in this thesis [44]. Further reading and detailed applications of TD-DFT and Linear Response Theory, can be found in Refs. [172–178].

4.1 Runge-Gross theorem

Considering N non-relativistic electrons, interacting via Coulomb repulsion, in a time-dependent external potential, the Runge-Gross theorem [170] states that the densities $n(\mathbf{r}, t)$ and $n'(\mathbf{r}, t)$ evolving from a common initial state $|\Psi_0\rangle = |\Psi(t=0)\rangle$ under the influence of two external potentials $\nu(\mathbf{r}, t)$ and $\nu'(\mathbf{r}, t)$ (both Taylor expandable about the initial $t = 0$) are

always different provided that the potentials differ by more than a purely time-dependent (r-independent) function:

$$\nu(\mathbf{r}, t) - \nu'(\mathbf{r}, t) \neq c(t) \quad (4.1)$$

The one-to-one mapping between densities and potentials is then established, and we say that the time-dependent potential is a functional of the time-dependent density (and the initial state). This statement implies that, if the only information we have about the system is its density, we can obtain the external potential that produced this density. With the external potential the TDSE can be solved, and all properties of the system obtained. One can conclude that the electronic density determines all other properties of the quantum system. It is important to notice that this is true for a fixed initial state, that is, in addition to the knowledge of $n(\mathbf{r}, t)$ we also need to know the initial many-body state $|\Psi_0\rangle$.

4.2 The Time-Dependent Kohn-Sham Equations

The role that is played by the energy functional in stationary Density Functional Theory is played by the *action functional* in Time-Dependent Density Functional Theory:

$$A[\Psi] = \int_{t_0}^t \langle \Psi(t) | i \frac{\partial}{\partial t} - \hat{H}(t) | \Psi(t) \rangle \quad (4.2)$$

where $\Psi(t)$ is an N-body function. In their original paper, Runge and Gross [170] offered a derivation of the Kohn-Sham equations starting from the action Eq.(4.2) and constructing an auxiliary system of noninteracting electrons subject to an external local potential ν_{KS} . This potential is unique (by the Runge-Gross theorem applied to the noninteracting system) and is chosen such that the density of the Kohn-Sham electrons is the same as the density of the original interacting system. Thus, the total time-dependent electronic density is given by:

$$n(\mathbf{r}, t) = \sum_{i=1}^N |\psi_i(\mathbf{r}, t)|^2 \quad (4.3)$$

Each of the non-interacting Kohn-Sham particles satisfies the following TDSE:

$$i \frac{\partial}{\partial t} |\psi_i(\mathbf{r}, t)\rangle = \hat{H}_{HK}(\mathbf{r}, t) |\psi_i(\mathbf{r}, t)\rangle \quad (4.4)$$

where the time-dependent Kohn-Sham Hamiltonian is defined as:

$$\hat{H}_{HK}(\mathbf{r}, t) = -\frac{1}{2} \nabla_i^2 + \nu_{KS}(\mathbf{r}, t) \quad (4.5)$$

and the time-dependent Kohn-Sham potential as:

$$\nu_{KS}(\mathbf{r}, t) = \nu(\mathbf{r}, t) + \nu_J(\mathbf{r}, t) + \nu_{XC}(\mathbf{r}, t) \quad (4.6)$$

that is, as in the Kohn-Sham scheme for the ground state, the time-dependent Kohn-Sham potential is normally written as the sum of the external, Hartree and exchange-correlation terms. The first one gathers the potential created by the electron-nuclei interaction plus any other time-dependent potential applied. The time-dependent Hartree potential can be written as:

$$\nu_J(\mathbf{r}, t) = \int \frac{n(\mathbf{r}_2, t)}{r_{12}} d\mathbf{r}_2 \quad (4.7)$$

Finally last term of Eq.(4.6), the exchange-correlation potential, includes all nontrivial many-body effects, and has an extremely complex (and essentially unknown) functional dependence on the density. This dependence is clearly nonlocal, both in space and in time, i.e., the potential at time t and position $\mathbf{r}$ can depend on the density at all other positions and all previous times. As it happens in ground-state density the results obviously depends on the quality of the approximation. Explicit density functionals, like the adiabatic LDA, only retain the density dependence.

In the original paper of Runge and Gross [170], the exchange-correlation potential was simply the functional derivative. Years after, it was discovered that this formulation suffered from fundamental problems [179]. Using the Keldysh formalism [180] and introducing a new action functional that does not explicitly contain $\partial/\partial t$, a final expression was obtained:

$$\nu_{XC}(\mathbf{r}, t) = \left. \frac{\delta A_{XC}}{\delta n(\mathbf{r}, \tau)} \right|_{n=n(\mathbf{r}, t)} \quad (4.8)$$

where τ is the Keldysh pseudotime.

4.3 Adiabatic Approximation

As we have already noticed, the exact exchange-correlation potential depends on the entire history of the density, and also on the initial wavefunctions of both the interacting and non-interacting systems.

The dependence on the initial wavefunction disappears whenever a non-degenerate ground state, for both interacting and noninteracting, is considered: the initial wavefunctions themselves are functionals of the initial density. However, the time and spatial dependences are meaningful since exchange-correlation potential at $\mathbf{r}$ and t depends not just on $n(\mathbf{r}, t)$ but on all $n(\mathbf{r}, t')$ for $0 \leq t' \leq t$, and for arbitrary points $\mathbf{r}$ in space. Thus the potential remembers the density's past, and it is said it has a *memory*.

In the adiabatic approximation the exchange-correlation potential is approximated as being local in time, that is, all dependence on the past is ignored and the dependence is then reduced to the instantaneous density¹:

$$\nu_{XC}[n](\mathbf{r}, t) \approx \nu_{XC}[n(t)](\mathbf{r}) \quad (4.9)$$

If the time-dependent potential changes very slowly (adiabatically), this approximation will be valid. However, the electrons will remain always in their instantaneous ground state, and in practice, the spatial nonlocality of the functional is also approximated:

$$\nu_{XC}[n](\mathbf{r}, t) \approx \nu_{XC}^{GS}[n_{GS}](\mathbf{r})|_{n_{GS}(\mathbf{r}')=n(\mathbf{r}', t)} \quad (4.10)$$

where $\nu_{XC}^{GS}[n_{GS}](\mathbf{r})$ is the exact ground-state exchange-correlation potential of the density $n_{GS}(\mathbf{r})$. Due to the resemblance to the argument made to determine the function used in LDA calculations for the ground-state energy, the approximation receives the name of *Adiabatic LDA approximation* (ALDA).

¹Notation: For clarity we consider the dependence of ν_{XC} on the electron density, $\nu_{XC}[n]$. This statement is also applicable to $\nu_J[n]$

4.4 Outline of molecular dynamics simulation based on time propagation

In order to know the time-dependent Kohn-Sham wavefunctions at certain time t one can solve the time-dependent Kohn-Sham equations (4.4) or, alternatively, the electronic orbitals can be propagated in time in the so-called *time propagation scheme*:

$$|\psi_i(t)\rangle = \hat{U}(t, t_0)|\psi_i(t_0)\rangle \quad i = 1, \dots, N_e \quad (4.11)$$

where linear propagator $\hat{U}(t, t_0)$ is defined as:

$$\hat{U}(t, t_0) = \hat{T} \exp \left(- \int_{t_0}^t \hat{H}_{KS}(\tau) d\tau \right) \quad (4.12)$$

and $\hat{T}$ is the time ordering operator [181]. The strategy for solving time-dependent problems is to divide the total evolution operator into short segments in which the Hamiltonian does not change significantly:

$$\hat{U}(t, t_0) = \prod_{i=0}^{N-1} \hat{U}(t_i + \Delta t_i, t_i) \quad (4.13)$$

where $t_0 = 0$, $t_{i+1} = t_i + \Delta t_i$ and $t_N = t$. In this way, the evolution from $t = t_0$ to t , is broken into $N - 1$ intervals separate a constant time Δt . Using the product of Eq.(4.13), the set of equations (4.11) can be written as function of the intermediate times:

$$|\psi_i(t + \Delta t)\rangle = \hat{U}(t + \Delta t, t)|\psi_i(t)\rangle = \hat{T} \exp \left(- \int_t^{t+\Delta t} \hat{H}_{KS}(\tau) d\tau \right) |\psi_i(t)\rangle \quad (4.14)$$

If the Hamiltonian is time-independent, $\hat{T}$ can be omitted. This results in great simplification:

$$|\psi_i(t + \Delta t)\rangle \approx \exp \left(-i\Delta t \hat{H}_{KS}(t) \right) |\psi_i(t)\rangle \quad (4.15)$$

where the exponential can be approximate by polynomial expansions or splitting techniques between others [182]. For time-dependent problems many approximations for $\hat{U}(t + \Delta t, t)$ have been proposed [183], and particularly, the Ehrenfest TD-DFT MD implementation used in this thesis is based on the Crank-Nicholson (CN) method:

$$\hat{U}_{CN}(t + \Delta t, t) = \frac{1 - \frac{i}{2}\Delta t \hat{H}_{KS}(t + \Delta t/2)}{1 + \frac{i}{2}\Delta t \hat{H}_{KS}(t + \Delta t/2)} \quad (4.16)$$

combined with a two-step Runge-Kutta scheme [44]. The solutions of the time-dependent Schrödinger equations for both half and full steps is accomplished by iterating until convergence the set of integral equations:

$$|\psi_i^{(n)}(t_0 + \Delta t)\rangle = |\psi_i^{(0)}(t_0 + \Delta t)\rangle - i \int_{t_0}^{t_0 + \Delta t} \hat{H}_{KS}(\{\psi^{(n-1)}(\tau)\}, \tau) \psi_i^{(n-1)}(\tau) d\tau \quad (4.17)$$

A general outline can be given by the following steps:

1. In a initial step, the following elements are read:

- A set of atoms, initial positions and velocities $\{\mathbf{r}_0\}$, $\{\mathbf{v}_0\}$

- The basis cut-off and the approximation for the exchange-correlation potential.
 - A time-step, Δt , for the atomic and electronic movement.
 - An initial time-dependent electron density $n_0(\mathbf{r}, t)$. (In our special case of double ionizations this density is built up by all the KS orbitals of the neutral species except the one removed).
2. A first Runge-Kutta step is initiated, the KS orbitals are propagated until a middle point, $t = t_0 + \Delta t/2$, using $\nu_{KS}(\mathbf{r}, t_0)$. Then, at the middle time, the effective potential $\nu_{KS}(\mathbf{r}, t_0 + \Delta t/2)$ is calculated.
 3. In the second step, the full time evolution is achieved by evolving the wavefunctions for the full time-step Δt , using the approximated KS Hamiltonian computed at the half step, $\hat{H}_{KS}(\mathbf{r}, t_0 + \Delta t/2)$.
 4. With the new density $n_0(\mathbf{r}, t + \Delta t) = \sum_{i=1}^N |\psi_i(\mathbf{r}, t + \Delta t)|^2$, a new set of forces $\{\mathbf{f}\}$ on the atoms are obtained using Hellmann-Feynman theorem.
 5. Regarding $\{\mathbf{f}\}$, $\{\mathbf{r}_0\}$ and $\{\mathbf{v}_0\}$; the new positions and velocities, $\{\mathbf{r}\}$ and $\{\mathbf{v}\}$, are straightforwardly updated by the Verlet algorithm.
 6. All steps are repeated until the total time of the simulation is achieved.

Chapter 5

Modelling double ionization

5.1 Computational details

As it was introduced, the aim of this thesis is to simulate the first ≈ 100 fs after the double ionization of biomolecules as a consequence of their interaction with ionizing radiation. The intention of this Chapter is to expose the hypothesis assumed on such simulations and to explain the way they have been performed.

All simulations have been done using the CPMD code on its latest version v 3.13.2 [168] with the special Ehrenfest TD-DFT MD implementation of Tavernelli et al. [44]. This code is based on DFT using plane waves basis sets and allows us to perform *ab-initio* molecular dynamics simulations under two different approximations: Born-Oppenheimer (BO) MD and Car-Parrinello (CP) MD. As it was detailed on previous chapters, both dynamics are essentially adiabatic since the electrons are assumed to be in the same stationary state. On the contrary, the Ehrenfest TD-DFT MD of Tavernelli et al. (not included in the official release) is suitable to treat excited states as it takes into account the time dependency of the electronic wavefunction and the non-adiabatic electronic couplings. In the modelling of the double charged ions and their Coulomb explosions we have used a combination of TD-DFT MD and BO MD, while CP MD have been just used to provide the desired temperature before the ionization occurs.

All simulations have been performed in the (N,V,T) ensemble, including the generalized gradient corrections of Becke, Lee, Yang and Parr (BLYP) for the exchange and correlation terms respectively [129,133]. In the case of TD-DFT MD, the ALDA approximation has been assumed. The forces on the nuclei are computed according to the Hellman-Feynman theorem as the analytic derivative of the expectation value of the energy. Core electrons have been replaced by pseudopotentials of the standard Troullier-Martins form [160] and valence electron wavefunctions have been expanded in a plane-wave basis set with a 70 Ry cutoff. The simulation cell has been chosen to be a cubic box of size varying from $L = 12.5$ to 21 Å depending on the studied system (sizes will be given for each case when corresponding). Systems in gas phase have been treated as isolated molecules [184]; in the particular case of uracil in water, where the solvent has been explicitly included, the simulation box has been periodically repeated in space using Periodic Boundary Conditions.

After these common characteristics, let's describe how the initial double charged state has been obtained. The process starts by fitting the temperature of the system to a certain value that is experimentally determined. The so-called *thermalization* is carried out by a CP MD run of ≈ 5 ps during which the kinetic energy of the ions is controlled by rescaling. In the CP approximation, the electronic movement is given by a fictitious classical system

that avoids to solve the KS equations each time-step. As a result, the computational cost of the simulation is reduced (i.e. 32 processors are enough for the studied systems). Before running the thermalization we have tested different fictitious electronic masses at different time-steps of integration. Setting the electron mass to 100 a.u. and the time-step to 1 u.a. ensures a good control of the adiabaticity and the conserved quantities throughout the simulations.

Once the system has reached the desired temperature, a conformation is chosen (velocities and positions). We assume that the double ionization occurs vertically, it takes place over the geometry of the neutral species and involves two electrons of the same double occupied KS orbital. Our purpose in the following is to argue in favor of these assumptions:

1. Vertical ionization: In the LET range of swift heavy ions relevant to our work, the experimental conditions (see, for instance, Refs. [185, 186] for liquid water) are such that the path of ions is about 2.5 mm and the energy loss of the colliding ions is about 25% of the incident energy, which results in a 12% loss in velocity (velocity typically decreases from 0.38c to 0.34c). An associated short collision time of about 0.1 fs can thus be estimated. In that range of velocities, capture cross sections are negligible, so that the charge state of the colliding ions remains unchanged. Moreover, swift heavy ions projectiles have both velocity and mass high enough so that elastic collisions can be safely neglected, and consequently the energy of the ions upon collision is released into the target molecules (i.e. water molecules or biomolecules) through electronic excitation and ionization. Thus, the vertical ionization of the target molecule can be viewed as a consequence of these ultra-short time-scales of the collision process.

It is important to emphasize that we do not model the colliding ion in our simulations. Actually, in such high collision energies the velocity of the incoming ion and of the extracted electrons are so high that ions and electrons are swiftly evacuated from the collision zone. Both are not localized anymore in the vicinity of the ionized molecule within 1 as (atto-second) subsequent to the collision and ionization events. Indeed, this time scale is three orders of magnitude lower than the femto-scale where the fragmentations occurred and in which we are interested.

2. Double ionization: As we pointed out in the Introduction of this thesis, double ionization arising from irradiation by swift heavy ions are about 10% of primary ionization events. Although less probable, they lead to more lethal events, for instance: In the case of pure liquid water, Gervais et al. [34] demonstrated that the multiple ionization hypothesis is indeed able to correctly interpret the experimental HO_2/O_2^- formation yield [185] for LET above 100 keV/ μ as obtained with swift heavy ions. Furthermore, combining multi-scale theoretical modeling has shown that the production of HO_2 can only be achieved as a consequence of an initial double ionization of one H_2O molecule of the liquid [35].
3. Ionization from the same orbital: This particular restriction is imposed by the fact that ionizations from different orbitals are unaffordable from computational point of view. The implementation of the Ehrenfest dynamics TD-DFT used here can handle systems within the Local Spin Density (LSD) approximation, however as the number of spin-orbitals becomes double, the CPU time required increases by 80 %. Same reasoning can be done for the hypothetical study of a single charged species fragmentation. In this case not only an expensive TD-DFT LSD calculation would be required, moreover it would need much more computational resources as the fragmentation of a mono-charged species is slower than the double ionized ones.

Once the double ionized molecule has been obtained, the fragmentation dynamics properly begins. The first tens of fs are obtained by TD-DFT MD, starting from the positions and velocities of the neutral species. This calculation also requires an initial electron density that is built up by all the KS orbitals of the neutral species, except the one we have decided to remove. Concerning the double ionization of water molecules, we have chosen a Wannier representation instead, as the KS orbitals are much more delocalized.

It is particularly important to set-up an optimal time-step for the time propagation of the KS orbitals in order to preserve the total energy of the system (KS energy + Kinetic energy of the ions) as constant as possible. In next sections we will give more details about this problem and also the computational effort invested in such simulations.

Since the complete dissociation of the ionized molecule requires in general longer times (a few ps), after certain time of TD-DFT MD, we assume that the dynamics becomes adiabatic. At that moment, the simulation is switched to a ground-state BO MD keeping the positions and velocities from the time-dependent outcome. The BO MD are much less CPU demanding than the TD-DFT MD, therefore the simulation can be extended up to ≈ 100 fs with minor CPU cost. Concerning the time-step used for the BO MD, we have checked that a 1 a.u. time-step leads to excellent conservation of the total energy.

The suitable time to switch from one method to the other is determined by running several BO MD, which are started from conformations taken at different times of the TD-DFT MD. When both dynamics lead to the same trajectory, we assume that the system has relaxed to its ground-state and the dynamics is continued using a BO MD simulation. This issue will be tackled in detail in Chapter 7 using uracil²⁺ in the gas phase as example. The time needed to reach the ground-state is approximately 60 fs for uracil²⁺ and pyrimidine²⁺ in the gas phase, and 25 fs for uracil²⁺ in the liquid phase. It is also worth mentioning that the total-energy difference between the last step of the TD-DFT MD and the first one performed by the BO MD is ≈ 1.5 eV.

Combination of both TD-DFT MD and BO MD leads to a deep understanding of the fragmentation process. The simulations provide us the positions and velocities of the atoms each time-step of dynamics, allowing us to identify the kinetic energy of the mass centers of the fragments during whole time of the trajectory. On the other hand, since the density is given each few steps of simulation (saved as grid cube files) it is also possible to monitor the charges of the fragments generated.

5.2 Choosing the time-step

The choice of the optimal time-step to propagate the KS orbitals in the TD-DFT MD, requires a special attention. Fig.(5.1a) shows the total energy for a TD-DFT simulation where a uracil molecule in the gas phase has been double ionized from its deepest KS orbital. The propagation have been done using three different time-steps: 0.05, 0.025 and 0.01 a.u. As it can be seen, the total energy of the system is not completely conserved due to numerical errors introduced by the algorithm of propagation. A smaller step assures a better conservation of the energy although it also considerably increases the computational time.

As consequence of the extra energy introduced by the numerical errors, the results of the simulations can be biased. For instance, in Fig.(5.1b) follows the N1-C6 distance during the three previous TD-DFT MD simulations. The one performed with a time-step of 0.05 a.u. leads to the formation of an extra fragment as N1-C6 is definitely broken. On the contrary, in the simulations with 0.025 and 0.01 a.u. time-steps, the two atoms remain bonded. Same tests have been repeated with smaller times of integration, 0.005 and 0.0025 a.u., without noticing changes into the fragmentation trajectories in comparison with the

ones obtained using 0.01 a.u. Hence, we have concluded that 0.01 a.u. is an optimal time-step that combines an affordable CPU time and keeps the energy gained, due to numerical propagation, at a reasonable value of 0.1 eV and without affecting the simulation final results as for example the number of fragments produced. It has to be notice, however, that at short times of dynamics no big changes are detected in the distances of the ionized molecule or in the energy.

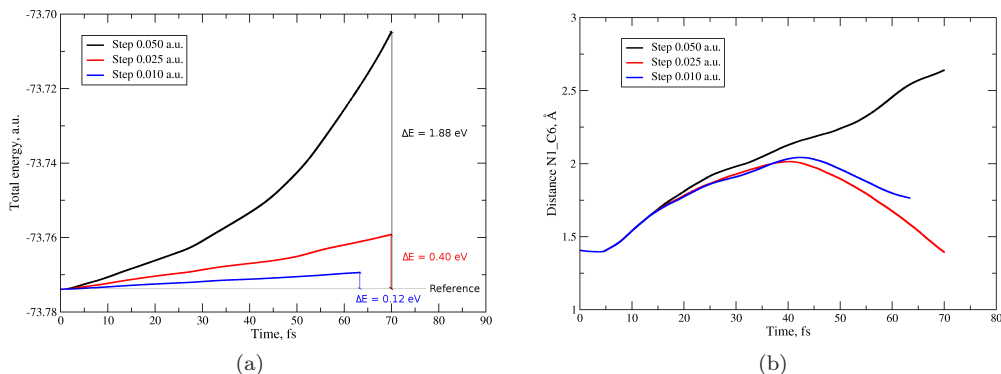


Figure 5.1: a) The total energy of three TD-DFT runs for uracil²⁺ fragmentation, using three different time-steps: 0.05, 0.025 and 0.01 a.u. b) N1-C6 distance is followed during the three simulations of different time-steps

For the case of the double ionization of a neighboring water molecule of uracil (studied in Chapter 8) we have run several TD-DFT MD tests in which a molecular orbital 2a1 is ionized and the positions of the nuclei are fixed. The total energy has been calculated with time-steps decreasing from 1 a.u. to 0.00625 a.u. and we concluded that the numerical convergence is good for a time-step of 0.05 a.u.

5.3 Typical CPU times

Differences between Ehrenfest and Born-Oppenheimer MD are also found in the computational cost. Table (5.1) gathers the comparison between three simulations referred to the same system (uracil²⁺) and extended during 60 fs, using 64-IBM SP Power6 processors and the CPMD v 3.13.2, compiled with Portland and IBM compilers with Lapack v3.2.1 and Essl v4.4 libraries.

Table 5.1: Comparisons of CPU times between a BO MD and TD-DFT MD

	Time-step	Total steps	Steps/min	Total CPU time
BO MD gas	1.00	2500	2.70	0.9 kh
TD-DFT MD gas	0.01	250000	1.30	205 kh
TD-DFT MD liq.	0.01	250000	0.65	410 kh

The BO MD simulation only requires 1 kh of CPU time while the time needed for the TD-DFT MD $\approx$ 200 times more (becoming double if the solvent is included). Considering the simulations in gas phase, we can clearly see that the major difference between both approximations is not the time invested in each time-step (Ehrenfest propagation needs

approximately the double of time to perform one step) but the number of steps needed to accomplish a given time of simulation. Indeed, this is a consequence of the small step required to integrate the electronic motion in TD-DFT MD.

MD simulations presented in this thesis have been run in Barcelona Supercomputing Center (BSC), Institut du Développement et des Ressources en Informatique Scientifique (IDRIS), Centre Informatique National (CINES), Málaga University Supercomputing Center, le Centre de Calcul Recherche (CCR- Université Pierre et Marie Curie) and Centro de Computación Científica (CCC- UAM).

Chapter 6

Observables

Computer simulation allows us to study properties of many-particle systems. However, not all properties can be directly measured in a simulation, and also, most of the quantities given by a simulation do not correspond to properties that are measured in real experiments. For example, in a MD simulation of liquid water, we could measure the instantaneous positions and velocities of all molecules in the liquid. However, this kind of information cannot be compared to experimental data because no real experiment provides such detailed information. Rather, a typical experiment measures an average property, averaged over the time of the measurement.

The connection between microscopic simulations and macroscopic properties is made via statistical mechanics which provides the rigorous mathematical expressions that relate macroscopic properties to the distribution and motion of the atoms and molecules of the N-body system; molecular dynamics simulations provide the means to solve the equations of motion of the particles and evaluate these mathematical formulas. In the next sections we will introduce some basic definitions on this field that have been handled in the MD simulation performed in this thesis. More details can be found in references [61, 63, 64, 187–189].

6.1 Ensembles

Statistical mechanics is the branch of physical sciences that studies macroscopic systems from a molecular point of view. The goal is to understand and to predict macroscopic phenomena from the properties of individual molecules and atoms. The system could range from a collection of solvent molecules to a solvated protein-DNA complex, for instance. In order to connect the macroscopic system to the microscopic system, time independent statistical averages are often introduced. We start this discussion by introducing a few definitions:

- The thermodynamic state of a system is usually defined by a small set of parameters, for example, the temperature, T , the pressure, P , and the number of particles, N . Other thermodynamic properties may be derived from the equations of state and other fundamental thermodynamic equations.
- The mechanical or microscopic state of a system is defined by the atomic positions, $\mathbf{r}$, and momenta, $\mathbf{p}$; these can also be considered as coordinates in a multidimensional space called phase space. For a system of N particles, this space has $6N$ dimensions. A single point in phase space, denoted by $\mathbf{G}$, describes the state of the system.

- An ensemble generally is defined as a collection of all possible systems which have different microscopic states but have an identical macroscopic or thermodynamic state. In MD simulations, a sequence of points in the phase space is generated as a function of time. These points represent microscopic states of the system and are imposed to belong to the same ensemble, that is, they must satisfy the conditions of a particular thermodynamic state. There exists different ensembles with different characteristics:
 - Microcanonical ensemble (NVE) : The thermodynamic state characterized by a fixed number of atoms, N , a fixed volume, V , and a fixed energy, E . This corresponds to an isolated system and is the one used in the MD performed in this work.
 - Canonical ensemble (NVT): This is a collection of all systems whose thermodynamic state is characterized by a fixed number of atoms, N , a fixed volume, V , and a fixed temperature, T .
 - Grand canonical ensemble (μ VT): The thermodynamic state for this ensemble is characterized by a fixed chemical potential, μ , a fixed volume, V , and a fixed temperature, T .

6.2 Time averages

An experiment is usually made on a macroscopic sample that contains an extremely large number of atoms or molecules, sampling an enormous number of conformations. In statistical mechanics, the average values are defined as ensemble averages given by:

$$\langle A \rangle_{ensemble} = \int \int A(\mathbf{p}, \mathbf{r}) \rho(\mathbf{p}, \mathbf{r}) d\mathbf{p} d\mathbf{r} \quad (6.1)$$

where $A(\mathbf{p}, \mathbf{r})$ is the observable of interest and it is expressed as a function of the momenta, $\mathbf{p}$, and the positions, $\mathbf{r}$, of the system. The integration is over all possible variables of $\mathbf{r}$ and $\mathbf{p}$.

The probability density of the ensemble is given by:

$$\rho(\mathbf{p}, \mathbf{r}) = \frac{1}{Q} \exp[-H(\mathbf{p}, \mathbf{r})/k_B T] \quad (6.2)$$

where $H(\mathbf{p}, \mathbf{r})$ is the Hamiltonian of the system, T is the temperature, k_B is Boltzmann's constant and Q is the partition function:

$$Q = \int \int \exp[-H(\mathbf{p}, \mathbf{r})/k_B T] d\mathbf{p} d\mathbf{r} \quad (6.3)$$

This integral is generally extremely difficult to calculate because one must determine all possible states of the system. In a molecular dynamics simulation, the points in the ensemble are calculated sequentially in time, so to get to know an ensemble average, the simulation must pass through all possible states corresponding to the particular thermodynamic constraints. The way to tackle this problem is to determine a *time average* of A , which is expressed as:

$$\langle A \rangle_{time} = \frac{1}{t} \int_{t_0}^{t_0+t} A(\mathbf{p}(\tau), \mathbf{r}(\tau)) d\tau \quad (6.4)$$

where t is the simulation time and $A(\mathbf{p}, \mathbf{r})$ is the instantaneous value of A . The dilemma appears because one can calculate time averages by molecular dynamics simulation, but the

experimental observables are supposed to be ensemble averages. In a practical way, if the equilibrium is assumed, the interval average of Eq.(6.4) is reliably approximated by the time average of $\langle A \rangle_{ensemble}$. Thus the later would be obtained from a measurement performed over an infinite duration:

$$\langle A \rangle_{ensemble} = \lim_{t \rightarrow \infty} \frac{1}{t} \int_{t_0}^{t_0+t} A(\mathbf{p}(\tau), \mathbf{r}(\tau)) d\tau \approx \frac{1}{S} \sum_{t=1}^S A(\mathbf{p}, \mathbf{r}) \quad (6.5)$$

Eq.(6.5) states one of the most fundamental axioms of statistical mechanics, the ergodic hypothesis. The basic idea is that if one allows the system to evolve in time indefinitely, that system will eventually pass through all possible states. The goal, therefore, of a molecular dynamics simulation is to generate enough representative conformations such that this equality is satisfied. If this is the case, experimentally relevant information may then be calculated if one is certain to sample a sufficient amount of phase space.

6.3 Temperature in Molecular Dynamics

One situation in which the ergodic assumption is obviously valid is when the integrand $A(\mathbf{p}, \mathbf{r})$ is a constant of the motion. In that case A does not vary along the phase-space trajectory and any interval average given by Eq.(6.4) is equal to the time average of Eq.(6.5). However as the phase point journeys along the hypersurface of constant energy E , most quantities are not constant: instead their values fluctuate.

One example of a fluctuating quantity is the translational kinetic energy, whose average tends to be stable about a value proportional to the kinetic temperature:

$$\langle E_k \rangle = \lim_{t \rightarrow \infty} \frac{1}{t} \int_{t_0}^{t_0+t} E_k(\mathbf{p}(\tau)) d\tau = \frac{k_B T}{2} (3N - N_c) \quad (6.6)$$

where:

$$E_k = \sum_{i=1}^n \frac{1}{2m_i} \mathbf{p}_i^2 \quad (6.7)$$

and N is the number of degrees of freedom, N_c the number of constraints, n the total number of particles, m_i and $\mathbf{p}_i$ are the mass and the momenta of the i -particle respectively.

Sometimes, it is interesting to check the effects that the increasing or decreasing of temperature generates in our simulated system. In other cases, such as phase transitions, the control has to be rigorous since this variable has to remain as a constant. Many schemes have been developed in this sense, starting from the scale of the velocities $\{\mathbf{v}\}_i$ [190], in which the temperature of the system at certain time $T(t)$ is multiplied by a factor λ . The associated change can be calculated as follows:

$$\Delta T = \frac{1}{2} \sum_{i=1}^N \frac{2m_i(\lambda \mathbf{v}_i)^2}{(3N - N_c)k_B} - \frac{1}{2} \sum_{i=1}^N \frac{2m_i \mathbf{v}_i^2}{(3N - N_c)k_B} = (\lambda^2 - 1)T(t) \quad (6.8)$$

$$\lambda = \sqrt{T_{new}/T(t)} \quad (6.9)$$

The simplest way to control the temperature is thus to multiply the velocities at each time-step by the factor $\lambda = \sqrt{T_{req}/T_{curr}}$ where T_{curr} is the current temperature as calculated from the kinetic energy and T_{req} is the desired temperature.

An alternative way to maintain the temperature is to couple the system to an external heat bath that is fixed at the desired temperature [191]. The bath acts as a source of

thermal energy, supplying or removing heat from the system as appropriate. Extended system methods, originally introduced by Nosé [192] and subsequently developed by Hoover [193], consider the thermal reservoir to be an additional degree of freedom with a certain potential energy and fictitious mass.

6.4 Charge analysis on DFT calculations

In spite of the special relevance that the atomic charges have for chemists, assigning charges to a molecule in a cluster, or to an atom into a molecule is not a trivial task. The atomic charges are not observable and therefore, not defined by quantum mechanical theory. Once we get, for instance, a converged electronic density from a quantum calculation it is not clear how we should partition it amongst the fragments of the system in order to extract charges. Many different partitioning schemes have been proposed and they can roughly be divided into two main categories: orbital-based methods, and methods based on electrostatic potential or electron density.

In the former category, the earliest of such schemes used in *ab-initio* framework was the Mulliken population analysis [194, 195]. This method partitions, in a somewhat arbitrary way, the contribution of each occupied spin-orbital to the total population of each basis function, and also the contribution to the overlap population between each pair. There are other population analysis based on electronic orbitals, for instance, the *Natural Atomic Orbitals* suggested by Weinhold [196] based on the previous work of Löwdin on the *Natural Orbitals* [197]. However, it is worth noting that atomic basis sets are overcomplete and such decomposition is not unique. In principle, a calculation could be done where all the basis functions are located on one of the atoms in the system, which would then lead to an assignment of all the electrons in the system to that atom.

One less arbitrary fashion of approaching this problem is the decomposition of the charge density proposed by Richard Bader’s theory of Atoms in Molecules (AIM) [198, 199] based on the electron density. This theory has been demonstrated to be very adequate with methods using DFT and specially with those using plane wave basis functions that are not associated with any particular atom in the system.

Within this context, each point of the space is assigned to one subsystem separated by surfaces in which *the gradient vector field of the electron density* $\nabla n(\mathbf{r})$ is orthogonal to the *normal vector of the density surface* $z(\mathbf{r})$ (they do not have normal component of the density gradient). These normal *zero-flux surfaces* $S(\Omega, \mathbf{r})$ are mathematically defined by:

$$\nabla n(\mathbf{r})z(\mathbf{r}) = 0 \quad \forall \mathbf{r} \in S(\Omega, \mathbf{r}) \quad (6.10)$$

The dominant topological property of a molecular density is that it exhibits maxima at the positions of the nuclei. The zero flux surfaces can be considered as the separation surfaces of the volume containing one maximum.

The boundary condition imposed by Bader leads to the partitioning of a molecular system into a set of disjoint spatial regions, Ω_A , each region containing a maximum (although is not necessary [200]) that corresponds to a single atom. The electronic population of a particular atom A, is thus given by:

$$N_A = \int_{\Omega_A} n(\mathbf{r}) d\mathbf{r} \quad (6.11)$$

It can be also shown that the atomic subsystems obey a local virial relation. Moreover, other properties further than atomic charge, like atomic dipole moment, atomic kinetic or potential energies, are shown to be well-defined by quantum mechanics if they are also

integrated over these atomic subvolumes [201]. The properties of the gradient vector field also contain the information needed for a definition of molecular structure and its stability by using the mathematics of qualitative dynamics [202, 203]. So that, a topological study of the density by means of the gradient density $\nabla n(\mathbf{r})$ indeed reveals a lot of information about the system.

The mapping of this vector is done by, first, localizing the *Critical Points* (local maxima, minima and saddle points) achieved by the calculation of its roots, and second, the characterization of these points via associated Hessian matrix $\mathbf{H}$:

$$\mathbf{H} = \begin{pmatrix} \frac{\partial n(\mathbf{r})}{\partial^2 x} & \frac{\partial n(\mathbf{r})}{\partial x \partial y} & \frac{\partial n(\mathbf{r})}{\partial x \partial z} \\ \frac{\partial n(\mathbf{r})}{\partial y \partial x} & \frac{\partial^2 y}{\partial n(\mathbf{r})} & \frac{\partial y \partial z}{\partial n(\mathbf{r})} \\ \frac{\partial n(\mathbf{r})}{\partial z \partial x} & \frac{\partial n(\mathbf{r})}{\partial z \partial y} & \frac{\partial n(\mathbf{r})}{\partial^2 z} \end{pmatrix} \quad (6.12)$$

Diagonalization of $\mathbf{H}$ in the critical points of the density $n(r_c)$ gives the eigenvalues h_{xx} , h_{yy} and h_{zz} and with this knowledge the different points can be characterized by their “rank” and “signature”, symbolically written as “(r,s)”. The rank r is defined as the number of non-zero eigenvalues of the Hessian and the signature s is the sign of the determinant evaluated at the location of the critical point.

The result is a theory of atoms, bonds, structure, and structural stability, where the method does not require the assignment of each basis function to a specific atom. This makes AIM particularly advantageous in situations where this assignment would be difficult or impossible, such as non-localized basis sets.

6.4.1 Finding the zero flux surfaces

Several methods have been proposed to analyze the atoms charges following the topological definition of Bader. The idea can be viewed as a “simple” two-step process in which, first, the Bader volumes are defined and after the electron density $n(\mathbf{r})$ is integrated within each region. Unfortunately, one of the biggest problems of the AIM theory is the fast and accurate calculation of these regions or equivalently their non-zero surfaces. It is indeed a mathematical challenge, specially in systems with sudden changes in the density.

Attending to how the surfaces are calculated one can find different techniques. For instance, in the early algorithms implemented for small molecules the gradient of the charge density is calculated from derivatives of an analytic wavefunction [198, 204]. These methods find stationary points in the charge density and after follow trajectories along the density gradient from these points to map out their connectivity and the zero-flux dividing surfaces as Fig.(6.4.1) shows. With the dividing surfaces represented in this way, the charge in each Bader volume can be integrated radially from the charge density maximum to the surface. However, while this approach works for small molecules, a high density of descent trajectories is needed to accurately represent the surface away from the critical points. For this reason, the method results computationally expensive for large systems and has been harshly criticized, giving rise to new approaches.

To perform the charge calculations in this thesis we have used a free software program (available in <http://theory.cm.utexas.edu>) based on the *fast and robust algorithm* of Henkelman et al. [205]. In the next pages we will describe the method, in its first approach, and we will also comment the biggest improvement [206] of the latest *Version 0.25c*.

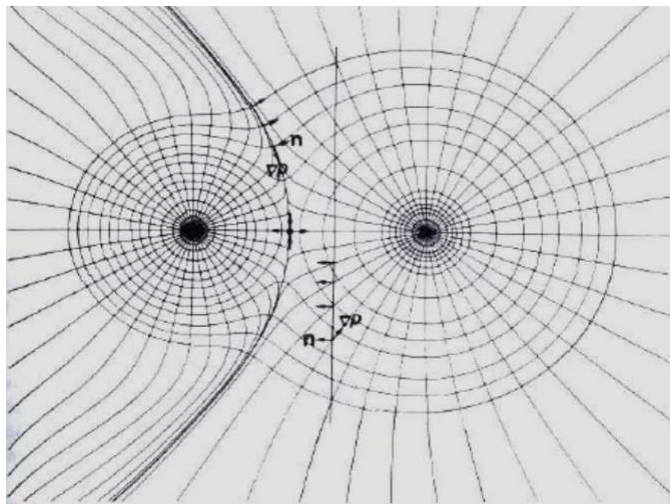


Figure 6.1: Mapping of $\nabla n(\mathbf{r})$ between two atoms. The isodensities surround the atoms while the calculated trajectories are the “radial” lines that start in the nuclei and go away. The non-zero surface is represented by the bold line

6.4.2 Henkelman’s grid-based algorithm

Typically, the output of DFT-based calculation is a continuous electronic charge density function $n(\mathbf{r})$ stored in a file a three-dimensional regular Cartesian mesh of values defined over a rectangular volume of space V . The size of this volume should be large enough so that the value of $n(\mathbf{r})$ is negligible outside this region. The original grid-based Bader analysis presented here [205] takes this density file as starting point and subdivides the volume V into M smaller volumes δV_i , $i = 1, 2 \dots M$ centered around each mesh point.

Rather than explicitly representing the interatomic surfaces, the spirit of the algorithm is to consider the problem of finding the Bader volumes under a different point of view, i.e., each volume element δV_i is assigned to a Bader region by tracing a trajectory from the center of the volume to an atomic nucleus (a maximum). What is followed during the trajectory is the direction that maximizes an approximated charge density gradient projected over the direction that links the initial grid point to one of its 26 neighbors. In this sense, all the grid points that make up the trace to the maximum are assigned to that atom. Once the whole assignation is accomplish, the total charge is given by integrating the charge density over the grid points of that region. Since the trajectory true density gradient is constrained to the mesh directions this method is also referred as “on-grid” algorithm.

The Fig.(6.4.2) has been taken from Ref. [206] and shows a visual idea of the process. On the left picture some ascent paths are shown. As an example, starting from the first point on the left of the upper row, a maximum is reached in a four step trajectory. After, a second point is chosen, for instance, the second one of the first row. In this case the trajectory leads to an already assigned point in two steps. Then, a third point is chosen and so on. As it can be seen all the points converge in two different maximums that are separated by a zero flux surface represented by the red line in the figure on the right.

A more detailed view of the procedure brakes it down into the following steps:

- (i) First an initial grid point is chosen, located by a position vector $\mathbf{r} = l\Delta x\mathbf{i} + m\Delta y\mathbf{j} + n\Delta z\mathbf{k}$; where l, m , and n are integers that indicate the discrete numbering of the mesh points and $\Delta x, \Delta y$ and Δz are the mesh spacings along each axis.

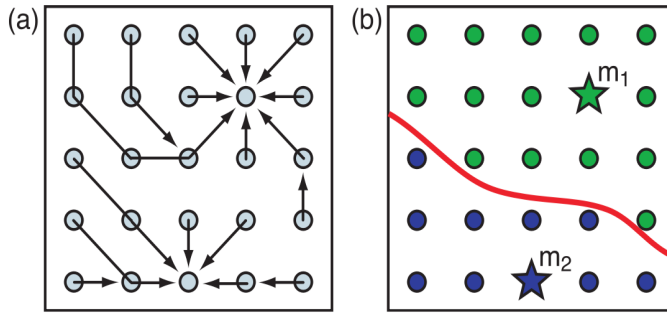


Figure 6.2: Drawing a), illustration of the steepest ascent paths traced by the Henkelman's grid-based algorithm. After all grid points are assigned (b), the set of points which terminate at each maximum (green to m_1 and blue to m_2) constitute that Bader volume. The Bader surfaces (red) separate the volumes.

- (ii) The projection of the charge density gradient along the direction to each of the 26 neighboring grid points is approximated using finite differences:

$$n_{\Delta\mathbf{r}} \equiv \frac{n(\mathbf{r} + \Delta\mathbf{r}) - n(\mathbf{r})}{|\Delta\mathbf{r}|} \quad (6.13)$$

Where $\Delta\mathbf{r}$ is the position vector to the neighbor reached by the grid step (di, dj, dk) . The integers di , dj , and dk can each take the values 1, 0, -1 (excluding $di = dj = dk = 0$). Notice that expression (6.13) is not a vector, but a scalar. It means an approximated value of the gradient density projection over the specified direction, but not a projection vector in the mathematical sense.

- (iii) The neighbor which posses the maximum projection from Eq.(6.13) is determined as the next point along the ascent path. If the point has been already assigned to a maximum then the trajectory is stopped and step (i) is re-initiated. Otherwise, the current position vector is stored in an array.
- (iv) If non of the projections are positive, the current mesh point is identified as a maximum. All the stored positions that contitute the trace to this point are attributed to it.
- (v) The process is repeated until all the volumes are assigned to a maximum.

We must insist that the algorithm does not calculate such surfaces, only their location can be infer after the assignment of grid points. The method is highly efficient, scales linearly with system size, and is also robust to complex bonding topology found in condensed systems. However, it introduces severe lattice bias specially when the true dividing surface runs at a small angle to the grid over long distances.

6.4.3 Fixing the lattice bias

The partialities on the “on-grid method” were firstly pointed out by the work of Sanville et al [207]. The gradient can follow any of the infinite directions of the whole space, however, in the algorithm its projection is constrained to only 26 of them. As a consequence, the trajectories accumulate error whenever the gradients lines do not run parallel to the lattice directions and it leads to wrong assignations, particularly near of the zero-flux surfaces. In these zones the gradient barely changes and one mistake done in a grid point is easily translated to next one. These bias can be solved increasing the grid points until each

projection finds the correct direction to project, circumstance that unfortunately occurs only in a ideal infinite grid.

Recently, W. Tang and et [206] have worked out the problem by introducing a correction vector $\Delta \mathbf{r}$ that adjusts the trajectories. The vector is defined as:

$$\Delta \mathbf{r} = \Delta \mathbf{r} + (\mathbf{r}_{grad} - \mathbf{r}_{grid}) \quad (6.14)$$

and takes into account the difference between the direction of the gradient vector, $\mathbf{r}_{grad}$, and the direction of the step just performed, $\mathbf{r}_{grid}$. The gradient vector is given by the direction that maximizes the change in the gradient density:

$$\mathbf{r}_{grad} = c(\nabla n_x, \nabla n_y, \nabla n_z) \quad (6.15)$$

evaluated at the six closet neighbors using a central finite difference:

$$\begin{aligned} \nabla n_x &= \frac{n(i+1, j, k) - n(i-1, j, k)}{\mathbf{r}(i+1, j, k) - \mathbf{r}(i-1, j, k)} \\ \nabla n_y &= \frac{n(i, j+1, k) - n(i, j-1, k)}{\mathbf{r}(i, j+1, k) - \mathbf{r}(i, j-1, k)} \\ \nabla n_z &= \frac{n(i, j, k+1) - n(i, j, k-1)}{\mathbf{r}(i, j, k+1) - \mathbf{r}(i, j, k-1)} \end{aligned} \quad (6.16)$$

with the constant equal to:

$$c = \min(dx/|\nabla n_x|, dy/|\nabla n_y|, dz/|\nabla n_z|) \quad (6.17)$$

The algorithm is particularly well suited to DFT calculations of large molecules or materials such as surfaces, alloys or molecules in solution [208–210]. It can be resume in the following steps:

- (i) Starting from an initial point, the charge density gradient is projected over the six closest neighbors following Eq.(6.16).
- (ii) The gradient vector $\mathbf{r}_{grad}$ is then calculated according to Eq.(6.17). This vectors points in the direction of the gradient density and normally runs off the grid. Since we desire to retain the trajectory on the grid, a jump to the nearest point is done. The step between them is labeled by the vector $\mathbf{r}_{grid}$ and the deviation from the true trajectory is calculated by means of the correction vector of Eq.(6.14) (initially zero).
- (iii) When the length of any component of $\Delta \mathbf{r}$ is larger than half of the grid spacing, a correction step is taken in that direction. The correction vector is then recalculated by subtracting the correction step. In this way, the true trajectory is never more than half a grid point from the current grid point in any direction.
- (iv) The ascent trajectory is ended if a charge density maximum is reached or, at some point, there is no non-assigned neighbors.

Depending upon the order in which grid points are analyzed, the grid point adjacent to the Bader surface can be assigned to one of the two volumes on either side of the dividing surface. This ambiguity is due to the fact that the trajectory between grid points deviates from the true trajectory by up to half a grid step. Thus, when all the points are assigned, a final refinement of the grid points is required to identify all grid points on the boundary of a volume.

Part III

Results

Introduction

In the present Chapter we will summarize the results obtained by applying the combination of TD-DFT MD and BO MD to the study of double ionizations of different biomolecules. The writing have been divided attending to the systems investigated:

- In Section 7: We will start by describing the fragmentation in gas phase of uracil²⁺, comparing the results with the experimental data obtained by the proton collision experiment of Moretto-Capelle et al. in Toulouse (France) [211].
- In Section 8: The theoretical study of uracil²⁺ in liquid phase will be presented. The comparison with gas phase results will allow us to estimate how the presence of water affects the fragmentation processes. Besides, we will analyze the effects of the ionization when it is produced in a surrounding water molecule near the uracil.
- In Section 9: We will investigate the double ionizations of small water clusters compared with the experimental data obtained by the COLTRIMS experiment [212] carried out in Caen (France). As we will see our theoretical calculations are able to describe the fast mobility occurring in the clusters after the collision.
- In Section 10: Preliminary results of double ionizations of pyrimidine will be presented.

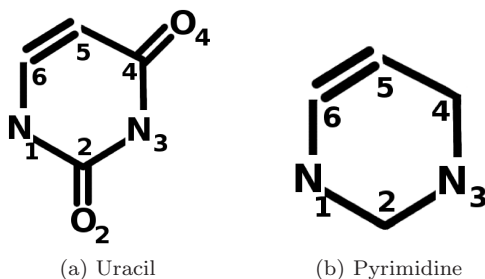


Figure 6.3: Labels used for uracil and pyrimidine

The labels used for atoms of the systems studied are displayed in Fig.(6.3). All dynamics have been run using the Car-Parrinello Molecular Dynamics code CPMD 3.13.2 [168] with the special TD-DFT MD Ehrenfest implementation of Tavernelli et al. [44]. Other conditions such as supercell size, time-step or cutoff basis set will be given for each particular case. We have also carried out all-electron calculations using Gaussian09 package [213] using standard Pople basis sets [214–218] in order to compare results and extract conclusions. In both, plane waves and gaussian-based calculation, the BLYP exchange-correlation functional [129, 133] have been used.

Chapter 7

Fragmentation of Uracil²⁺ in the gas phase

Abstract *We apply a combination of Time-Dependent Density Functional theory and Born-Oppenheimer molecular dynamics methods to investigate fragmentation of doubly-charged gas-phase uracil produced in collisions with 100 keV protons. The results are in good agreement with ion/ion coincidence measurements performed in parallel to this study. Fragmentation mainly arises from inner shell two-electron vacancies. Orbitals of similar energy and/or localized in similar bonds can lead to very different fragmentation patterns, thus showing the importance of the intramolecular chemical environment. The observed fragments almost never correspond to the energetically most favorable dissociation paths, which is due to dynamical effects occurring in the first few femtoseconds after electron removal.*

7.1 Introduction

To unravel the mechanisms of the early stages of fragmentation after the radiation of biological tissues by highly charged ions, physicists have performed numerous experiments in gas phase (see e.g. [48, 49, 219]), in which swift charged ionic projectiles impinge on DNA/RNA bases, sugars, nucleotides or even biomolecular clusters. In contrast with experiments performed in solution or directly on living systems, gas phase experiments provide direct and precise information on *single* collision events. Thus they allow one to unambiguously identify fragmentation channels associated with a given biomolecule and not with the environment. Such detailed information can be achieved by combining techniques that are state-of-the-art in gas phase chemistry, e.g., high resolution mass spectrometry, and in collision physics, such as multicoincidence detection techniques that provide the correlation between different charged fragments as well as their relative kinetic energies and momenta.

Fragmentation results from relaxation of the excess electronic energy associated with vacancies created in the different electronic shells of the molecule. Since, in these collisions, electrons can be removed from many of these shells, experiments cannot tell us how fragmentation depends (i) on the shape/energy of the molecular orbital (MO) in which the electron vacancies are created, and (ii) on the intramolecular environment, i.e., on the neighboring functional groups. This information can only be obtained from *ab-initio* MD calculations such as those based on TD-DFT methods [170], which have been already applied to investigate the radiolysis of bulk water [45] and water clusters [212]. In this work we apply

TD-DFT MD to investigate the fragmentation of doubly ionized uracil. The study has been done in close collaboration with the experimental group of Patrick Moretto-Capelle in Toulouse who has studied the fragmentation of uracil after collision with 100 keV protons. We will center the discussion in the theoretical results, although the experiments done in the group in Toulouse will be also explained.

Comparison between theory and experiments is important to show that not only theoretical results are compatible with experimental ones, but also that fragmentation depends critically on the energy, shape and intramolecular chemical environment of the ionized MO. This suggests that, by targeting specific MOs, damage can be induced rather selectively.

The choice of uracil (an RNA basis) follows a rich tradition in experimental collision physics [46–52]. Uracil was the first biomolecule to be investigated in this context and is often considered as a benchmark due to its simplicity and high stability in gas phase (which prevents, e.g., its thermal degradation). The study of uracil in gas phase is also important to understand the early presence of RNA bases on Earth (e.g., the fact that uracil can be formed in outer space suggests an exogenous origin of life [220]). The most detailed experiments on uracil have been performed with proton projectiles of 20–150 keV [46, 49, 50]. In particular, Tabet et al. [49] have reported relative cross sections for electron capture and direct ionization, and the corresponding fragmentation branching ratios. Therefore, H^+ /uracil collisions are ideal (i) to check the accuracy of existing theoretical methods and (ii) to obtain, with the help of the latter, dynamical information that experiments alone cannot provide.

The initial double charged state is obtained as it was explained in Chapter 5, i.e., by removing a double occupied Kohn-Sham (KS) orbital of the neutral species. We consider the ionization from inner orbitals, namely, the three lowest ones obtained in the KS pseudopotential representation of neutral uracil described in Fig.(7.1) (9a', 10a' and 11a', hereafter called KS1, KS2 and KS3), and from the valence highest occupied molecular orbital 5a'' (HOMO). The corresponding energies of KS1, KS2, KS3 and HOMO orbitals are -27 eV, -26 eV, -24 eV and -6 eV respectively. KS1 and KS2 orbitals are associated with localized σ -like C_2 -O and C_4 -O bonds in a different intramolecular chemical environment: while KS1 lies between two N atoms, KS2 lies between N and C atoms. KS3 is delocalized over the C_2 -O bond and part of the uracil ring. The HOMO is even more delocalized since it is extended over half of the ring and the two O atoms bonded to it. This orbital sampling allows us to investigate not only the influence of the energy and shape of the ionized MO on the fragmentation processes, but also the effect of the chemical environment on a particular bond.

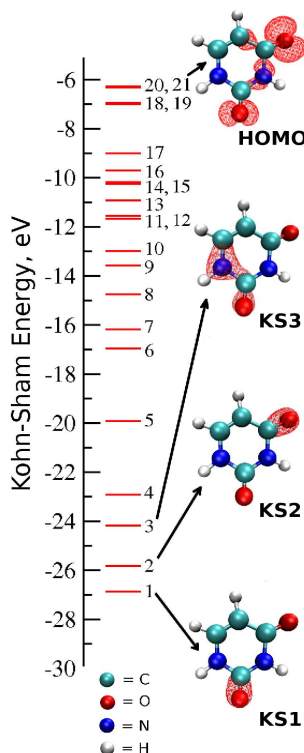


Figure 7.1: Energies of the occupied KS orbitals of neutral uracil and electron densities associated with the KS1, KS2, KS3 and HOMO orbitals (0.018 isovalue).

Once the initial electron density is built up, the fragmentation dynamics is performed as it was explained in Chapter 5. Briefly, due to the very small time step used in the real time propagation (of the order of attoseconds), the total simulation time is necessarily limited to less than 100 fs. Since the complete dissociation of the ionized molecule requires in general longer times (a few ps), we then switch to ground-state BO MD when this dynamics becomes essentially adiabatic (i.e., when Ehrenfest and BO MD produce the same trajectory). Fragmentation associated with HOMO vacancies is studied by using standard Car-Parrinello MD from the very first instant after ionization and thermal equilibration by velocity scaling (in this case, fragmentation is so slow that TD-DFT MD is prohibitively expensive). For all simulations presented here we use a cubic box of size $L=19$ Å and a plane wave basis with an energy cutoff of 70 Ry.

For the discussion of the results, we will present first the TD-DFT MD simulations corresponding to KS1, KS2 and KS3 ionization of uracil. Afterwards, we will summarize the data obtained while the dynamics were extended to longer times using BO MD. At this point, we will also comment HOMO ionization since it has been performed using the same ground-state techniques. We will end the section by comparing the theoretical results with the experimental ones obtained in proton collision experiments.

7.2 Dynamics of fragmentation

First of all we will analyze how the total electronic energy (Kohn-Sham energy) and the temperature of the systems evolve during the the fragmentation. Fig.(7.2) displays these variations for KS1, KS2 and KS3 TD-DFT MD. Additionally, the dashed line in the lower panel gives the total energy for KS1 ionization.

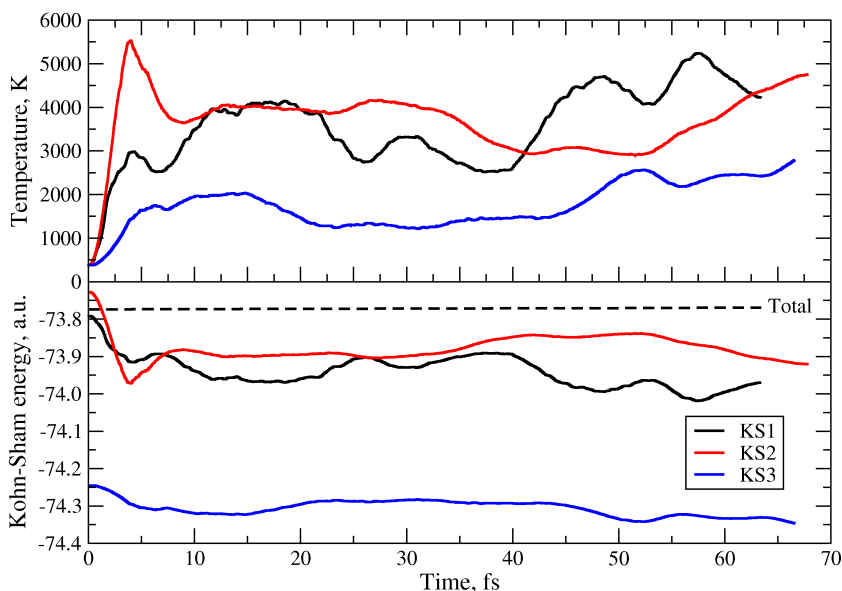


Figure 7.2: Kinetic energy of the nuclei and Kohn-Sham energy of the system during the KS1 (black), KS2 (red) and KS3 (blue) TD-DFT MD. The dashed line gives the total energy of the system for KS1 ionization.

KS1 and KS2 ionizations represent a much more energetic processes as their initial electronic energies are ≈ 0.5 a.u. (13.6 eV) above the energy of KS3. In both cases the Kohn-Sham energy undergoes a change of 0.12 and 0.24 a.u. (3.35 and 6.63 eV) in the first 4 fs of the dynamics for KS1 and KS2 respectively. These sudden variations really require the small time-step of integration used, otherwise the total energy of the system would not be conserved and the dynamics might be biased. Regarding the temperatures of the atoms, they increase due to the exchange of potential energy into kinetic energy during the first fs of dynamic (noticing however that the total energy remains constant). Particularly, KS2 ionization has a maximum peak of 5500 K at about 4 fs that, as we will see later, is related to the oxygen loss.

These three fragmentations dynamics are schematically summarized in Figures (7.3), (7.4) and (7.5). For all of them, first frame is taken at the beginning of the dynamics with the ionized orbital represented by the red mesh. Main distances are indicated in black and mass and charge of fragments are given in blue.

Fragmentations are different even for shape-equivalent orbitals like KS1 and KS2. Dynamics of the deepest orbital ionization (KS1, Fig.(7.3) leads to three fragments: CO, HNCCH₂ and OCNH of masses 28/41/43; and charges 0.3/1.1/0.6. In this case, the uracil²⁺ ion undergoes the first bond breaking of N₁-C₂ at 14 fs. Bond C₄-C₅ is dissociated at ≈ 32 fs, and soon after, C₂-N₃ also breaks (≈ 34 fs). At 35 fs the three fragments are formed and they remain unaltered until the end of the dynamics. As can be seen, the original CO bond from which the electrons were removed is not broken.

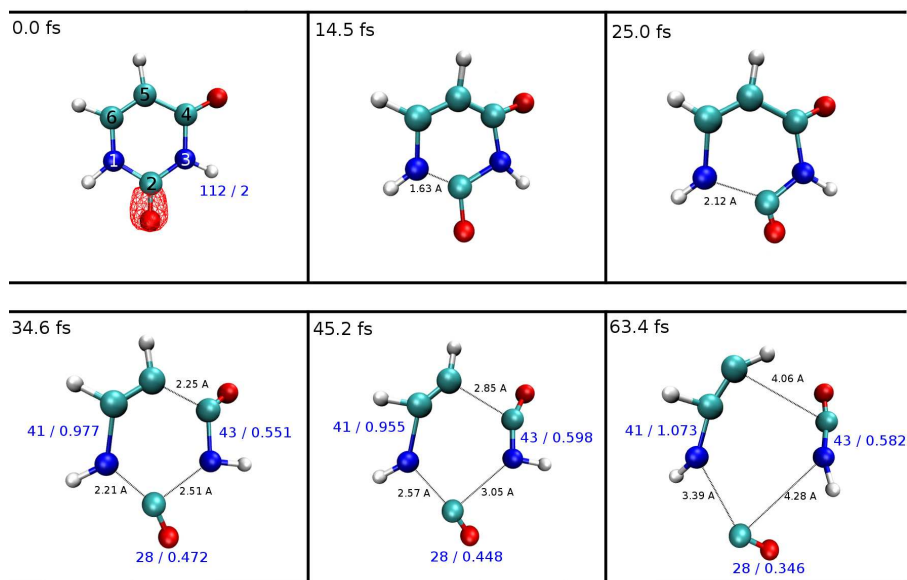


Figure 7.3: Snapshots of the TD-DFT MD trajectory after the KS1 ionization

Concerning the charges, fragment 41 preserves a charge of 1+ from almost the beginning. Fractional charges of 28 and 43 fragments, however, must be taken with some caution. As it is well known, DFT methods can lead to unphysical fractional charges. An example is H₂⁺, for which TD-DFT predicts dissociation into H^{+0.5} + H^{+0.5}, instead of into either H + H⁺ or H⁺ + H. Thus, in general, calculated fractional charges of the order of +0.5 are an indication of the existence of two dissociation paths in which the charge is asymmetrically distributed between two fragments. That suggests the formation of a singly charged fragment with mass 28 and a neutral fragment with mass 43, and the other way around.

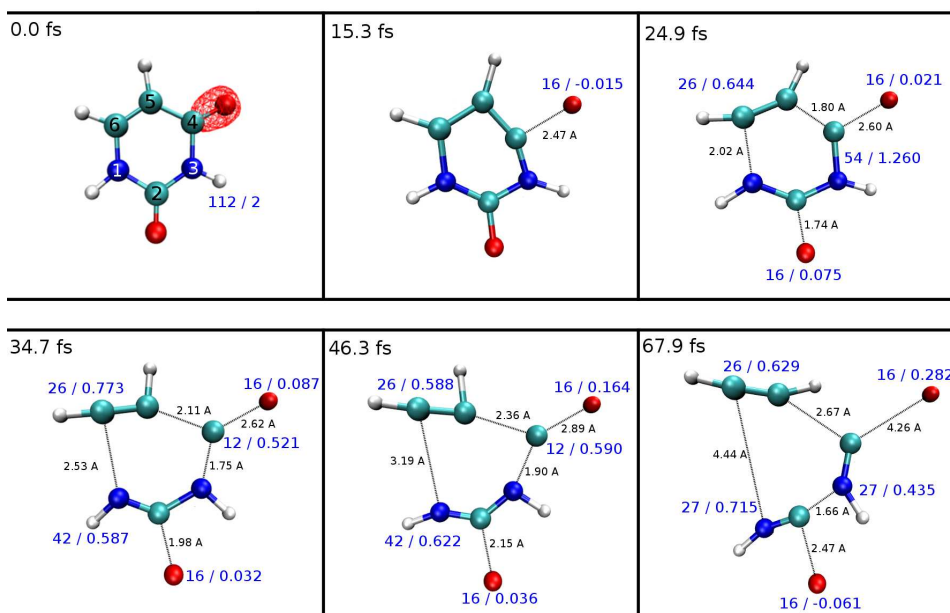


Figure 7.4: Snapshots of the TD-DFT MD trajectory after the KS2 ionization

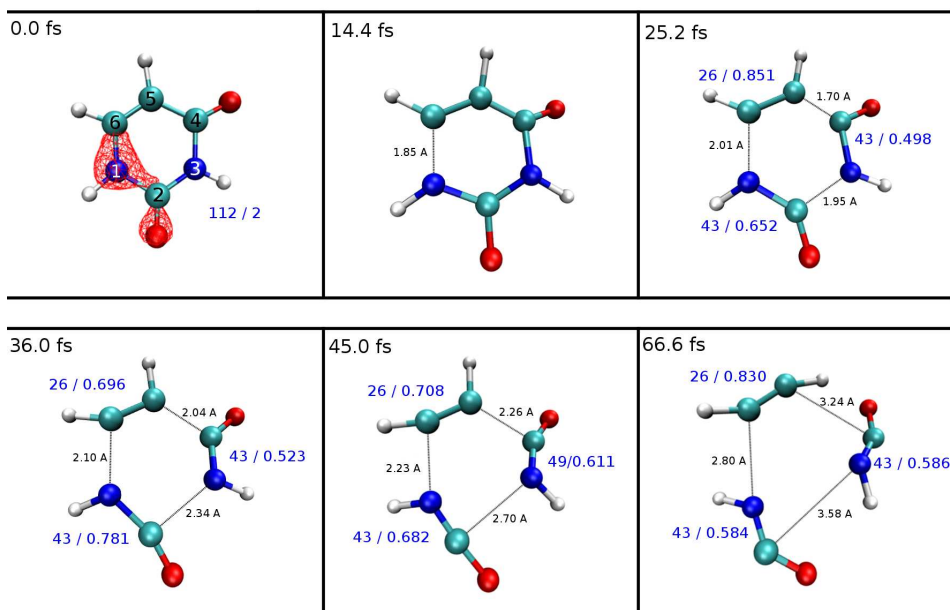


Figure 7.5: Snapshots of the TD-DFT MD trajectory after the KS3 ionization

The dynamics starting from the KS2 ionization, see Fig.(7.4), represents the most fragmented case since it leads to two oxygen atoms and three fragments C_2H_2 , and two CNH of masses 26/27/27. The C_4 -O bond is broken at 4 fs and the O4 atom is consequently detached with a charge that fluctuates between -0.1 and +0.3 during the trajectory. At 24 fs, a fragment of mass 26 (C_2H_2) can be considered formed since the bonds N_1 - C_6 and C_4 - C_5 are broken, and it has a charge of +0.5. The fragment that remains loses a neutral oxygen atom at 28 fs. Finally the big fragment of mass 54 is broken leading to two equivalent fragments 27/27 sharing equally the charge.

Ionization of KS3 produces one fragment of mass 26 and two equivalent fragments of 43. First bond breaking of mass N_1 - C_6 occurs at 13 fs while C_4 - C_5 and N_1 - C_6 break almost simultaneously at about 22 fs. The charge +2 is distributed amongst the fragments in fractional charges.

In our approach, time-dependent simulations are performed approximately during the first 60 fs that follow the ionization of the target molecule. After this time, a ground-state dynamics is introduced starting from last configuration (positions and velocities) coming from the end of the TD-DFT MD. The change from one method to the other is done once the system evolve with the same trajectory under both schemes. Fig.(7.6) shows one example of this process in the case of KS1 ionization. Distances C_2 - N_3 and C_2 - N_1 are monitored during the TD-DFT MD simulation and are given by solid and dashed black lines respectively. At different times of the TD-DFT simulation the position and velocities of the atoms have been taken as an input for different BO MD. The respective C_2 - N_3 and C_2 - N_1 distances measured during these ground-state dynamics are represented by different colors in the figure. The upper set of color lines, that represents the C_2 - N_3 distance, follows almost the same trend as TD-DFT MD (given by the black line). However, the lower one (C_2 - N_1 distance) shows different behavior than the TD-DFT MD when the change is done at early times (leading for instance to the rearrangement of two fragments). When the BO MD is started at 62 fs, both distances follow the same trend as TD-DFT MD keeping the same number of fragments.

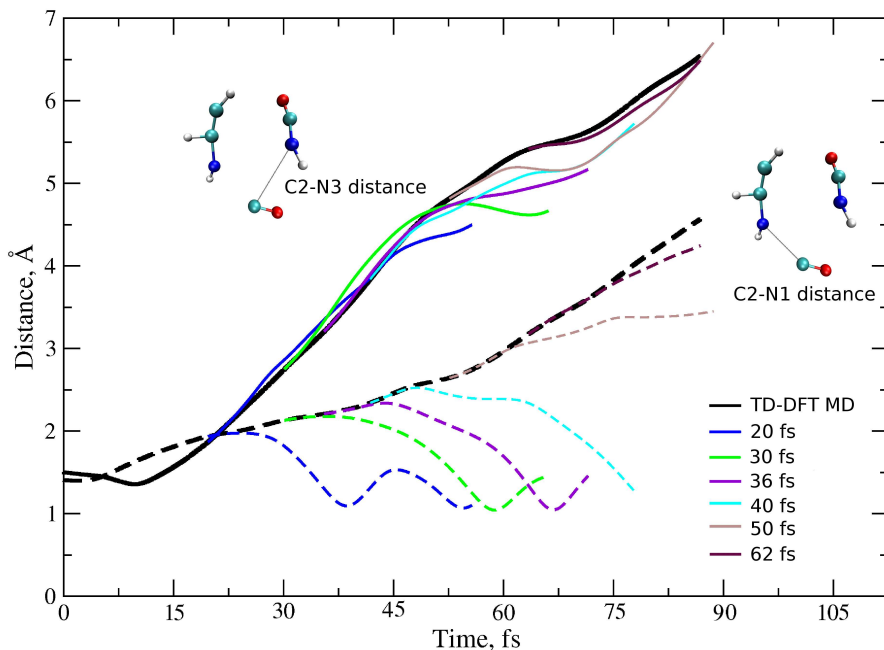


Figure 7.6: C_2 - N_3 and C_2 - N_1 distances are followed during the TD-DFT simulation and several BO MD started at different times (see text for details).

The results of the theoretical calculations at the end of the TD-DFT MD + BO MD simulation are summarized in Fig.(7.7). During the BO dynamics of KS1 ionization, the fragment of mass 41 has a charge of about +1 during the trajectory (as it happened during the TD-DFT MD simulation), nevertheless, the atom H6 of C₆ is transferred to C₅. Optimizations using all-electron calculations revealed that fragment 41 charged 1+ is stable only if H6 is bonded to C₅, otherwise it does not represent a minimum. The rest of the charge is shared between fragments of mass 28 and 43 that have charges close to +0.5. Thus, KS1 ionization is compatible with the 28/41 and 41/43 coincidences.

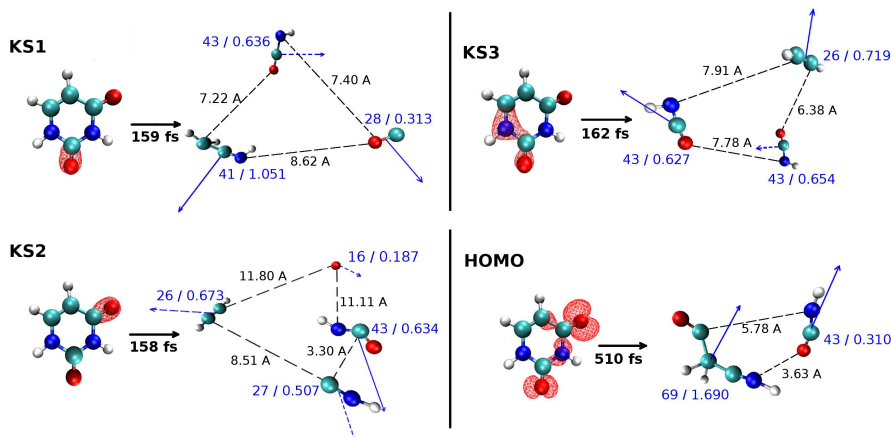


Figure 7.7: Fragments observed at the end of the TD-DFT MD + BO MD of KS1, KS2, KS3 and HOMO (2800K) ionizations. Their corresponding M/Z ratios and center-of-mass velocities (dashed arrows point inwards); fragment distances (dashed lines) are given in angstroms.

It has to be indicated that switching from TD-DFT MD to BO MD decreases the total energy of the system. In the particular case of KS1 ionization this decreasing is 1.2 eV. This variation is reflected in the change of the slope of the kinetic energy of the fragments as Fig.(7.8) shows.

In the case of KS2 ionization, one of fragments of 27 mass generated during the TD-DFT MD simulation rotates in such a way that C₄ faces the second oxygen atom detached (O₂). In consequence O₂ is finally bonded to C₄. Fragments with masses 26, 27 and 43, and charge of the order of +0.5 are produced. None of these fragments are seen in KS1 fragmentation, showing that chemical environment plays a crucial role. Thus, results for KS2 are compatible with coincidence signals at 26/27, 26/43 and 27/43.

Finally, for KS3 ionization, former TD-DFT MD fragments are maintained: A charged fragment with mass 26 plus two fragments with mass 43 and fractional charges. Therefore a 26/43 coincidence event should be observed.

HOMO fragmentation have been performed by standard CP MD since as it is less energetic and, consequently, takes much longer time. For instance, no fragmentation is observed in CP MD simulations at 350 K extended up to 5 ps. For this reason, we have also performed calculations at 2300 and 2800 K, so that the initial average kinetic energy per atom is comparable to the energy available when two electrons are removed from an inner shell. In both cases, the fragmentation occurs in ≈ 0.5 ps. The dissociation products are the same for both temperatures leading to a 43/69 coincidence. Finally, Table (7.1) gathers all the relevant information extracted from the simulations to compare with the experimental results.

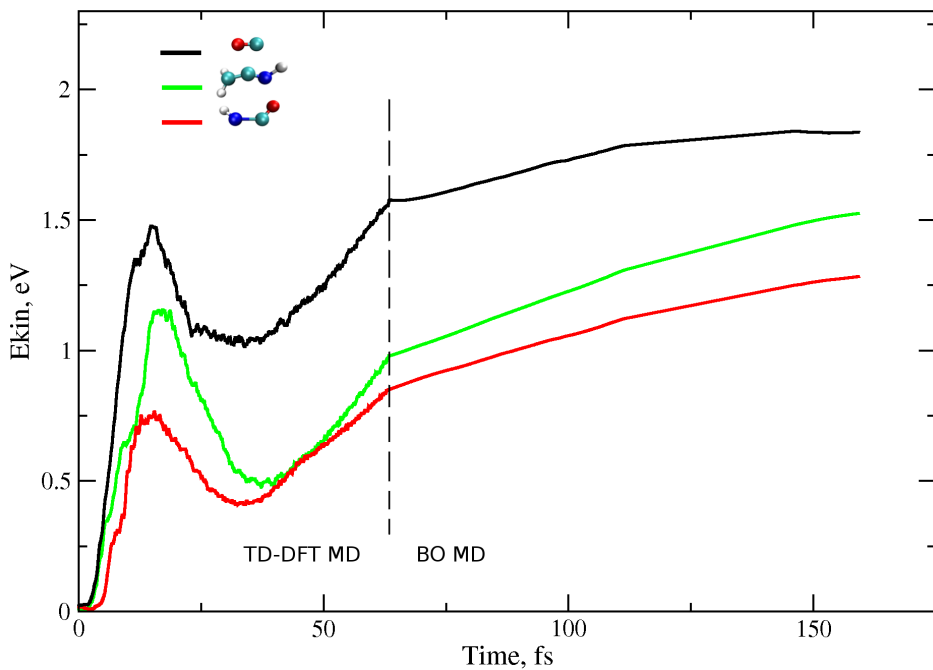


Figure 7.8: Kinetic energy of the fragments during the fragmentation dynamics of KS1 ionization. The dashed line points out the time at the change between both methods is produced.

Table 7.1: Mass, charge and kinetic energy of the fragments obtained at the end the dynamics performed in this work.

Orbital	Fragment	Mass	Charge	Ekin (eV)
KS1	CO	28	0.313	1.836
	HNCCH ₂	41	1.051	1.526
	OCNH	43	0.636	1.284
KS2	O	16	0.187	2.011
	HCCH	26	0.673	1.997
	CNH	27	0.507	0.856
	OCNH	43	0.634	0.601
KS3	HCCH	26	0.719	1.385
	OCNH	43	0.654	1.228
	OCNH	43	0.627	1.591
HOMO	HNCCH ₂ CO	69	1.690	0.214
	OCNH	43	0.310	0.309

7.3 Comparison with experiment

The experimental setup used has been described in detail in Ref. [221]. Briefly, a pulsed proton beam of 100 keV and 2 mm diameter collides an effusive uracil gas jet produced by heating commercial powder purchased at Sigma-Aldrich at a temperature of 120-150 °C. Charged fragments produced in the collision are mass-over-charge (M/Z) analyzed by means of a time-of-flight mass spectrometer operating in second order space focusing with pulsed extraction.

The fragments are detected by a high efficiency multi channel plate (MCP) assembled with a secondary electron repeller grid in front of the first MCP. The time of flight (TOF) of all fragments created by one ion pulse is measured by a multistop time device and stored in an event-by-event mode. The TOF is derived from $T = T_0 \sqrt{p/(qE)}$, where T_0 (time of flight of the fragment with no initial velocity) is proportional to M/Z , p is the projection of the momentum along the cell axis, and E is the extraction field.

Fig.(7.9) shows the time-of-flight spectrum of singly charged fragments and of two singly charged fragments measured in coincidence. Apart from the obvious difference in intensity, both spectra exhibit similar trends, with pronounced peaks for M/Z (in amu) 28, 42, and 69. The spectrum for non coincident singly charged fragments is very similar to that obtained with less energetic projectiles [222, 223]. This suggests that bond breakage is only efficient at specific locations within the molecule.

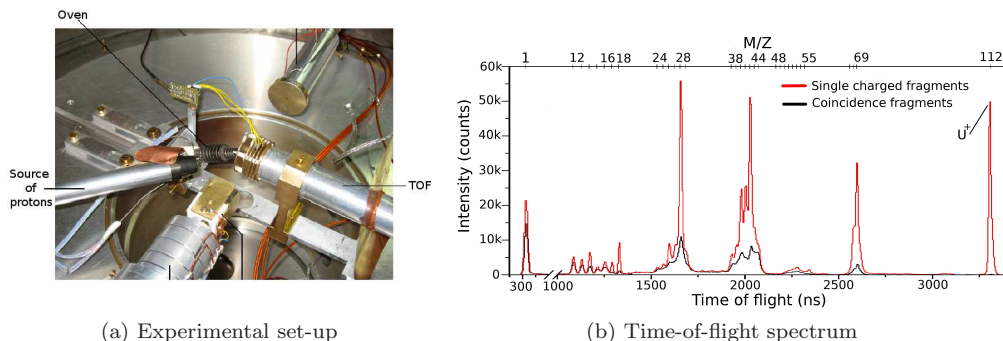


Figure 7.9: On the left part (a), picture of the experimental set-up in which the oven, source jet and TOF detector are pointed out. On the right side (b), time-of-flight spectra of singly charged uracil fragments (red line) compared to that of charged fragments measured in coincidence (black line). The peak at 112 amu corresponds to uracil⁺.

Comparison with theoretical and experimental results have been done through two figures. The first one, Fig.(7.10), shows the agreement between the mass 2D ion/ion coincidence spectrum with theoretical mass coincidences of Table (7.1) given as junction of lines. The second figure, Fig.(7.11), shows the TOF 2D ion/ion coincidence spectrum where the signature of uracil²⁺ fragmentation has been zoomed (apart from H⁺ correlations, coincidences for $M/Z \leq 16$ are much less intense). Combining the spectrum with the theoretical velocities for each fragment predicted on Table (7.1), we have generated a theoretical 2D TOF spectrum that is superimposed by color symbols. In order to obtain this spectrum, all the momenta of charged fragment are extracted directly from the theoretical values and are rotated in space randomly because of the isotropy (this is done many times in order to generate “all” the molecular orientations via Monte Carlo simulation). Only component along the axis cell is kept giving the a “spot” that, for the shake of clarity, it is limited to its boundary given as color symbols in the figure.

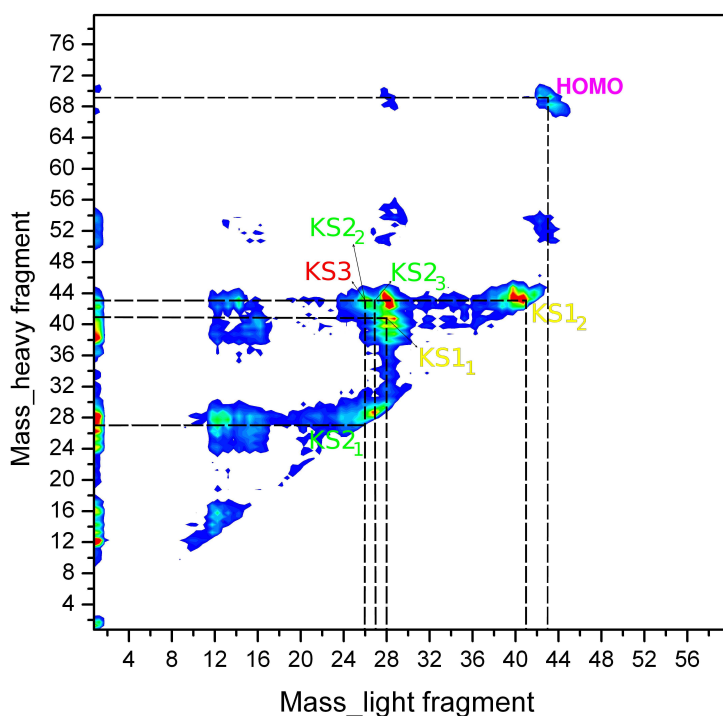


Figure 7.10: 2D coincidence mass spectrum of charged fragments after the proton-uracil collision. Results for KS1, KS2, KS3 and HOMO are pointed out by color letters.

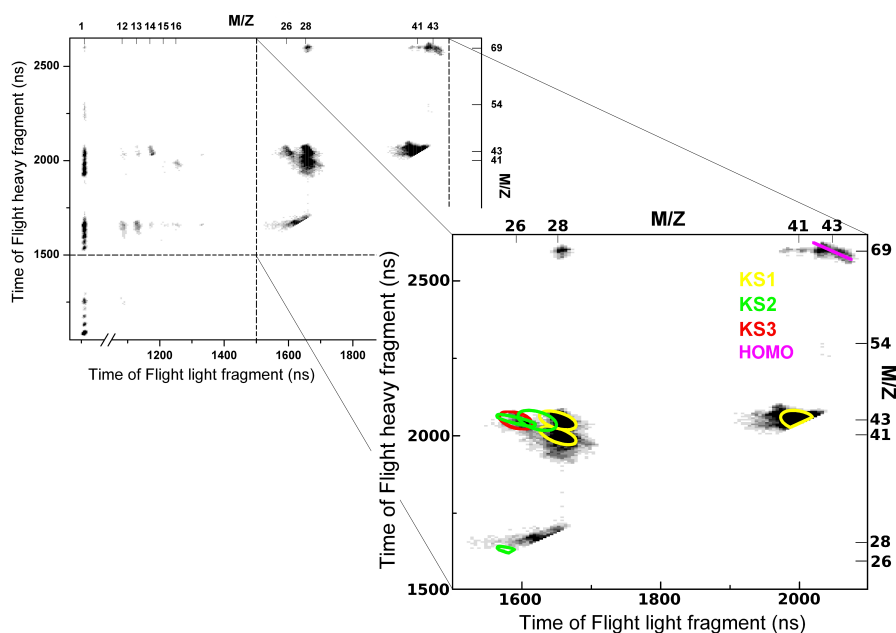


Figure 7.11: 2D TOF coincidence spectrum of charged uracil. In the inset, fragments for $M/Z \geq 22$. Theoretical results for KS1, KS2, KS3 and HOMO fragmentations are represented by open symbols.

The calculations predict the most intense coincidence signals, as well as (in most cases) their shape in the TOF distributions. This suggests that not only the fragments but also their relative velocities are correctly described by theory.

In order to know which is the most thermodynamical stable path, we have analyzed several fragmentation channels with all electron calculations. The information is presented in the Table (7.2) where the geometry of the fragments have been separately optimized at BLYP/6-311+G(d,p) and final energies evaluated at CCSD(T)/6-311++G(3df,2pd) level. The total energy of each channel has been determined by adding the electronic energies of the fragments and compared with that of uracil²⁺.

Table 7.2: Energies of different fragmentation channels calculated at CCSD(T)/6-311++G(3df,2pd)//BLYP/6-311+G(d,p) level. The total energy of each channel has been determined by adding the electronic energies of the fragments and compared with that of uracil²⁺ (E_{rel}).

	Fragments	Signal	E (eV)	E_{rel} (eV)
0	Uracil ²⁺	—	-11257.94989	—
1	HNCCH ₂ ¹⁺ + OCNH ¹⁺ + CO	41/43	-11246.77234	11.18
2	HNCCH ₂ ¹⁺ + OCNH + CO ¹⁺	41/28	-11244.36876	13.58
3	HNCCH ₂ + OCNH ¹⁺ + CO ¹⁺	43/28	-11241.53764	16.41
4	HCCH ¹⁺ + CNH ¹⁺ + O(NH)C + O	26/27	-11232.92048	25.03
5	HCCH ¹⁺ + CNH + O(NH)C ¹⁺ + O	26/43	-11233.80076	24.15
6	HCCH ¹⁺ + CNH + O(NH)C + O ¹⁺	26/16	-11231.46252	26.49
7	HCCH + CNH ¹⁺ + O(NH)C ¹⁺ + O	27/43	-11233.16764	24.78
8	HCCH + CNH ¹⁺ + O(NH)C + O ¹⁺	27/16	-11230.82940	27.12
9	HCCH + CNH + O(NH)C ¹⁺ + O ¹⁺	43/16	-11231.70968	26.24
10	C ₂ H ₂ ¹⁺ + OCNH ¹⁺ + OCNH	26/43	-11243.72892	14.22
11	C ₂ H ₂ + OCNH ¹⁺ + OCNH ¹⁺	43/43	-11243.50313	14.45
12	HNCCH ₂ CO ¹⁺ + OCNH ¹⁺	69/43	-11258.20567	-0.26
13	HNCCH ₂ CO ²⁺ + OCNH	—	-11244.60809	13.34
14	HCCH ¹⁺ + CO ¹⁺ + OCNH + NH	26/28	-11237.55032	20.40
15	HCCH ¹⁺ + CO + OCNH ¹⁺ + NH	26/43	-11239.95389	18.00
16	HCCH ¹⁺ + CO + OCNH + NH ¹⁺	26/15	-11238.06003	19.89
17	HCCH + CO ¹⁺ + OCNH ¹⁺ + NH	28/43	-11237.32454	20.63
18	HCCH + CO ¹⁺ + OCNH + NH ¹⁺	28/15	-11235.43068	22.52
19	HCCH + CO + OCNH ¹⁺ + NH ¹⁺	43/15	-11237.83425	20.12
20	HNC ₃ H ₂ ¹⁺ + OCNH ¹⁺ + O	53/43	-11241.50292	16.45
21	HNC ₃ H ₂ ¹⁺ + OCNH + O ¹⁺	53/16	-11239.63762	18.31
22	HNC ₃ H ₂ + OCNH ¹⁺ + O ¹⁺	43/16	-11236.69055	21.26
23	HNCCH ₂ CO ¹⁺ + CNH ¹⁺ + O	69/27	-11240.50512	17.44
24	HNCCH ₂ CO ¹⁺ + CNH + O ¹⁺	69/16	-11239.04716	18.90

Regarding the four ionization studied, the energetically most favorable dissociation path is HNCCH₂CO⁺ + OCNH⁺. However, this channel is only seen in the HOMO fragmentation dynamics. Thus, inner shell vacancies favor dissociation channels that lie much higher in energy but are dynamically favored. KS1 ionization corresponds to 1, 2 channels in Table (7.2); KS2 ionization to 4, 5 and 7 channels and KS3 to channel 10.

These channels are associated with fission barriers of the order of 2-5 eV (see Fig.(7.12)), but the barriers play a minor role due to the large amount of available energy (of the order of 50 eV). Therefore, one can conclude that the experimental observations are mainly due to sophisticated electronic dynamical effects taking place during the first few fs after removal of inner shell electrons.

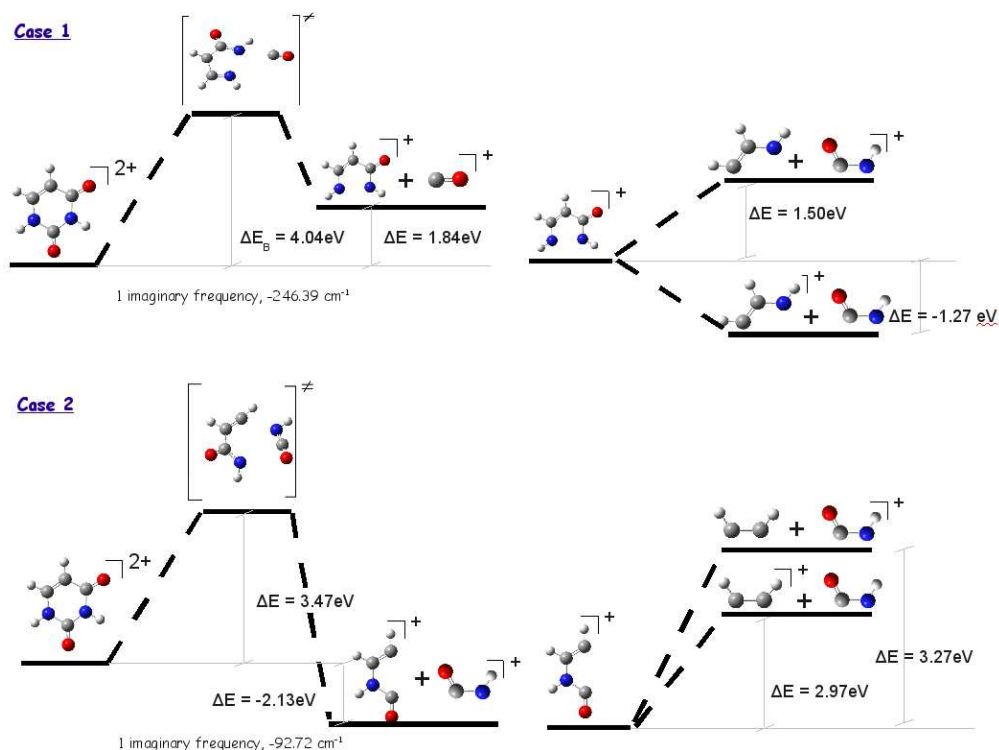


Figure 7.12: Calculations of transitions states and energies of the fragments for the different fragmentations observed in KS1 and KS2 ionization. Calculations have been done at the BLYP/6-311+G(d,p) level of theory

In summary, a combination of TD-DFT MD and BO MD methods shows that two-electron removal from uracil leads to results in good agreement with ion/ion coincidence experiments in which gas-phase uracil is bombarded with 100 keV protons. The calculations show that fragmentation induced by vacancies in MOs of similar energy and/or localized in similar bonds strongly depends on chemical environment, and cannot be understood in terms of energetic stability of the fragments.

Chapter 8

Double ionization of Uracil in the liquid phase

Abstract *We apply a combination of Time-Dependent Density Functional Theory and Born-Oppenheimer molecular dynamics to study the early stages that follow the double ionization of uracil in liquid water. Ionization from equivalent Kohn-Sham orbitals than in gas phase lead to different fragmentation patterns. The solvent is revealed as an active participant in the dissociation process driving electron density from outer hydration shells to uracil. We have also studied the double ionization of a surrounding water molecule of uracil. Ionizations either from one water molecule or the uracil lead to the formation of an atomic oxygen as a direct consequence of the molecule Coulomb explosion.*

8.1 Introduction

We investigate the double ionization process of uracil immersed in water in two cases: Direct ionization of the uracil, and ionization of the solvent. To that end, we set-up a system composed by one uracil molecule surrounded by 49 water molecules (total number of 159 atoms) in a cubic box of $L=11.50$ Å. A system of this size reproduces the density of liquid water at standard thermodynamics conditions, introducing two complete layers of solvent and part of the third hydration shell around the solute [224].

The exact number of molecules have been carefully considered following the preparation procedure of Ref. [224, 225]. Briefly, the cubic cell was first filled with water molecules only at ambient density of bulk liquid and the pressure was determined in a 500 ps run at a temperature of 300 K using standard classical force field MD. Afterwards, a cluster of H_2O molecules was replaced by the solute (uracil) and the pressure recomputed. Taking the pressure of the pure water sample as the target, the number of water molecules was adjusted until this pressure was recovered. The final configuration of this classical simulation was then used as the starting point for the TD-DFT Molecular Dynamics simulations

Four cases are studied removing two electrons from different molecular orbitals, namely: two of them involve the removal of two different orbital localized on the uracil molecule; in the other two, the ionization concerns orbitals of a neighboring water molecule situated 2.97 Å away from uracil (indirect effect). We will start off describing the two direct cases leaving the water ionizations to the end of this section.

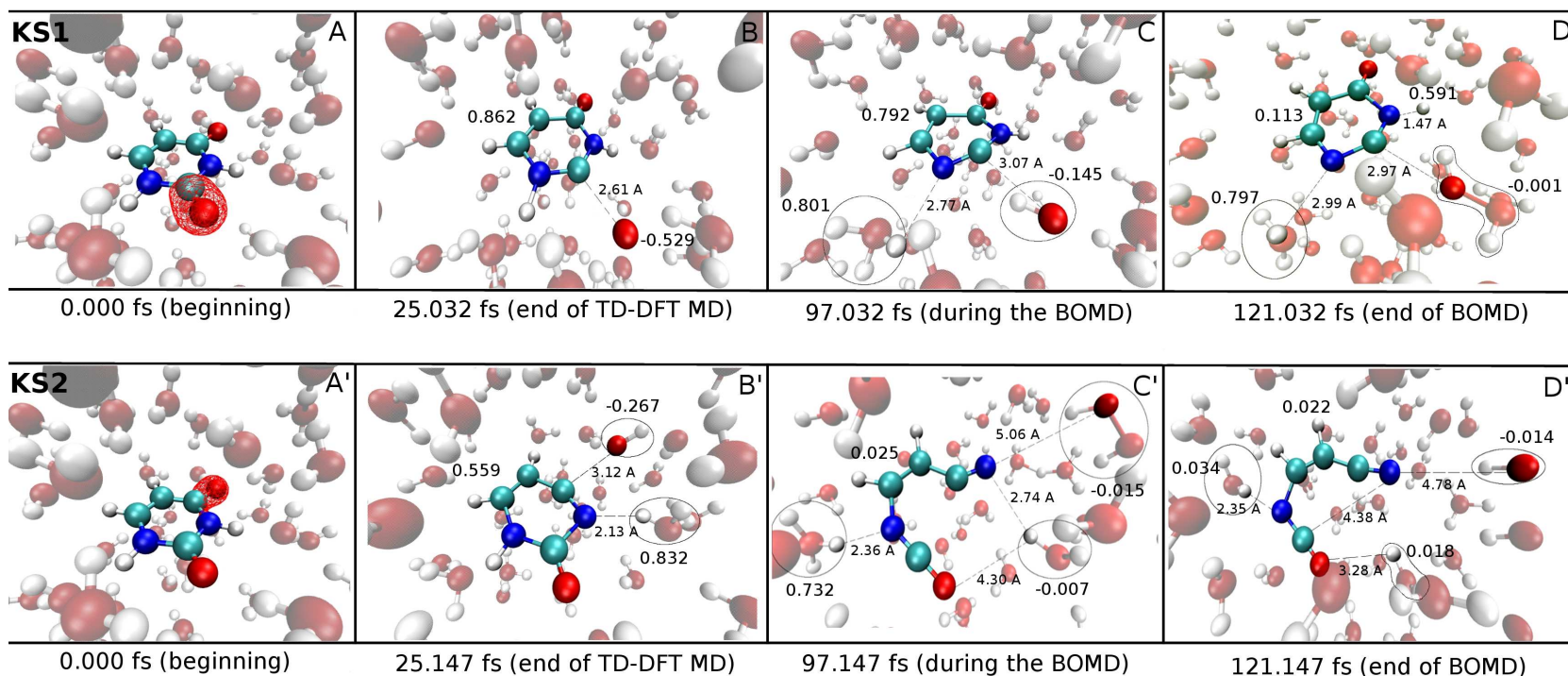


Figure 8.1: Snapshots corresponding to KS1 (upper strip) and KS2 (lower strip) dynamics in liquid phase. **A** panels correspond to the beginning of the simulation in which the isodensity associated to each of the removed orbitals is represented (isovalue 0.01). **B** panels are taken from the last step of the TD-DFT MD simulations showing in both cases a release of atomic oxygen. **C** and **D** panels represent intermediate and final steps of the BO MD. Some distances (in Angstrom) and charges of the most representative species are given. For further details see the text.

For all simulations, the wavefunction is represented using Periodic Boundary Conditions and a cutoff of 70 Ry. Core electrons are replaced by pseudopotentials of the standard Troullier-Martins form [160], giving a total number of 436 electrons described explicitly. The Kleinman-Bylander [226] integration scheme is used for all atoms.

As it was shown in Chapter 5, calculations in the liquid phase are highly demanding. By including the solvent the TD-DFT MD propagation becomes 2 times more expensive than in the gas phase. Nevertheless, in the liquid we expect shorter relaxation times to the ground-state than in the gas phase since the excited uracil²⁺ is able to exchange energy with the surrounding molecules. Considering both factors, time-dependent density propagations are extended to 25 fs in the case of uracil ionizations, and to 10 fs in the case of water ionizations. For the time scale of the later we have also considered the previous work of Tavernelli et al. [45], where it has been shown that double ionization of one water molecule in liquid water leads to the formation of one oxygen atom and two protons in a ultra-fast process that lasts less than 4 fs. Therefore, we consider that 10 fs are enough to model the Coulomb explosion of the H₂O²⁺ species.

Concerning the time-steps of integration, they have been chosen according to the tests presented in Chapter 5, namely: For uracil ionization we have chosen a time-step of 0.01 a.u. (0.00024 fs), while for the ionization of the water molecule the time-step has been set-up to 0.05 a.u. (0.0012 fs). Last conformations at the end of time-dependent simulations have been followed by BO MD dynamics (1 a.u. time-step). They are extended until no significant changes in the fragments produced are observed: total time of 120 fs for uracil ionization dynamics, and 58 fs for water ionization.

8.2 MD simulations subsequent to ionization of uracil

To clarify how the solvent affects the fragmentation patterns of uracil²⁺, we must ionize the molecule in the same sites as we did in the gas phase. For this reason we looked in our system for two equivalent orbitals to KS1 and KS2 of the isolated uracil molecule that were centered, see Fig. (7.1), on the two C-O bonds. The KS representation of our liquid phase gives a deepest orbital (-27.87 eV) similar to the one obtained in the gas phase (KS1). On the contrary, the second deepest KS orbital is not center on a C-O bond any more. Indeed, this orbital is partially delocalized between two water molecules making it useless for our purposes. Turning to Wannier representation 3.4.3 we find that orbital #20 (-26.56 eV) is centered over the C₄-O bond and is shape-equivalent to the KS2 orbital of isolated uracil. For clarity, these two orbitals chosen will be referred as KS1 and KS2 according to their resemblance to gas phase uracil orbitals.

Fig.(8.1) gathers some snapshots of the TD-DFT MD and BO MD subsequent to double ionization from KS1 and KS2 orbitals of uracil (upper and lower strips respectively). Panels **A** and **A'** represent the beginning of the dynamics with the ionized density plotted in red. **B** and **B'** panels show the end of the TD-DFT molecular dynamics simulation at 25 fs. Finally, **C** and **D** panels of Fig.(8.1) display the situation after 72 fs and 96 fs of BO MD started from the end of the TD-DFT MD simulation. A detailed description of both processes will be given in the next sections but, before continuing, let's take a look at the variations of the KS and the kinetic energies during the TD-DFT MD simulations.

As Fig.(8.2) shows, the same behavior is found in both dynamics, i.e., there is a decreasing in the potential energy ≈ 8.5 eV during the first 5 fs. The temperature of the system reaches its maximum around ≈ 750 K at about this time. This similarity can be correlated with the release of an atomic oxygen that occur in both dynamics.

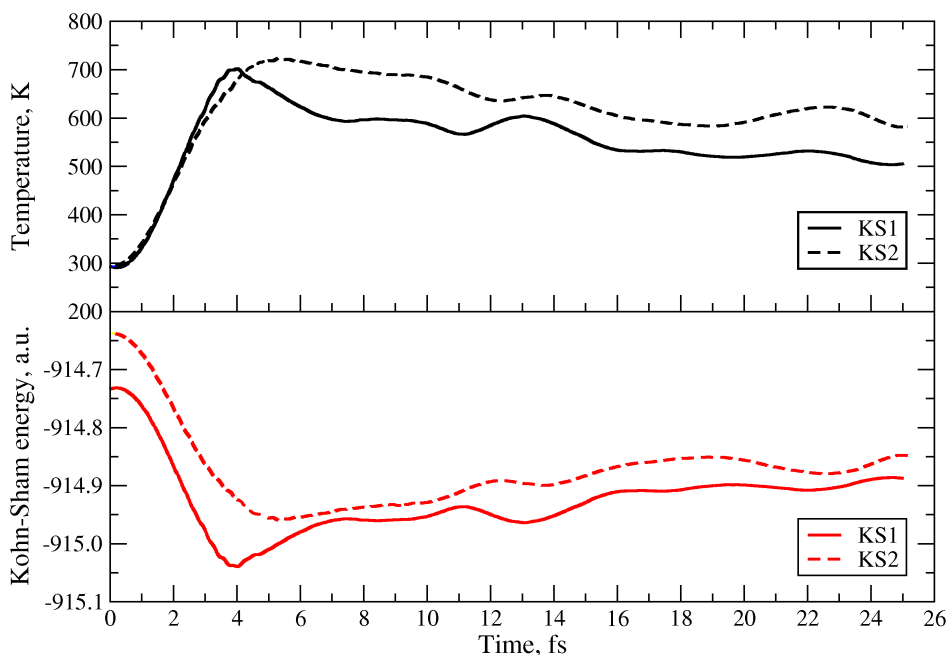


Figure 8.2: Time evolution of kinetic energy and Kohn-Sham energy of the system during the KS1 (solid lines) and KS2 dynamics (dashed lines)

In the following we will describe both ionizations, KS1 and KS2, separately. To give more insights into the modifications that the intermediate fragments undergo, several all-electron optimizations at BLYP/6-311+G(d,p) level have been performed using Gaussian09 [213]. These calculations have been done considering the isolated species, and also modelling the solvent effect by the Tomasi's Polarized Continuum Model (PCM) [227, 228]. The results are summarized in Table (8.1).

8.2.1 Ionization of KS1 orbital of uracil

Fig.(8.3) shows the most important steps occurring during the TD-DFT MD and BO MD after the ionization of the KS1 orbital of uracil. At 5.5 fs, there is a bond breaking of C₂-O where the hole was initially created. Thus we can separate the description of the process into two parts: On one hand, we will describe what happens with the ring that remains after the oxygen loss, and on the other, the combinations of such oxygen with the solvent.

After the oxygen atom (O2) is violently shot out into the solvent, the remaining ring rearranges to accommodate the charge and the excess of energy. There are, for instance, large N-H vibrational movements in the first fs of the dynamics (≈ 18 fs). Thus, at 25 fs, the structure has a charge of +0.862. The H atom (H1) attached to N₁ is transferred to a water molecule at 85 fs. Last change is occurred at 116 fs when the ring loses the hydrogen atom (H3) of the N₃-H bond. Thus, a final six-membered structure is obtained, with an almost neutral charge of +0.113 at 121 fs. Concerning the released protons, H3 remains in the medium and H1 forms an H₃O⁺ with a surrounding water molecule (W2).

The released O2 atom remains unbonded and negatively charged (for instance it has a charge of -0.529 at 25 fs). At 97 fs, it is combined with a neighboring water molecule (W1 in the picture) forming two OH neutral species. These species merge into a neutral H₂O₂ at 118 fs.

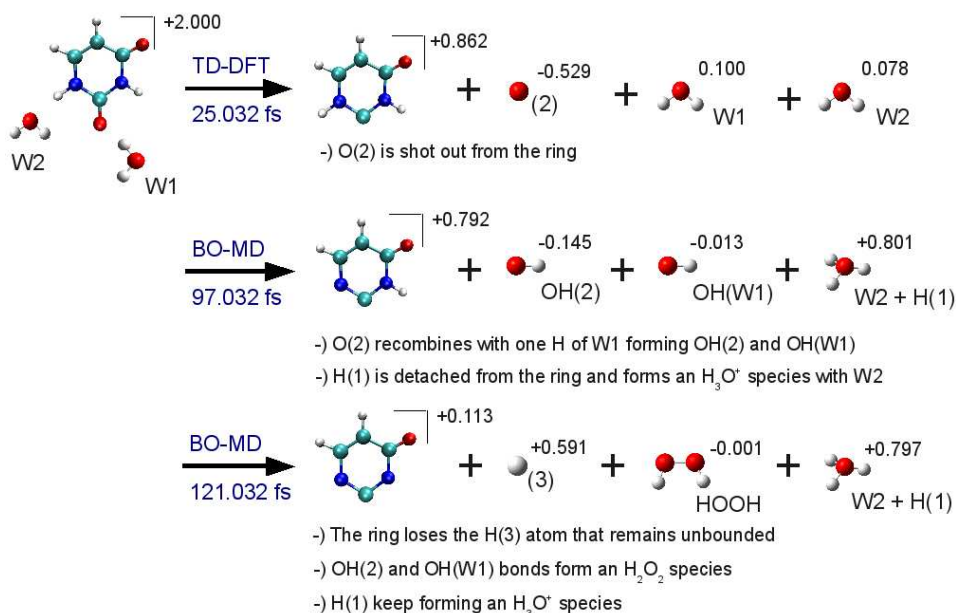


Figure 8.3: Summary of the ionization of the KS1 orbital of uracil in terms of intermediate and final species. The two most interacting molecules of solvent during the fragmentation are indicated and labeled by W1 and W2. These molecules are initially neutral.

These findings are in agreement with all-electron calculations presented in Table (8.1). The first six-membered ring obtained after the O2 atom is released, *KS1-Ring1* in the table, represents a minimum for both gas phase and PCM all-electron calculations. The charge of the structure that remains after the H1 loss, *KS1-Ring2*, fluctuates between +0.6 and +0.8 during the dynamics. Attending to the all-electron calculations, if the structure has a charge of +1, is unstable as it does not represent a minimum (the optimization leads to *KS1-Ring2-open*). Indeed, as it was shown above, *KS1-Ring2* loses the H3 leaving a finally six-membered structure (*KS1-Ring3*) with an almost neutral charge. This structure results a minimum for gas phase and PCM all-electron calculations.

8.2.2 Ionization of KS2 orbital of uracil

Fig. (8.4) analyzes the TD-DFT MD and BO MD subsequent to the ionization of the KS2 orbital of uracil. At ≈ 5.2 fs there is an almost simultaneous loss of the oxygen atom (O4) of the C₄-O bond, and the H atom (H3) attached to N₃. Thus, we can separate this description into two parts, as we did in the former KS1 ionization.

Regarding the six-membered ring formed at 5.2 fs, it remains unaltered and positively charged until 37 fs when the H1 is detached from N₁. The structure that remains has neutral charge. Soon after, ≈ 41 fs, there is a bond breaking between C₂ and N₃ that leads to an opened structure that remains without changes until the end of the simulation. It is worth to mention that the released H1 and H3 atoms are combined with two neighboring water molecules (W3 and W2 respectively) forming transient H_3O^+ species. At the end of the simulation both H atoms are forming part of two new water molecules, W4 and W5, as the H_3O^+ formed lose different H atoms.

The O4 atom released in the early stages of the simulation is linked to an H atom of a close water molecule (W1) in a process that takes ≈ 25 fs. The neutral OH species generated remains stable until it is combined to form a neutral H_2O_2 at 77.7 fs. The new peroxide

decomposes at ≈ 100 fs leading to the former two OH species.

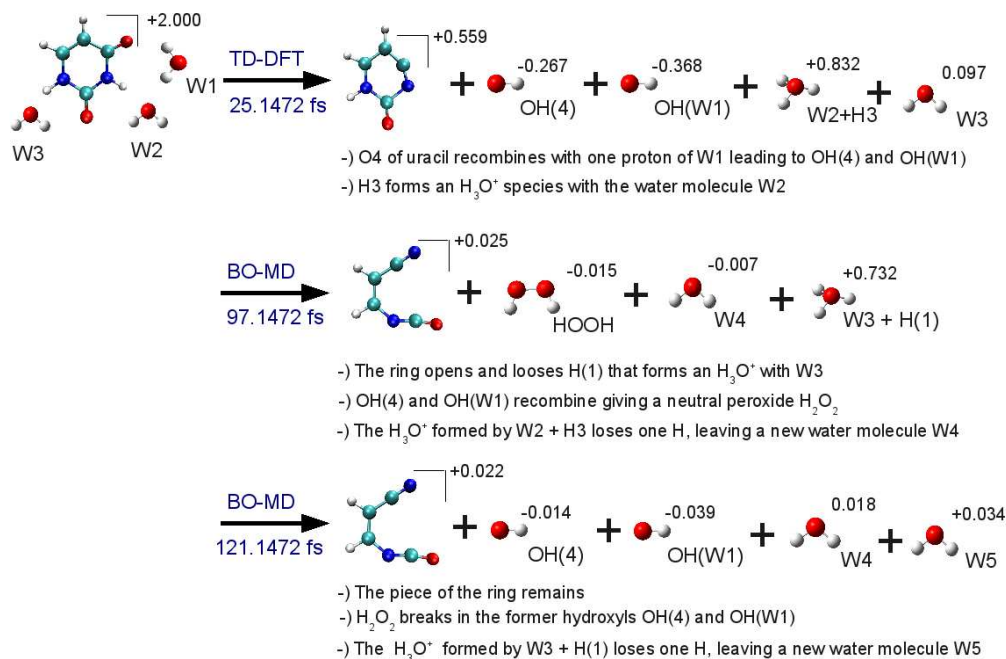
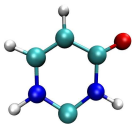
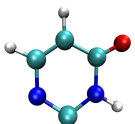
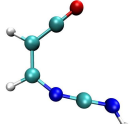
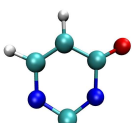
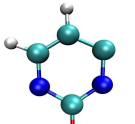
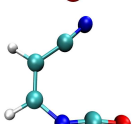


Figure 8.4: Summary of the ionization of the KS2 orbital of uracil in terms of intermediate and final species. Most interacting molecules of solvent during the fragmentation are indicated and labeled by W1, W2 and W3. These molecules are initially neutral.

Analyzing the results by all-electron calculations, the uracil structure that remains after the O4 and H3 are released (labeled as *KS2-Ring1* in Table (8.1)) represents a minimum for both gas phase and PCM. However this structure still changing during the simulation, after the released of H1, the structure formed (*KS2-Ring2*) is not a minimum, neither in gas phase nor in PCM calculations. The optimization in this case gives an opened structure (*KS2-Ring2-open*) in agreement with the one obtained in the dynamical simulation after the bond breaking of $\text{C}_2\text{-N}_3$.

Table 8.1: Gaussian09 optimizations at BLYP/6-311+G(d,p) level, of the intermediate and final structures of uracil appeared during the TD-DFT MD/BO MD dynamics. Total energies of the optimized species in gas phase and using the PCM model phase are given in last columns. The third column indicates if the structure represents a minimum or not.

Name		Charge	Minimum?	E (a.u.)	E^{PCM} (a.u.)
<i>KS1-Ring1</i>		Neutral	YES	-339.5292286	-339.5450390
		1+	YES	-339.2370942	-339.3303684
<i>KS1-Ring2</i>		Neutral	YES	-338.9032326	-338.9129471
		1+	NO!	it opens to	<i>KS1-Ring2-open</i>
<i>KS1-Ring2-open</i>		1+	YES	-338.6527598	-338.7373960
<i>KS1-Ring3</i>		Neutral	YES	-388.2994210	-338.3070175
		1+	YES	-337.9819790	-338.0797323
<i>KS2-Ring1</i>		Neutral	YES	-338.9026939	-338.9187016
		1+	YES	-338.6246340	-338.7199912
<i>KS2-Ring2</i>		Neutral	NO!	it opens to	<i>KS2-Ring2-open</i>
		1+	NO!	it opens to	<i>KS2-Ring2-open</i>
<i>KS2-Ring2-open</i>		Neutral	YES	-338.3489367	-338.3580786
		1+	YES	-338.0021390	-338.0979813

8.3 MD simulations subsequent to ionization of water

If a cell is exposed, the probability of the radiation to interact with the DNA molecule is very small because these critical components are assumed part of the cell. Certainly, the highest probability is given by the interaction with water since each cell is mostly composed of it. When radiation affects the water, it may break the bonds that hold the molecule together, producing fragments such as hydrogen and oxygen species. These products may recombine or may interact with other fragments or ions to form compounds which would harm the cell.

The experimental and theoretical study of water radiolysis in swift ion tracks has received a lot of attention in the recent years [34]. It provides remarkable information about how the secondary species generated after the ionization could combine to form toxic substances. Unfortunately, the ionization by itself and the subsequent chemical processes that occur at the femtosecond time-scale are experimentally inaccessible so far. The only way to truly approach these first events is by using theoretical methods based on time-dependent molecular dynamics.

As it has been referred previously, previous investigations of our group using TD-DFT MD simulations concerned the double ionizations of a selected water molecule in sample liquid water [45]. We are interested now in evaluating the effects of such ionization when it takes places in a surrounding water molecule located in the first hydration shell of an uracil molecule. Our interest is to check if the protons arisen from the Coulomb explosion induce any variation on the structure of the uracil, and also to study the physical properties of oxygen species generated.

With this purpose on mind, we run two dynamics ionizing a water molecule situated 2.97 Å to the nearest atom of the uracil (specially the oxygen atom of the C₂-O bond). This molecule belongs to the second hydration shell and it is not bonded to the biomolecule.

The ionizations are done by removing two electrons from a different molecular orbital of the chosen water molecule each time. As it is shown in Fig.(8.5) each water molecule has five occupied molecular orbitals referred following their symmetry properties as: 1a₁, 2a₁, 1b₂, 3a₁ and 1b₁. In our pseudo-potential representation the core orbital 1a₁ is not simulated, leaving only four orbitals (eight electrons) treated explicitly. For their representation, Wannier Orbitals centered on each O atom are used. This type of orbitals enable to separate molecular orbitals of each water molecule from the ones of the biomolecule.

For the water molecule chosen, the corresponding molecular orbital energies are: -25.32 eV, -7.34 eV, -4.57 eV and -2.72 eV respectively. In this work, we have created holes from the two deepest orbitals, i.e. 2a₁ and 1b₂. The gap of ≈ 18 eV existing between them will allow us to estimate the influence of the energy deposited in the initial H₂O²⁺ ions.

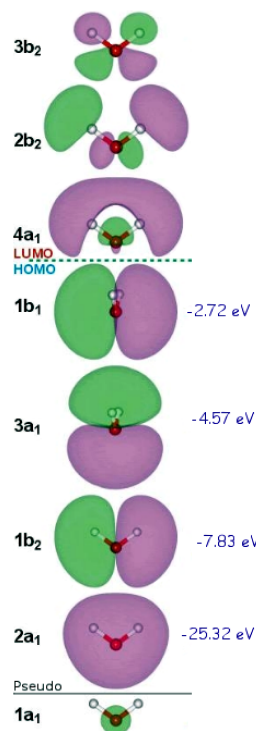


Figure 8.5: Molecular orbitals of a water molecule. Energies are given for those orbitals treated explicitly in the simulations.

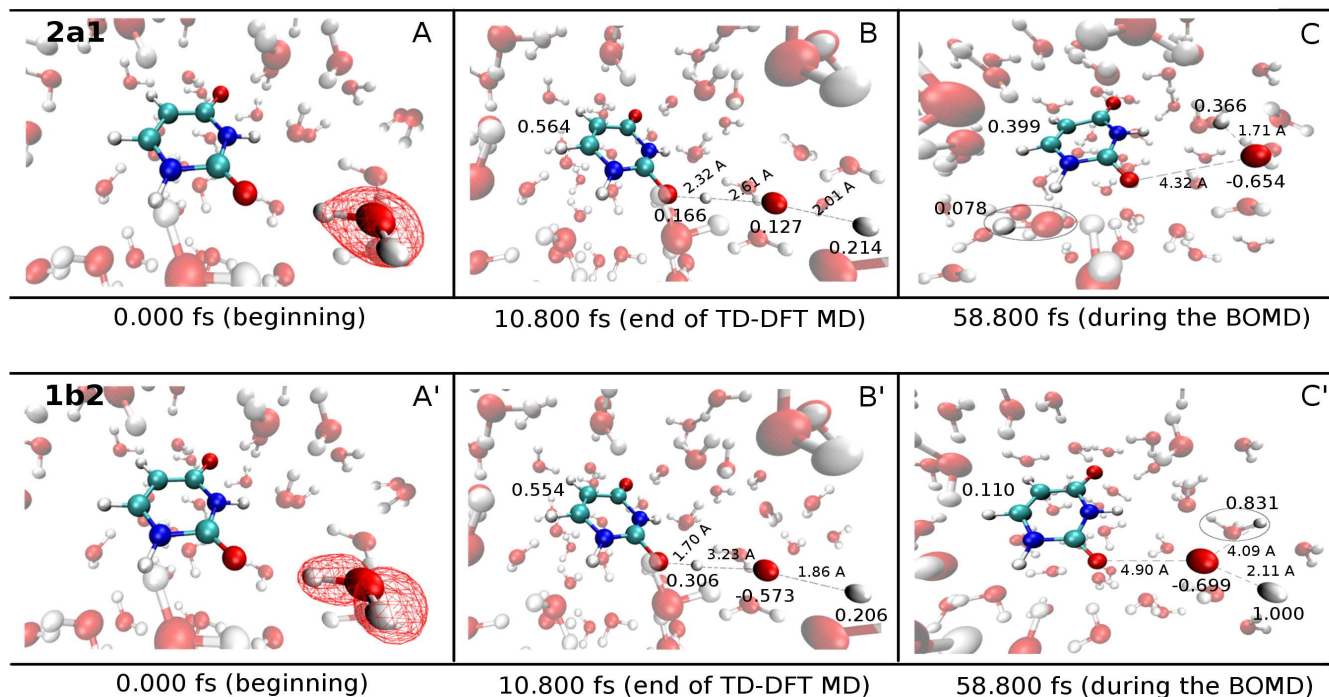
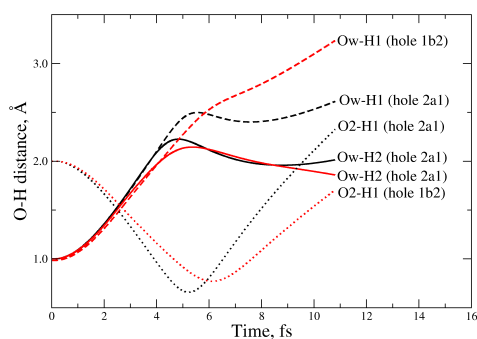


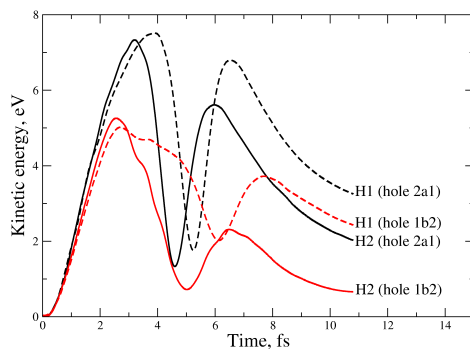
Figure 8.6: Snapshots corresponding to 2a1 (upper strip) and 1b2 (lower strip) water double ionizations. **A** panels refer to the beginning of the simulation and represent the isodensity associated to each of the removed Wannier orbitals (isovalue 0.01). **B** panels are taken from the last step of the TD-DFT MD simulations, and **C** panels are taken from the final step of the BOMD. For all, some distances (in Angstrom) and charges of the most representative species are given.

A summary of the results is given in Fig.(8.6). After 10.8 fs of TD-DFT MD both cases give rise to almost the same situation: At ≈ 3 fs, there is a Coulomb explosion of the H_2O^{2+} ion that leads to the formation of one oxygen atom and two protons (panel **B** and **B'** on the figure). At the end of the simulations (**C** panels) no big changes occur in the uracil that remains unaltered and slightly charged. The released oxygen atoms have charges of -0.654 and -0.605. The characteristics of these atoms will be discussed in the next section.

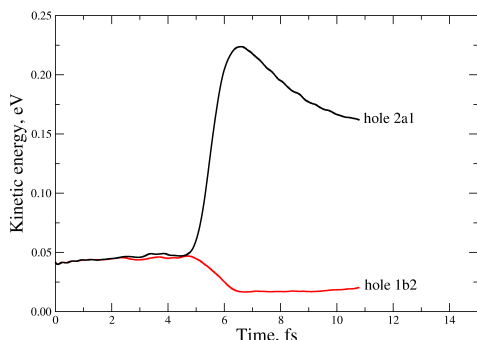
Concerning the ejected protons, most relevant information during the TD-DFT MD simulation is depicted in Fig.(8.7). In both simulations there is one proton that collides with the uracil (from now on H1) and the other one (labeled as H2) is released in the medium. As it is showed in panel (a), the O-H bond breaking is produced in similar ways at the same time (≈ 3 fs) for the two ionizations. Small differences are found, however, in the trajectories. For both dynamics the proton H1 that goes towards the uracil covers a bigger distance than H2 proton. The later is slowed down by interactions with neighboring water molecules forming a transient H_3O^+ species. The protons that target the biomolecule collide with it and hit the oxygen atom of the $\text{C}_2\text{-O}$ group. No attachment is detected, the atom collides and bounces in a process that lasts 7.5 fs for the case of 2a1, and a bit longer, 9.0 fs, for 1b2 ionization.



(a) O-H distances



(b) Kinetic energies of the ejected protons



(c) Kinetic energy of the center of mass of the uracil molecule

Figure 8.7: Data obtained from the TD-DFT MD simulation of 2a1 (black lines) and 1b2 (red lines) water ionizations. Panel (a) shows by solid and dashed lines the O-H distances between the oxygen (Ow) and the two protons of the ionized water molecule. Dotted lines represent the distances between the O2 of the uracil and the colliding protons. Panel (b) displays the kinetic energies of the released protons and panel (c) the kinetic energies of the center of mass of the uracil.

Panel (b) displays the evolution with time of the kinetic energies of the ejected protons during the dynamics. The released protons coming from the 2a1 water ionization present higher values than the ones liberated in the 1b2 dynamics. As it is reflected on the figure, protons shot out after the ionization of the deepest water molecular orbital, get maximum kinetic peaks of ≈ 7.5 eV. On the contrary, the second deepest orbital protons reach maximum values of ≈ 5 eV. These differences are related the initial energy of the double ion at the beginning of the dynamics. In all cases, sudden variations are present mainly due to the

interactions with water and uracil molecules. Nevertheless, the major difference is found in the red dashed curve since the proton H1 coming from 1b2 hole interacts with the uracil in a different way than when coming from 2a1 hole. In fact, the collision angles described by the trajectories are very different: 42° and 83° for 2a1 and 1b2 holes respectively. It turns out that the collision in 2a1 is more direct. For instance, the proton ejected after the 2a1 ionization reaches a closest distance to uracil equal to $0.66 \text{ \AA}$, at 5.2 fs, while less energetic case, 2b1 hole, gets a distance of $0.77 \text{ \AA}$ at 6.1 fs.

Fig.(8.7) gathers the kinetic energy of the center of mass of the uracil during the time-dependent dynamics. As the figure shows, the energy loose by the protons during the collision with the uracil is not transmitted as kinetic energy to the biomolecule. The process is far from being elastic, for instance, the proton coming from the 2a1 ionization (black line) loses 5 eV during the collision, while the kinetic energy of the uracil is only ≈ 0.2 eV increased.

We can conclude that the ionization of a surrounding water molecule of uracil steers to the formation of two protons and an oxygen atom that becomes negatively charged after interacting with other solvent molecules. The energy of the ejected protons are between 5 and 7.5 eV in accordance to the energy of the ionized orbital. Some of the protons collide with the uracil, but no significant changes are produced in the structure of the biomolecule. The possible biological damage of such ionization is then reduced to the indirect effect that the released oxygen atoms might cause on uracil.

8.4 Liquid vs gas phase

We have shown that different fragmentation patterns are found in the liquid in comparison to the gas phase. Ionizations from the two deepest KS1 and KS2 orbitals of the isolated system cause the fragmentation into three 43/41/28 (CO, HNC and HNCHCH) and four 43/27/26/16 (O(NH)C, CNH, HCCH and O) fragments respectively. The ion embedded in water, on the contrary, undergoes softer fragmentations that help to keep bigger fragments intact. For instance KS1 ionization preserves the six-membered ring of the uracil, Fig.(8.3), and liberates two H and one O atom into the solvent. On the other hand, KS2 ionization in the liquid (Fig.(8.4)) leads to a neutral opened structure, one OH species and two H atoms.

These dissimilarities are closely related to the role of solvent, specifically with the charge transfers that occur with the solute. The water-uracil interactions prevent the deslocalization within the uracil of the initial $2+$ charge, and also help to migrate the positive charge of the biomolecule to outer hydration shells.

In this sense, Fig.(8.8) and (8.9) recap the atomic charges as direct measure of the electronic density evolution. The first figure is divided into two panels (a)-gas and (b)-liquid representing the charge of the uracil during TD-DFT MD starting from the deepest ionization (KS1) of uracil for both phases. To make it simpler, the uracil has been divided into three fragments that correspond to the ones obtained in the gas phase dynamics: 28/43/41. Although graph (b) represents fake fragments not observed in the liquid phase, they are useful for our purposes.

Looking at the figure, both phases represent different charge distributions because the way of fragmentation is different. Indeed, the first bond breaking in gas takes a longer time (N_1 -C₂ at ≈ 14 fs, see Fig.(7.3)) than in the liquid, where after 5.5 fs an O atom is definitely ejected. Thus, to compare the charge mobility that occurs into the biomolecule we should regard only the time where molecule does not fragment, that is, times below 5.5 fs.

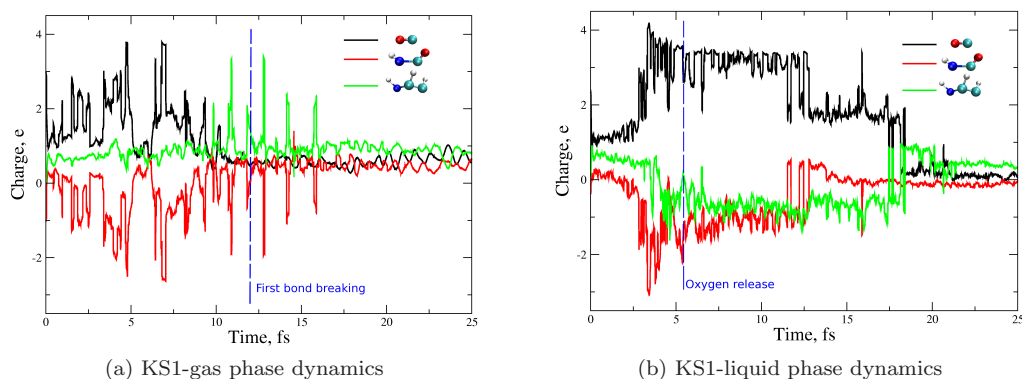


Figure 8.8: Comparison of the uracil charge in terms of fragments for gas (a) and liquid (b) TD-DFT MD. The charge over the ionized C-O piece is given by the black line. The rest of the ring is separated into two fragments: HCNO, red line, and HNCHCH, green line.

For both and just at the beginning of the dynamics, the charge $2+$ set-up on the CO as a consequence of the ionization, is quickly reduced by charge exchange with the other two fragments. Afterwards, in the gas phase, big fluctuations among the fragments are observed: The HNCHCH fragment remains almost neutral and the charge of the CO is suddenly and directly exchanged with HNCN. The charges keep fluctuating until the molecule is broken, indicating large charge mobility inside the molecule. For liquid phase, the charge over the CO keeps localized around $+1$ during the first femtoseconds, without sudden changes until the oxygen atom starts to be ejected. These differences are attributed to the presence of water since it is the only element that differs from one dynamics to the other.

As we have already said, the influence of the solvent is not only reflected in the charge mobility inside the uracil molecule, but also it plays a role as a drain of positive charge. To measure this effect we have analyzed in the KS1 ionization the charges of the hydration shells of our system. They have been defined by the distances that separate the center of mass of the uracil and the O atoms of the water molecules: If the distance R is ≤ 4 Å the molecule is included in the first shell (3 molecules), if 4 Å $< R \leq 5$ Å (11 molecules), they are in second shell. Finally, molecules with 5 Å $< R$ (35 molecules) are counted together.

Thus, black line of Fig.(8.9) represents the charge of the atoms of the uracil as time evolves. It has to be reminded that until 25 fs the only fragment produced is an oxygen atom. The charge lost by uracil until 7 fs is transferred to the surrounding water molecules. The abrupt changes observed at 7.5 and 13.5 fs correspond to interactions with protons of the surrounding water molecules. At the end of the dynamics, the molecule has a charge of $+0.3$ as the sum of a negative oxygen atom (-0.5) and a positive fragment ($+0.8$) formed by all the other atoms of uracil.

In the meantime, while the uracil is getting neutral, the charge rises in the solvent. Specially, the outer shells of hydrations are the ones that become more positively charged. As it is shown in the figure: The first hydration shell (red line) barely changes its initial charge, the second shell (blue) gets a half of a charge and the outermost (green) has a charge of $+1$.

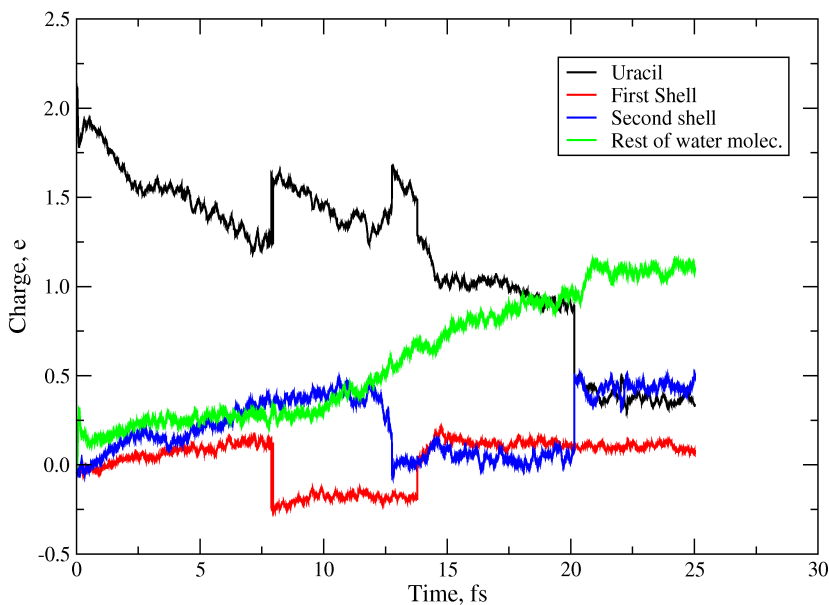


Figure 8.9: Charges of uracil and different hydration shells during the TD-DFT MD subsequent to the KS1 ionization.

About the similarities between the two phases, we have found that there can be an oxygen release in both. Namely, the ionization of the KS2 orbital in gas and all ionizations produced in the liquid phase lead to the formation of one atomic oxygen. Attending to their physical properties, the meaning of Fig.(8.10) is to show how different these species are. Thus, left panel (a) shows the kinetic energies of the released oxygen atoms. The ones arisen from the ionization of a water molecule (black lines) have a much less energy. In fact, the kinetic energy peaks are below 1 eV. Notice that for the case of the deepest ionization 2a1, the energy is even less than the ionization coming from the 1b2 orbital. As it was depicted in the second graphic of Fig.(8.7), this fact is due to a major energy transfer to protons. Surprisingly, looking at the oxygen charges in panel (b) of Fig.(8.10), both oxygens do not carry the same charge: the one released by the 2a1 ionization gets an almost neutral charge while the one coming from the 1b2 ionization is negatively charged.

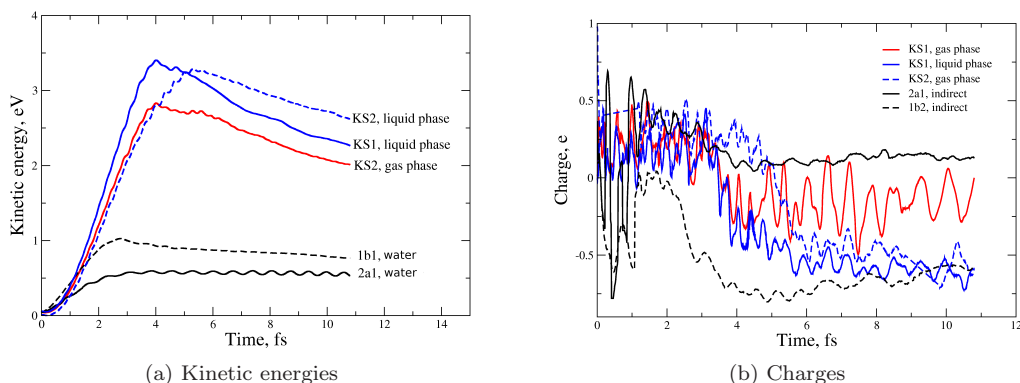


Figure 8.10: Kinetic energy (a) and charges (b) of different oxygen atoms released during the TD-DFT MD

The oxygens ejected by uracil ionizations in gas (red line) and liquid (blue lines) also have characteristics in common. For instance, their maximum kinetic energies are in between 2.5-3.5 eV at 4-6 fs. The two cases in liquid, KS1 and KS2 ionizations, lead to negative oxygens while the KS2-gas oxygen also gets negatively charged, but varying between charges of 0 and -0.25. This particular difference is due to the absence of solvent. In the liquid once the oxygen is released, it starts to drain negative charge from the nearby water molecules.

Chapter 9

Double ionization of small water clusters

Abstract *Fragmentation of double charged small water clusters have been investigated by CarParrinello Molecular Dynamics and Time-dependent Density Functional Theory simulations. The results are compared with recent experiments of small water clusters collisions with 12 MeV/u Ni^{25+} ions combined with cold target recoil ion momentum spectroscopy (COLTRIMS).*

9.1 Introduction

The past decade has brought tremendous advances in the study of molecular clusters. One of the goals has been to understand how the properties of these molecular clusters vary with size and, in particular, how condensed phase attributes develop as the size increases [229–233]. Most investigations have focused on either weakly bonded van der Waals systems, like rare gases, or strongly covalently or metallically bonded systems. Water clusters belong to a different category in which the individual molecules interact through hydrogen bonds whose strength lies between that of weak van der Waals and strong covalent clusters [234].

The dynamics of water clusters is interesting for a number of reasons. In atmospheric chemistry, it is important to understand the early stages of cloud and droplet formation [235]. Water clusters have also been suggested as transient intermediates in liquid water [236], which is crucial to understand many chemical and physical processes [234]. In particular, ionization of water clusters is relevant to radiation damage studies because water is the natural environment of biomolecules. It has been shown, e.g., that such an environment may efficiently protect the biomolecules against ionizing radiations [219] or, in contrast, lead to secondary electrons and radicals (water radiolysis) that react with the biomolecules.

In this work, we will study the fragmentation dynamics of double charged small water clusters using Car-Parrinello Molecular Dynamics (CP-MD) and Time-Dependent Density Functional Theory Molecular Dynamics (TD-DFT MD). Our aim is to identify which are the products of such fragmentations and compare them with the experimental measurements performed in the GANIL facility in Caen (France). In these experiments, small water clusters are collided with highly ionizing 12 MeV/u Ni^{25+} by using imaging techniques with coincident measurement of the full momentum vector of the charged fragments. This technique gives simultaneous information on stability, energetics and charge mobility inside the cluster. In the following, we will first introduce the experimental results that will be

compared with our theoretical results in the second part of the chapter.

9.2 Experimental results

Water clusters are produced using a supersonic expansion. At the start of the experiment, approximately 20 ml of distilled water is loaded into a heated reservoir (stagnation chamber). Its temperature is controlled by a thermocouple and is typically kept at $T_S \approx 80^\circ\text{C}$. The pressure of the water vapour p_s is determined by the temperature of the stagnation chamber, and is around 400 mbar. A 1 bar Ar seeding gas is mixed with the water vapour at the exit of the sealed oven to increase the total stagnation pressure. The vapour flows to the nozzle ($30\ \mu\text{m}$ diameter, temperature 90°C) through a stainless steel tube. The temperature of this tube is controlled separately, and is held slightly above the reservoir temperature to avoid condensation. The gas mixture undergoes isentropic expansion and the temperature drops rapidly with an increasing distance from the nozzle. This leads to the supersaturation of the water vapour and subsequent clustering [237]. Note that clusters produced by such an adiabatic expansion are usually considered to be close to the cluster melting temperature [238], but the seeding with argon may result in colder clusters. Under the present conditions, small size clusters are expected to be produced [239]. 12 MeV/u Ni^{25+} projectiles produced by the GANIL facility in Caen intersect the $(\text{H}_2\text{O})_n$ supersonic cluster beam. The same projectile has been used in our previous study of dissociative ionization of a single water molecule [240–243]. The charged fragments are extracted by a uniform electric field with 4π solid angle acceptance and directed onto an 80 mm diameter position sensitive detector. As in the case of molecules [241,244–246], the energy distribution of the fragments is determined from both time of flight and coordinates of the impact position of each particle detected. Such coincidence experiments are necessary to unravel the repartition of the initial multiple ionization on the subsequent fragments.

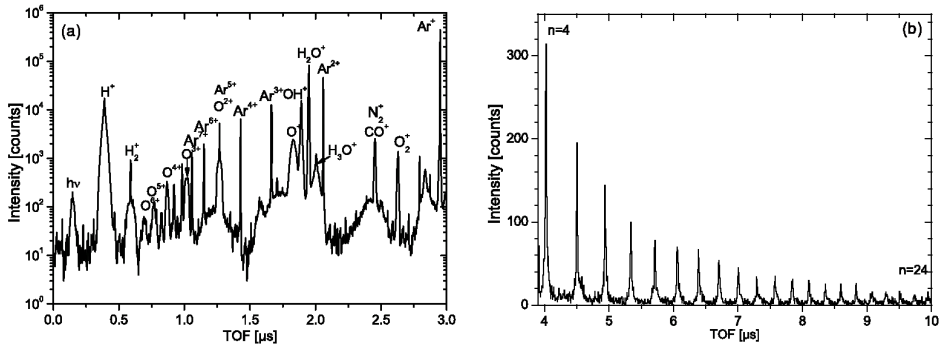


Figure 9.1: Time-of-flight spectrum for 12 MeV/u Ni^{25+} on $(\text{H}_2\text{O})_n$ (a) For short time of flight (less than $3\ \mu\text{s}$) and (b) for long time-of-flight (more than $4\ \mu\text{s}$) where protonated $(\text{H}_2\text{O})_n\text{H}^+$ are detected.

Time-of-flight spectra are presented in Fig.(9.1). They allow us to identify the species produced in the collision. For short time of flight (less than $3\ \mu$, Fig.(9.1)(a)), the spectrum is very similar to that obtained for a single water molecule. This is not surprising because, under the present experimental conditions dominated by the formation of small clusters, individual water molecules are also produced in the supersonic expansion. The experiments detect multiply charged oxygen fragment O^q+ ions (with q ranging from 1 to 6) coming from dissociation of H_2O^{8+} molecular ions produced in a single collision event (O^{6+} ions are always emitted in coincidence with two protons). They also detect Ar^q+ ions coming from

the seeding gas ionization.

Cluster ions appear clearly for longer time of flight (more than 4 μs , Fig.(9.1)(b)). In this case, the mass spectrum is dominated by protonated cluster ions $(\text{H}_2\text{O})_n\text{H}^+$ (or $(\text{H}_2\text{O})_{n-1}\text{H}_3\text{O}^+$) and reflects, in some way, the initial size distribution produced in the supersonic expansion. Indeed, the formation of $(\text{H}_2\text{O})_n\text{H}^+$ is dominated by large impact parameters, which produce mainly singly ionized clusters. Doubly charged cluster fragments are detected in the experiment. Previous photoionization experiments have determined that a critical cluster size of $n = 37$ has to be reached to observe series of the form $(\text{H}_2\text{O})_n\text{H}^{2+}$. This critical size is smaller in collisions with highly charged ions [247]. Such a difference is generally attributed to the fact that, in a collision, double electron removal takes place at a large impact parameter, thus leading to “cold” dications that can survive more easily. In the present experiment, the critical cluster size is clearly above $n = 20$.

Fig.(9.2) represents the relative intensities of the peaks shown in Fig.(9.1)(b) as functions of size n . To a good approximation, the intensities follow an exponential behavior, with the notable exception of $(\text{H}_2\text{O})_4\text{H}^+$. Similar exponential behavior has been observed in cluster growth studies [239]. The unusually high stability of $(\text{H}_2\text{O})_4\text{H}^+$ has also been reported in the case of free jet expansion of liquid water [248] and for clusters sputtered from frozen H_2O under keV He^+ [249] and MeV/u Ar ion impact [250].

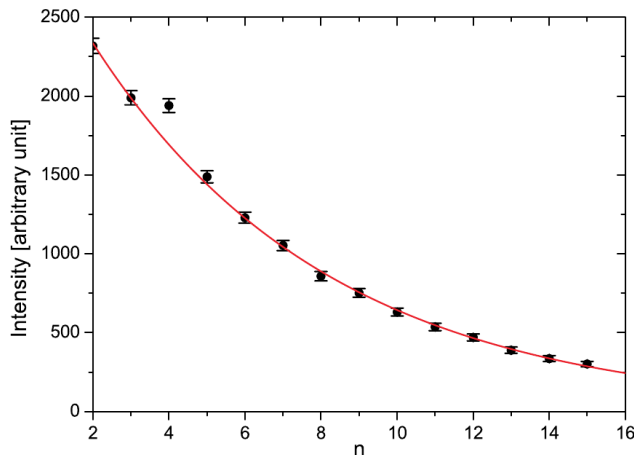


Figure 9.2: Relative intensity of the different protonated $(\text{H}_2\text{O})_n\text{H}^+$ species as a function of n . The line is an exponential data fit.

To our knowledge, this is the first time that such enhanced stability is seen from gas-phase water clusters that are subsequently ionized. $(\text{H}_2\text{O})_4\text{H}^+$ clusters have been extensively studied [251] since they are possible building blocks for large proton hydrate species. There is continued speculation about the location of the excess proton in these hydrates and, by extension, in the bulk [252]. Most chemists consider that the hydrated proton appears in the form of a hydronium ion, H_3O^+ . But this picture was refined by Eigen [253] and Zundel and Metzger [254], who advocated larger complexes such as H_9O_4^+ and H_5O^+ , respectively. In the “Eigen cation”, the central hydronium ion is strongly hydrogen-bonded to three H_2O molecules forming $\text{H}_3\text{O}^+(\text{H}_2\text{O})_3$. In the “Zundel ion“, the proton lies midway between two water molecules $\text{H}_2\text{O} - \text{H}^+ - \text{H}_2\text{O}$. The most appropriate picture to describe the hydrated proton is still a subject of debate. The relatively high stability of the $(\text{H}_2\text{O})_4\text{H}^+$ fragments observed in this work seems to support the Eigen cation as the preferential configuration.

In order to explore the charge repartition on the cluster fragments, Fig.(9.3) shows the coincidences between the different charged species. A relevant region of the coincidence spectrum is shown in Fig.(9.3). It can be unambiguously seen coincident emission of singly charged protonated cluster fragments. In general, the dominant coincidence events involve at least one H_3O^+ . High resolution energy and angular distributions of the charged fragments are determined from the measured full momentum vectors. Energy distribution measurements are still scarce in the case of protonated water clusters [255, 256]. However, precise measurements are necessary as dissociation energies can vary depending on the dissociation path [257]. As an illustration, the insets in Fig.(9.3) show energy distributions for H_3O^+ and $\text{H}_3\text{O}^+/\text{H}_5\text{O}_2^+$ coincident ions and for two $\text{H}_3\text{O}^+/\text{H}_3\text{O}^+$ coincident ions (in the latter case, a single distribution is shown because both are almost identical). From the analysis of the mass spectra for a given number of charged particles per event, we can conclude that the two coincident ions mainly come from the dissociation of a doubly charged parent in two (and only two) singly charged fragments plus eventually some neutrals. The H_3O^+ and H_5O_2^+ coincident distributions peak around 900 meV and 500 meV, respectively, while the H_3O^+ and H_3O^+ ones peak around 900 meV. In the latter case, the measured angular distributions show that the two H_3O^+ fragments are preferentially ejected at 110° - 120° , thus evidencing that another neutral fragment(s) contributes significantly to the total momentum balance.

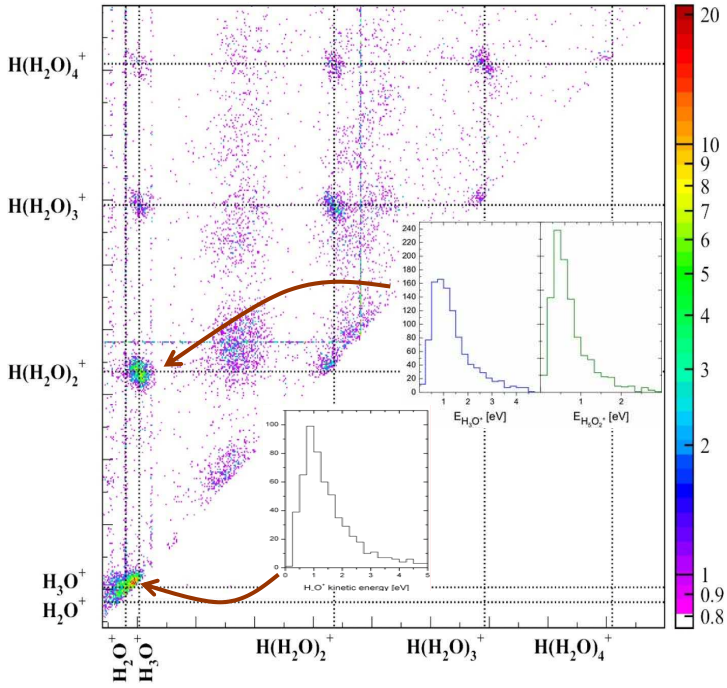


Figure 9.3: Coincidence spectrum between singly charged fragments. The insets show the kinetic energy distribution of coincident H_3O^+ and H_5O_2^+ , and of two coincident H_3O^+ cluster fragments for 12 MeV/u $\text{Ni}^{25+} + (\text{H}_2\text{O})_n$ collisions.

These findings lead us to explore the question of charge localization and mobility within the multiply charged water clusters. In this respect, multiply charged oxygen ions (even O^{2+}) in coincidence with cluster fragments are never detected, even after five-fold ionization of the cluster. This suggests a high charge mobility similar to that observed in multiply ionized van der Waals clusters of fullerenes [258]. It also agrees with synchrotron radiation ionization experiments in which resulting holes are rather delocalized with a preference for molecules lying in the cluster surface [259].

9.3 TD-DFT MD simulation of $(\text{H}_2\text{O})_3^{2+}$ and $(\text{H}_2\text{O})_{11}^{2+}$

In order to explain some of the findings observed in the experiments we have performed different simulations. The first issue to be solved is to know whether the clusters are stable at the temperature they are produced in the experiment. For this purpose, we have carried out CP MD simulations of clusters containing 3 and 11 water molecules at different temperatures. All simulations were performed with the plane wave KohnSham-based DFT code CPMD [94, 168]. We have followed similar procedure as in study of uracil $^{2+}$ in the gas phase of Chapter 7. However, in the present work, the water clusters have been enclosed in a larger cubic box of length $L = 21 \text{ \AA}$ to avoid interaction between clusters in neighboring boxes. The energy cut-off of the plane wave basis is also larger: 90 Ry. The optimized initial cluster geometry for $(\text{H}_2\text{O})_3$ is in good agreement with that reported in [260]. The results of these calculations show that, after 0.7 ps, $(\text{H}_2\text{O})_3$ is stable below 90 K while $(\text{H}_2\text{O})_{11}$ is stable below 300 K. Since clusters produced in the supersonic expansion are rather cold ($T < 80 \text{ K}$) [256], we infer from thermal stability of the neutral clusters that the ionization process occurs before fragmentation processes. Some small clusters are therefore present in the jet, which could be further ionized and their product detected.

As stated in the previous section, it is difficult to establish a direct link between this spectrum of Fig.(9.1) and the initial size distribution. To check whether H_3O^+ fragments can also arise from singly charged $(\text{H}_2\text{O})_n^+$ water clusters produced in the collision, we have performed CP MD calculations for $(\text{H}_2\text{O})_3^+$ and $(\text{H}_2\text{O})_{11}^+$ during 0.7 ps for several temperatures. We have found that the larger cluster is stable up to $T = 200 \text{ K}$, while the smaller cluster breaks even at $T = 10 \text{ K}$. When the latter breaks, a smaller $(\text{H}_2\text{O})_n^+$ fragment is produced but never $(\text{H}_3\text{O})^+$. This suggests that a minimal size is required to observe such a fragment.

The analysis of the correlation spectra of Fig.(9.3) suggests a charge mobility of the fragments after the ionization. To elucidate the mechanisms behind such mobility, we have performed TD-DFT MD calculations of the fragmentation of the doubly charged clusters $(\text{H}_2\text{O})_3^{2+}$ and $(\text{H}_2\text{O})_{11}^{2+}$ produced by removal of two electrons from the corresponding neutral clusters. We have chosen the 11 water cluster because it is the smallest size with a cubic closed cage and an attached cycle. Such a configuration allows three different coordination numbers depending on the position of the water molecule in the cluster. This has allowed us to study the effect of ionizing a tri-coordinated water molecule (A in Fig.(9.4)) or a tetra-coordinated one (B in Fig.(9.4)). In these TD-DFT MD calculations, we have used the same procedure as in [35, 45] for electron removal. We have also evaluated the Bader charge of the different species produced in the fragmentation process [206]. For the larger cluster, the main results of the simulations are presented in Fig.(9.4). The left panel shows the initial geometry of the $(\text{H}_2\text{O})_{11}^{2+}$ cluster, which is the same as that of the corresponding neutral cluster for a Frank-Condon transition. Two cases have been considered: both electrons are removed from either the A or the B molecule (see the left panel). In both cases, two electrons were removed from the 1b1 orbital which is one of the most likely processes. It is worth noting that direct two-electron removal or the Auger process will lead to very similar two-

valence hole configuration. The middle and right panels in Fig.(9.4) show the corresponding cluster geometries after 12 fs.

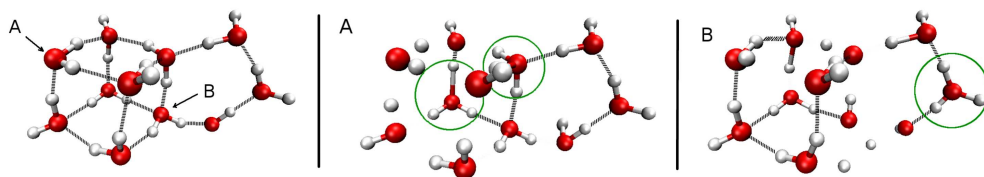


Figure 9.4: Geometry of the $(\text{H}_2\text{O})_{11}^{2+}$ cluster obtained from TD-DFT MD simulations at $t = 0$ (left panel) and $t = 12$ fs (middle and right panels) when two electrons are removed from the $1b_1$ orbital of molecules A and B. The circles enclose H_3O^+ units. Oxygen atoms are in red, hydrogen atoms are in grey.

In case A, two H_3O^+ ions are clearly produced (in both cases, the calculated charge $> +0.8$). These ions lead to one of the most intense peaks in Fig.(9.3). The H_3O^+ ions do not arise from the water molecule that loses the two electrons, but from water molecules that “capture” protons ejected from the ionized molecule. The capture proceeds through several steps that involve neighboring water molecules in which the protons are successively captured and released. This is very similar to what has been found for liquid water [35, 45] and support the hypothesis of atomic O formation to explain the formation of HO_2 and O_2^- radicals in water radiolysis by swift ions. The dynamics is very similar when electrons are removed from molecule B. In this case, only one H_3O^+ ion is produced (calculated charge $> +0.8$), but again it does not come from the water molecule that was originally ionized.

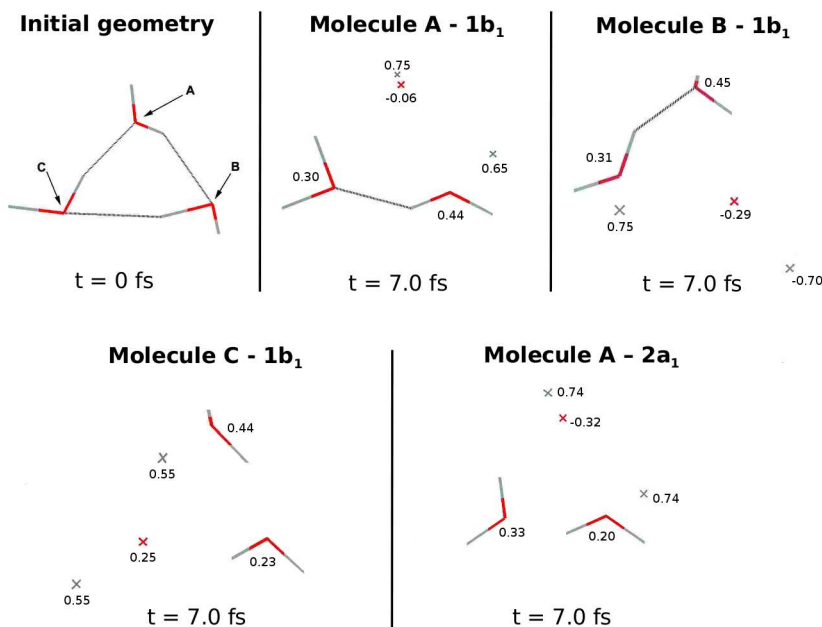


Figure 9.5: Geometry of the $(\text{H}_2\text{O})_3^{2+}$ cluster obtained from TD-DFT MD simulations at $t = 0$ (first panel) and $t = 7$ fs when two electrons are removed from the $1b_1$ orbital of molecules A, B and C (panels from 2 to 4). Additionally, panel 5 shows the ionization of the $2a_1$ orbital of molecule A. Numbers refer to charges of the species and grey and red crosses mean H and O atoms respectively.

For the study of the smaller doubly charged cluster $(\text{H}_2\text{O})_3^{2+}$, we ionized two different types of orbitals i.e., 1b1 and 2a1, centered on different water molecules. Final structures after 7 fs of TD-DFT MD simulation are shown in Fig.(9.5). For all cases the initial H_2O^{2+} ion fragments into an oxygen atom and two protons in a fast process that lasts about 4 fs. This result is independent of the initial energy of the ionized orbital and of the molecule were double charge is initially set. None of the studied cases lead to H_3O^+ fragments thus implying that there is a critical size for the production of these fragments.

Chapter 10

Fragmentation of pyrimidine²⁺ and Auger effect

Abstract *TD-DFT molecular dynamics are used to investigate fragmentation of doubly-charged gas-phase pyrimidine produced by photoelectron-ion coincidence experiments. Unlike collision techniques, in which ionization can occur from very different orbitals without control, photo-ionization experiments present a major selectivity of the process since there is much more information of the deposited energy in the target molecule. Thus, combination of both, experimental and theoretical tools, is aimed to a better correspondence of the ionized orbital and the mass spectrum signal obtained.*

10.1 Introduction

Pyrimidines are an important class of biomolecule mainly due to the fact that the pyrimidine ring forms the base structure of three nucleic acids, i.e. uracil, cytosine and thymine (see Fig.(10.1a)). Furthermore, halogenated pyrimidine bases present an efficient Auger electron emission, and have found applications as radiosensitisers in radiotherapy. For example, it was discovered more than 40 years ago that 5-bromo-deoxyuridine (an analogue of thymidine) radiosensitises cells [261,262].

During the recent years electron-impact and photoelectron-ion coincidence experiments have provided a great understanding of the electronic structure of such compounds [53–55]. However, so far, experiments are unable to provide insights into the processes that occur at the femtosecond scale after the electron removal and, therefore, to establish the mechanisms through which the fragments are formed. In the present work we have carried out a theoretical study of the fragmentation of pyrimidine using TD-DFT MD simulations. One of our future goal is to compare with ongoing experiments in which the fragments of pyrimidine are measured in coincidence with the emitted Auger electrons. By measuring the kinetic energy of the emitted electrons it could be possible to identify in which orbitals the electronic hole was created before fragmentation. This may have important practical consequences since by targeting specific orbitals, one can significantly affect the biological damage.

Before starting let's take a deeper look at the Auger effect and explain how, by theoretical calculations, the kinetic energies of the emitted electrons can be estimated. In essence, the Auger effect is a physical phenomenon that can be viewed as a two-step process schematically

summarized in Fig.(10.1b): in a first step a core electron is expelled from the molecule (single ionization) due to the interaction with X-ray radiation.

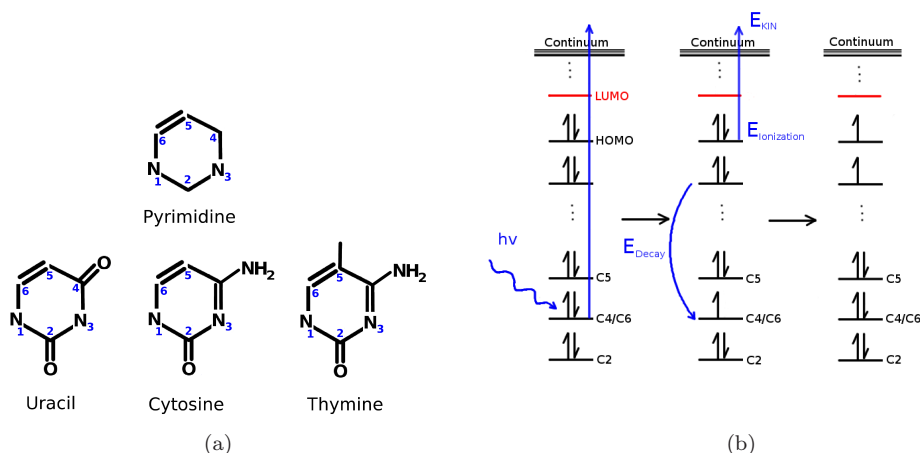


Figure 10.1: a) Structures of pyrimidine, uracil, cytosine and thymine; b) Scheme of the Auger effect.

In a second step, the K-hole hence generated is filled by the decay of an electron located in an upper molecular orbital (MO). The energy associated with the decay to the K-shell can be released through the emission of a photon or it can be transferred to a second electron, called the “Auger electron”, which is emitted with a measurable kinetic energy. At the end of the process the molecule is in a doubly ionized state with two holes located in the same or different MO.

10.2 Theoretical calculation of the Auger spectra

Accurate simulations of the kinetic energy of the Auger electrons can be obtained by using the Green function base second order diagrammatic construction method (ADC(2)) [263, 264]. This method has been successfully applied recently to simulate the Auger Spectra of pyrimidine and halogenated pyrimidines [53]. Our goal in this section is not to reproduce these spectra, but to have a simple method to estimate the kinetic energy of the Auger electrons by means of all-electron calculations.

The process starts by evaluating the first and second vertical ionization potentials (PI and PI2 respectively). To calculate them, we have fully optimized the geometry of the neutral pyrimidine at the MP2/6-311++G(d,p) level of theory. This geometry has been used to calculate the total electronic energy at higher level of theory, CCSD(T)/6-311++G(3df,2pd), of the neutral, single and double charged species. From these three values we have obtained the first ionization potential, 9.81 eV (the experimental one is 9.64 eV) and the second ionization potential, 16.78 eV. These calculations also provide the energy of the MOs shown in Fig.(10.2).

The energy released by the decay to an inner orbital K will depend on the energy of the orbital from which the electron originates. For an electron decaying from the orbital j to a core orbital, the energy is given by:

$$\Delta E_{decay}^j = E_{core} - E_j \quad (10.1)$$

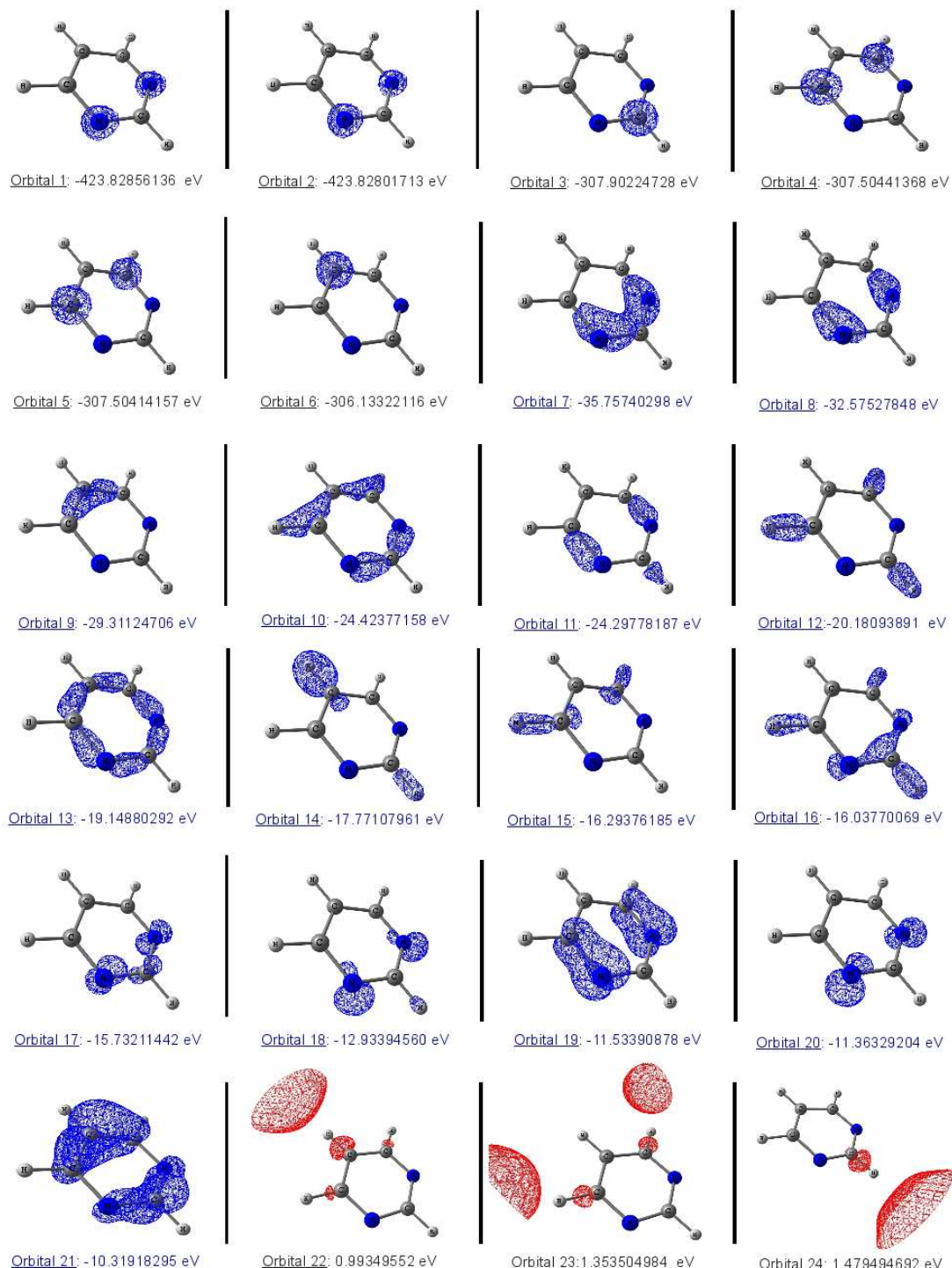


Figure 10.2: Molecular orbitals of neutral pyrimidine calculated using Gaussian09 at HF/6-311++G(3df,2pd) level. In blue occupied orbitals, in red virtual orbitals. The orbitals with black font letters are those ones that in a electronic wavefunction calculation using the CPMD code are replaced by pseudopotentials

In electron impact experiment of 1000 eV, core ionizations are the most probable process to occur [53]. For our particular study, we have only considered three core ionizations of the K-orbitals centered on C2, C4/C6 and C5 atoms (corresponding to orbitals 3, 4 and 6 from the Fig.(10.2)).

On the other hand, the energy needed by an electron i to be ionized is:

$$\Delta E_{ioniz}^i = (E_{HOMO} - E_i) + (PI2 - PI) \quad (10.2)$$

where the subtraction of the second and first ionization potentials approximates the energy needed to take the electron to the continuum.

The Auger process occurs only if:

$$\Delta E_{decay}^j > \Delta E_{ioniz}^i \quad (10.3)$$

The possible values of kinetic energies of the Auger electron ionized from an orbital i and associated with a decaying from an orbital j are given by:

$$E_K^{ij} = \Delta E_{decay}^j - \Delta E_{ioniz}^i \quad (10.4)$$

Using Eqs.(10.1) and (10.2) into Eq.(10.4) we obtain a final expression:

$$E_K^{ij} = E_{core} - E_j - (E_{HOMO} - E_i) - (PI2 - PI) \quad (10.5)$$

where i and j runs over all the pair of orbitals that fulfill the condition (10.3). Particularly in our model, E_i and E_j are the energies of different molecular orbitals of the pyrimidine 1+; that is, we assume that the Auger electron is emitted from the single charged species. As the spin state is doublet we have considered both the alpha and beta molecular orbitals from the UHF calculation.

Gathering all the possibilities for each of the K-holes (C2, C4/C6 and C5) we obtain three theoretical spectra shown in Fig.(10.3a). In order to compare with the experimental results [53] the spectra have been convoluted with gaussian functions ($\sigma^2=2.5$) and shifted by 20 eV to better reproduce the experimental spectra of Fig.(10.3b).

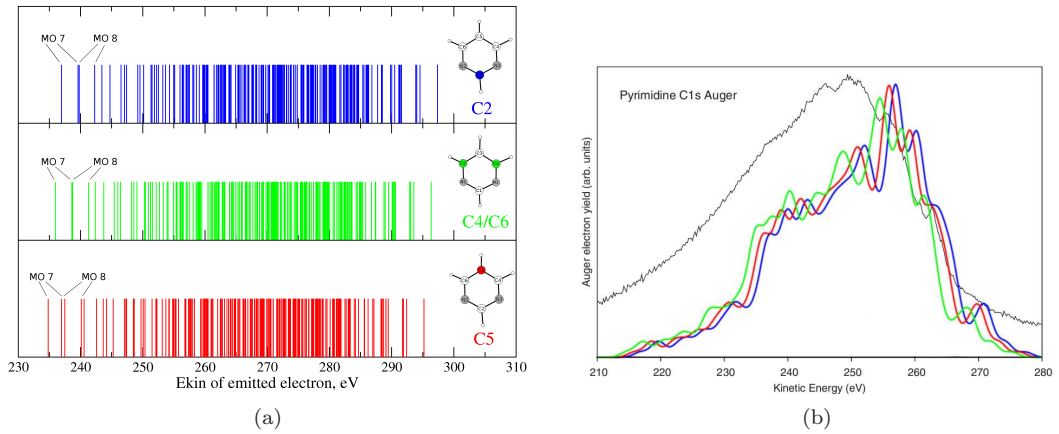


Figure 10.3: a) Theoretical E_K^{ij} values of the Auger electrons emitted after three different carbon 1s core ionizations. Lines corresponding to electrons that leave double vacancies in the MO 7 and 8 and are highlighted. b) The solid line represents the experimental Auger spectrum of pyrimidine taken from Ref. [53]; blue, green and red curves are the convoluted spectra.

The predicted kinetic energy values of Fig.(10.3a) are in accordance with the order of magnitude of the experimental ones even considering the simplicity of our model. In fact we have made quite a few rough approximations: The calculation of the kinetic energy of the Auger electron have been based on the energy of the Hartree-Fock MO of the singly charged pyrimidine evaluated in its ground state, i.e., without taking into account that the electron hole after ionization is in a core orbital that should be represented by an excited state. We have also assumed that energies necessary to extract one electron from an specific orbital is given by the energy of the MO in an independent particle mode. Despite its simplicity, the model allows us to: a) reveal the zones in which the signals are more separated, and b) to locate the kinetic energies of all the Auger electron that leave a double vacancy in the same MO.

Regarding the first issue, we must recall that the aim of this work is to predict the fragmentation pattern produced by the double ionization of a selected molecular orbital, and to determine the kinetic energy of the associated Auger electron in order to compare with a forthcoming experiment in which the fragments and the kinetic energy of the Auger electron can be determined in coincidence. In such experiments the resolution is limited, that is, the Auger electrons are selected within a certain range of energy. Therefore, we need to restrict our theoretical calculations to those regions in which the signals are as much separated as possible. By analyzing the Auger spectra of Fig.(10.3a) only the regions of lower and higher kinetic energies are within this requirement. The later involve the ionization of valence orbitals whose fragmentation will take so long that they can be dismissed beforehand (the simulations would be unaffordable in terms of CPU time).

The second issue b), listed above, is directly related to the restriction of our model. As it has been previously commented in this thesis, TD-DFT MD simulations of double charged species obtained by removing two electrons from different orbitals are prohibitive in terms of CPU time. Using calculations like the ones shown in Fig.(10.3a) we can estimate the kinetic energy of the Auger electrons produced under the desired circumstance, i.e., when a double vacancy has been created in the same orbital. Such signals are spread over all the range of the spectra (235-295 eV) but, in the region of lower kinetic energies the lines are well separated from the rest of the spectra and they should be easier to isolate in the experimental spectra.

For all these reasons, we have decided to investigate by TD-DFT MD the first lines appearing in the spectra of Fig.(10.3a) that correspond to double vacancies created in the MO 7 and 8 of Fig.(10.2). For the shake of clarity we have pointed out these signals on the spectra (10.3a), where each of the MOs leaves two signals since the kinetic energy of the Auger electrons have been calculated taking into account the energies of alpha and beta orbitals to estimate the error.

It can be seen that independently of the core orbital considered (C2, C4/C6 or C5) the kinetic energy of the Auger electron corresponding to the double hole in the MO 7 appear at a range of 235-237 eV, while for MO 8 appear at 238-240 eV, meaning that there should be no overlap between those lines.

10.3 TD-DFT MD simulation of pyrimidine²⁺

The first point before starting the simulation is to identify the HF MO 7 and 8 from an all-electron calculation to the Kohn-Sham orbitals given by a CPMD plane-wave based calculation [168]. To that end we have examined the orbitals of a neutral pyrimidine geometry obtained after an equilibration at 300 K. Fig.(10.4) shows the density of these two KS orbitals drawn with an isovalue of 0.01. As it can be seen both represent an electron density symmetrically delocalized between the two N-C bonds and, for the case of MO 7, the density is also distributed over C2. Both orbitals look like the molecular wavefunctions 7 and 8 of Fig.(10.2).

However, we have to check that the resemblances are not only restricted to the shape but to the electronic densities they represent. To that end, we have proceed to rest the densities of the orbitals 7 and 8, (stored as grid cube files) between both representations. As a result, the differences are almost negligible, indeed the subtraction functions can be only viewed if very diffuse isovalues $\approx 10^{-6}$ are used. We can conclude that the correspondence between the KS and HF orbitals is then justified.

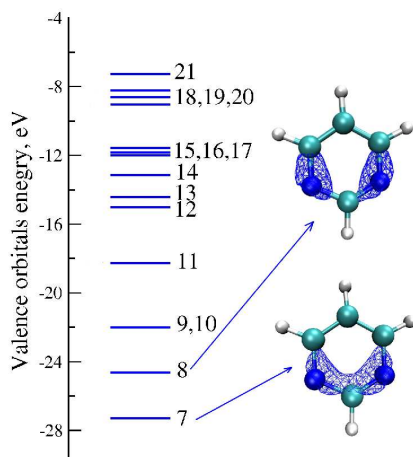


Figure 10.4: Schematic representation of the Kohn-Sham orbital energies obtained by a DFT-base CPMD calculation. The drawing corresponds to the two ionized orbitals

Once we are sure that the KS representation of the orbitals is similar to the previous calculated MOs we have proceed with the TD-DFT simulations. We have run two independent dynamics starting by the same neutral geometry cited above. The initial double charged state have been obtained by removing two electrons from MO 7 and MO 8 independly. Similar conditions have been used for both simulations, namely: the box size have been chosen to be 20Å , the basis set have been cut at 70 Ry and we have used a step of integration of 0.01 a.u. (0.00024 fs) that assures an energy conservation of ≈ 0.1 eV of the total energy of the system.

Fig.(10.5) briefly shows the results of TD-DFT MD simulation at a final time of 47.5 fs. Bonds N1-C6 and C2-N3 are broken almost at the same time ≈ 12 fs leading to two fragments of 27 and 53 masses. The bigger one breaks soon after (at 20 fs), leading to one fragment of 26 mass and two fragment of 27 mass that remain unaltered until the end of the simulation. Thus, as a result of the Coulomb explosion we have two equivalent NCH charged 0.593 and 0.646, and C₂H₂ fragment, 0.761 charged. In a mass coincidence spectrum they will lead to a univocal 26/27 signal.

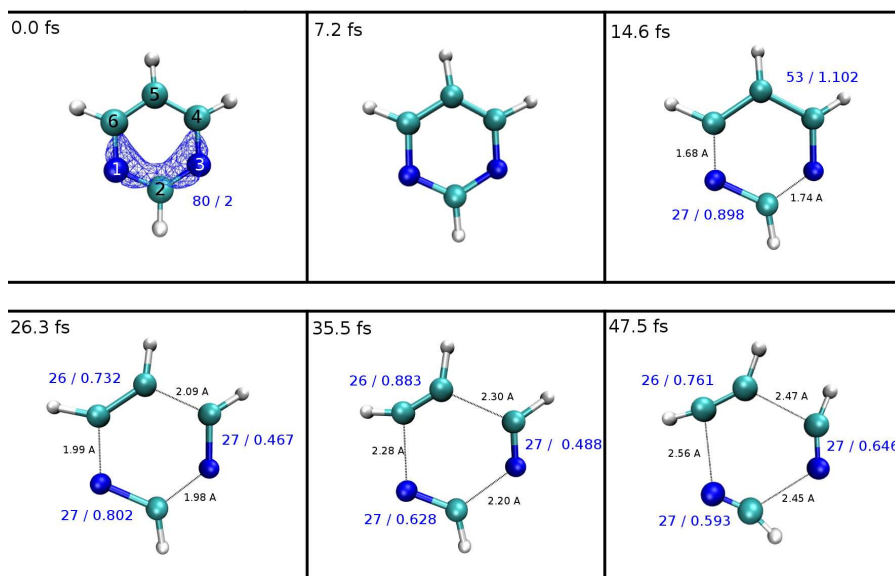


Figure 10.5: TD-DFT MD fragmentation of pyrimidine²⁺ generated after the removal of the MO 7

Fig.(10.6) corresponds to the TD-DFT simulation of the first ≈ 20 fs after the removal of the MO 8 (second deepest KS orbital). As it is shown, the first bond ring breaking is produced in N1-C6 at ≈ 9 fs. After that the N3-C4 bond breaks at 13 fs, leading to two charged fragments of 39/41 masses. Even if it is a preliminary simulation that should be extended to longer times, we can unambiguously state that MO 7 and 8 will lead to different fragmentation patterns.

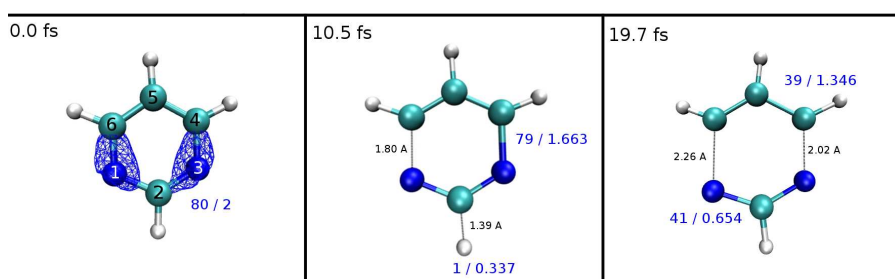


Figure 10.6: TD-DFT MD fragmentation of pyrimidine²⁺ generated after the removal of the MO 8

From this study we can conclude the emitted Auger electrons at lowest kinetic energies involve the double ionization of the two deepest valence orbitals. Two different fragmentation patterns should be observed (26/27 and 39/41) in the fragmentation spectra corresponding to the creation of a double hole in different orbitals and different kinetics energies of the associated Auger electrons.

Part IV

Conclusions

Conclusiones

En esta tesis hemos presentado una combinación de Dinámica Molecular TD-DFT y Born-Oppenheimer para estudiar la etapas iniciales (≈ 100 fs) que se suceden tras la doble ionización de moléculas de interés biológico tales como uracilo o pirimidina y pequeños agregados de agua. Los resultados obtenidos han dado lugar a las siguientes conclusiones:

- La simulación de los procesos de fragmentación del ion uracilo²⁺ en fase gas, obtenido mediante la ionización de dos electrones internos, ha mostrado estar en buen acuerdo con los experimentos de colisión de protones de 100 keV. La teoría no sólo ha sido capaz de predecir las señales obtenidas por el espectro de masas en coincidencia, sino también de dar los valores correctos de energía cinética de los fragmentos producidos.
- Se ha puesto de manifiesto la relevancia del ambiente químico intramolecular ya que orbitales de misma energía o localizados en enlaces similares, han dado lugar a fragmentaciones muy diferentes.
- Los fragmentos observados casi nunca corresponden a los canales de disociación más favorables. Este hecho está relacionado con los efectos dinámicos que tienen lugar en los primeros femtosegundos después de la ionización.
- La fragmentación del uracilo doblemente cargado, generado mediante la ionización de orbitales equivalentes, ha producido diferentes fragmentaciones tanto en fase gas como en fase líquida. Los efectos del solvente evitan la movilidad de la carga en el anillo durante los primeros femtosegundos de la fragmentación, y también conducen densidad electrónica de capas de hidratación más externas a la biomolécula.
- Hemos mostrado que la producción de oxígeno atómico puede ser consecuencia tanto de la ionización directa de la molécula de uracilo como de la ionización de una molécula de agua del disolvente. Los oxígenos producidos presentan diferentes cargas y energías cinéticas que posiblemente influenciarán el daño biológico que puedan causar sobre la biomolécula.
- En el estudio de los pequeños agregados de agua, nuestras simulaciones han corroborado la alta movilidad de carga sugerida por el experimento COLTRIMS de colisión. Hemos mostrado que los iones H_3O^+ no son producidos por la molécula de agua que pierde los dos electrones, pero sí por aquellas que capturan los protones liberados de las propias moléculas ionizadas. Por otra parte, fragmentaciones de $[\text{H}_2\text{O}]_3^{2+}$ han llevado a la formación de un oxígeno y dos protones, mostrando que el ion H_3O^+ nunca es producido por la fragmentación del trimero doblemente cargado.
- Las simulaciones del ion pirimidina²⁺ generado tras la doble ionización de los dos orbitales de valencia más profundos han dado lugar a diferentes patrones de fragmentación. La baja energía cinética del electrón Auger emitido desde ambos orbitales hace pensar en una futura combinación experimental que nos proporcione una mayor correspondencia entre el orbital ionizado y los fragmentos generados.

Conclusions

In this thesis we have presented a combination of TD-DFT Molecular Dynamics and BO Molecular Dynamics to study the first stages (≈ 100 fs) that follow the double ionization of molecules of biological interest, namely: uracil, pyrimidine and small water clusters. The results obtained have led to the following conclusions:

- Simulation of the fragmentation process of uracil²⁺ in the gas phase has shown a good agreement with the experimental results of collisions with protons of 100 keV. The theory has been able to predict the most intense signals in the coincidence mass spectrum as well as their shapes in the TOF distributions. Therefore not only the fragments but also their relative velocities seems to be correctly described by the theoretical simulations.
- The fragmentation patterns predicted for uracil in the gas phase shown the importance of the intramolecular chemical environment since orbitals of similar energy and/or localized in similar bonds have led to very different fragmentation patterns.
- The observed fragments almost never correspond to the energetically most favorable dissociation paths, pointing out that fragmentation is mainly driven by the dynamical effects occurring in the first femtoseconds after the electron removal.
- Comparison of the fragmentation process between the double ionized uracil in the gas phase and the liquid phase shows that the fragmentation patterns are quite different when electrons are removed from similar orbitals. The solvent is revealed as an active participant in the dissociation process. Its effect seems to prevent the charge mobility inside the molecule during the first femtoseconds of dynamics, and to drive electron density from outer hydration shells to the uracil.
- We have found that the atomic oxygen release can be produced by the ionization of the uracil, or from a water molecule of the solvent. The oxygen atoms produced have different charge and kinetic energy that will influence the biological damage they might cause on the biomolecule.
- In the study of small water clusters, our simulations have supported the proton mobility suggested by COLTRIMS experiments. We showed that H_3O^+ ions do not arise from the water molecule that loses the two electrons, but from the water molecules that capture protons ejected from the ionized ones. On the other hand, fragmentations of $[\text{H}_2\text{O}]_3^{2+}$ led to one oxygen atom and two protons, showing that H_3O^+ ion is never produced by the Coulomb explosion of the doubly charged trimer.
- Fragmentation patterns predicted for pyrimidine²⁺ strongly depends on the orbitals from which the electrons are extracted. In the case of the two most inner valence orbitals, not only the fragments produced are different but also the Auger electron has quite different kinetic energy. This opens the possibility of directly observe the correspondence between the orbital ionized and fragments produced.

Part V

Appendixes

Appendix A

Appendix

A.1 Ehrenfest theorem

The Ehrenfest theorem gives the variation with time of the mean value of an observable $\langle A \rangle$ varies with time:

$$\frac{d\langle A \rangle}{dt} = \left\langle \frac{\partial A}{\partial t} \right\rangle + \frac{i}{\hbar} \langle [H, A] \rangle \quad (\text{A.1})$$

where $\langle [H, A] \rangle$ is the average of the Hamiltonian commutator defined as $\langle HA - AH \rangle$. The proof starts by considering a time time-dependent wavefunction, $|\Phi(\mathbf{r}, \mathbf{R}, t)\rangle$, that satisfy the TDSE (2.1). Then, the variation of time of the average observable $\langle A \rangle$ is given by:

$$\begin{aligned} \frac{d\langle A \rangle}{dt} &= \left\langle \frac{\partial}{\partial t} \Phi(\mathbf{r}, \mathbf{R}, t) \right| \hat{A} |\Phi(\mathbf{r}, \mathbf{R}, t)\rangle + \langle \Phi(\mathbf{r}, \mathbf{R}, t) | \frac{\partial \hat{A}}{\partial t} | \Phi(\mathbf{r}, \mathbf{R}, t) \rangle + \\ &\quad + \langle \Phi(\mathbf{r}, \mathbf{R}, t) | \hat{A} | \frac{\partial}{\partial t} \Phi(\mathbf{r}, \mathbf{R}, t) \rangle \end{aligned} \quad (\text{A.2})$$

where we can identify the second term on the right as the partial derivate of the mean value of $\hat{A}$ respect to time. Considering now the postulated time variations of the wavefunction:

$$\left\langle \frac{\partial}{\partial t} \Phi(\mathbf{r}, \mathbf{R}, t) \right| = -\frac{1}{i\hbar} (\hat{H} |\Phi(\mathbf{r}, \mathbf{R}, t)\rangle)^* \quad \text{and} \quad \left| \frac{\partial}{\partial t} \Phi(\mathbf{r}, \mathbf{R}, t) \right\rangle = \frac{1}{i\hbar} \hat{H} |\Phi(\mathbf{r}, \mathbf{R}, t)\rangle \quad (\text{A.3})$$

First and third term of right-hand side of Eq. (A.2) can be written as:

$$\begin{aligned} \left\langle \frac{\partial}{\partial t} \Phi(\mathbf{r}, \mathbf{R}, t) \right| \hat{A} |\Phi(\mathbf{r}, \mathbf{R}, t)\rangle + \langle \Phi(\mathbf{r}, \mathbf{R}, t) | \hat{A} | \frac{\partial}{\partial t} \Phi(\mathbf{r}, \mathbf{R}, t) \rangle &= \\ &= -\frac{1}{i\hbar} \langle \hat{H} \Phi(\mathbf{r}, \mathbf{R}, t) | \hat{A} | \Phi(\mathbf{r}, \mathbf{R}, t) \rangle + \frac{1}{i\hbar} \langle \Phi(\mathbf{r}, \mathbf{R}, t) | \hat{A} \hat{H} | \Phi(\mathbf{r}, \mathbf{R}, t) \rangle \end{aligned} \quad (\text{A.4})$$

and considering the hermiticity of both operators, Eq.(A.4) leads to:

$$\begin{aligned} \left\langle \frac{\partial}{\partial t} \Phi(\mathbf{r}, \mathbf{R}, t) \right| \hat{A} |\Phi(\mathbf{r}, \mathbf{R}, t)\rangle + \langle \Phi(\mathbf{r}, \mathbf{R}, t) | \hat{A} | \frac{\partial}{\partial t} \Phi(\mathbf{r}, \mathbf{R}, t) \rangle &= \\ &= \frac{i}{\hbar} \langle \Phi(\mathbf{r}, \mathbf{R}, t) | \hat{H} \hat{A} | \Phi(\mathbf{r}, \mathbf{R}, t) \rangle - \frac{i}{\hbar} \langle \Phi(\mathbf{r}, \mathbf{R}, t) | \hat{A} \hat{H} | \Phi(\mathbf{r}, \mathbf{R}, t) \rangle \end{aligned} \quad (\text{A.5})$$

corresponding with the definition of the commutator $\langle [H, A] \rangle$.

A.2 The BO ansatz

Substituting the Born-Oppenheimer ansatz of Eq.(2.53) into the equation Time-dependent Schrödinger Equation (2.1), we get:

$$i\hbar \frac{\partial}{\partial t} \sum_{l=0}^{\infty} |\Psi_l(\mathbf{r}; \mathbf{R})\rangle |\chi_l(\mathbf{R}, t)\rangle = \hat{H} \sum_{l=0}^{\infty} |\Psi_l(\mathbf{r}; \mathbf{R})\rangle |\chi_l(\mathbf{R}, t)\rangle \quad (\text{A.6})$$

From the left part of the Eq.(A.6):

$$\begin{aligned} i\hbar \frac{\partial}{\partial t} \sum_{l=0}^{\infty} |\Psi_l(\mathbf{r}; \mathbf{R})\rangle |\chi_l(\mathbf{R}, t)\rangle &= \\ &= i\hbar \sum_{l=0}^{\infty} \frac{\partial |\Psi_l(\mathbf{r}; \mathbf{R})\rangle}{\partial t} |\chi_l(\mathbf{R}, t)\rangle + i\hbar \sum_{l=0}^{\infty} |\Psi_l(\mathbf{r}; \mathbf{R})\rangle \frac{\partial |\chi_l(\mathbf{R}, t)\rangle}{\partial t} \end{aligned} \quad (\text{A.7})$$

where the first term is equal to zero. Thus, replacing Eq.(A.7) and Eq.(2.3) into Eq.(A.6):

$$\begin{aligned} i\hbar \sum_{l=0}^{\infty} |\Psi_l(\mathbf{r}; \mathbf{R})\rangle \frac{\partial |\chi_l(\mathbf{R}, t)\rangle}{\partial t} &= \\ &= \hat{H}_e(\mathbf{r}, \mathbf{R}) \sum_{l=0}^{\infty} |\Psi_l(\mathbf{r}; \mathbf{R})\rangle |\chi_l(\mathbf{R}, t)\rangle + \hat{T}_N(\mathbf{R}) \sum_{l=0}^{\infty} |\Psi_l(\mathbf{r}; \mathbf{R})\rangle |\chi_l(\mathbf{R}, t)\rangle \end{aligned} \quad (\text{A.8})$$

Multiplying both sides of Eq.(A.8) by $\langle \Psi_k(\mathbf{r}; \mathbf{R}) |$:

$$\begin{aligned} i\hbar \sum_{l=0}^{\infty} \langle \Psi_k(\mathbf{r}; \mathbf{R}) | \Psi_l(\mathbf{r}; \mathbf{R}) \rangle \frac{\partial |\chi_l(\mathbf{R}, t)\rangle}{\partial t} &= \sum_{l=0}^{\infty} \langle \Psi_k(\mathbf{r}; \mathbf{R}) | E_l(\mathbf{R}) | \Psi_l(\mathbf{r}; \mathbf{R}) \rangle |\chi_l(\mathbf{R}, t)\rangle + \\ &+ \sum_{l=0}^{\infty} \langle \Psi_k(\mathbf{r}; \mathbf{R}) | \hat{T}_N(\mathbf{R}) | \Psi_l(\mathbf{r}; \mathbf{R}) \rangle |\chi_l(\mathbf{R}, t)\rangle \end{aligned} \quad (\text{A.9})$$

The second term in the right side of Eq.(A.9) is equal to:

$$\begin{aligned} \sum_{l=0}^{\infty} \langle \Psi_k(\mathbf{r}; \mathbf{R}) | \hat{T}_N(\mathbf{R}) | \Psi_l(\mathbf{r}; \mathbf{R}) \rangle |\chi_l(\mathbf{R}, t)\rangle &= \\ &= \sum_{l=0}^{\infty} \langle \Psi_k(\mathbf{r}; \mathbf{R}) | \chi_l(\mathbf{R}, t) \rangle \hat{T}_N(\mathbf{R}) | \Psi_l(\mathbf{r}; \mathbf{R}) \rangle + \sum_{l=0}^{\infty} \langle \Psi_k(\mathbf{r}; \mathbf{R}) | \Psi_l(\mathbf{r}; \mathbf{R}) \rangle \hat{T}_N(\mathbf{R}) | \chi_l(\mathbf{R}, t) \rangle - \\ &- \sum_{l=0}^{\infty} \sum_{\alpha=1}^N \frac{\hbar^2}{M_{\alpha}} \langle \Psi_k(\mathbf{r}; \mathbf{R}) | \nabla_{\alpha} | \Psi_l(\mathbf{r}; \mathbf{R}) \rangle \nabla_{\alpha} | \chi_l(\mathbf{R}, t) \rangle \end{aligned} \quad (\text{A.10})$$

Knowing this expression, taking into account that $|\Psi_k(\mathbf{r}; \mathbf{R})\rangle$ are orthogonal, Eq.(A.8) can be expressed as:

$$\begin{aligned} i\hbar \frac{\partial |\chi_k(\mathbf{R}, t)\rangle}{\partial t} &= \\ &= [\hat{T}_N(\mathbf{R}) + E_k(\mathbf{R})] |\chi_k(\mathbf{R}, t)\rangle + \\ &+ \sum_{l=0}^{\infty} \left(\langle \Psi_k(\mathbf{r}; \mathbf{R}) | \hat{T}_N(\mathbf{R}) | \Psi_l(\mathbf{r}; \mathbf{R}) \rangle - \sum_{\alpha=1}^N \frac{\hbar^2}{M_{\alpha}} \langle \Psi_k(\mathbf{r}; \mathbf{R}) | \nabla_{\alpha} | \Psi_l(\mathbf{r}; \mathbf{R}) \rangle \nabla_{\alpha} \right) |\chi_l(\mathbf{R}, t)\rangle \end{aligned} \quad (\text{A.11})$$

where the terms in parenthesis correspond to the definition of coupling operator defined as:

$$C_{kl} = \langle \Psi_k(\mathbf{r}; \mathbf{R}) | \hat{T}_N(\mathbf{R}) | \Psi_l(\mathbf{r}; \mathbf{R}) \rangle - \sum_{\alpha=1}^N \frac{\hbar^2}{M_\alpha} \{ \langle \Psi_k(\mathbf{r}; \mathbf{R}) | \nabla_\alpha | \Psi_l(\mathbf{r}; \mathbf{R}) \rangle \} \nabla_\alpha$$

A.3 Semiclassical equations

Substituting Eq.(2.62) into the TISE Eq.(2.61) we obtain:

$$[\hat{T}_N(\mathbf{R}) + E_k(\mathbf{R})]A_k(\mathbf{R}, t) \exp\left(\frac{i}{\hbar}S_k(\mathbf{R}, t)\right) = i\hbar \frac{\partial}{\partial t} A_k(\mathbf{R}, t) \exp\left(\frac{i}{\hbar}S_k(\mathbf{R}, t)\right) \quad (\text{A.12})$$

That can be written as:

$$\begin{aligned} & - \sum_{\alpha=1}^N \frac{\hbar^2}{2M_\alpha} \nabla_\alpha^2 A_k(\mathbf{R}, t) \exp\left(\frac{i}{\hbar}S_k(\mathbf{R}, t)\right) + E_k(\mathbf{R})A_k(\mathbf{R}, t) \exp\left(\frac{i}{\hbar}S_k(\mathbf{R}, t)\right) = \\ & = i\hbar \exp\left(\frac{i}{\hbar}S_k(\mathbf{R}, t)\right) \frac{\partial A_k(\mathbf{R}, t)}{\partial t} - A_k(\mathbf{R}, t) \exp\left(\frac{i}{\hbar}S_k(\mathbf{R}, t)\right) \frac{\partial S_k(\mathbf{R}, t)}{\partial t} \end{aligned} \quad (\text{A.13})$$

Expanding the left term of the equation above:

$$\begin{aligned} & - \sum_{\alpha=1}^N \frac{\hbar^2}{2M_\alpha} A_k(\mathbf{R}, t) \exp\left(\frac{i}{\hbar}S_k(\mathbf{R}, t)\right) + E_k(\mathbf{R})A_k(\mathbf{R}, t) \exp\left(\frac{i}{\hbar}S_k(\mathbf{R}, t)\right) = \\ & = - \sum_{\alpha=1}^N \frac{\hbar^2}{2M_\alpha} \exp\left(\frac{i}{\hbar}S_k(\mathbf{R}, t)\right) \nabla_\alpha^2 A_k(\mathbf{R}, t) - \sum_{\alpha=1}^N \frac{\hbar^2}{2M_\alpha} A_k(\mathbf{R}, t) \nabla_\alpha^2 \exp\left(\frac{i}{\hbar}S_k(\mathbf{R}, t)\right) - \\ & - \sum_{\alpha=1}^N \frac{\hbar^2}{M_\alpha} \nabla_\alpha A_k(\mathbf{R}, t) \nabla_\alpha \exp\left(\frac{i}{\hbar}S_k(\mathbf{R}, t)\right) + E_k(\mathbf{R})A_k(\mathbf{R}, t) \exp\left(\frac{i}{\hbar}S_k(\mathbf{R}, t)\right) \end{aligned} \quad (\text{A.14})$$

Regarding that:

$$\nabla_\alpha \exp\left(\frac{i}{\hbar}S_k(\mathbf{R}, t)\right) = \frac{i}{\hbar} \nabla_\alpha S_k(\mathbf{R}, t) \exp\left(\frac{i}{\hbar}S_k(\mathbf{R}, t)\right) \quad (\text{A.15})$$

and also the second derivative:

$$\begin{aligned} \nabla_\alpha^2 \exp\left(\frac{i}{\hbar}S_k(\mathbf{R}, t)\right) & = \nabla_\alpha \left(\frac{i}{\hbar} \nabla_\alpha S_k(\mathbf{R}, t) \exp\left(\frac{i}{\hbar}S_k(\mathbf{R}, t)\right) \right) = \\ & = \frac{i}{\hbar} \exp\left(\frac{i}{\hbar}S_k(\mathbf{R}, t)\right) \nabla_\alpha^2 S_k - \frac{1}{\hbar^2} \exp\left(\frac{i}{\hbar}S_k(\mathbf{R}, t)\right) (\nabla_\alpha S_k(\mathbf{R}, t))^2 \end{aligned} \quad (\text{A.16})$$

and dividing the expression Eq.(A.14) by Eq.(2.62), we obtain:

$$\begin{aligned} & \sum_{\alpha=1}^N \frac{\hbar^2}{2M_\alpha} A_k(\mathbf{R}, t) \exp\left(\frac{i}{\hbar}S_k(\mathbf{R}, t)\right) + E_k(\mathbf{R})A_k(\mathbf{R}, t) \exp\left(\frac{i}{\hbar}S_k(\mathbf{R}, t)\right) = \\ & = - \sum_{\alpha=1}^N \frac{\hbar^2}{2M_\alpha} \frac{\nabla_\alpha^2 A_k(\mathbf{R}, t)}{A_k(\mathbf{R}, t)} - \sum_{\alpha=1}^N \frac{i\hbar}{2M_\alpha} \nabla_\alpha^2 S_k(\mathbf{R}, t) + \sum_{\alpha=1}^N \frac{1}{2M_\alpha} (\nabla_\alpha S_k(\mathbf{R}, t))^2 - \\ & - \sum_{\alpha=1}^N \frac{i\hbar}{M_\alpha} \frac{\nabla_\alpha A_k(\mathbf{R}, t) \nabla_\alpha S_k(\mathbf{R}, t)}{A_k(\mathbf{R}, t)} + E_k(\mathbf{R}) \end{aligned} \quad (\text{A.17})$$

Finally, matching this expression with the right hand side of Eq.(A.13), previously divided by Eq.(2.62), we get:

$$\begin{aligned} & - \sum_{\alpha=1}^N \frac{\hbar^2}{2M_\alpha} \frac{\nabla_\alpha^2 A_k(\mathbf{R}, t)}{A_k(\mathbf{R}, t)} - \sum_{\alpha=1}^N \frac{i\hbar}{2M_\alpha} \nabla_\alpha^2 S_k(\mathbf{R}, t) + \sum_{\alpha=1}^N \frac{1}{2M_\alpha} (\nabla_\alpha S_k(\mathbf{R}, t))^2 - \\ & - \sum_{\alpha=1}^N \frac{i\hbar}{M_\alpha} \frac{\nabla_\alpha A_k(\mathbf{R}, t) \nabla_\alpha S_k(\mathbf{R}, t)}{A_k(\mathbf{R}, t)} + E_k(\mathbf{R}) = i\hbar \frac{1}{A_k(\mathbf{R}, t)} \frac{\partial A_k(\mathbf{R}, t)}{\partial t} - \frac{\partial S_k(\mathbf{R}, t)}{\partial t} \end{aligned}$$

where real and imaginary terms can be separated, leading straightforwardly to Eq.(2.63) and Eq.(2.64).

A.4 Hartree-Fock Approximation

A.4.1 Variational theorem

Assuming that we know the exact solutions to the Schrödinger equation, being E_0 the lowest energy value:

$$\hat{H}|\Psi(\mathbf{r}, \mathbf{R})_k\rangle = E_k|\Psi(\mathbf{r}, \mathbf{R})_k\rangle \quad k = 0, 1, 2, \dots, \infty \quad (\text{A.18})$$

If $|\phi(\mathbf{r}, \mathbf{R})\rangle$ is any well-behaved function of the same variables as $|\Psi_k(\mathbf{r}, \mathbf{R})\rangle$ and satisfies the same boundary conditions, then:

$$\xi = \frac{\langle \phi(\mathbf{r}, \mathbf{R}) | \hat{H} | \phi(\mathbf{r}, \mathbf{R}) \rangle}{\langle \phi(\mathbf{r}, \mathbf{R}) | \phi(\mathbf{r}, \mathbf{R}) \rangle} \geq E_0 \quad (\text{A.19})$$

Thus, the variational theorem establish an upper energy limit for the ground state of the system, which allow us to successively compare our energy guess in order to obtain the lowest possible value of ξ .

A.4.2 Slater determinant

The wave function that represents a system formed by n equal non-interacting electrons is:

$$|\Psi(\chi_1, \chi_2, \chi_3, \dots, \chi_n)\rangle = \prod_{i=1}^n |\chi_i(\mathbf{x})\rangle \quad (\text{A.20})$$

where functions $|\chi_i(\mathbf{x})\rangle$ are called spin-orbitals and each one represents a non-interacting electron. In turn, they are made up of the product of two functions, one to describe the spatial distribution $|\psi(\mathbf{r})\rangle$, and other, $|\alpha(\omega)\rangle$ or $|\beta(\omega)\rangle$, to describe its spin:

$$|\chi_i(\mathbf{x})\rangle = \begin{cases} |\psi(\mathbf{r})\rangle |\alpha(\omega)\rangle \\ |\psi(\mathbf{r})\rangle |\beta(\omega)\rangle \end{cases} \quad (\text{A.21})$$

The *antisymmetry principle of Pauli* [265] establishes that a many-electron wave function must be antisymmetric with respect to the interchange of the coordinates (both space and spin) of any two electrons:

$$|\Psi(\chi_1, \chi_2, \chi_3, \dots, \chi_n)\rangle = -|\Psi(\chi_1, \chi_3, \chi_2, \dots, \chi_n)\rangle \quad (\text{A.22})$$

Thus, the antisymmetry condition must be imposed on many-electron wave function (A.20) to be solution of the Schrödinger equation (2.45):

$$|\Psi_e\rangle = (N!)^{-1/2} \sum_{r=1}^{N!} \delta_r \hat{P}_r |\Psi(\chi_1, \chi_2, \chi_3, \dots, \chi_n)\rangle \quad (\text{A.23})$$

where the operator $\hat{P}_r$ is any one of the $N!$ operators, including the identity operator, that permute a given order of particles to another order. The summation is taken over all $N!$ permutation operators. The quantity δ_r is +1 or -1 if the permutation operator $\hat{P}_r$ involves the exchange of an even or odd number of particle pairs. The factor $(N!)^{-1/2}$ normalizes $|\Psi_e\rangle$ if $|\Psi(\chi_1, \chi_2, \chi_3, \dots, \chi_n)\rangle$ is normalized.

The equation (A.23) corresponds to the Slater Determinant, usually denoted as $|\chi_1 \chi_2 \chi_3 \dots \chi_n|$

A.4.3 Hartree-Fock method

The Hartree-Fock method or Self-Consistent Field (SCF) method [266] represents an approximate variational method to determine the ground-state wave function and ground-state energy of a quantum many-electron system where the wave function is assumed to be a single Slater determinant $|\Psi_e\rangle$ of n spin-orbitals (A.23).

The Hartree-Fock approximation constitutes the simplest *ab-initio* approximation to solve the time-independent Schrödinger equation and is the starting point of more accurate methods that include higher electronic correlation. Briefly, it consists in minimize the variational integral with the constraint that the spin-orbitals remain orthonormal:

$$L[\{\chi_i\}] = \xi[\{\chi_i\}] + \sum_{i=1}^n \sum_{j=1}^n l_{ij}(\langle \chi_i(\mathbf{1}) | \chi_j(\mathbf{1}) \rangle - \delta_{ij}) \quad (\text{A.24})$$

Process in which the iterative resolution of the “eigenvalue equations” known as *Hartree-Fock equations*:

$$\begin{aligned} \hat{f}(\mathbf{1})|\chi'_1(\mathbf{1})\rangle &= \epsilon_1|\chi'_1(\mathbf{1})\rangle \\ \hat{f}(\mathbf{1})|\chi'_2(\mathbf{1})\rangle &= \epsilon_2|\chi'_2(\mathbf{1})\rangle \\ &\vdots \\ \hat{f}(\mathbf{1})|\chi'_n(\mathbf{1})\rangle &= \epsilon_n|\chi'_n(\mathbf{1})\rangle \end{aligned} \quad (\text{A.25})$$

where ϵ_i term represents each time the energy of an electron in the orbital-orbital $|\chi_i\rangle$ subject to interaction with all the other electrons:

$$\epsilon_i = H_{ii}(\mathbf{1}) + \sum_{j=1}^n (J_{ij}(\mathbf{1}) - K_{ij}(\mathbf{1})) \quad (\text{A.26})$$

where the H_{ii} integral is:

$$H_{ii} = \langle \chi_i(\mathbf{1}) | \hat{h}(\mathbf{1}) | \chi_i(\mathbf{1}) \rangle \quad (\text{A.27})$$

The J_{ij} term is called “coulomb integral” and represents the electrostatic repulsion between an electron in the spin-orbital $|\chi_i\rangle$ and one in the $|\chi_j\rangle$, furthermore is related with the coulomb operator $\hat{J}_i$ as follows:

$$J_{ij} = \langle \chi_j(\mathbf{1}) | \hat{J}_i(\mathbf{1}) | \chi_j(\mathbf{1}) \rangle = \langle \chi_j(\mathbf{1}) \chi_i(\mathbf{2}) | \frac{1}{r_{12}} | \chi_j(\mathbf{1}) \chi_i(\mathbf{2}) \rangle = \langle \chi_j \chi_i | \chi_j \chi_i \rangle \quad (\text{A.28})$$

The “exchange integral” K_{ij} , however, does not have a physical meaning, but mathematically can be regarded as a correction of J_{ij} . We can write this integral as a function of the exchange operator $\hat{K}_i$:

$$K_{ij} = \langle \chi_j(\mathbf{1}) | \hat{K}_i(\mathbf{1}) | \chi_j(\mathbf{1}) \rangle = \langle \chi_j(\mathbf{1}) \chi_i(\mathbf{2}) | \frac{1}{r_{12}} | \chi_i(\mathbf{1}) \chi_j(\mathbf{2}) \rangle = \langle \chi_j \chi_i | \chi_i \chi_j \rangle \quad (\text{A.29})$$

Thus, in the Hartree-Fock approximation the complicated multielectron problem is replaced by many monoelectronic problems in which the electrons are described as independent particles interacting with each other through a mean-averaged repulsion potential.

Appendix B

List of Publications

- Journal: *In preparation*
Authors: P. López-Tarifa, M.A. Hervé du Penhoat, R. Vuilleumier, I. Tavernelli, M.-P. Gaigeot, M. Alcamí, F. Martín, and M.-F. Politis Title: “Ultrafast dissociation of double ionized uracil immersed in water”
- Journal: *In preparation*
Authors: P. López-Tarifa, M. Alcamí and F. Martín Title: “DFT calculations of small gold clusters with S, Se and Te compounds”
- Journal: Phys. Rev. Lett. submitted
Authors: P. López-Tarifa, M.A. Hervé du Penhoat, R. Vuilleumier, I. Tavernelli, M.-P. Gaigeot, J.P. Champeaux, A. Le Padellec, M. Alcamí, P. Moretto-Capelle, F. Martín, and M.-F. Politis
Title: “Ultrafast non-adiabatic fragmentation dynamics of doubly charged uracil in gas phase”
- Journal: Mutation Research-Reviews in Mutation Research (2010)
Authors: M-P. Gaigeot, P. López-Tarifa, F. Martín, M. Alcamí, R. Vuilleumier, I. Tavernelli, M. A. Hervé du Penhoat and M-F. Politis
Volume/Pages: 704, 45
Title: “Theoretical Investigation of the Ultrafast Dissociation of Ionised Biomolecules Immersed in Water: Direct and Indirect Effects”
- Journal: J. Phys. B: At. Mol. Opt. Phys. (2009)
Authors: L. Adoui, A. Cassimi, B. Gervais, L. Guillaume, S. Legendre, M. Tarisien, P. López-Tarifa, M-F. Politis, M. A. Hervé du Penhoat, R. Vuilleumier, M-P. Gaigeot, I. Tavernelli, M. Alcamí , and F. Martín
Volume/Pages: 42, 075101
Title: “Ionization and fragmentation of water clusters by fast highly charged ions”
- Journal: Croatica Chemica Acta (2009)
Authors: P. López-Tarifa, F. Martín, M. Yañez, M. Alcamí
Volume/Pages: 82, 129
Title: Theoretical study of doubly charged X-H₂O and X-NH₃ (X = Si, Ge, Sn, Pb) molecular ions

List of Figures

1.1	Bragg peak	10
1.2	Diagram of time scales	11
5.1	Energy conservation in TD-DFT MD	58
6.1	Mapping of $\nabla n(\mathbf{r})$	66
6.2	Steepest ascent paths	67
6.3	Labels of biomolecules	71
7.1	Ionized molecular orbitals of uracil gas	74
7.2	Kinetic and KS energies of gas phase dynamics	75
7.3	KS1 TD-DFT MD of uracil in gas phase	76
7.4	KS2 TD-DFT MD of uracil in gas phase	77
7.5	KS3 TD-DFT MD of uracil in gas phase	77
7.6	Switching to a BO MD	78
7.7	BO MD gas phase dynamics	79
7.8	Kinetic energy of the fragments	80
7.9	TOF spectrum of single charged fragments	81
7.10	2D mass coincidence spectrum of charged uracil	82
7.11	2D TOF coincidence spectrum of charged uracil	82
7.12	Transition states of uracil ²⁺ fragmentations	84
8.1	Summary of KS1 and KS2 dynamics in liquid phase	86
8.2	Kinetic and KS energies of liquid phase dynamics	88
8.3	Ionization of the KS1 of uracil in the liquid phase	89
8.4	Ionization of the KS2 of uracil in the liquid phase	90
8.5	Molecular orbitals of a water molecule. Energies are given for those orbitals treated explicitly in the simulations.	92
8.6	Summary of 2a1 and 1b2 dynamics in liquid phase	93
8.7	O-H distances and kinetic energies in 2a1 and ab2 dynamics in liquid phase	94
8.8	Comparison of charges gas/liquid	96
8.9	Charges of hydration shells	97
8.10	Kinetic energies and charges of O atoms	97

9.1	TOF spectrum for 12 MeV/u Ni^{25+} on $(\text{H}_2\text{O})_n$	100
9.2	Intensities of the different protonated $(\text{H}_2\text{O})_n\text{H}^+$ species	101
9.3	Coincidence spectrum of water clusters	102
9.4	TD-DFT MD simulations of $(\text{H}_2\text{O})_{11}^{2+}$	104
9.5	TD-DFT MD simulations of $(\text{H}_2\text{O})_3^{2+}$	104
10.1	Scheme of the formation pyrimidine $^{2+}$ ions by Auger effect	108
10.2	Molecular orbitals of pyrimidine	109
10.3	Theoretical and experimental Auger spectra of pyrimidine	110
10.4	Ionized orbital of pyrimidine	112
10.5	KS1 TD-DFT MD of pyrimidine in gas phase	113
10.6	KS2 TD-DFT MD of pyrimidine in gas phase	113

List of Tables

5.1	Comparison of CPU times	58
7.1	Theoretical data of fragments obtained in gas phase	80
7.2	Energies of uracil ²⁺ fragmentation channels	83
8.1	Stability of intermediate species	91

References

- [1] J. Ward, Prog. Nucleic Acid Res. Mol. Biol. **35**, 95 (1988).
- [2] G. A. Jeffrey and W. Saenger, *Hydrogen Bonding in Biological Structures* (Springer-Verlag, New York, 1991).
- [3] S. Gregoli, S. M. Olast, and A. Bertinchamps, Radiat. Res. **89**, 238 (1982).
- [4] D. Becker and M. Sevilla, Adv. Radiat. Biol. **17**, 121 (1993).
- [5] T. L. Vere, D. Becker, and M. D. Sevilla, Radiat. Res. **145**, 673 (1996).
- [6] M. Debije, M. Strickler, and A. Bernhard, Radiat. Res. **154**, 163 (2000).
- [7] N. F. Mott and H. S. W. Massey, *The Theory of Atomic Collisions*, 2nd ed. (Oxford University Press, Oxford at the Clarendon Press, 1952).
- [8] M. R. C. M. Dowell and J. Coleman, *The Theory of Atomic Collisions* (Introduction to the Theory of the Ion-Atom Collision, North Holland, Amsterdam, 1970).
- [9] C. Sonntag, *The Chemical Basis for Radiation Biology* (Taylor and Francis, London, 1987).
- [10] C. Sonntag, *Free-Radical-Induced DNA damage and its Repair* (Taylor and Francis, Springer, 2006).
- [11] B. Fayard *et al.*, Radiat. Res. **157**, 128 (2002).
- [12] M.A.Huels, I.Hahndorf, E. Illenberger, and L.Sanche, J. Chem. Phys. **108**, 1309 (1998).
- [13] M.Gromova *et al.*, Int. J. Radiat. Biol. **74**, 81 (1998).
- [14] K.M.Prise, C.H.L.Pullar, and B.D.Michael, Carcinogenesis **20**, 905 (1998).
- [15] A. Chetoui *et al.*, Int. J. Radiat. Biol. **65**, 511 (1994).
- [16] M. Watanabe *et al.*, Int. J. Radiat. Biol. **61**, 161 (1992).
- [17] M.-A. H. du Penhoat *et al.*, Radiat. Res. **151**, 649 (1999).
- [18] S. Saigusa, Y. Ejima, K. Kobayashi, and M. Sasaki, Int. J. Radiat. Biol. **61**, 785 (1992).
- [19] F.-N. Gobert *et al.*, Int. J. Radiat. Biol. **80**, 135 (2004).
- [20] A. Eschenbrenner *et al.*, Int. J. Radiat. Biol. **83**, 687 (2007).
- [21] M.-A. H. du Penhoat *et al.*, Int. J. Radiat. Biol. **86**, 205 (2010).

- [22] B. Boudaiffa *et al.*, Science **287**, 1658 (2000).
- [23] P. Agrawala *et al.*, Int. J. Radiat. Biol. **48**, 1536 (2008).
- [24] K. Fujii and A. Yokoya, Radiat. Phys. and Chem. **78**, 1188 (2009).
- [25] K. Akamatsu, K. Fujii, and A. Yokoya, Radiat. Res. **161**, 442 (2004).
- [26] K. Akamatsu, K. Fujii, and A. Yokoya, Int. J. Radiat. Biol. **80**, 849 (2004).
- [27] K. Fujii, K. Akamatsu, and A. Yokoya, Radiat. Res. **161**, 435 (2004).
- [28] K. Fujii, K. Akamatsu, and A. Yokoya, Int. J. Radiat. Biol. **80**, 909 (2004).
- [29] A. Yokoya, K. Akamatsu, K. Fujii, and M. Ukai, Int. J. Radiat. Biol. **80**, 833 (2004).
- [30] Y. Gauduel, *Ultrafast Dynamics of Chemical Systems* (Kluwer Publisher, ADDRESS, 1994).
- [31] Y. Gauduel *et al.*, Biochemistry **27**, 2509 (1988).
- [32] Y. Gauduel, H. Gelabert, and F. Guillaud, J. Am. Chem. Soc. **112**, 5082 (2000).
- [33] W. Chun-Rong, J. Nguyen, and L. Qing-Bin, J. Am. Chem. Soc. **113**, 11320 (2009).
- [34] B. Gervais, M. Beuve, G. Olivera, and M. Galassi, Radiat. Phys. Chem. **75**, 493 (2006).
- [35] M.-P. Gaigneot *et al.*, Phys. B: At. Mol. Opt. Phys. **40**, 1 (2007).
- [36] C. J. Mundy, M. E. Colvin, and A. A. Quong, J. Phys. Chem. A **106**, 10063 (2002).
- [37] Y. Wu, C. Mundy, M. Colvin, and R. Car, J. Phys. Chem. A **108**, 2922 (2004).
- [38] F. Ban, K. Rankin, J. Gauld, and R. Boyd, Theor. Chem. Acc. **108**, 11 (2002).
- [39] S. Naumov and C. von Sonntag, Radiat. Res. **169**, 355 (2008).
- [40] N. Jena and P. C. Mishra, J. Phys. Chem. B. **109**, 14205 (2005).
- [41] A. Kumar and M. Sevilla, *Radiation effects on DNA: theoretical investigations of electron hole and excitation pathways to DNA damage* (Springer-Verlag, Berlin, 2008).
- [42] X. Li and M. Sevilla, Adv. Quant. Chem. **52**, 59 (2007).
- [43] R. H. D. Lyngdoh and H. S. III, Acc. Chem. Res. **42**, 563 (2009).
- [44] I. Tavernelli, U. F. Rohrig, and U. Rothlisberger, Mol. Phys. **103**, 963 (2005).
- [45] I. Tavernelli *et al.*, Chem. Phys. Chem. **9**, 2099 (2008).
- [46] P. Moretto-Capelle and A. L. Padellec, Phys. Rev. A **74**, 062705 (2006).
- [47] T. Schlatholter, R. Hoekstra, and R. Morgenstern, Int. J. Mass Spectrom. **233**, 173 (2004).
- [48] J. de Vries, R. Hoekstra, R. Morgenstern, and T. Schlatholter, Phys. Rev. Lett. **91**, 053401 (2003).
- [49] J. Tabet *et al.*, Phys. Rev. A **81**, 012711 (2010).
- [50] B. Coupier *et al.*, Eur. Phys. J. D **20**, 459 (2002).

- [51] J. de Vries, R. Hoekstra, R. Morgenstern, and T. Schlatholter, *J. Phys. B: At. Mol. Opt. Phys* **35**, 4373 (2002).
- [52] J. de Vries, R. Hoekstra, R. Morgenstern, and T. Schlatholter, *Phys. Scr.* **T110**, 336 (2004).
- [53] L. Storchi *et al.*, *J. Chem. Phys.* **129**, 154309 (2008).
- [54] O. Plekan *et al.*, *Phys. Scr.* **78**, 058105 (2008).
- [55] M. Schwell, H.-W. Jochims, H. Baumgrtel, and S. Leach, *Chem. Phys.* **353**, 145 (2008).
- [56] B. J. Alder and T. E. Wainwright, *J. Chem. Phys.* **27**, 1208 (1957).
- [57] B. J. Alder and T. E. Wainwright, *J. Chem. Phys.* **31**, 459 (1959).
- [58] A. Rahman, *Phys. Rev. A* **136**, 405 (1964).
- [59] F. H. Stillinger and A. Rahman, *J. Chem. Phys.* **60**, 1545 (1974).
- [60] J. A. McCammon, B. R. Gelin, and M. Karplus, *Nature* **267**, 585 (1977).
- [61] J. M. Haile, *Molecular dynamics simulations: Elementary Methods* (Eilry Professional Paperback Series, New York, 1997).
- [62] D. C. Rappaport, *The art of molecular dynamic simulation*, 2nd ed. (Cambridge University Press, UK, 2004).
- [63] D. Frenkel and B. Smit, *Understanding molecular simulation*, 2nd ed. (Academic press, California, 2002).
- [64] A. R. Leach, *Molecular modeling: Principles and applications*, 2nd ed. (Pearson Education, England, 2001).
- [65] G. Sutmann, in *Quantum simulation of complex many-body systems: from theory to algorithms, Lecture Notes*, Vol. 10 of *NIC Series*, edited by J. Grotendorst, D. Marx, and A. Maramatsu (Jon von Neumann Institute for Computing, ADDRESS, 2002), p. 211.
- [66] P. Balbuena and J. M. Seminario, *From Classical to Quantum methods*, 1st ed. (Elsevier Science, Netherlands, 1999).
- [67] M. E. Tuckerman, *J. Phys. Chem. B* **104**, 159 (2000).
- [68] M. C. Payne *et al.*, *Rev. Mod. Phys.* **64**, 1045 (1992).
- [69] R. Vuilleumier, *Lect. Notes Phys.* **703**, 217 (2006).
- [70] *Multidimensional Quantum Dynamics*, 1st ed. (Wiley-VCH, Germany, 2009).
- [71] *Introduction to Quantum Mechanics: a time-dependent perspective*, 1st ed. (University Science Books, California, 2007).
- [72] E. Deumens, A. Diz, R. Longo, and Y. Ohrn, *Rev. Mod. Phys.* **66**, 917 (1994).
- [73] *Introduction to Quantum Mechanics: a time-dependent perspective*, 1st ed. (Cambridge University Press, Cambridge, UK, 2009).
- [74] J. C. Tully, in *Moder Methods for multidimensional dynamics computations in Chemistry*, edited by B. J. Berne, G. Ciccotti, and D. F. Coker (World Scientific, London, 1998), Chap. II.

- [75] E. Schrodinger, Phys. Rev. **28**, 1049 (1926).
- [76] D. S. Sholl and J. A. Steckel, *Density functional theory* (Cambridge University Press, New York, 2002).
- [77] K. Gottfried, *Quantum Mechanics: Fundamentals* (Springer, New York, 2003).
- [78] A. Messiah, *Quantum Mechanics*, 12th ed. (Dover Publications, INC., New York, 1991).
- [79] P. Guttinger, Z. Phys. **73**, 169 (1932).
- [80] W. Pauli, *Handbuch der Physik* (Springer, Berlin, 1933), p. 162.
- [81] H. Hellmann, *Einführung in die Quantenchemie* (Franz Deuticke, Leipzig, 1937), p. 285.
- [82] R. Feynman, Phys. Rev. **56**, 340 (1939).
- [83] P. Pulay, *Modern Theoretical Chemistry* (Plenum, New York, 1977), Vol. 4.
- [84] B. M. Deb, Rev. Mod. Phys. **45**, 22 (1973).
- [85] M. D. Ventra and S. T. Pantelides, Phys. Rev. B **61**, 16207 (2001).
- [86] J. I. Steinfeld, *Molecules and Radiation: An Introduction to Modern Molecular Spectroscopy* (MIT Press, Cambridge, UK, 2005).
- [87] M. Born and R. Oppenheimer, Annalen der Physik **84**, 457 (1927).
- [88] M. Baer, *Beyond Born-Oppenheimer* (John Wiley and Sons, INC., New York, 2006), Chap. II.
- [89] W. Domcke, D. R. Yarkony, and H. Koppel, *Canonical intersections* (World Scientific Publishing Co., London, 2004), Chap. I.
- [90] P. A. M. Dirac, *Principles of Quantum Mechanics*, 3rd ed. (Oxford University Press, Oxford, 1947).
- [91] D. Durr and S. Teufel, *Bohmian Mechanics: The Physics and Mathematics of Quantum Theory*, 1st ed. (Springer, Berlin, 2009).
- [92] J. T. Bruer, Synthese **50**, 167 (1982).
- [93] H. Goldstein, C. Poole, and J. Safko, *Classical Mechanics*, 3rd ed. (Addison Wesley, San Francisco, 2002).
- [94] R. Car and M. Parrinello, Phys. Rev. Lett. **22**, 2471 (1985).
- [95] J. C. Tully, in *Classical and Quantum Dynamics in Condensed Phase Simulations*, edited by B. J. Berne, G. Ciccotti, and D. F. Coker (World Scientific, Singapore, 1998), Chap. XXI.
- [96] P. A. M. Dirac, Proc. Cambridge Philos. Soc. **26**, 376 (1930).
- [97] E. J. Heller, J. Chem. Phys. **64**, 63 (1976).
- [98] R. Harris, J. Chem. Phys. **72**, 1776 (1980).
- [99] R. B. Gerber, V. Buch, and M. A. Ratner, J. Chem. Phys. **77**, 3022 (1982).

- [100] N. Makri and W. H. Miller, J. Chem. Phys. **87**, 5781 (1987).
- [101] R. B. Gerber and M. A. Ratner, Adv. Chem. Phys. **70**, 97 (1988).
- [102] M. Messina and R. D. Coalson, J. Phys. Chem. **90**, 4015 (1989).
- [103] K. Haug and H. Metiu, J. Phys. Chem. **99**, 6253 (1993).
- [104] J. Kohanoff, *Electronic structure calculations for solids and molecules*, 1st ed. (Cambridge, UK, 2006).
- [105] P. Ehrenfest, Zeitschrift fur Physik **45**, 455 (1927).
- [106] N. L. Doltsinis and D. Marx, J. Theor. Comp. Chem. **2**, 319 (2002).
- [107] J. Tully, M. Gomez, and M. Head-Gordon, J. Vac. Sci. Technol. **A11**, 1914 (1993).
- [108] S. Klein *et al.*, Chem. Phys. Lett. **292**, 259 (1998).
- [109] A. P. Horsfield, D. R. Bowler, and A. J. Fisher, J. Phys.: Condens. Matter **16**, L65 (2004).
- [110] X. Li, J. C. Tully, H. B. Schlegel, and M. J. Frisch, J. Chem. Phys. **123**, 084106 (2005).
- [111] C. M. Isborn, X. Li, and J. C. Tully, J. Chem. Phys. **126**, 134307 (2007).
- [112] X. Andrade *et al.*, J. Chem. Theo. Comp. **5**, 728 (2009).
- [113] W. Liang *et al.*, J. Phys. Chem. A **114**, 6201 (2010).
- [114] W. Koch, M. C. Holthausen, and E. J. Baerends, *A Chemist's Guide to Density Functional Theory* (WILEY-VCH, Weinheim, 2001).
- [115] A. J. Cohen, P. Mori-Sánchez, and W. Yang, Science **321**, 792 (2008).
- [116] I. N. Levine, *Quantum chemistry* (Prentice Hall, New Jersey, 1999), Chap. I.
- [117] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
- [118] R. M. Dreizler and E. K. U. Gross, *Density Functional Theory* (Berlin Heidelberg, Weinheim, 1990).
- [119] W. Kohn and L. J. Sham, Phys. Rev **140**, A1133 (1965).
- [120] R. G. Parr and W. Yang, *Density-functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989).
- [121] R. M. Martin, *Electronic Structure : Basic Theory and Practical Methods* (Cambridge University Press, ADDRESS, 2004).
- [122] J. Slater, *Quantum Theory of Molecular and Solids. Vol. 4: The Self-Consistent Field for Molecular and Solids* (McGraw-Hill, New York, 1974).
- [123] J. Slater, Phys. Rev. **81**, 385 (1951).
- [124] S. Vosko, L. Wilk, and M. Nusair, Can. J. Phys. **58**, 1200 (1980).
- [125] R. O. Jones and O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989).
- [126] U. von Barth, *Methods of Electronic Structure Calculations* (World Scientific, Singapore, 1994).

- [127] F. Jensen, *Introduction to Computational Chemistry* (Wiley, New York, 1999).
- [128] J. P. Perdew and W. Yang, Phys. Rev. B **33**, 8800 (1986).
- [129] A. D. Becke, Phys. Rev. A **38**, 3098 (1988).
- [130] P. Gill, Mol. Phys. **89**, 433 (1996).
- [131] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).
- [132] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **78**, 1396 (1997).
- [133] C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).
- [134] R. Colle and O. Salvetti, Theor. Chim. Acta **37**, 4135 (1975).
- [135] A. D. Corso, A. Pasquarello, A. Baldereschi, and R. Car, Phys. Rev. B **53**, 1180 (1996).
- [136] M. Sprik, J. Hutter, and M. Parrinello, J. Chem. Phys. **105**, 1142 (1996).
- [137] A. D. Becke, J. Chem. Phys. **104**, 1040 (1996).
- [138] A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).
- [139] P. J. Stephens, F. J. Devlin, C. F. Chablowski, and M. J. Frisch, J. Phys. Chem. **98**, 11623 (1994).
- [140] A. Alavi *et al.*, Phys. Rev. Lett. **80**, 3650 (1998).
- [141] W. Pollard and R. Friesner, J. Chem. Phys. **99**, 6742 (1993).
- [142] A. Alavi, J. Kohanoff, M. Parrinello, and D. Frenkel, Phys. Rev. Lett. **73**, 2599 (1994).
- [143] J. Cullum and R. A. Willoughby, *Lanczos algorithms for large symmetric eigenvalue computations, Vol. 1: Theory of Progress in Scientific Computing*, 3rd ed. (Birkhauser, Boston, 1985).
- [144] C.-O. Almbladh and U. von Barth, Phys. Rev. B **31**, 3231 (1985).
- [145] M. P. Teter, M. C. Payne, and D. C. Allen, Phys. Rev. B **40**, 12255 (1989).
- [146] P. Pulay, Chem. Phys. Lett. **73**, 393 (1980).
- [147] C. Csaszar and P. Pulay, J. Mol. Struct. **114**, 31 (1984).
- [148] J. Hutter, H. P. Luthi, and M. Parrinello, Comput. Mat. Sci. **2**, 244 (1994).
- [149] N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, 1st ed. (Brooks Cole, USA, 1976).
- [150] M. P. M. Boero, K. Terakura, T. Ikeshoji, and C. C. Liew, Phys. Rev. Lett. **226403**, 90 (2003).
- [151] C. Nicolas, R. Spezia, A. Boutin, and R. Vuilleumier, Phys. Rev. Lett. **91**, 208304 (2003).
- [152] R. Spezia *et al.*, Mol. Sim. **30**, 749 (2004).
- [153] J. C. Phillips, Phys. Rev. **112**, 685 (1958).
- [154] M. L. Cohen and V. Heine, Solid State Phys. **24**, 37 (1970).

- [155] M. T. Yin and M. L. Cohen, Phys. Rev. B **25**, 7403 (1982).
- [156] P. Fuentealba, H. Preussa, H. Stolla, and L. V. Szentpály, Chem. Phys. Lett. **89**, 418 (1982).
- [157] M. T. Yin and M. L. Cohen, Phys. Rev. B **52**, 11793 (1995).
- [158] R. Hamann, M. Schluter, and C. Chiang, Phys. Rev. Lett. **43**, 1494 (1979).
- [159] G. B. Bachelet, D. R. Hamann, and M. Schluter, Phys. Rev. Lett. **43**, 1494 (1979).
- [160] N. Troullier and J. L. Martins, Phys. Rev. B **43**, 1993 (1991).
- [161] J. Pipek and P. G. Mezey, J. Chem. Phys. **90**, 4916 (1989).
- [162] D. A. Kleier, T. A. Halgren, J. H. H. Jr, and W. N. Lipscomb, J. Chem. Phys. **61**, 3905 (1974).
- [163] J. M. Foster and S. F. Boys, Rev. Mod. Phys. **32**, 300 (1960).
- [164] N. Marzari and D. Vanderbilt, Phys. Rev. B **56**, 12847 (1997).
- [165] L. Verlet, Phys. Rev. **159**, 98 (1967).
- [166] L. Verlet, Phys. Rev. **165**, 201 (1967).
- [167] R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics, Vol. 1*, 2nd ed. (Addison-Wesley, New-York, 2005).
- [168] CPMD V3.13, copyright IBM Corp (1990-2008), copyright MPI fuer Festkoerperforschung Stuttgart (1997-2001).
- [169] H. C. Andersen, J. Comp. Phys. **52**, 24 (1983).
- [170] E. Runge and E. K. U. Gross, Phys. Rev. Lett. **52**, 997 (1984).
- [171] M. A. L. Marques and E. K. U. Gross, Annu. Rev. Phys. Chem. **55**, 427 (2004).
- [172] M. A. L. Marques *et al.*, *Lecture Notes in Physics: Time-Dependent Density Functional Theory* (Springer, Germany, 2005).
- [173] K. Burke, J. Werschnik, and E. Gross, J. Chem. Phys. **123**, 062206 (2005).
- [174] F. Furche and K. Burke, Annual Reports in Computational Chemistry **1**, 19 (2005).
- [175] N. Maitra, A. Wasserman, and K. Burke, *What is time-dependent density functional theory? Successes and Challenges* (Springer, Germany, 2003).
- [176] P. Elliott, K. Burke, and F. Furche, *Reviews in Computational Chemistry* (Wiley, Hoboken, NJ, 2009), p. 91.
- [177] R. V. Leeuwen, Int. J. Mod. Phys. B **14**, 1969 (2001).
- [178] M. E. Casida, *Recent Advances in Density Functional Methods* (World Scientific, Singapore, 1995), p. 155.
- [179] R. van Leeuwen, Phys. Rev. Lett. **80**, 1280 (1998).
- [180] L. V. Keldysh, Sov. Phys. **20**, 1018 (1965).
- [181] J. Sakurai, *Modern Quantum Mechanics* (Addison Wesley Longman, USA, 1994), Chap. II.

- [182] R. Kosloff, J. Phys. Chem. **92**, 2087 (1988).
- [183] A. Castro, M. A. L. Marques, J. A. Alonso, and A. Rubio, J. Chem. Phys. **121**, 3425 (2004).
- [184] R. W. Hockney, Methods Comput. Phys. **9**, 136 (1970).
- [185] G. Baldacchino *et al.*, Nucl. Instr. Meth. Phys. Res. B. **146**, 528 (1998).
- [186] G. Baldacchino *et al.*, Radiat. Res. **149**, 128 (1998).
- [187] D. McQuarrie, *Statistical Mechanics* (Harper and Row, New York, 1976).
- [188] D. Chandler, *Introduction to Modern Statistical Mechanics* (Oxford University Press, New York, 1987).
- [189] R. E. Wilde and S. Singh, *Statistical Mechanics, Fundamentals and Modern Applications* (John Wiley and Sons, Inc, New York, 1998).
- [190] L. V. Woodcock, Chem. Phys. Lett. **10**, 257 (1971).
- [191] H. J. C. B. et al, J. Chem. Phys. **81**, 3684 (1984).
- [192] S. Nosé, Mol. Phys. **53**, 255 (1984).
- [193] W. G. Hoover, Phys. Rev. A **31**, 1695 (1985).
- [194] R. S. Mulliken, J. Chem. Phys. **23**, 1833 (1955).
- [195] R. S. Mulliken, J. Chem Phys. **36**, 3428 (1962).
- [196] A. E. Reed, L. A. Curtiss, and F. Weinhold, Chem. Rev. **88**, 899 (1988).
- [197] P.-O. Lwdin, Adv. Quant. Chem. **5**, 185 (1970).
- [198] R. F. W. Bader, *Atoms in molecules: A quantum theory* (Oxford University Press, Oxford, 1990).
- [199] R. F. W. Bader, Chem. Rev **91**, 893 (1991).
- [200] G. K. H. Madsen *et al.*, Phys. Rev. B **59**, 12359 (1999).
- [201] R. F. W. Bader, Phys Rev B **49**, 13348 (1994).
- [202] R. F. W. Bader, T. T. N. en Dang, and Y. Gal, Rep. Prog. Phys. **44**, 893 (1981).
- [203] R. F. W. Bader, Y. Tal, S. G. Anderson, and T. T. N. en Dang, Israel J. Chem. **19**, 8 (1980).
- [204] B. B. Stefanov and J. Cioslowski, J. Comput. Chem **16**, 1394 (1995).
- [205] G. Henkelman, A. Arnaldsson, and H. Jnsson, Comp. Mat. Sci. **36**, 354 (2005).
- [206] W. Tang, E. Sanville, and G. Henkelman, J. Phys.: Condens. Matter **20**, in press (2008).
- [207] E. Sanville, S. D. Kenny, R. Smith, and G. Henkelman, J. Comput. Chem. **28**, 899 (2007).
- [208] H.-J. Li and J.-J. Ho, J. Phys. Chem. C **114**, 1194 (2010).

- [209] G. Giorgi, M. V. Schilfgaarde, A. Korkin, and K. Yamashita, *Nanoscale. Res. Lett.* **5**, 469 (2010).
- [210] G. Giorgi, M. V. Schilfgaarde, A. Korkin, and K. Yamashita, *J. Phys. Chem. A* **114**, 4486 (2010).
- [211] López-Tarifa *et al.*, submitted –, (2010).
- [212] L. Adoui *et al.*, **48**, 075101 (2009).
- [213] Gaussian 09 Revision A.1, gaussian Inc. Wallingford CT 2009.
- [214] J. S. Binkley and J. A. Pople, *J. Am. Chem. Soc.* **102**, 939 (1980).
- [215] W. J. Hehre, R. Ditchfiel, and J. A. Pople, *J. Chem. Phys.* **56**, 2257 (1972).
- [216] R. Krishnan, J. S. Binkley, and R. S. an J. A. Pople, *J. Chem. Phys.* **72**, 650 (1980).
- [217] M. J. Frisch, J. A. Pople, and J. S. Binkley, *J. Chem. Phys.* **80**, 3265 (1984).
- [218] M. M. Francl *et al.*, *J. Chem. Phys.* **77**, 3654 (1982).
- [219] B. Liu *et al.*, *Phys. Rev. Lett.* **97**, 133401 (2006).
- [220] M. Nuevo *et al.*, *Astrobiology* **9**, 683 (2009).
- [221] P. Cafarelli *et al.*, in *Radiation Damage in Biomolecular Systems*, edited by K. Tokési and B. Sulik (Springer, Debrecen, Hungary, 2008), Vol. 1080, p. 71.
- [222] S. Denifl *et al.*, *Int. J. Mass Spectrom.* **238**, 47 (2004).
- [223] H.-W. Jochims, M. Schwell, H. Baumgrtel, and S. Leach, *Chem. Phys.* **314**, 263 (2005).
- [224] M.-P. Gaijeot and M. Sprick, *J. Phys. Chem. B* **108**, 7458 (2004).
- [225] R. Vuilleumier and M. Sprik, *J. Chem. Phys.* **115**, 3454 (2001).
- [226] L. Kleinman and D. M. Bylander, *Phys. Rev. Lett.* **48**, 1425 (1982).
- [227] S. Miertus, E. Scrocco, and J. Tomasi, *J. Chem. Phys.* **55**, 117 (1981).
- [228] S. Miertus and J. Tomasi, *J. Chem. Phys.* **65**, 239 (1982).
- [229] Z. Bacic and R. E. Miller, *J. Phys. Chem.* **100**, 12945 (1996).
- [230] I. V. Hertel, C. Huglin, C. Nitsch, and C. P. Schulz, *Phys. Rev. Lett* **67**, 1767 (1991).
- [231] F. N. Keutsch and R. J. Saykally, *Proc. Natl. Acad. Sci.* **98**, 10533 (2001).
- [232] A. W. J. Castleman and K. H. J. Bowen, *J. Phys. Chem.* **100**, 12911 (1996).
- [233] R. N. Barnett and U. Landman, *Phys. Rev. Lett.* **70**, 1775 (1993).
- [234] O. Bjorneholm, F. Federmann, S. Kakar, and T. Moller, *J. Chem. Phys.* **111**, 546 (1999).
- [235] Z. Sternovsky, M. Horányi, and S. Robertson, *Phys. Rev. A* **64**, 023203 (2001).
- [236] F. H. Stillinger, *Science* **209**, 451 (1980).
- [237] D. R. Miller, *Atomic and Molecular Beam Methods* (Oxford University Press, New York, 1988).

- [238] J. Farges, M. F. de Feraudy, B. Raoult, and G. Torchet, J. Chem. Phys. **78**, 5067 (1983).
- [239] C. Bobbert, S. Schutte, C. Steinbach, and U. Buck, Eur. Phys. J. D **19**, 183 (2002).
- [240] S. Legendre, Master's thesis, University of Caen Basse-Normandie, 2006.
- [241] S. Legendre *et al.*, J. Phys. B: At. Mol. Opt. Phys. **38**, L233 (2005).
- [242] L. A. et al, Nucl. Instrum. Method B **245**, 94 (2006).
- [243] L. A. et al, J. Phys: Conf. Ser. **88**, 012023 (2007).
- [244] J.-P. G. C. Caraby, L. Adoui and A. Cassimi, Eur. Phys. J. D **2**, 53 (1998).
- [245] M. T. et al, J. Phys. B: At. Mol. Opt. Phys. **33**, L11 (2000).
- [246] G. Laurent *et al.*, Phys. Rev. Lett. **70**, 1775 (2006).
- [247] F. Chandezon *et al.*, Phys. Rev. Lett. **87**, 153402 (2001).
- [248] O. Echt, D. Kreisle, M. Knapp, and E. Recknagel, Chem. Phys. Lett. **108**, 401 (1984).
- [249] G. M. Lancaster, F. Honda, Y. Fukuda, and J. W. Rabalais, J. Am. Chem. Soc. **101**, 1951 (1979).
- [250] T. Matsuo *et al.*, J. Chem. Phys. **101**, 5356 (1994).
- [251] J. M. H. et al, Science **308**, 1765 (2005).
- [252] R. Ludwig, Chem. Phys. Chem. **5**, 1495 (2004).
- [253] M. Eigen, Angew. Chem. **75**, 489 (1964).
- [254] G. Zundel and H. Metzger, Z. Phys. Chem. **58**, 225 (1968).
- [255] E. S. Wisnieswski, J. R. Stairs, and A. W. J. Castleman, Int. J. Mass. Spectrom. **212**, 273 (2001).
- [256] F. Dong, S. Heinbuch, J. J. Rocca, and E. R. Bernstein, J. Chem. Phys. **124**, 224319 (2006).
- [257] T. Wroblewski, Radiat. Phys. Chem. **68**, 313 (2003).
- [258] B. Manil *et al.*, Phys. Rev. Lett. **91**, 215504 (2003).
- [259] G. O. et al, J. Chem. Phys. **123**, 054310 (2005).
- [260] O. Mó, M. Yaez, and J. Elguero, J. Chem. Phys. **97**, 1992 (6628).
- [261] W. Dewey and R. Humphrey, Radiat. Res. **26**, 538 (1965).
- [262] K. Sano, T. Hoshino, M. Nagai, and J. Neurosurg, Radiat. Res. **28**, 530 (1968).
- [263] J. Schirmer and A. Barth, Z. Phys. A **317**, 267 (1984).
- [264] F. Tarantelli, Chem. Phys. **329**, 11 (2006).
- [265] W. Pauli, Nobel Lecture (1946).
- [266] A. Szabo and N. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory* (Dover Publications, INC., Mineola, New York, 1989).