Numerical simulation of ultrafast response of biological molecules subjected to light fast ions and XUV pulses

Karwan Omar

université PARIS-SACLAY

Numerical simulations of ultrafast response of biological molecules subjected to light fast ions and XUV pulses

Simulations numériques des réponses ultra-rapides de molécules biologiques soumises à des ions légers rapides et à des impulsions UV extrême

Thèse de doctorat de l'université Paris-Saclay

École doctorale n°571 : sciences chimiques : molécules, matériaux, instrumentation et biosystèmes (2MIB)
Spécialité de doctorat : Chimie
Graduate School : Chimie. Référent : Faculté des sciences d'Orsay

Thèse préparée à l'**Institut Chimie Pysique** (Université Paris-Saclay, CNRS), sous la direction de **Aurélien de la LANDE**, Directeur de recherche

Thèse soutenue à Paris-Saclay, le 22 novembre 2023, par

Karwan Ali OMAR

Composition du Jury

Membres du jury avec voix délibérative

Mme Sophie LE CAËR Directrice de recherche, CEA-CNRS Saclay (NIMBE)	Présidente
Mme Thi Phuong Mai SÈVE-DINH Professeure, Université Paul Sabatier (LPT)	Rapportrice & Examinatrice
M. Mathias RAPACIOLI Chargé de recherche, CNRS Toulouse (LCPQ)	Rapporteur & Examinateur
Mme Valérie VENIARD Directrice de recherche, CNRS Palaiseau (LSI)	Examinatrice
M. Franck LEPINE Directeur de recherche, CNRS Lyon (ILM)	Examinateur
M. Fabien BRUNEVAL Directeur de recherche, CEA Saclay (SRMP)	Examinateur

THESE DE DOCTORAT
NNT : 2023UPASF055

چیبچێنی ئەوە ئەدوریتەوە

Acknowledgements

Expressing gratitude to individuals who wholeheartedly supported me throughout my PhD journey is challenging. I extend my sincerest thanks to:

Firstly, my supervisor Aurélien de la LANDE. He's more than a supervisor; he's been a true friend. Our acquaintance began during my master's internship in 2014, and he has consistently offered guidance and unwavering support to this day. His mentorship extends beyond the realms of science, encompassing everyday life. A remarkable scientist and educator, he imparted invaluable knowledge, for which I am immensely grateful for your continuous support.

I am deeply grateful to Carine CLAVAGUERA and Fabien CAILLIEZ, my invaluable secondary supervisors and advisors. They provided unwavering guidance and advice from the inception to the culmination of my thesis. Their support extended beyond academic matters to encompass everyday challenges, akin to that of family. I am truly appreciative of their assistance.

My heartfelt thanks to my wife, Khabat, and my daughters Varya and newborn Bavos. Particularly, my wife's unwavering support has been my cornerstone. Her continuous encouragement and dedication to my scientific journey, often making significant sacrifices, have been remarkable.

Additionally, my sincere appreciation extends to my parents, brother, and sisters, whose unwavering support has been a constant throughout my life, shaping me into who I am today.

I wish to extend my gratitude to Karim Hasnaoui for implementing GPU in deMon2K and Xiaojing WU for implementing RT-TD-DFT in deMon2K. I also want to express my appreciation to our PhD group: Feven A. Korsaye, Damien Tolu, Jean Deviers, Angela Parise, Julia Attard, Jesus Nain Pedroza-Montero, Baptiste Etcheverry, and Mosbah El-hayek for their engaging discussions during our group meetings.

My thanks also go out to our research team, ThéoSim, especially David Lauvergnat, Federica Agostini, Ridard Jacqueline, Pascal Pernot, Bernard Lévy, Bernard Rousseau, Michèle Desouter-Lecomte, Dominik Domin, Jessica Bowles, Rika Tandiana, and Valeriu Cemortan.

Acknowledgment is also owed to the members of ICP, particularly Jessica Martinez for logistical support and administrative personnel Wafa Zaied, Carine Roger-Roulling, and Séverine Bourguignon for their assistance during my Ph.D. I'd also like to thank Ali Dolamari, representing Kurdistan region in Paris, for his continuous support from my Master's to the present.

Lastly, I express gratitude to the people of France for their rich culture, remarkable tolerance, and warm welcome. I've felt no discrimination or racism during my stay, experiencing only inclusivity and acceptance.

Table of Contents

General introduction

Ionizing radiations (IoR) inflict detrimental damage on genetic material, proteins, and the signaling pathways that typically govern cellular functions[1]. This damage can result in adverse outcomes, such as premature aging[2], the development of tumors and cancers[3], but also in therapeutic benefits in the realm of radiation therapy[4]. Ionizing radiation can consist of high-energy photons, such as extreme ultraviolet rays (XUV), X-rays, or gamma rays, or charged particles like protons (H^+), alpha particles (He^{2+}), electrons (e^-), or muons (μ^-). The sources of ionizing radiations can be natural, such as radioactive minerals, oceans, solar radiation, and cosmic rays, or of anthropogenic nature, originating from medical procedures, nuclear power plant waste, nuclear catastrophes, or nuclear weapons testing[5].

The responses of biomolecules to IoR are multi-timescales[6,7]. In this thesis, our focus is on the physical stage, which takes place within a timescale ranging from 10^{-18} to 10^{-15} seconds. The physical stage is initiated by the deposition of energy from IoR to the target, leading to ionizations and excitations, and creating a superposition of electronic states. This stage ends with charge redistribution before the atomic nuclei respond. The ultrafast events occurring at the physical stage are of utmost importance as they set the stage for all subsequent events. The excited or ionized biomolecule experiences rapid structural transformations from 10^{-15} to 10^{-9} seconds, during which reactive chemical species such as free radicals can be produced. These reactive species can then interact with nearby molecules, such as water or other biomolecules, from 10^{-9} to 10^{-6} seconds, leading to a cascade of chemical reactions. Finally, the altered biomolecules and chemical species engage with larger biochemical networks in the cell over extended periods ranging from seconds to days (or more). This engagement encompasses signaling pathways, DNA repair mechanisms, and potentially programmed cell death if the damage is extensive[8]. This multi-timescale nature is essential to understand the detailed mechanisms of radiation interaction with biological systems and the potential consequences, either harmful or therapeutic.

Energy deposition is a critical parameter for controlling the dosage and penetration of IoR. For instance, in particle therapy, it is essential to precisely target tumor cells while sparing healthy tissues. In the case of photons, for example, it is important to determine the dose liberated during medical imaging[9]. Stopping power, or attenuation in the case of photons, is a term that refers to the amount of energy transferred from ionizing radiation to the target material (energy deposition) per unit distance. The stopping power or energy deposition generally depends on the energy and types of IoR. Charged particles deposit their energy in a highly localized region within matter, with the amount depending on the kinetic energy of the particles. A sharp and significant transfer of energy occurs at the end of the particle's trajectory, known as the Bragg peak. This characteristic makes charged particles especially useful in radiotherapy. On the other hand, photons can deposit energy through various processes, such as photoabsorption, photoelectric effect, Compton scattering, or pair production, depending on their energy[10]. In general, the energy deposition by photons is spread over molecules and attenuates with penetration distance.

The creation of a superposition of electronic states in biomolecules as a result of IoR interaction can lead to an ultra-fast redistribution of charge. This can occur as a result of charge migration[11], Interatomic Coulombic Decay (ICD)[12], and auto-ionization[13]. Consequently, the charge does not remain localized at the initial position but is transferred to another part of the molecules, leading to characteristic fragmentation and processes in biological systems. For instance, the decarboxylation of acidic amino acids induced by X-rays[14] or ionization at specific sites in polypeptide molecules can lead to fragmentation elsewhere within the molecule[15].

Recently, the advent of ultrafast spectroscopies has enabled the direct investigation of the earliest mechanisms in the physical stage, in real-time. Experimental techniques employing pump-probe XUV lasers, generated through high harmonic generation[16–18], offer exceptional approaches for probing matter's responses with attosecond resolution. However, time-resolved pulsed radiolysis using beams of charged particles has been mainly constrained to the picosecond range. Beyond experiments, theoretical access to the physical stage for large biomolecules is now possible. Pure Real-Time Time-Dependent Density Functional Theory (RT-TD-DFT) allows for the step-by-step simulation of the response of the electronic cloud in large molecular systems

subjected to strong, ionizing radiations. This is achieved by discretizing time into small increments, typically of a few attoseconds. Notably, it provides theoretical insight into the real-time interaction of IoR with matter and the superposition of electronic states.

Energy deposition by charged particles depends on both the kinetic energy of the particles and the characteristics of the target molecules. For particles with an energy above 1 MeV, the relativistic effect is incorporated into the contribution of energy deposition. This effect is also responsible for the increase in energy deposition seen in the stopping power curve at high-energy particles. With respect to target materials, energy deposition relies on the electronic structure or density, which in turn depends on the types of atoms, as well as the chemical and physical bonds within the material.

Accurate calculation of energy deposition or stopping power presents a significant challenge for both experimental and theoretical approaches. Experimental measurements suffer from a high degree of uncertainty[19]. Empirical [20,21] and semi-empirical models[22–25] are quite precise for calculating the stopping power of homogeneous systems. However, for heterogeneous systems like biomolecules, errors in stopping power can accumulate due to Bragg's additive rule. This rule considers that the stopping power of a compound can be calculated from the linear combination of the stopping power of its individual atoms. It neglects the effect of chemical and physical bonds between atoms in the molecules. The effect of hydrogen bonds on energy deposition is well known for water, corresponding to about 10% at low-energy particles[26]. However, for biomolecules, this effect is not clear, even though hydrogen bonding in biomolecules is known to be essential.

RT-TD-DFT is a promising model for calculating stopping power[27] and electronic relaxation[10] in real systems. However, to date, simulations using RT-TD-DFT are restricted to low-energy particles (below 2 MeV) due to the absence of the relativistic effect of charged particles in the simulation. It is not clear how the relativity of charged particles affects energy deposition in RT-TD-DFT simulations. When dealing with large biological systems, simulating the entire system using RT-TD-DFT becomes impossible. Instead, the technique of combining quantum mechanics

with molecular mechanics (QM/MM) proves useful. In this approach, the atoms of interest are described at the RT-TD-DFT level, while the effects of the molecular environment are taken into account using molecular mechanics. However, the impact of the size of the QM region and the type of force field in MM on energy deposition and electronic relaxation induced by IoR remains elusive.

The attosecond XUV-pulse plays an impressive role in revealing the ultra-fast mechanisms that occur in small and relevant biomolecules[17,28] in the gas phase. However, studying large biomolecules like peptides or proteins presents a significant challenge for both experimental and theoretical approaches. More recently, our experimental collaborators (team of F. Lepine at ILM, Lyon) have developed a technique using XUV-IR pump-probe lasers to investigate large biomolecules up to the size of entire proteins[29]. For us, the question remains: how does the secondary structure in a biomolecule affect the interaction with XUV-pulse, ionization of the molecule, or electron relaxation? This aspect is still unclear.

This thesis consists of five Chapters:
Chapter One lays the necessary background to understand the subsequent Chapters. It begins by introducing the processes involved in radiation damage, followed by an exploration of the characteristics of the physical stage, including energy deposition and electronic responses. Finally, the chapter introduces the theoretical background for RT-TD-ADFT as a workhorse model in our work.

Chapter Two focuses on the effect of some relativistic effects on energy deposition when the irradiating ion approaches the speed of light. We consider a Watson-Crick guanine-cytosine nucleobase pair, with two water molecules hydrogen-bonded to the guanine nitrogen atoms. These molecules are essential components of the DNA structure.

In Chapter Three, we investigate the effect of hydrogen bonds on the energy deposition for the same system that is used in Chapter Two.

In Chapter Four, we investigate the physical stage of energy deposition and charge migration in protein/DNA complexes exposed to α-particles. We utilize RT-TD-DFT within the framework of Quantum Mechanics/Molecular Mechanics (QM/MM) approaches. In this chapter, we examine the impact of the QM-region size and the electrostatic induction of the MM-region on energy deposition and post-irradiation charge migrations. Finally, we introduce a new analytical technique for studying charge migration in large systems.

Finally, Chapter Five is dedicated to the ultrafast response of the $Insulin^{+6}$ protein and the Substance PH^{+} peptide to XUV-pulses in the gas phase. We begin by assessing the effect of atomic center basis sets on the simulation of XUV-pulses for N_2 and model peptides. This chapter includes the study and preparation of the structures of $Insulin^{+6}$ and Substance PH^{+} in the gas phase, achieved through molecular dynamics simulations. Additionally, we present the results from RT-TD-DFT concerning the interaction of XUV-pulse with $Insulin^{+6}$ and Substance PH^{+}.

REFERENCES:

(1) Spitz, D. R.; Azzam, E. I.; Jian Li, J.; Gius, D. Metabolic Oxidation/Reduction Reactions and Cellular Responses to Ionizing Radiation: A Unifying Concept in Stress Response Biology. *Cancer and Metastasis Reviews* **2004**, *23* (3), 311–322. https://doi.org/10.1023/B:CANC.0000031769.14728.bc.

(2) Dizdaroglu, M.; Jaruga, P. Mechanisms of Free Radical-Induced Damage to DNA. *Free Radical Research* **2012**, *46* (4), 382–419. https://doi.org/10.3109/10715762.2011.653969.

(3) Lomax, M. E.; Folkes, L. K.; O'Neill, P. Biological Consequences of Radiation-Induced DNA Damage: Relevance to Radiotherapy. *Clinical Oncology* **2013**, *25* (10), 578–585. https://doi.org/10.1016/j.clon.2013.06.007.

(4) Levin, W. P.; Kooy, H.; Loeffler, J. S.; DeLaney, T. F. Proton Beam Therapy. *Br J Cancer* **2005**, *93* (8), 849–854. https://doi.org/10.1038/sj.bjc.6602754.

(5) Dartnell, L. R. Ionizing Radiation and Life. *Astrobiology* **2011**, *11* (6), 551–582. https://doi.org/10.1089/ast.2010.0528.

(6) Platzman, RL. The Physical and Chemical Basis of Mechanisms in Radiation Biology. *Radiation Biology and Medicine. Selected Reviews in the Life Sciences* **1958**, 15–72.

(7) Azzam, E. I.; Jay-Gerin, J.-P.; Pain, D. Ionizing Radiation-Induced Metabolic Oxidative Stress and Prolonged Cell Injury. *Cancer Letters* **2012**, *327* (1–2), 48–60. https://doi.org/10.1016/j.canlet.2011.12.012.

(8) Rodriguez Goyes, J. C.; Gomez, N. J.; Restrepo, V. J.; Gaviria, S.; Hincapie, A.; Arias, E. Biologic Effect of Ionizing Radiation. *Interventional-Cardiology* **2018**, *10* (5). https://doi.org/10.4172/Interventional-Cardiology.1000618.

(9) Einstein, A. J.; Moser, K. W.; Thompson, R. C.; Cerqueira, M. D.; Henzlova, M. J. Radiation Dose to Patients From Cardiac Diagnostic Imaging. *Circulation* **2007**, *116* (11), 1290–1305. https://doi.org/10.1161/CIRCULATIONAHA.107.688101.

(10) Hubbell, J. H. Review of Photon Interaction Cross Section Data in the Medical and Biological Context. *Phys. Med. Biol.* **1999**, *44* (1), R1–R22. https://doi.org/10.1088/0031-9155/44/1/001.

(11) Cederbaum, L. S.; Zobeley, J. Ultrafast Charge Migration by Electron Correlation. *Chemical Physics Letters* **1999**, *307* (3–4), 205–210. https://doi.org/10.1016/S0009-2614(99)00508-4.

(12) Jahnke, T.; Hergenhahn, U.; Winter, B.; Dörner, R.; Frühling, U.; Demekhin, P. V.; Gokhberg, K.; Cederbaum, L. S.; Ehresmann, A.; Knie, A.; Dreuw, A. Interatomic and Intermolecular Coulombic Decay. *Chem. Rev.* **2020**, *120* (20), 11295–11369. https://doi.org/10.1021/acs.chemrev.0c00106.

(13) Codling, K.; Madden, R. P.; Ederer, D. L. Resonances in the Photo-Ionization Continuum of Ne I (20-150 EV). *Phys. Rev.* **1967**, *155* (1), 26–37. https://doi.org/10.1103/PhysRev.155.26.

(14) Bury, C.; Garman, E. F.; Ginn, H. M.; Ravelli, R. B. G.; Carmichael, I.; Kneale, G.; McGeehan, J. E. Radiation Damage to Nucleoprotein Complexes in Macromolecular Crystallography. *Journal of Synchrotron Radiation* **2015**, *22* (2), 213–224. https://doi.org/10.1107/S1600577514026289.

(15) Weinkauf, R.; Schanen, P.; Metsala, A.; Schlag, E. W.; Bürgle, M.; Kessler, H. Highly Efficient Charge Transfer in Peptide Cations in the Gas Phase: Threshold Effects and Mechanism. *J. Phys. Chem.* **1996**, *100* (47), 18567–18585. https://doi.org/10.1021/jp960926m.

(16) Palacios, A.; Martín, F. The Quantum Chemistry of Attosecond Molecular Science. *WIREs Comput Mol Sci* **2020**, *10* (1). https://doi.org/10.1002/wcms.1430.

(17) Nisoli, M.; Decleva, P.; Calegari, F.; Palacios, A.; Martín, F. Attosecond Electron Dynamics in Molecules. *Chem. Rev.* **2017**, *117* (16), 10760–10825. https://doi.org/10.1021/acs.chemrev.6b00453.

(18) Sato, S. A. First-Principles Calculations for Attosecond Electron Dynamics in Solids. *Computational Materials Science* **2021**, *194*, 110274. https://doi.org/10.1016/j.commatsci.2020.110274.

(19) Siiskonen, T.; Kettunen, H.; Peräjärvi, K.; Javanainen, A.; Rossi, M.; Trzaska, W. H.; Turunen, J.; Virtanen, A. Energy Loss Measurement of Protons in Liquid Water. *Physics in Medicine & Biology* **2011**, *56* (8), 2367. https://doi.org/10.1088/0031-9155/56/8/003.

(20) Ziegler, J. F. Comments on ICRU Report No. 49: Stopping Powers and Ranges for Protons and Alpha Particles. *Radiation Research* **1999**, *152* (2), 219. https://doi.org/10.2307/3580097.

(21) Janni, J. F. Energy Loss, Range, Path Length, Time-of-Flight, Straggling, Multiple Scattering, and Nuclear Interaction Probability: In Two Parts. Part 1. For 63 Compounds Part 2. For Elements $1 \leqslant Z \leqslant 92$. *Atomic Data and Nuclear Data Tables* **1982**, *27* (2), 147–339. https://doi.org/10.1016/0092-640X(82)90004-3.

(22) Vagena, E.; Androulakaki, E. G.; Kokkoris, M.; Patronis, N.; Stamati, M. E. A Comparative Study of Stopping Power Calculations Implemented in Monte Carlo Codes and Compilations with Experimental Data. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* **2020**, *467*, 44–52. https://doi.org/10.1016/j.nimb.2020.02.003.

(23) Seltzer, S. Stopping-Powers and Range Tables for Electrons, Protons, and Helium Ions, NIST Standard Reference Database 124, 1993. https://doi.org/10.18434/T4NC7P.

(24) Friedland, W.; Dingfelder, M.; Kundrát, P.; Jacob, P. Track Structures, DNA Targets and Radiation Effects in the Biophysical Monte Carlo Simulation Code PARTRAC. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **2011**, *711* (1–2), 28–40. https://doi.org/10.1016/j.mrfmmm.2011.01.003.

(25) Ziegler, J. F.; Ziegler, M. D.; Biersack, J. P. SRIM – The Stopping and Range of Ions in Matter (2010). *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* **2010**, *268* (11), 1818–1823. https://doi.org/10.1016/j.nimb.2010.02.091.

(26) Matteson, S.; Powers, D.; Chau, E. K. L. Physical-State Effect in the Stopping Cross Section of $\mathrmH_2$O Ice and Vapor for 0.3-2.0-MeV $\ensuremath\alpha$ Particles. *Phys. Rev. A* **1977**, *15* (3), 856–864. https://doi.org/10.1103/PhysRevA.15.856.

(27) Gu, B.; Cunningham, B.; Muñoz Santiburcio, D.; Da Pieve, F.; Artacho, E.; Kohanoff, J. Efficient *Ab Initio* Calculation of Electronic Stopping in Disordered Systems via Geometry Pre-Sampling: Application to Liquid Water. *J. Chem. Phys.* **2020**, *153* (3), 034113. https://doi.org/10.1063/5.0014276.

(28) Calegari, F.; Trabattoni, A.; Palacios, A.; Ayuso, D.; Castrovilli, M. C.; Greenwood, J. B.; Decleva, P.; Martín, F.; Nisoli, M. Charge Migration Induced by Attosecond Pulses in Bio-Relevant Molecules. *J. Phys. B: At. Mol. Opt. Phys.* **2016**, *49* (14), 142001. https://doi.org/10.1088/0953-4075/49/14/142001.

(29) Hervé, M.; Boyer, A.; Brédy, R.; Allouche, A.-R.; Compagnon, I.; Lépine, F. On-the-Fly Investigation of XUV Excited Large Molecular Ions Using a High Harmonic Generation Light Source. *Sci Rep* **2022**, *12* (1), 13191. https://doi.org/10.1038/s41598-022-17416-4.

CHAPTER ONE

RADIATION DAMAGE AND THEORETICAL BACKGROUND

Contents

1 Radiation damage and theoretical background

The responses of biomolecules interacting with ionizing radiations is a complex, multi-stage process that begins with the ultra-fast response of the electron cloud and may culminate in biological damage over several days or years. In this thesis, our interest lies in the ultra-fast events that occur within a range from attoseconds to a few femtoseconds in biomolecules when induced by fast-charged particles and XUV-pulses. This chapter is dedicated to introducing the foundational elements required to understand the subsequent chapters.

The first section introduces the processes involved in the radiation damage of biomolecules. In the second section, we delve into the physical stages of radiation damage, encompassing the general mechanisms related to energy deposition by ionizing radiation and ultra-fast electron relaxation. The third section introduces the theoretical background employed throughout this thesis, specifically highlighting ground state density functional theory and Real-Time Time-Dependent Auxiliary Density Functional Theory. The chapter concludes with a summarizing section.

This chapter draws inspiration from two review articles of which I am a co-author. The first article, conducted during the COVID shutdown, is published in the *Annual Review of Physical Chemistry* under the title "First-Principles Simulations of Biological Molecules Subjected to Ionizing Radiation"[1]. The second article has been published more recently in *The European Physical Journal Special Topics*, under the title "Current status of deMon2k for the investigation of the early stages of matter irradiation by time-dependent DFT approaches"[2]. In this article, I contribute by composing two sections on the effects of complex absorbing potential on RT-TD-ADFT simulations of ionizing radiation and the effect of force field polarizability on energy deposition and charge migration during the interaction of charged particles with DNA/protein complexes.

1.1 Radiation damage of biomolecules

When high energetic particles (α, β particles, H^+,n ….) and photons (X-rays, γ-rays ….) penetrate biomolecules, they interact strongly with the electron cloud, carrying enough energy to excite molecules or detach electrons. This primary interaction initiates successive physical-chemistry processes, leading to complex formation and breaking of chemical bonds along the ionizing radiation (IoR) track. As a result of chemical lesions, biomolecules, such as DNA, proteins, enzymes, or lipids, malfunction leading to the death of the vital cell[3], accelerates age in human or raises cancerous tumors[4]. In this section, we attempt to define the key processes involved in radiation damage to biological systems.

The response of a biomolecule subjected to IoR is exceedingly complicated and of multiscale processes. It is customary to classify the responses into four major successive stages according to the kind of processes involved, namely the physical, the physicochemical, the chemical and, finally, the biological stages[5,6]. Understanding the mechanisms involved in each stage aids in unraveling the consequences of irradiation. As a complement to experimental approaches, such as time-resolved pulsed radiolysis or pump-probe spectroscopy, numerical simulations are employed to reveal the damage mechanism at the atomic level. To select sophisticated models, we must first understand the characteristics and features of each stage. Figure 1-1 schematically summarizes the multiscale responses of biomolecules subjected to irradiation[1].

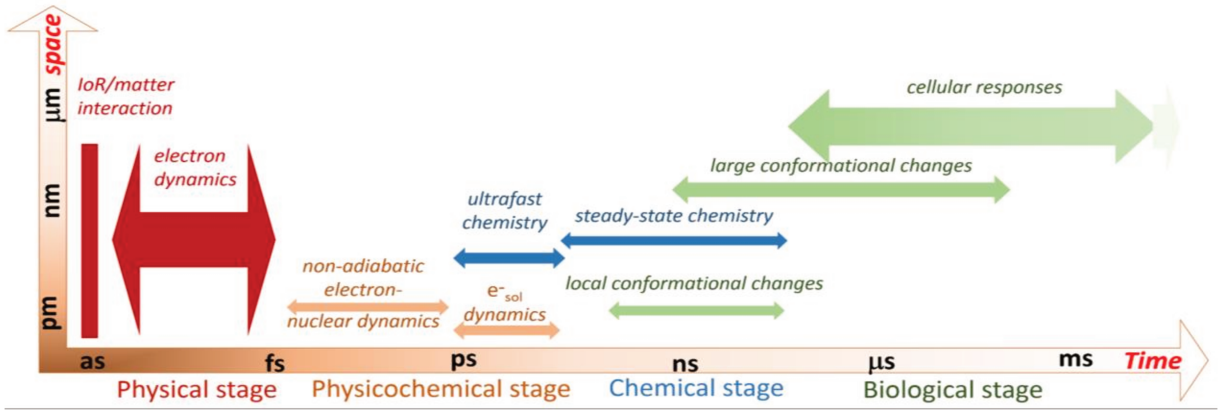

Figure 1-1: Interaction of IoR with biological matter induces multiscale responses in time and space. In the bottom we indicate, not exhaustively, some of the methodologies found in the literature to address the different scales. The figure is taken and adopted with permission[1].

The physical stage begins when an IoR penetrates matter, depositing a large amount of its energy on a time scale of a few attoseconds (10^{-18} s), this energy leads to excited and ionizing target molecules and ejects energetic electrons in the continuum. The excited and ionized molecules undergo pure electronic dynamics processes like, intra- and inter-molecular charge and energy migration, Coulomb Decay (ICD[7]) and Auger electron relaxation[8] etc. Secondary electrons generated as a result of the ionization process, can have energies ranging from a few electron volts (eV) to several kiloelectron volts (keV), depending on the type and energy of the incident radiation. They thermalize by collisions in the environment and may excite other molecules. Finally, very dense track of excited, radical cation molecules and secondary electrons are formed along the IoR trajectories. This stage continues until the nuclear response becomes significant in general in a few femtoseconds (10^{-15} s). The pure electrons dynamic model is sufficient to simulate the physical stage. We concentrate on the physical stage in this work, therefore we go over it in greater detail in following section.

Following the physical stage, the physicochemical stage appears and continues for a few picoseconds (10^{-12} s). The reactive species created during the physical stage are unstable, they start to expand and initiate complex nuclear and electron coupling dynamics. The excited radical cation undergoes an ultra-fast process by means of a sequence of electron or proton transfers from neighboring molecules, or it dissociates by dissipating energy into vibrational modes. Secondary electrons will be trapped in molecular cavities and solvated[9]. They can also be caught in molecules, which would then activate the electron attachment dissociation pathways[10]. Because the electronic system is still in an excited state at the beginning of the physicochemical stage, we have to provide a non-adiabatic model to simulate this time range. After a few ps, the system returns to the ground state, an adiabatic model might be enough.

The non-homogeneous chemical stage extends from few ps to the time when the reactants are distributed homogeneously (μs, 10^{-6} s). The radiolytic species are highly reactive and non-homogeneously distributed at the end of the physical-chemical stage. Depending on the radiation type and reaction energy barriers, the reactive species diffuse and react with one another or with the surrounding molecules, until the tracks vanish completely. In this stage, the system is entirely

in its ground state and expands, so that the system is governed by thermodynamic equilibrium and diffusion rate. So, kinetic models and diffusion models are recommended.

Finally, at a time scale of ca. 10^{-3} s or longer, the cells respond to the irradiation resulting from the products formed in the preceding stages. This is called the biological stage. As a consequence of the alteration of molecules, irregularities in the biochemical process happen at the subcellular level, for instance modification of the shape of proteins, enzyme inactivation, or DNA disorder, etc. As a result, the cells lose their specific properties. At the end, the cells may be repaired, mutated, or die[11].

Biological damages are classified as being the consequences of direct or indirect effects[12,13] as show in Figure 1-2. The direct effect refers to damages triggered by the direct deposition of energy into the biomolecules[14]. The indirect effect[15] refers to the damage caused to molecules not by the direct action of the IoR itself, but through the intermediaries produced when radiation interacts with other molecules, primarily water, in the vicinity of the target molecule. Because a large quantity of water exists in cells (~70%-85%), the indirect effect plays an important role in radiotherapy. The energy is largely deposited in water, which leads to the production of reactive oxygen species (ROS) via water radiolysis for instance, hydroxyl radical, hydrogen peroxide, or superoxide, etc. These species diffuse and interact with biomolecules within the cells, to produce chemical alterations and consequently deleterious effects. Furthermore, quasi-direct effects were proposed by Sevilla and coworkers[16] as a frontier between direct and indirect effects. They observed the yield of OH radicals formed during γ-irradiation of DNA for the different shells of hydration at low temperatures. They found that when the DNA is in contact with water molecules, the amount of OH radicals is not significant. They suggested that the yield of OH radicals decreases due to the ultra-fast hole transfer from water radical cations to DNA, which occurs before the production of OH radicals.

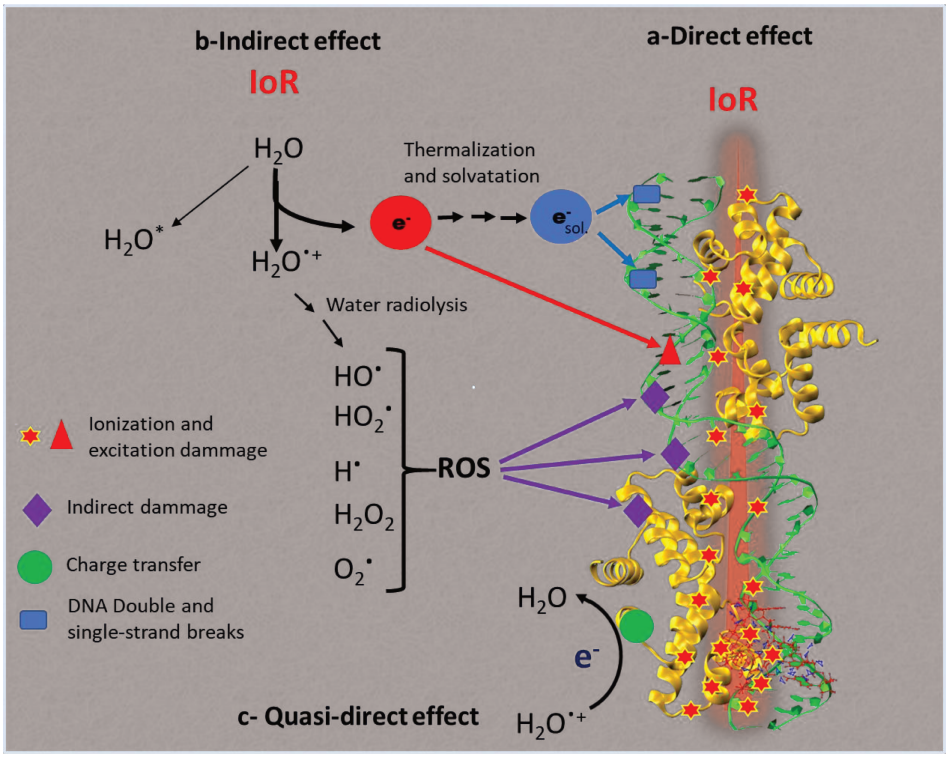

Figure 1-2: the main mechanisms involve in biological damage subjected to IoR. (a) Indirect effect under water radiolysis produces ROS and solvated electrons that attack biomolecules after diffusion. (b) Direct effect corresponds to the direct irradiation of DNA-protein complex that lead to the multiple excitation and ionization of biomolecule. (c) Quasi-direct effect lies at the frontier between the direct and indirect effects.

The competition of biological damage by direct and indirect effects highly depends on the energy deposition of the radiation source and the concentration of water[17]. The indirect effect is considered to be the dominant damage mechanism due to the abundance of water *in vivo*. However, biological systems are not homogeneous; the concentration of water diminishes around macromolecules (such as DNA, RNA, proteins, etc.). For instance, DNA in a cell is densely packaged into nucleosomes, where it is wrapped around a histone protein core[18]. When IoR is applied, biological damage is primarily caused by direct effects.

In this work, we aim to investigate the physical stage of large biomolecules, for instance, the guanine-cytosine nucleobase pair with two water molecules in Chapters two and three, and the DNA-protein complex in Chapter four, as well as large peptides and proteins in the gas phase in Chapter five. Therefore, our focus is on the physical stage and the direct/quasi-direct effects.

1.2 Physical stage

The physical stage corresponds to the pure electronic response of matter to IoR. This includes energy deposition, primary excitation, ionization, and electronic relaxation. The processes that occur in the physical stage highly depend on the type of IoR. In this section, we briefly discuss the main processes of the electronic response, depending on the kind of IoR. We focus on the evaluation of energy deposition and electronic relaxation.

1.2.1 Physical stage depending on the types of ionizing radiation

a. Energy deposition and ionization process

The energy deposition and ionization process in target electrons by γ-rays (which are produced by the decay of excited nuclei and by nuclear reactions), and X-rays (which are produced when upper electrons in high orbital levels descend to a core hole or when fast-moving particles are deflected and decelerated), occur through four mechanisms, depending on the energy of the photon. These mechanisms are coherent (Rayleigh) scattering, the photoelectric effect, Compton scattering, and pair production[19]. For the purposes of our discussion, we will neglect the nuclear reactions and nuclear disintegration caused by IoR, as these occur at very high energy.

- Coherent (Rayleigh) Scattering[20]: in this process, a photon interacts with an entire atom, causing it to become momentarily polarized. The photon is then scattered without a loss of energy or ionization of the target, though there may be a change in direction. This process becomes less significant as the photon's energy increases.
- Photoelectric effect[21]: here, a photon interacts with an inner-shell electron in an atom, completely absorbing the photon's energy and ejecting the electron from the atom. The kinetic energy of the ejected electron equals the energy of the incoming photon minus the binding energy of the electron in its shell. This process is particularly significant at lower photon energies, such as in the range of 0.5-700 keV for carbon graphene [22].
- Compton scattering[23]: this occurs when a photon interacts with a loosely bound outer-shell electron, transferring a part of its energy to the electron and being scattered in a different direction. The scattered photon has lower energy, and its change in

wavelength is thus described. Compton scattering is a dominant interaction process for moderate-energy photons.

- Pair production[24]: this takes place when a photon with energy greater than 1.022 MeV interacts with the nucleus of an atom. The photon's energy is converted into an electron-positron pair, with the excess energy of the photon above 1.022 MeV being shared between the kinetic energies of the electron and positron.

Extreme Ultraviolet (XUV) is another type of electromagnetic wave, with an energy range of approximately 10-130 eV. This kind of IoR deposits its energy specifically by interacting with the valence electrons of the target, leading to its excitation and ionization. XUV pulses can be created with durations on the attosecond scale, allowing for the study of ultra-fast electron dynamics. This is essential for revealing the physical stage.

Charged particles (such as fast protons, α-particles, or heavy ions) deposit their energy via Coulomb interaction with the electrons or nuclei of the target[25], depending on their energy. The contribution of nuclear energy deposition, resulting from elastic collisions with target nuclei, is dominant at low energy (< 0.01 MeV), as shown in Figure 1-3. This type of interaction leads to rotational and vibrational excitations. Above 0.01 MeV, the nuclei do not have enough time to respond during the collision; the only interaction of the charged particle with electrons contributes to energy deposition, referred to as inelastic collision. Depending on the kinetic energy of the particle, the energy deposition sharply increases until it reaches a peak, known as Bragg's peak, and then decreases. At relativistic velocities of charged particles, the energy deposition starts to rise again. Unlike photons, charged particles deposit their energy locally along their trajectory, and the energy deposition decreases with distance from the target according to Coulomb interaction. The energy transfer from particles leads to the local excitation and ionization of several electrons in the valence and core levels of the target. The contribution of core excitation increases with particle energy and decreases with the collision distance from target electrons.

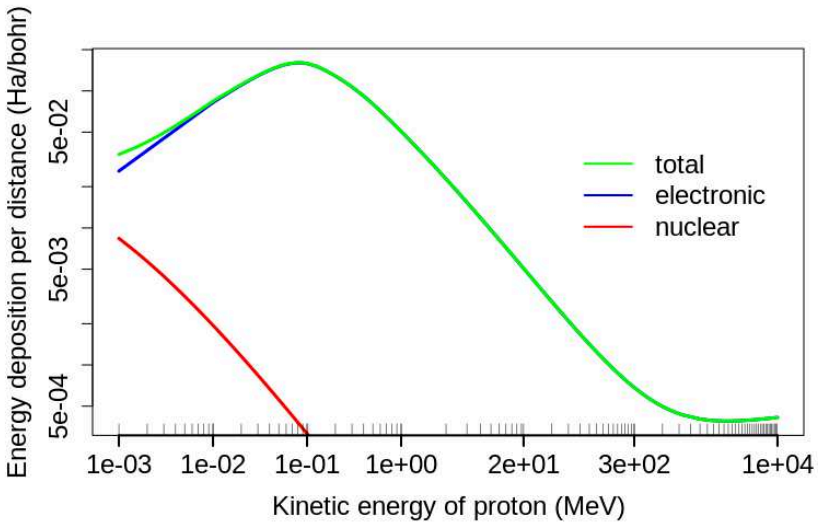

Figure 1-3: Contribution to energy deposition per distance for a proton in liquid water. The energy deposition is calculated using the PSTAR database[26]. The x and y axes are scaled logarithmically.

b. Penetration distance

As mentioned before, the interaction of IoR with matter depends on several factors, such as the type of radiation (*e.g.* charged particles, or uncharged particles like photons), the energy of the radiation, and the type of matter it interacts with. As a result, the penetration distance of radiation through matter differs significantly between photons and charged particles[27].

The intensity of photon beams attenuates exponentially with increasing penetration depth, which depends on the absorption cross-section. The absorption cross-section is influenced by both the energy of the photons and the type of material they interact with. In the case of XUV radiation, the attenuation primarily depends on the strong absorption cross-section of the valence electrons, resulting in a relatively short penetration depth for XUV beams. On the other hand, for γ-rays and X-rays, photon absorption is dependent on the cross-sections of three processes: the photoelectric effect, Compton scattering, and pair production[22]. The likelihood of absorption via these mechanisms is generally lower because they primarily occur through interactions with core electrons (in the case of the photoelectric effect) or atomic nuclei (in the case of pair production). That is why these types of photons typically have a larger penetration depth into matter. One crucial characteristic of photon penetration is the spreading of energy deposition over target molecules along their path, decreasing exponentially with penetration distance. As shown in Figure 1-4, the energy deposition expands over a short distance due to high-energy photons initiating an electromagnetic cascade through pair production, which generates more photons,

electrons, and positrons of lower energy. Following this, energy deposition decreases with distance.

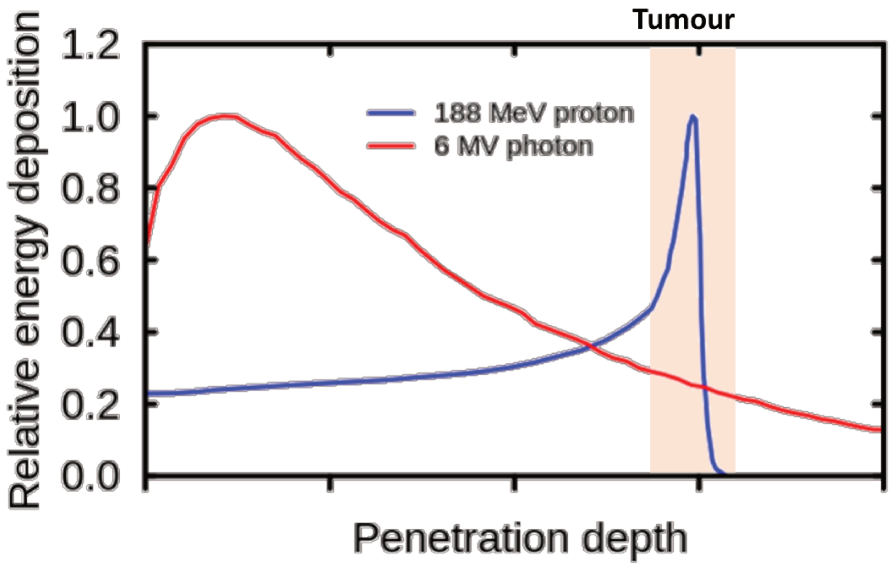

Figure 1-4: schematic of relative energy deposition of a proton (blue line) and a photon (red line) as a function of penetration distance in matter.

For charged particles, the energy deposition per unit of penetration distance in matter - known as stopping power - essentially describes how effectively a specific medium can decelerate or stop a charged particle. In contrast to the exponential decrease observed with photons, the penetration depth for charged particles more closely resembles a 'Bragg peak'. The energy deposition gradually increases with penetration distance as the kinetic energy of the ion beam decreases due to energy loss from collisions. Consequently, they transfer the majority of their energy to a very localized position just before the particle comes to a halt, as shown in Figure 1-4. This characteristic makes charged particles particularly attractive for radiotherapy, as they can be utilized to effectively target deep-seated tumor cells.

The difference in energy deposition with penetration distance is one reason why different types of IoR are used for different applications. For example, X-rays are often used for imaging deep inside the body, while proton therapy can be used to target tumors more locally with less damage to surrounding tissue. In our work, we particularly focus on the energy deposition of charged particles.

c. Electron relaxation

As a result of energy deposition from IoR, target molecules can become excited and ionized. These molecules may then undergo complex ultra-fast electronic relaxation through several channels before the atomic nuclei significantly move[7]. The electronic relaxation is purely electronic and relies on electron correlation. It depends on the initial ionization, the excited state, molecular environments, and the electronic structure of the target. Here, we mention some of the common relaxation processes.

The creation of a hole in the core levels, as a result of the photoelectric effect, for example, may undergo Auger relaxation. When an electron is removed from a core level, it leaves behind a vacancy. This vacancy can then be filled by an electron from a higher energy level, transitioning into the lower energy state. This transition releases energy, which can either be emitted as a photon (in X-ray fluorescence) or transferred to another electron in the system, often one in a higher energy state or in the outer shell. When the energy is transferred to an electron, it gives it enough energy to escape from the atom or molecule. This ejected electron is known as an Auger electron, and the entire process is termed Auger relaxation or the Auger effect[28,29]. The Auger effect is a non-radiative process, meaning it does not result in the emission of a photon. As a result of emitting electrons, the molecules can become doubly ionized, producing a cascade of secondary electrons along the radiation passage, leading to the efficient damage of biomolecules. On the other hand, the low-energy electrons produced from Auger decay can be used in radiotherapy[30,31] by injecting Auger electron-emitting agents, such as I^{123}. Low-energy electrons deposit their energy over nanometer to micrometer distances. Auger electron emitters accumulate in the target tumor cells and kill them while protecting healthy cells.

The interaction of IoR with molecules may lead to an electron transition to a highly excited state. This excited electron may then de-excite and return to the ground state, releasing energy. If an electron in a higher energy level absorbs this released energy, it may lead to the ejection of the electron and simultaneous ionization of the atom, a process known as auto-ionization[32,33]. This is similar to the Auger effect, but the key difference lies in the nature of the interaction: auto-ionization typically involves electrons in excited states, whereas Auger relaxation involves a core vacancy. Auto-ionization can lead to specific spectral lines in the emission or absorption

spectrum. This channel of relaxation becomes crucial when the system's radiation process includes high excited states and Rydberg excitations[34], such as those resulting from the interaction of XUV rays and charged particles.

Interatomic and intermolecular Coulomb decay (ICD)[35] is another channel of relaxation that occurs when the inner shell of a molecule is ionized. An electron from an upper shell fills the vacancy and liberates energy. This energy is absorbed by an electron of neighboring atoms or molecules, which generally have a weak bond with the host atom or molecule, leading to the ejection of a second low-energy electron. The energy transfer is mediated by the Coulomb interaction between the two atoms or molecules. The main distinction between ICD and the Auger effect is that ICD is a non-local ionization process. In other words, ionization on one side of a molecule can lead to ionization of another molecule on the other side. ICD is a fast process, typically occurring on the femtosecond time scale, and it plays a significant role in various contexts, such as in biological systems[36,37].

Charge migration[38] between atoms and molecules is another ultra-fast electronic relaxation process induced by IoR. Upon ionization or excitation of a molecule, certain electronic states may become temporarily populated. Due to the interaction and correlation between electrons, this sudden change in electronic configuration can induce a rapid reorganization of the electron density within the molecule. This reorganization is referred to as "charge migration," where the charge effectively "moves" from one part of the molecule to another. This phenomenon is essentially an electronic response and occurs on an extremely short time scale, typically within attoseconds to a few femtoseconds. The electronic superposition of states and the overlap between them are required to trigger charge migration.

Charge migration is distinct from charge transfer[39], even though both processes correspond to the motion of charge from one molecule to another. The main difference lies in the nature of the processes: charge migration has an oscillatory[40] or electron delocalization nature, while in charge transfer, the coupling between electronic and nuclear dynamics is essential. Nuclear motion leads to electronic dephasing or decoherence, breaking the periodic nature in time, and localizing the charge permanently on a side of molecule.

The delocalization of electrons through charge migration is particularly significant when inner-valence electrons are ionized. The removal of an electron from a specific inner-valence orbital, where strong correlation typically exists, leads to coupling with several other states. Consequently, the charge migrates to various other orbitals, and by the end of the process, it spreads more or less uniformly across the entire system. Even if the initial charge is localized, it rapidly becomes delocalized[41,42].

Charge migration is essential for biological systems irradiation, especially when it involves a change in the initial location of a charge. The first observation of charge migration in a biomolecule was indirectly observed by Schlag and colleagues[43] with nanosecond pump-probe experiments. Though nanosecond time intervals are large for observing charge migration, they discovered that after ionizing chains of peptides at a specific molecular site, subsequent instantaneous probed fragmentation occurred at the other end of the chain. This led them to conclude that the initially created hole had migrated through the molecule. These observations were later theoretically examined by Cederbaum and Zobeley[44], who first introduced the concept of ultrafast charge migration. The first experimental clearly evidence of charge migration was reported by Calegari and co-workers[45], who utilized attosecond pump-probe experiments to investigate phenylalanine. They employed 300 as XUV pulses to ionize the molecules and used 4 fs near-infrared (NIR) pulses for probing. They observed a typical oscillatory fragmentation yield with a period of 4.3 fs, which they attributed to periodic charge migration in phenylalanine. This charge migrated from the amine functional group to the phenyl group and then back to the amino acid. Ikeura-Sekiguchi and Sekiguchi[46] employed core hole spectroscopy to explore electron delocalization in the conduction band through the phosphate backbone of DNA. Their aim was to discern the periodic and aperiodic charge transfers in DNA. Their findings indicated that electron delocalization occurs in approximately 740 as. These experimental results highlight the significance of charge migration in biomolecules, triggered by IoR, manifesting in both oscillatory and delocalization patterns.

For the interaction of IoR with large biomolecules, the probability of various relaxation channels may be incorporated, with a specific probability assigned to each relaxation process. In our case, where we are interested in the electron dynamics of low-energy charged particles and XUV-pulse interactions, charge migration, ICD, and auto-ionization are expected to be dominant when

ionization and excitation of valence electrons are significant. Consequently, considering the probability of these electron relaxations, the initial charge in our systems is expected to increase over time (due to ICD and auto-ionization), spatially delocalize, and rearrange over the molecules (through the processes of charge migration, ICD, and auto-ionization).

1.2.2 Energy deposition of charged particles

The accuracy in quantifying the energy deposition by charged particles, which is experimentally measured as stopping power, is of high importance in assessing biological damage and particularly in radiotherapy. This significance is due to the highly concentrated energy deposition at the Bragg peak. Any error in determining the Bragg peak position can result in extensive damage to healthy tissue. Both the measurement and calculation of stopping power represent a challenge for the realms of experiment and theory alike. Here, we will briefly discuss this challenge in both contexts.

a. Experimental measurement of stopping power

The stopping power can be directly measured by passing well-defined mono-energetic charged particles through a known thin target. The stopping power (S) is calculated by comparing the kinetic energy of charged particles before and after passing through (ΔE^{proj}) the thickness of the target (ΔX^{targ})[47].

$$S = \frac{\Delta E^{proj}}{\Delta X^{targ}} \tag{1.1}$$

This type of measurement may seem simple, but it represents a significant challenge, especially for the condensed phase, as it may have large uncertainty. For example, accurate experimental stopping power data for liquid water, which are highly important in assessing biological damage, do not exist. In modern experimental setups, the uncertainties for the measurement of the stopping power of liquid water correspond to 5% below 15 MeV protons[48], when an accuracy of ± 2% would be desirable for radiotherapy[49]. The sources of these uncertainties mainly come from various factors related to the radiation source and target. These include the mono-energetic nature of the ion beam, inhomogeneity of the ion beam over the target, the thickness of the film used to reduce and select the desired ion beam energy, and the thickness and precise

determination of the target density. Additionally, the detectors used to determine the kinetic energy of the ion beam before and after passing through the target also contribute to the uncertainty. Therefore, alongside experiments, accurate theoretical calculations and numerical simulations are valuable.

b. Theoretical approach for calculating stopping power

The Bethe-Bloch semi-empirical model is a crucial model used in many codes[26,50–52] and data[53,54] for calculating electronic stopping power, in particularly for high-energy particles. Hans Bethe[55], in 1932, and later Bloch[56], included the relativity of charged particles, and proposed a model for calculating the stopping power based on first-order approximation. To simplify his formalism, Bethe made two main assumptions. The first assumption is the electron rest assumption, in which the electrons of the target do not have time to move during the collision. The second assumption is that the interaction of charged particles with the target electrons occurs as binary collisions. This means that the charged particles interact with individual electrons one by one in a many-electron target, neglecting the effect of correlation interactions between electrons. As a result, many corrections and parameters are included in the Bethe-Bloch equation. The modern equation[57] can be expressed as follows:

$$S_e = \frac{\kappa\, \rho_e}{\beta^2}\, Z^2[L_0(\beta) + L_1(\beta) + \cdots] \tag{1.2}$$

Where Z is the charge of the particles. $\beta = c/\upsilon$ where c is speed of light and υ is the velocity of the charged particle. ρ_e is the electron density of the target. $\kappa = 4\pi r_0^2 m_e c^2$, is a constant equal to 0.0005099 where m_e and r_0 are mass and bohr radius of an electron, respectively . The term in the brackets corresponds to the logarithmic part of the equation that includes the parameters and corrections in the Bethe-Bloch equation. The first-order term can be read as:

$$L_0 = \ln\left(\frac{2m_e c^2 \beta^2 \Delta E_{max}}{1-\beta^2}\right) - \beta^2 - ln\langle I\rangle - \frac{C}{Z} - \frac{\delta}{2} \tag{1.3}$$

ΔE_{max} corresponds to the maximum energy transfer from charged particles. $\langle I\rangle$ is the mean excitation energy of the target, which takes into account the electronic structure of the target. The mean excitation energy refers to the minimum energy transfer in a collision.

The term $\frac{C}{Z}$ is a shell correction that arises due to the neglect of many-body electronic effects in the Bethe assumption. Shell correction is more important for heavy atoms. For biologically interesting atoms (with atomic numbers less than 20), this correction represents about a 10% reduction of stopping power at 1 MeV proton energy[57], and it tends to be less than 1% at 100 MeV.

The term $\frac{\delta}{2}$ refers to the density effect correction or Fermi density effect. When the velocity of a charged particle increases, the relativistic effect causes the electric field of the charged particle to shrink and expand in the perpendicular direction of propagation. As a result, when this particle passes through the target, it strongly polarizes the target, and some portion of the long-distances electric field is screened by polarization. Consequently, this correction decreases the stopping power. This correction is more important for condensed phase and high energy particles (above rest mass energy of particles). For stopping power of liquid water for proton, this effect starts raise slow from 1 MeV and reaches about 2% at 1000 MeV[58].

The second term L_1, within the logarithmic bracket in Eq. (1.2), refers to Barkas correction, which corrects some errors from the electron rest approximation. When charged particles, especially those moving at low velocity, pass through a target, some electrons are faster than the particle. As a result, the faster electrons move toward the charged particle during the collision. If the particle carries a positive charge, this causes a decreased distance in the collision between charged particles and electrons, which in turn leads to an increase in energy deposition. Conversely, if the charged particles are negative, they push electrons away during their passage, resulting in a decrease in energy deposition. This phenomenon was discovered by Barkas[59]. This effect is more significant for low-energy particles, as the number of electrons that are faster than the projectile increases. For the stopping of protons in aluminum, this correction in the Bethe-Bloch equation corresponds to a 10% effect at 1 MeV, and this effect can continue up to 300 MeV, as some of the core electrons are still faster even at high energy[57].

Besides the corrections mentioned for the Bethe-Bloch equation, an additional correction is required for low-energy particles, referred to as the effective charge of the projectile. At low energies, particularly those less than 1 MeV[26], the electrons of the target may accumulate around

the projectile during the collision. As a result, the charge of the charged particles is reduced, leading to a decrease in stopping power.

The Bethe-Bloch model is valid for calculating the electronic stopping power for individual elements. For a mixture or molecule, the stopping power should be calculated using Bragg's additive rule, which was proposed by Bragg and Kleeman[60]. According to Bragg's additive rule, the mass stopping powers of a mixture or compound can be determined by a linear combination of the stopping powers of their individual atoms.

$$s_{molecule}(\beta) = \sum_i n_i \, S_i(\beta) \qquad (1.4)$$

Where, n_x is the ratio of number density of the atom i to the total density of all atoms in the target. S_i is the stopping power of atom i as function of velocity of projectile.

According to Bragg's rule, the stopping power of molecules depends solely on the density of their atoms, completely neglecting the effects of chemical and physical bonds within the molecules. This omission must be corrected using experimental data. Bragg's additive rule will be discussed in more detail in Chapter Three.

Overall, due to the high level of correction and parameters in the Bethe-Bloch equation, as well as the model's dependence on Bragg's additive rule, the accurate calculation of stopping power for biomolecules represents a substantial challenge.

To reduce the corrections and parameters in stopping power calculations, several models have been proposed. In their seminal works, Lindhard *et al.* [61,62] proposed a model for electronic stopping power based on the linear perturbation approximation form of the linear dielectric response of a homogeneous electron gas. The Lindhard equation's dielectric function encodes information about quantum mechanical, excitation energies, and many-body effects. The stopping power calculated using Lindhard's formula, without parameters, resembles the Bethe-Bloch equation at high energies but describes the stopping power at low energies more accurately. Lindhard's theory is specifically designed for homogeneous systems, limiting its application to real systems. However, there are extensions for non-homogeneous electron systems through the use of a local-density approximation[63]. In this approach, local stopping is treated as a function of density, with the density values obtained from ground-state calculations.

This method accounts, to some degree, for variations in the electron density. It is important to note that Lindhard's theory, based on first-order perturbation theory, cannot capture features that depend on higher-order effects, such as the Barkas effect and the effective charge of particles. Corrections for these aspects will be necessary. Lindhard's model serves as the basis for several Monte Carlo track structure (MCTS) algorithms used to calculate stopping power at low energy (less than 1 MeV)[64–66], especially when sophisticated experimental data are lacking. MCTS is a common method found in the literature to simulate the physical and physicochemical stages[67,68]. It relies on sets of parameterized elementary cross-sections (such as excitation/ionization, electron scattering, electron attachment, and others) to stochastically simulate the succession of physical and physicochemical events in the medium. MCTS is valuable for dealing with track structure and energy deposition of charged particles at the microscopic level in homogeneous media. However, up to the present date, for biomolecules, the simulation of stopping power has been restricted to water as a way to mimic the biological system[51,69], due to the lack of accurate data.

On this path, advancements in Density Functional Theory (DFT), specifically Real-Time Time-Dependent Density Functional Theory (RT-TD-DFT)[70,71], based on non-linear perturbation, allow for the calculation of energy deposition or stopping power for large and real systems. It is achieved step by step in real time by interaction of charged particles with target materials, without the need for specific parameters for stopping power. RT-TD-DFT can accurately describe the physical processes involved in charged particle interaction.

RT-TD-DFT, which is the principal model in our work, is a promising approach used for calculating energy deposition or stopping power. For instance, in solid-state crystalline structures, Maliyov *et al.* [72,73] investigated stopping power of proton and anti-proton in lithium and aluminum. Correa *et al.* [74] also calculated the stopping power of protons in aluminum. In the condensed phase, Kohanoff *et al.*[75] explored the proton-stopping power of liquid water. Kanai *et al.* [70] calculated the stopping power of DNA for proton and α-particle. All authors found a good agreement between the result of RT-TD-DFT and stopping power data. Up to the present day, the calculation of energy deposition and stopping power by RT-TD-DFT has been limited to energies below 2 MeV, due to the relativistic effect of charged particles. In this thesis, we aim to explore the effect of charged

particle relativity and hydrogen bonding in biomolecules on energy deposition across a wide range of energies, particularly in Chapters Two and Three.

1.2.3 Attoseconds electron relaxation

In order to explicitly investigate the dynamics at the physical stage, where the time range spans between a few attoseconds and femtoseconds, an understanding of the time evolution of electronic relaxation is necessary. Over the past decade, advancements in quantum mechanics have enabled access to this time scale. Beyond theory, experimental techniques have progressively evolved to investigate events occurring at the physical stage. More recently, attosecond pump-probe spectroscopy has successfully provided explicit real-time insights into electron dynamics. It is difficult to summarize all experimental and theoretical models here. For more detailed information, we refer readers to review articles in this domain[76–78]. Instead, we will briefly introduce the attosecond XUV pump-IR probe pulse spectroscopy and focus on the RT-TD-DFT method.

a. Experimental approach

The pump-probe laser approach is an essential technique for revealing ultrafast events at the molecular level using two short laser pulses. The first pulse, referred to as the 'pump,' clocks ultrafast events, while the second pulse reveals time-dependent events through a time-delay with the first pulse. The time window for investigation using the pump-probe laser approach highly depends on the duration of the laser pulses. To access the physical stage, very short pulses in the range of attoseconds and femtoseconds are required.

In the case of charged particles, preparing short packets of ion beams is limited to the picosecond time scale[79] due to the high repulsion between charged particles, which prevents a compact ion beam in a short space. On the other hand, after investigating high harmonic generation (HHG)[80,81], it becomes possible to create coherent photon pulses on a very short-time scale. When a strong laser field interacts with a gas of atoms or molecules, it can pull electrons away from their parent ions and then drive them back. Upon re-collision with the parent ion, the electron may recombine, releasing its kinetic energy as a highly coherent photon in the range of XUV rays.

The XUV pulse, generated through HHG, offers exceptional approaches for probing matter's responses with attosecond resolution. For instance, in XUV pump-IR probe spectroscopy, after placing the molecule in the gas phase, the XUV pump pulse ionizes and excites the valence electrons of the molecule. Meanwhile, an IR laser pulse induces a second ionization state and characteristic fragmentation that can be detected by mass spectroscopy. By introducing a delay of a few attoseconds between the XUV and IR pulses, specific electron dynamics events can be indirectly investigated by observing positive or negative signals of fragmentation yield.

The investigation of large biomolecules through the XUV-IR approach represents a great challenge. Our experimental collaborators in the Lépine group in Lyon have developed a new XUV-IR approach that allows for the investigation of large biomolecular systems, such as peptides and proteins[82]. This is achieved by combining an electrospray ionization source and a mass spectrometer with XUV-IR pump-probe laser pulses. The results of XUV-pulse interaction simulations for peptides and proteins, in collaboration with the Lépine group, are presented in Chapter Five.

Attosecond experiments have been successfully used to investigate the processes that occur in the physical stage, as mentioned in the previous sub-section. Examples include the investigation of ultra-fast charge migration in iodoacetylene[40], biomolecules like phenylalanine[45]. Beyond experiments, theoretical simulations are complement to interpret experimental results, especially when the experimental data come from indirect events through fragmentation.

b. Theoretical approach

Numerical simulation for investigating electron dynamics induced by IoR depends on two stages (I) preparing the initial ionization state, (II) tracking the evolution of electron dynamics.

I. Initial ionization state

IoR strongly interacts with molecules, creating a superposition of states, and as a result, several electrons at different levels with varying magnitudes can be ionized and excited. The generation of an ionization state in a molecule is essential for properly studying electron relaxation. There are two major approaches for preparing the initial ionization state.

The first approach is called simultaneous ionization or sudden ionization. In this method, the interaction of IoR is neglected, and the electron dynamics occurs directly by creating Koopman's hole[83] through removing an electron from the ground state orbital. This model is widely[44,84] used because it simplifies the ionization processes, particularly as it does not require the description of continuum states, which is a challenging task in simulation. However, the sudden ionization approach omits the electron relaxation that could happen during radiation interaction.

In the second approach, the interaction of IoR is directly introduced during the ionization processes. This can be achieved by interacting the electron cloud with the electric field of an electromagnetic wave, which can be executed under the assumption of dipole interaction. Furthermore, interactions with charged particles can be simulated via Coulomb scattering. This approach will be described in more detail in the next section. This approach is more accurate, and it is necessary, particularly in the case of charged particles, where the ionization is local along the trajectory and the dynamics induced by the particle during the collision is essential. On the other hand, Fernando Martin *et al.*[85] revealed that the two approaches produce the same result after a sub-femtosecond (sub-fs) time frame in the case of interaction a short XUV-pulse (300 as) with glycine. However, the clarity of the comparison remains uncertain in the case of longer pulses.

II. Electron dynamics

Time-dependent electron dynamics are required to capture electron relaxation. Since electron relaxation is purely governed by electron correlation, sophisticated quantum mechanics models that properly include the effects of electron correlation are necessary. Post-Hartree-Fock models[80,86,87] can successfully describe electron relaxation. However, these models are limited to small molecules due to computational cost. For larger molecules, TDDFT is often the method of choice. In particular, RT-TD-DFT directly obtains information about the electron relaxation step by step in real-time, whether a molecule is induced by photons or charged particles. RT-TD-DFT has been successfully used to investigate electron relaxation. For instance, it has been applied to the study of charge migration in nitrosobenzene induced by X-ray[88], charge migration in phenylalanine induced by XUV-pulse[45], and electron dynamics and charge migration in DNA induced by charged particles[71].

In most applications of RT-TD-DFT such as in our case, the nuclei are considered fixed, and the dynamics are purely electronic. This assumption is often made on the basis that the nuclei are heavier and require more time to respond. However, the electron dynamics may be affected by nuclear motions, leading to decoherence of the wave function.

For instance, the role of nuclear motion in charge migration was studied by Fernando Martin *et al.*[85] They found that nuclear dynamics only has a significant effect after approximately 8 femtoseconds (fs), but it does not destroy the coherent motion of the electronic wave packet for at least a few additional tens of fs. This suggests that while nuclear dynamics can play a role, it may not always be critical in the timescale of certain electronic processes. Therefore, the assumption of fixed nuclei may be valid in many cases, particularly for the short time scales often investigated in RT-TD-DFT simulations.

1.3 Theoretical background

In this section, we introduce the theoretical models used in this work. In the first part, we discuss the principles of ground-state Density Functional Theory (DFT), Linear Combination of Atomic Orbitals (LCAO), and Auxiliary Density Functional Theory (ADFT). In the second part, we delve into Real-Time Time-Dependent Auxiliary Density Functional Theory (RT-TD-ADFT), encompassing topics such as numerical propagation of electron densities, the introduction of external perturbations, the role of the exchange-correlation functional, the effect of Gaussian basis sets, and the utilization of Complex Absorbing Potential.

1.3.1 Ground-state Density Functional Theory

Density-functional theory (DFT) is a popular quantum mechanical method for applications in physics, chemistry and materials science to investigate the electronic structure of many-body systems such as atoms, molecules, or condensed phases based on the spatial functional of the electron density. DFT has emerged as one of the most powerful theoretical approaches not only for quantitative prediction of the electronic structure and molecular properties, but also for a

rigorous foundation for many concepts widely used in chemistry. The main advantage of DFT is that it incorporates electronic correlation, and the predicted quantitative results are comparable to those obtained from various post-Hartree-Fock methods (models based on wave function including electronic correlation) at much less computational cost.

c. History and Kohn-Sham equation

DFT finds it roots in the works of Thomas and Fermi (1927)[89]. These authors used the electron density as a basic variable to calculate, quantum mechanically, the total energy of a uniform electron gas. The model developed by these researchers is straightforward to apply to real-systems. However, it cannot describe essential properties of atom and molecules, for example chemical bonds or electronic shells.

In 1964, Hohenberg and Kohn[90] laid the foundations of modern DFT. Hohenberg-Kohn theorem (HK) is a prof of the one-to-one mapping between the electron density and the external potential. This means that the total energy of a system is a functional of the electron density only. Furthermore, the uniqueness and electron density of the ground state can be achieved by the minimizing energy via the variational principle. Finally, this density allows us to find any observable of ground or excited states when the expectation value of ground state $\mathcal{O}$ is a unique functional of electron density. where Ψ is the wave function and $\hat{\mathcal{O}}$ is operator associated to an arbitrary observable.

$$\langle \Psi[\rho] | \hat{\mathcal{O}} | \Psi[\rho] \rangle = \mathcal{O}[\rho] \qquad (1.5)$$

However, The HK theorem tells us that the universal functional exists, but it does not provide us with an expression to find it. In 1965, Kohn and Sham (KS)[91] proposed to represent the N-electron system as a non-interacting electron system for which the energy expression is known. Since the electron density of the ground state is solely determined by the external potential (HK), the ground-state densities of both interacting and non-interacting systems are the same. This is useful in solving the problem of the main part of the kinetic energy by introducing non-interacting orbitals φ_i which are represented as one slater determinant Φ_n. According to KS, the total energy of N-electrons system E^{KS} is given as:

$$E^{KS}[\rho] = T_s[\rho] + E_{ext}[\rho] + J_{ee}[\rho] + E_{xc}[\rho] \quad (1.6)$$

ρ is the electron density. E_{ext} is the energy associated to the external potential, mainly that created by the atom nuclei. The external potential can be supplemented with additional components, such as interactions with an external electric field or a charged projectile. In hybrid Quantum Mechanics/Molecular Mechanics (QM/MM) simulations, the external potential also encompasses the potential generated by charged or polarized MM atoms. J_{ee} is the classical repulsion between electrons, E_{xc} is the exchange-correlation energy, and T_s is the kinetic energy of the reference non-interacting system :

$$T_s[\rho] = \sum_{i=1}^{N_e} \int \varphi_i^*(\boldsymbol{r}) \left(-\frac{1}{2}\nabla^2\right) \varphi_i(\boldsymbol{r}) \quad (1.7)$$

Where N_e is number of electrons in the system, φ is one electron wave functions.

We can rewrite the KS total energy as:

$$E^{KS}[\rho] = T_s[\rho] + \int V_{ext}\, \rho(\boldsymbol{r}) d\boldsymbol{r} + \frac{1}{2}\int \frac{\rho(\boldsymbol{r})\rho(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|}\, d\boldsymbol{r}\, d\boldsymbol{r}' + E_{xc}[\rho] \quad (1.8)$$

The electron density ρ is given by mono electronic Kohn-Sham Slater determinant. For a closed-shell system, it reads:

$$\rho(\boldsymbol{r}) = 2\sum_{i=1}^{N_e/2} |\varphi_i(\boldsymbol{r})|^2 \quad (1.9)$$

T_s, E_{ext}, and E_{ee} can be solved through the analytical expression by using a non-interacting electrons system. However, E_{xc} remains unknown and all features of the interacting system are hidden, especially, in the correlation energy E_c.

d. Exchange and correlation functionals

The exact expression for E_{xc} is currently unknown and must be approximated. This remains an interesting topic until now. The accuracy of a DFT calculation significantly depends on the model chosen for E_{xc}. The latter decomposes into an exchange and a correlation functionals. The exchange functional E_x corresponds to the difference between electron-electron repulsion energy of non-interacting system and the classical Coulomb electron energy.

$$E_x[\rho] = \langle \Phi_n | \widehat{W}_{ee} | \Phi_n \rangle - \frac{1}{2} \int \frac{\rho(\boldsymbol{r})\rho(\boldsymbol{r}')}{|\boldsymbol{r} - \boldsymbol{r}'|} \, d\boldsymbol{r} \, d\boldsymbol{r}' \quad (1.10)$$

The correlation functional E_c refers to the energy resulting from the influence of all electrons on each electron in a N-electron system which consists of an electron kinetic and potential correlation functional.

$$E_c[\rho] = T_c[\rho] + U_c[\rho] \quad (1.11)$$

The kinetic energy contribution represents the difference of the kinetic energy of the complete interacting system (wave function ψ_i) and that of the non-interacting system.

$$T_C[\rho] = \langle \psi_i | \hat{T} | \psi_i \rangle - \langle \Phi_n | \hat{T} | \Phi_n \rangle \quad (1.12)$$

By the same way, the potential energy contribution is defined by:

$$U_c[\rho] = \langle \psi_i | \widehat{W}_{ee} | \psi_i \rangle - \langle \Phi_n | \widehat{W}_{ee} | \Phi_n \rangle \quad (1.13)$$

There are numerous functionals available in the literature, each attempting to solve a particular problem while striking a balance between accuracy and computational cost. Before we introduce some common classes of approximations, we will summarize the main issues that may appear due to these approximations.

I. Asymptotic behavior of DFT. The exact exchange potential of an electron decays is $-\frac{1}{\boldsymbol{r}}$ in the asymptotic limit $\boldsymbol{r} \rightarrow \infty$, but in the most DFT functionals, the exchange potential decreases exponentially and decays faster at long distance. This has an impact on the prediction of some molecular properties that are dependent on long-distance interactions, for instance, electronic excitation energy[92], orbital bandgap[93], ionization energy, charge transfer, molecular dissociation energy, bond lengths[94], binding energies of weakly interacting systems[95] etc. However, Hartree-Fock (HF) exchange functional exactly shows the correct decay($-\frac{1}{\boldsymbol{r}}$) for large distances (***r***).

II. Second, the problem known as the derivative discontinuity[96] (DDC). For systems with fractional electrons, when the number of electrons passes through the integer, the exact exchange-correlation potential undergoes a jump by a constant and the total energy of the system varies by a straight-line connecting integer electron number. DFT cannot

capture this behavior when there is a non-integer number of electrons. This phenomenon directly relates to the exchange-correlation approximation and will impact the system that depends on the fraction of electrons[97] , for instance, partial ionization, electron affinity, band gaps, etc[98,99].

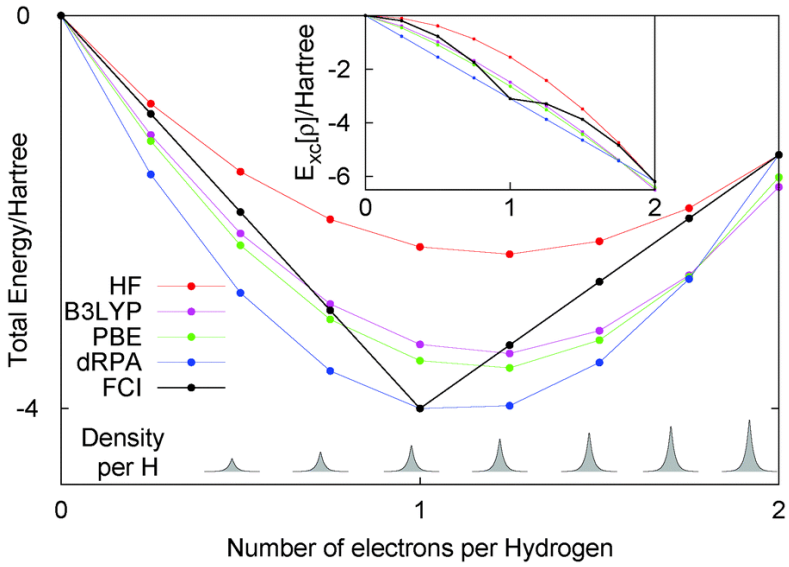

Figure 1-5: The energy of the closed shell of H_8 with fractional numbers of electrons per hydrogen atom, calculated using various methods including Hartree-Fock (HF), direct random phase approximation (dRPA), hybrid functional (B3LYP), GGA (PBE), and exact (FCI). The figure illustrates the lack of discontinuous behavior in both the total energy and the exchange-correlation part, as shown in the inset. The figure is taken from Reference[97] with permission.

III. Self-interaction error (SIE). It corresponds to the interaction of residual electron with itself. For any system with one electron or less, there are no electron-electron interactions, leading to the condition that the exchange energy exactly cancels the Coulomb energy, and the energy of correlation disappears.

$$-\frac{1}{2}\int \frac{\rho(\boldsymbol{r})\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\, d\boldsymbol{r}\, d\boldsymbol{r}' + E[\rho,0] = 0 \qquad (1.14)$$

This condition is not reached by common XC functional. The problem is still a challenge of approximating the exchange-correlation term in DFT, but it is exactly satisfied in HF and all wave function theories. Many qualitative and quantitative failures in fractional charge systems are caused by SIE. It is significantly greater for radical systems[100] than for neutral systems[101]. The SIE mainly leads to increased charge delocalization[102], decreased total energy of systems, reduced reaction barriers[103,104], and overestimated electron attachment [105].

The three types of errors are correlated[106], but each of them requires a different strategy to overcome. As a result, there are numerous approximations and functionals available.

A first approximation assumes that the electron density varies slowly, so that, the electron density at a point in space is constant. Consequently, the exchange-correlation energy for electrons in the small volume in space can be replaced by a uniform gas for the constant density at this volume. That is called the local density approximation (LDA).

$$E_{xc}^{LDA}[\rho] = \int d\boldsymbol{r}\rho(\boldsymbol{r})\,\varepsilon_{xc}^{unif}(\rho(\boldsymbol{r}) \tag{1.15}$$

According to the LDA, the exchange contribution could be analytically calculated as uniform gas density. However, the correlation part cannot be achieved analytically, instead of that the parameterized function exists by fitting it with the value obtained from quantum Monte Carlo calculation. LDA is highly simple and computationally faster than other DFT approximations, with reasonable accuracy even for slightly perturbed systems. However, due to the poor performance of these functionals in predicting molecular properties for real-system, several other XC functionals have been proposed.

One of them included gradient corrections, this is called the Generalized Gradient Approximation (GGA). The method has been developed to consider the density variation by expressing the exchange and correlation energies as a function of the density, and also of its gradient.

$$E_{xc}^{GGA}[\rho] = \int d\boldsymbol{r}\, f_{xc}[\rho(\boldsymbol{r}), \nabla\rho(r)], \tag{1.16}$$

GGA is semi-local because the function f_{xc} depends on the gradient of the density as well. It improves the quality of the XC functional for systems whose electron density is inhomogeneously distributed. In practice, GGA is more accurate than LDA. The main errors in DFT (noticeably smaller in GGA compared to LDA) are still inside. In this class of functionals, Perdew, Burke, and Ernzerhof (PBE)[107] is one of the most widely-used functionals within the physics and chemistry communities. It does not contain any empirical parameters. This functional is a simplified version of the Perdew and Wang functional (PW91)[108]. It is often considered the best in the GGA class, and sometimes even outperforms functionals from superior classes, for instance, in term of self-interaction error[109], band gap energy[110], electronic structure[111], static polarizabilities[112], molecular electron affinity[113], cohesive energies in sold[114], ligand dissociation energies[115] etc.

Meta-GGA functionals go beyond GGA by incorporating not just the electron density and its gradient but also the Laplacian of the density or of the kinetic energy density.

$$E_{xc}^{Meta-GGA}[\rho] = \int d\boldsymbol{r}\, f_{xc}[\rho(\boldsymbol{r}), \nabla\rho(r), \nabla^2\rho(r), \tau(r)], \tag{1.17}$$

Incorporating the kinetic energy density (τ) generally makes meta-GGA functionals more flexible and accurate for certain systems, especially where the gradient of the electron density is not sufficient to capture the variations in the exchange-correlation energy. However, the additional complexity in meta-GGA functionals also means that calculations using these functionals are computationally more demanding than those using LDA or GGA functionals. Examples of popular meta-GGA functionals include TPSS (Tao-Perdew-Staroverov-Scuseria)[116].

Global hybrid functionals represent another class of XC-functionals. The idea behind these functionals is to include a percentage of exact exchange, calculated in the same manner as in the Hartree-Fock (HF) method, but using Kohn-Sham molecular orbitals. This procedure is motivated by the fact that the exchange part in DFT is responsible for several issues. This approximation brings significant improvements to DFT, depending on the exact exchange percentage and type of exchange-correlation (XC) functional, but it does come with an increased computational cost. An example of this class is PBE0 functional[117,118] that reads:

$$E_{xc}^{PBE0} = 0.25\, E_x^{exact} + 0.75\, E_x^{PBE} + E_c^{PBE} \tag{1.18}$$

where E_x^{exact} is the exact exchange energy, E_x^{PBE} is the PBE exchange energy, and E_c^{PBE} is the PBE correlation energy.

With hybrid functionals, the incorrect asymptotic behavior of exchange potential persists[119]. In order to overcome it, other classes of functionals have been developed, so call range-separated functionals. The idea of this approach is to vary the exact exchange percentage according to the inter-electronic distance. This approximation is the most accurate but also the most computationally expensive of the other classes mentioned[120]. The CAM-B3LYP functional[121] is one of the range-separated hybrid density functionals. It is an acronym for Coulomb-Attenuating Method Becke, 3-parameter, Lee-Yang-Parr functional. CAM-B3LYP was specifically designed to enhance the representation of charge-transfer excitations.

e. Linear combination of atomic orbitals

The linear combination of atomic orbitals (LCAO) approach is a commonly used method for constructing molecular orbitals (MO) in DFT and in quantum chemistry more broadly, notably in deMon2k[2,122] which is the DFT program mainly used in this thesis. Within the LCAO framework, we expand the non-electronic wavefunction as a linear combination.

$$\varphi_i(\boldsymbol{r}) = \sum_{\mu} c_{\mu i}\, \mu(\boldsymbol{r}) \tag{1.19}$$

Where the $c_{\mu i}$, are molecular orbital coefficients. μ represents an atomic wavefunction, often referred to as atomic orbital (AO) or atomic basis functions.

The atomic basis set can be formed with Gaussian type functions. In this approach, the radial components of atom-centered orbitals are represented as linear combinations of Gaussian Type orbitals (LCGTO).

$$\mu(\boldsymbol{r}) = Y_l^m(\hat{r}) r^l \sum_{b} c_b e^{-\alpha_b (r-R)^2} \tag{1.20}$$

The real spherical harmonics functions are represented by Y_l^m to describe the angular part of the orbital, while the decomposition coefficients, represented by c_b, are established for a given basis set and R denotes the position of a nucleus. In this equation, the Gaussian exponents α_b give the spatial range of the basis function. Basis functions with smaller α_b parameters are more diffuse compared to those with larger α_b parameters. This configuration allows for an adjustable level of precision in describing both core and valence electrons. In principle, the accuracy of the method improves as the number of atomic orbitals in the basis set increases.

The electron density (ρ) is written as a product of the density matrices and basis functions.

$$\rho(\boldsymbol{r}) = \sum_{\mu,\nu}^{N_{AO}} P_{\mu\upsilon}\, \mu(\boldsymbol{r}) \nu(\mathbf{r}) \tag{1.21}$$

μ , ν denote AO. N_{AO} is the number of AOs. The density matrices elements $P_{\mu\upsilon}$, are defined as follows:

$$P_{\mu\nu} = 2\sum_{i}^{occ} c_{\mu i}^{*} c_{\nu i} \tag{1.22}$$

Where occ is the number of occupied molecular orbitals.

Now, we can rewrite the KS energy expression as follows:

$$E^{KS}[\rho] = \sum_{\mu,\nu} P_{\mu\nu}\, H_{\mu\nu}^{core} + \frac{1}{2}\sum_{\mu,\nu}\sum_{\sigma,\tau} P_{\mu\nu}\, P_{\sigma\tau}\langle\mu\nu||\sigma\tau\rangle + E_{xc}[\rho] \tag{1.23}$$

$H_{\mu\nu}^{core}$ is the core Hamiltonian which collects one-electron operators, namely the external potential operator $H_{\mu\nu}^{ext}$ and kinetic operator $H_{\mu\nu}^{kin} = \left\langle \mu \middle| -\frac{1}{2}\nabla^2 \middle| v \right\rangle$. The symbol ‖ stands for the Coulomb operator $(1/|\boldsymbol{r} - \boldsymbol{r}'|)$. The KS potential is obtained by differentiating the energy with respect to the electron density. In matrix notation, the elements of the KS potential can be derived directly by differentiating with respect to the elements of the density matrix.

$$H_{\mu\nu} \equiv \frac{\partial E^{KS}}{\partial P_{\mu\nu}} = H_{\mu\nu}^{core} + \sum_{\sigma,\tau} P_{\sigma\tau}\langle\mu\nu\|\sigma\tau\rangle + \frac{\partial E_{xc}[\rho(\boldsymbol{r})]}{\partial P_{\mu\nu}} \tag{1.24}$$

The LCGTO method necessitates careful consideration due to its potential limitations in capturing some phenomena prompted by strong density perturbations, such as ionization. We will delve into this concern in a subsequent section. Local basis sets are selected owing to their efficient computational algorithms developed by the quantum chemistry community for electronic integral calculations[123,124].

f. Auxiliary Density Functional Theory

The computation of classical electron-electron repulsion involves four-center-electron-repulsion-integrals (ERIs) which imposes a significant computational burden. Once this challenge is managed, another computational obstacle presents itself in the form of evaluating the exchange-correlation contribution. To alleviate these issues, the deMon2k software implements the Auxiliary Density Functional Theory (ADFT) framework[125]. The ADFT approach capitalizes on a variational fitting method for the Coulomb potential, approximating four-center ERIs with simpler two- and three-center integrals[126]. This method utilizes an auxiliary density function ($\tilde{\rho}$), expressed as a linear combination of auxiliary basis functions $\bar{k}(\boldsymbol{r}) : \tilde{\rho} = \sum_{\bar{k}} x_{\bar{k}} \bar{k}\,(\boldsymbol{r})$, to replicate the Coulomb repulsion energy as precisely as possible. For the purpose of computational

efficiency, deMon2k comes with an in-built algorithm that auto-generates auxiliary function sets from a provided atomic orbital basis set. These auxiliary basis set functions are, in fact, atom-centered primitive Hermite-Gaussian functions that are assembled into groups with same exponents ζ_k at atom K[127].

$$\bar{k}(\boldsymbol{r}) = \left(\frac{\partial}{\partial K_x}\right)^{\bar{k}_x}\left(\frac{\partial}{\partial K_y}\right)^{\bar{k}_y}\left(\frac{\partial}{\partial K_z}\right)^{\bar{k}_z} e^{-\zeta_{\bar{k}}(\boldsymbol{r}-K)^2} \tag{1.25}$$

The $x_{\bar{k}}$ terms are density fitting coefficients. The coefficients are derived using a variational fitting method by minimizing a self-interacting error function.

$$\varepsilon = \frac{1}{2}\langle \rho - \tilde{\rho} \| \rho - \tilde{\rho} \rangle = \frac{1}{2}\langle \rho \| \rho \rangle - \langle \rho \| \tilde{\rho} \rangle + \frac{1}{2}\langle \tilde{\rho} \| \tilde{\rho} \rangle \tag{1.26}$$

The error introduced via the fitted density can be progressively minimized by enhancing the quality of the auxiliary basis set. In matrix notation, the KS energy derived from the ADFT energy expression is given as:

$$E^{ADFT}[\tilde{\rho}] = \sum_{\mu,\nu} P_{\mu\nu}\, H_{\mu\nu}^{core} + \sum_{\mu,\nu}\sum_{\bar{k}} P_{\mu\nu}\langle \mu\nu || \bar{k}\rangle x_{\bar{k}} - \frac{1}{2}\sum_{\bar{k},\bar{l}} x_{\bar{k}}\,\langle \bar{k} || \bar{l}\rangle x_{\bar{l}} + E_{xc}[\tilde{\rho}] \tag{1.27}$$

The KS potential, which is derived from the ADFT energy expression, can be reformulated as follows:

$$H_{\mu\nu}^{ADFT} = \left(\frac{\partial E^{ADFT}}{\partial P_{\mu\nu}}\right) = H_{\mu\nu}^{core} + \sum_{\bar{k}} \langle \mu\nu \| \bar{k} \rangle \left(x_{\bar{k}} + z_{\bar{k}}\right) \tag{1.28}$$

$$z_{\bar{k}} = \sum_{\bar{l}} \langle \bar{k} \| \bar{l} \rangle^{-1} \langle \bar{l} | v_{xc} \rangle, \text{ where } v_{xc} \equiv \frac{\partial E_{xc}[\tilde{\rho}]}{\partial P_{\mu\nu}} \tag{1.29}$$

In contrast to KS energy, there are no four-center ERIs, but only two- and three-center ERIs involved in the energy or potential calculations. This approximation brings down the formal computational complexity of the Coulomb integral computation from the fourth power of the number of atomic orbitals (N_{AO}^4) to a product of the number of atomic orbitals and auxiliary basis functions ($N_{AO}^2 . N_{AF}$). Typically, N_{AF} is three to five times the number of basis functions, but this is significantly less than the number of product terms involving atomic orbitals. It is also important to note that the auxiliary density enters in the exchange-correlation contribution (XC) as well. When using LDA or GGA exchange-correlation functionals, the auxiliary density ($\tilde{\rho}$) can be directly

used to evaluate E_{xc}. However, when it comes to hybrid functionals, the process is a bit more complex as it necessitates the variational fitting of the Fock potential[128].

1.3.2 Real-Time Time-Dependent Auxiliary Density Functional Theory

a. Numerical propagation of electron densities

The Runge-Gross theorem[129] provides the theoretical foundation for the application of density functional theory to time-dependent systems.

In essence, the Runge-Gross theorem states that for a system of interacting electrons in an external potential that varies over time, the time-dependent external potential is a unique functional of the time-dependent density of the system, up to a constant. In other words, the theorem tells us that if we know how the electron density of a system changes over time, we can determine the corresponding time-dependent external potential (up to a constant) that must have been applied to the system.

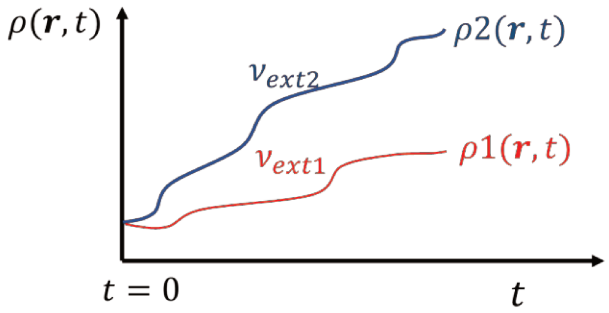

Figure 1-6: Conceptual representation of the Runge-Gross theorem: two densities, ρ1 and ρ2, under the influence of two distinct potentials, v_{ext1} and v_{ext2}, begin to diverge instantaneously post the time instance t = 0.

This uniqueness of the time-dependent potential-density mapping is the time-dependent generalization of the Hohenberg-Kohn theorem of ground state density functional theory, which states that the ground-state density determines the external potential, up to a constant. The time-dependent DFT equations can be formulated in terms of the Liouville-von Neumann equation[130,131]:

$$i\frac{\partial\rho(r,t)}{\partial t} = [H(r,t),\rho(r,t)] \tag{1.30}$$

The non-relativistic time-dependent Kohn-Sham (TDKS) operator H is a functional of the density. Here, the Hamiltonian operator is akin to that in the ground state, but it evolves not just spatially but also temporally. In addition, the same procedure used for the ground state can also be applied for density fitting in the context of time-dependent.

The v_{XC} is a density functional that, in theory, is non-local in both space and time. However, in many practical implementations, the time-dependency of v_{XC} is often disregarded, leading to what is referred to as the adiabatic approximation. This approach allows the use of approximate XC functionals designed for stationary DFT calculations to be reused in RT-TD-DFT without the need for specific reparametrizing. Although the adiabatic approximation can be adequate for many applications, it introduces an additional approximation to the propagation, on top of the inherent uncertainty regarding the exact form of the XC functional.

The issue with numerically propagating Eq. (1.29) is the intrinsic time-dependency of H, which is a functional of the density. As such, long propagation (for instance, for tens of attoseconds) is impractical. This difficulty is overcome, in practice by discretizing time into tiny increments, Δt, which are typically in the range of 0.1-2 as. There are two key properties that should be satisfied by this process: the propagator must be unitary and time-reversible[132,133]. It is also important to remember that there is a physical limit to the maximum allowable Δt, which is determined by the specific physical process being examined. If ω_{max} denotes the maximum frequency that needs to be resolved in the simulation, the maximum Δt should not exceed roughly $1/\omega_{max}$. Currently, deMon2k employs the second-order Magnus method[132,134] to fulfill these requirements.

$$\rho(t_n + \Delta t) = e^{-iH\left(t+\frac{\Delta t}{2}\right)*\Delta t}\rho(t_n)e^{-H\left(t+\frac{\Delta t}{2}\right)*\Delta t} \tag{1.31}$$

Eq. 1.31 facilitates the propagation of electron density from t_n to $t_n + \Delta t$, given that we know the Kohn-Sham potential at time $t_n + \frac{\Delta t}{2}$. deMon2k proposes two propagation schemes: one is based on an iterative solver[133], and the other on a predictor-corrector (PC) solver[135]. The iterative solver, while more robust and likely to ensure stable propagations with sufficiently small time-steps[136], is also more computationally demanding. The PC solver, on the other hand, is often stable in most simulations when used with time steps of around 1 as, and it requires only one evaluation of the KS potential per propagation step, thus offering substantial computational time savings. A challenging aspect of these calculations is the need to compute the exponential of the

KS matrix. In deMon2k, the user can choose from several options to perform this task: a straightforward diagonalization of the matrix, a Taylor expansion, a Chebyshev expansion, or a Baker–Campbell–Hausdorff expansion[131,132].

In all our RT-TD-ADFT simulations, we used PC solver and Taylor expansion, ensuring the stability of the simulation beforehand.

b. External perturbations

In RT-TD-ADFT, various types of perturbations can be used to stimulate dynamics in the electron cloud. These enable explicit simulation of the interaction with IoR, thereby directly quantifying energy deposition and ionization, as well as initiating the superposition of states and tracking the subsequent electron response. Generally, a simulation begins with a stationary Self-Consistent Field (SCF) calculation to determine the ground state electron density of the system of interest. Following this, Electron Dynamics (ED) simulations are carried out by introducing a perturbation. This is achieved by adding terms into the external potential (v_{ext}). One option is to simulate the interaction of the electron cloud with the electric field of an electromagnetic wave, which can be executed under the assumption of dipole interaction. For charged particles, the interactions can be incorporated via Coulomb interaction between charged particle and electrons. Both of approximations are implemented in deMon2k[137].

I. Electromagnetic wave

The homogeneous electric field component ($\boldsymbol{F_e}$) of the electromagnetic wave interacts with the molecular dipole moment ($\boldsymbol{\mu}$) (vectors are written in bold):

$$E^{pert} = -\boldsymbol{\mu}(t) \cdot \boldsymbol{F_e}(t) \tag{1.32}$$

$\boldsymbol{F_e}$ can either be the result of the external electric field of monochromatic light multiplied by a carrier function (such as a Gaussian pulse, a squared cosine pulse, or a linear ramp), or it could be an instantaneous electric kick. Where $\boldsymbol{\mu}$ is the molecular dipole moment.

$$\boldsymbol{\mu}(t) = \sum_A Z_A \boldsymbol{R}_A - \int \rho(\boldsymbol{r}, t).\boldsymbol{r}.d\boldsymbol{r} = \sum_A Z_A \boldsymbol{R}_A - \sum_{\mu,\nu} P_{\mu\nu}(t)\, \boldsymbol{D}_{\mu\nu} \tag{1.33}$$

Z_A and $\boldsymbol{R}_A$ are the charge and position of atom nuclei A. $\boldsymbol{D}_{\mu\nu}$ is an element of the dipole operator matrix : $\boldsymbol{D}_{\mu\nu} = \langle\mu|\boldsymbol{r}|\nu\rangle$.

The external potential that should be included in the KS potential during the propagation is derived by differentiating the interaction energy relative to the electron density. In matrix notation, it can be written as:

$$\frac{\partial E^{pert}}{P_{\mu\nu}} = \boldsymbol{F_e}(t).\boldsymbol{D}_{\mu\nu} \tag{1.34}$$

In our work, especially in Chapter Five, we employed an instantaneous electric kick to calculate the absorption spectrum and a squared cosine pulse to simulate XUV interactions using RT-TD-ADFT.

The instantaneous electric kick represents a wide range of energies that suddenly perturbs the system from its ground state, resulting in the formation of a superposition of many excited states. By monitoring the dipole moment, we can calculate a broad range of spectral features. Hence, an absorption spectrum can be calculated using RT-TD-ADFT by initiating three electron dynamics simulations from a static electron density, perturbed by a weak electric field with strength κ directed along either the x, y, or z axes (d). The Fourier transformation of the dipole moment (μ_j) captured during the simulation provides insights into the polarizability tensor α: $\alpha_{d,j} = \frac{1}{\kappa}\mu_{d,j}(\omega)$. Then, the absorption cross-section tensor σ and dipole strength function S in the frequency domain can be calculated using the formulas: $\sigma = \frac{4\pi\omega}{c}\mathrm{Im}[\alpha(\omega)]$, where c denotes the speed of light, and $S(\omega) = \frac{1}{3}Tr[\sigma(\omega)]$.

For monochromatic light, $\boldsymbol{F_e}$ can also be expressed as: $\boldsymbol{F_e}(t) = F_o.X(\mathrm{t}).\sin\omega t\,.\boldsymbol{d}$, Here, F_o represents the maximum field intensity, X indicates the pulse envelope, $\boldsymbol{d}$ is the direction vector of the pulse, and ω signifies the frequency of the light. Here, squared cosine pulse X can be expressed as:

$$X(t) = \begin{cases} 0 & if\ t > 2t_o \\ cos^2[\pi(t - t_o)/2t_o] & if\ t \leq 2t_o \end{cases} \tag{1.35}$$

The shape of the squared cosine pulse is depicted in Figure 1-7 where F_o and ω are 0.05 (Ha/e.bohr) and 0.057 Ha respectively. The half-duration of pulse t_o is 26.664 fs.

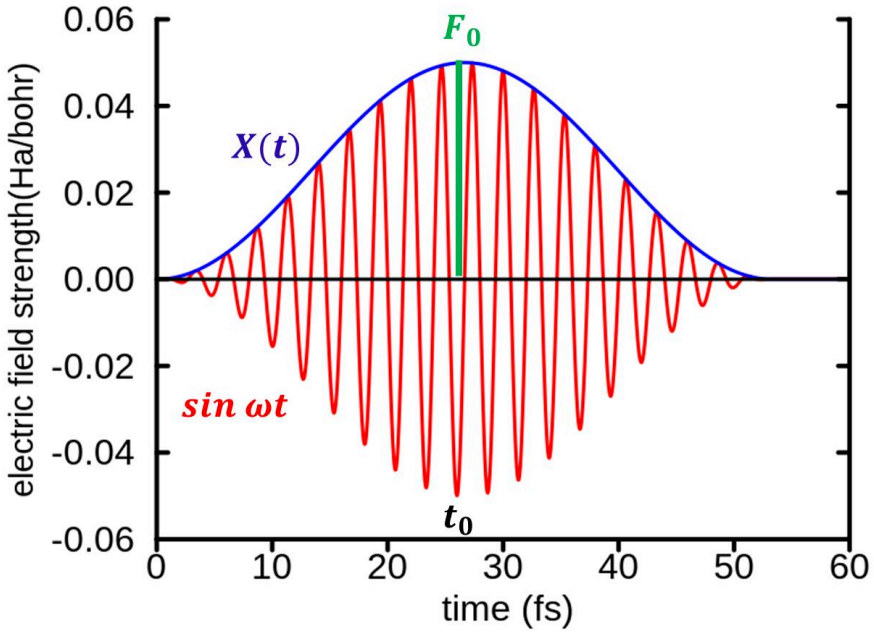

Figure 1-7: representation of the squared cosine pulse shape.

II. Fast ions

In deMon2k, another kind of perturbation that can be used is the interaction with a fast charged particle. This interaction allows the simulation of inelastic collisions, meaning collisions that do not alter the trajectory of the projectile. The energy resulting from the interaction between the molecule and the projectile is expressed as:

$$E^{mol/proj} = -\int \rho(\boldsymbol{r},t).\phi_{proj}(\boldsymbol{r},t).\boldsymbol{dr} + \sum_A Z_A \boldsymbol{R}_A.\phi_{proj}(\boldsymbol{r},t) \qquad (1.36)$$

ϕ_{proj} is the potential generated by the projectile with charge q_{proj}. For projectiles significantly slower than light, this potential is typically computed using the Coulomb's law, a common practice in numerous codes. However, when dealing with projectiles of which the velocity approaches the speed of light, such as protons with kinetic energies exceeding 1 MeV, the Liénard-Wiechert potential (LWP) becomes more appropriate as it incorporates relativistic effects[137]. In our research, we utilized the LWP since we aimed to investigate the interaction of charged particles across a broad energy spectrum, ranging from 0.1 to 600 MeV. The LWP can be expressed as follows:

$$\phi_{proj}(r,t) = \frac{q_{proj}}{r\,\gamma} \qquad (1.37)$$

In this equation, q_{proj} is the charge of the projectile, while r is the distance vector between an electron and the particle ($r = \boldsymbol{r} - \boldsymbol{R}_{proj}$). $\gamma = \left(1 - v_{proj}^2 sin^2\theta/c^2\right)^{-1/2}$ stands for the angle-dependent Lorentz factor, with v_{proj} and c being the velocity of the projectile and the speed of light, respectively. θ represents the angle between the electron and the propagation direction of the projectile. For projectiles moving at velocities substantially lower than the speed of light and $v_{proj}^2 \to c^2$, $\gamma \to 1$ if $\theta \to 0$, the LWP reduces to a standard Coulomb potential. However, if $v_{proj}^2 \to c^2$ and $\theta \to \pi/2$, γ tends towards infinity.

The total energy of a molecule/projectile system now reads:

$$E^{total}(t) = E_{kin}^{proj}(t) + E^{ADFT}(t) \tag{1.38}$$

E_{kin}^{proj} is the kinetic energy of the projectile. In our simulations, the projectile's kinetic energy is constant and the atomic nuclei are frozen. Therefore, the first term in the equation is directly removed from our calculations. The energy changed by perturbation is solely electronic, referred to as electronic energy deposition. E^{ADFT} is the sum of contributions from E^{mol} and $E^{mol/proj}$. E^{mol} is molecular energy, which reads as:

$$E^{mol}(t) = E_{core}(t) + E_{ee}(t) + E_{XC}(t) + E_{NR} \tag{1.39}$$

where E_{NR} is the atom nuclei repulsion energy.

The electronic energy deposition (ΔE) can be calculated in two ways: $\Delta E = \int_{t_0}^{t} \frac{\partial\left(E^{mol}+E^{mol/prj}\right)}{\partial \mathrm{t}}\, \partial \mathrm{t}$ or $\Delta E = \int_{t_0}^{t} \frac{\partial E^{mol}}{\partial \mathrm{t}}\, \partial \mathrm{t}$. These two methods coincide for finite systems and $E^{mol/prj}$ vanishes when the projectile is far from the molecular system of interest. This means that the two methods yield the same results for a complete collision. In some calculations, we are interested in the evolution of energy over time. $E^{mol/prj}$ includes the interaction of the charged particle with nuclei and electrons. As a result, the contribution from the nuclei appears in the calculation. Therefore, in all calculations concerning the electronic energy deposition of charged particles, we use ΔE^{mol}.

c. Exchange correlation

In addition to the standard error in the XC functionals mentioned in the previous sub-section, the adiabatic approximation in TDDFT presents another issue: it assumes that the XC potential

responds instantaneously to changes in the electron density. The adiabatic approximation has been included in all practical RT-TD-DFT simulations of real systems to date. The memory of the XC potential can impact the electron dynamics induced by strong perturbation, particularly for long-range charge transfer[138], and also in describing multiple electron/hole transition in attosecond pulse[139]. These examples have been demonstrated for the Hubbard model and small molecular systems; however, for larger molecular systems, it is not clear. Work continues to address this issue, but the path seems long and arduous[140–142]

In the case of simulation of IoR using RT-TD-DFT, it is expected that these issues need to be more effectively incorporated because the molecules will be partially ionized or excited during the simulation. The presence of radicals and fractional electrons, in particular, is likely to increase SIE and DDC. Moreover, the asymptotic behavior could influence the processes of ionization and excitation, particularly when the long-distance correlation between electrons becomes more important in these processes.

Despite these issues, on the positive side, XC-functionals are known to have a small impact on energy deposition and stopping power. Kanai and coworkers [143] investigated the stopping power of protons in silicon using RT-TD-DFT with both PBE and PBE0. They found that both functionals reproduced similar stopping curves, with minor differences at low velocities. The same result was found by Maliyov[72] in his thesis on the stopping power of protons in lithium and aluminum, by comparing the LDA functional with PBE0. This could be related to the nature of energy transfer from charged particles, which occurs very rapidly, in just a few attoseconds. During this period, the electrons do not move significantly, especially at high velocities. However, the conditions change for an electromagnetic pulse, as the interaction occurs over a longer time, depending on the pulse duration.

When considering charge migration, the influence of the exchange-correlation (XC) functional and memory could be quite significant. However, there are also encouraging findings. For example, the recent work by Chakraborty and colleagues[144]. They investigated the photoexcitation of C_{60} using non-adiabatic molecular simulation combined with TDDFT and semi-classical surface hopping approach, employing various XC-functionals such as PBE, PBE0, and B3LYP. In their study, they analyzed the dynamics of charge transfer and charge recombination in C_{60} under different

excitation conditions. Even though there are fundamental differences among the XC-functionals in long-term simulations, it seems that all functionals provide quite consistent results in the short term (approximately 20 fs), which is the timeframe of interest for us.

DFT is the principal method that we used in this work. Ground-state DFT was employed to optimize the geometry of the structures, and radiation simulations were conducted with real-time time-dependent auxiliary density functional theory (RT-TD-ADFT). Because the systems we want to study are relatively large, we had to choose a DFT functional that balances accuracy and computational cost. Therefore, we used the PBE[107] GGA-functional developed by Perdew, Burke, and Ernzerhof.

d. Gaussian basis and Complex absorbing potential

Most *ab initio* methods, which use local basis sets, are primarily designed to deal with bound state transitions. However, when a molecule is subjected to a strong perturbation, transitions near and above ionization become accessible, leading to several issues related to the description of continuum states and the presence of non-bound electrons (NBEs). For instance, high transitions are characterized by resonance states that decay rapidly and result in the auto-ionization of the molecule[145]. The presence of NBEs (in continuum states) can affect this process. Additionally, due to the use of finite basis sets unlike non-local basis set, continuum states become discrete[146], leading to artificially high ionization transitions. These transitions create spurious absorption at high energy spectrum, possibly resulting in unrealistic auto-ionization. Moreover, in the presence of a laser field, the NBEs may be reflected back to the ground state during the simulation as seen in HHG spectrum[81]. This occurs due to the phase shift between the dipole moment of the ionized molecule and the propagation of laser field, which creates a mix of destructive and constructive interference. Additionally, the presence of non-bound electrons (NBEs) introduces technical issues when one attempts to define atomic charges and analyze charge flows in the simulation outputs.

Consequently, an inadequate description of the continuum and the presence of NBEs in the simulation can influence the energy deposition, the probability and characteristics of ionization, as well as charge migration.

I. Gaussian basis

The wave functions of electrons in the continuum, characterized by their delocalized and oscillatory nature, are challenging to capture accurately using standard basis sets. Yet, non-local basis sets, like grid-basis sets and plane waves, manage to describe the continuum and valence of electrons well. In an effort to replicate non-local basis sets and enhance Gaussian basis sets for better representation of high-lying excited and continuum states, a number of strategies have been proposed in the literature. Here, we will mention a few of them.

Schlegel and colleagues[80] optimized diffuse functions for standard basis sets to improve the continuum and spatial contribution of electrons. Their findings demonstrated good agreement with grid-basis sets for the ionization of H_2^+ under a strong pulse.

Another approach was proposed by Shore[147], who combined Gaussian basis set functions, which are efficient at describing bound electrons, with B-splines to describe electrons in the continuum. B-splines have proven to be highly efficient in describing multiphoton ionization processes in atoms and molecules. Zapata and colleagues[148] illustrated that B-splines and grid-basis sets produce similar results for H_2^+ under a strong field.

Another strategy to improve the depiction of Rydberg and continuum states involves augmenting the standard basis set with Gaussian atomic orbitals fitted for the continuum. Various procedures have been proposed to accomplish this, including those suggested by Kaufmann and colleagues[149], and Nestmann and colleagues[150]. Nestmann and Peyerimhoff, with further extension by Faure and coworkers[151], proposed a method to derive Gaussians optimized for the continuum (AOC) within the context of electron-molecule scattering. These Gaussian-type orbitals have been deployed, for instance, in calculating cross-sections for electron or positron molecule collisions via the "R-matrix" approach[152,153]. The aforementioned procedure is implemented in the software GTOBAS, which empowers users to generate appropriate AOC for any atom of their interest. Furthermore, users can optimize the AOC based on their system and the physical process of interest by defining parameters such as the maximum angular momentum, the upper bound energy, and the finite range of the wave in space.

Improving the continuum representation in local basis sets inevitably increases computational costs, making it impractical for larger systems. we have demonstrated that the addition of AOC optimized by Nestmann procedure[154], can significantly improve the calculation of stopping power using RT-TD-ADFT, with less computational cost compared to sophisticated standard basis sets. We will further explore the addition of AOC to standard basis sets in Chapter five, specifically in the context of interactions with XUV-pulses. Additionally, we will discuss the enhancement of standard basis sets and their reliability when the system size is increased for this type of interaction.

II. Complex absorbing potential

As previously noted, the release of NBEs is crucial in real-time simulations of IoR. The most commonly adopted method to accomplish this involves the introduction of a complex absorbing potential (CAP) ($\Gamma(r)$) into the imaginary part of the Kohn-Sham potential. This transforms the Hamiltonian into a non-Hermitian Hamiltonian.

$$H(r,t) = H_0(r,t) - i\Gamma(r) \tag{1.40}$$

In deMon2k, two types of CAP are available[155]. The first type defines the potential according to distance criteria in real space, termed as "spatial CAP"[145]. The second type defines it according to the energies of the KS molecular orbitals (MOs) that describe NBEs.

The real-space CAP absorbs electrons that reach a threshold distance from the center of an atom (γ_a^{space}). The spatial CAP for molecular systems ($\Gamma^{space}(r)$) is achieved by superimposing atomic center absorbing potentials, effectively creating a cavity around the molecule.

$$\Gamma^{space}(r) = min_a \sum_{a=1}^{Natoms} \gamma_a^{space}(r) \tag{1.41}$$

$$\gamma_a^{space}(r) = \begin{cases} 0, & R \leq R_\circ, \\ V^{max} sin^2\left[\frac{\pi}{2W}(R - R_\circ\right], & R_\circ < R < R_\circ + W, \\ V^{max}, & R \geq R_\circ + W. \end{cases} \tag{1.42}$$

The strength of the CAP smoothly increases with the distance (R) of electrons from the atom and the width (W) of the CAP, reaching a maximum potential (V^{max}). In practice, the outcome is less

sensitive to V^{max} and W as compared to the threshold distance from the molecule $R_\circ$. The value of $R_\circ$ should be large enough to prevent interaction with ground state electrons or low-lying excited states, but close enough to effectively remove NBEs, implying it must be situated outside the Coulomb well. Therefore, the parameters $R_\circ$, W, and V^{max} need to be carefully optimized. For instance, one could run a RT-TD-ADFT simulation from the ground state without perturbation, and then check for conservation of energy and the number of electrons. If the energy and the number of electrons are not conserved in the system, it signifies that the CAP is too close, absorbing not only NBEs, but also bonded electrons.

Utilizing spatial CAPs with localized basis sets can indeed be challenging. First and foremost, the need for basis functions to represent electrons at large distances requires the use of extremely diffuse functions[80]. This in turn increases the computational cost. Secondly, spatial CAPs do not differentiate between unbound electrons and bound electrons in extremely diffuse states, such as Rydberg states. Consequently, a spatial CAP may not only absorb NBEs.

An alternative approach is to define the CAP based on another characteristic of NBEs, specifically their energies. One could choose to absorb electrons populating Molecular Orbitals (MOs) of high energy as these starts to populate during RT-TD-ADFT simulations. Lopata and colleagues[156] suggested applying the imaginary potential directly to the Fock matrix in the MO. The CAP energy (Γ^{ener}) is obtained by projecting a diagonal damping matrix (Γ) on the metric of orthonormal MO coefficients (C') at the onset of the simulation.

$$\Gamma^{ener} = C'(t)\Gamma C'^{*}(t) \tag{1.43}$$

$$\gamma_i = \begin{cases} 0, & \tilde{\varepsilon}_i \leq 0, \\ \gamma_\circ[\exp(\xi\tilde{\varepsilon}_i) - 1], & \tilde{\varepsilon}_i > 0. \end{cases} \tag{1.44}$$

The diagonal damping matrix is comprised of a damping parameter γ_i for the i^{th} MOs, which increases exponentially with MO energies. $\tilde{\varepsilon}_i = \varepsilon_i - \varepsilon_O$ represents the energies of the i MO shifted by the value of the vacuum energy cutoff (ε_O). $\gamma_\circ$ sets the energy scale while ξ specifies the rate at which electrons populating state i will be absorbed. Consequently, the lifetime of electrons in each MO equals $1/2\gamma_i$. In practice, due to errors in the exchange-correlation

function, ε_O is not exactly zero, so it can be approximated by using electron affinities, as described by Lopata and coworkers. In the case of weak perturbations, the CAP in energy domain may be evaluated just once from the SCF solution. However, for stronger perturbations, which are associated with substantial electron density perturbation, the electronic spectrum ($\tilde{\varepsilon}_i$) changes and one should re-evaluate the CAP on-the-fly. However, this is an option made recently available in deMon2k but that we have not used in this work[157].

The main advantage of an energy-space CAP is that it does not rely on the spatial distribution of electrons. Therefore, it can efficiently operate even with non-diffuse basis sets. This feature allows for an increase in the size of the simulation. Additionally, it effectively distinguishes non-bound electrons.

III. Test of Complex absorbing potential

To evaluate the efficiency of these two types of CAP, we conducted RT-TD-ADFT simulations using both CAPs for a di-nitrogen molecule (N_2). The geometry of the nitrogen molecule was aligned along the z-axis, with a bond distance of 1.090 Å. We irradiated N_2 molecule with an XUV-pulse and swift proton. This test was performed for a review article that was published in *The European Physical Journal Special Topics*.[2]

For the XUV irradiation, we used a squared cosine-shaped pulse along the z-axis. The maximum electric field strength and the energy of the XUV pulse are set at 0.005 (Ha/e.bohr) (equivalent to 3.5 x 10^{12} W/cm^2) and 30 eV respectively. The total duration of the pulse is 30 fs and it is centered at 15 fs. The simulation is performed using the PBE XC functional[107], and it ran for 50 fs with a time step of 1 attosecond. A diffuse basis set, constructed from the aug-cc-pVTZ and supplemented with 28 diffuse functions[158], is used.

For CAP in real-space, the threshold distance $R_\circ$, width W and maximum potential V^{max} are optimized at 15 Å, 5 Å and 15 Ha, respectively. For CAP in energy-space, we optimized the energy scale (γ_0) and the damping strength (ξ) to 0.2 Ha and 0.05 Ha^{-1}, respectively. The vacuum energy cut-off was approximated at 0.0318 Ha, which ensures the corrected energy of the lowest unoccupied MO equals the electron affinity, as calculated from two separate SCF calculations for the neutral and anionic N_2.

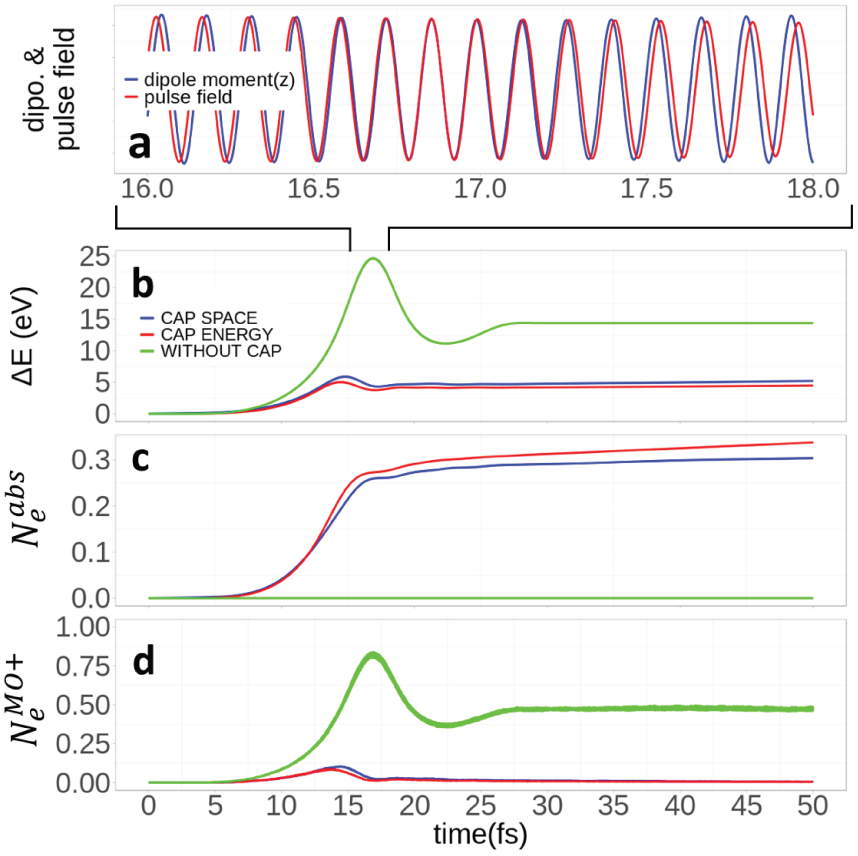

Figure 1-8: N_2 ionization by 30 eV XUV pulse (b) variation of total energy (c) number of electrons absorbed by the CAP and (d) the numbers of electrons still in positive energy KS MOs, all as function of the simulation time (fs). (a) corresponds to the propagation of the dipole moment on the z axis and the pulse field between 16 and 18 fs. In order to easily compare the amplitude of the pulsed field is scaled. This figure is reproduced with permission of Springer Nature, License Number 5595280327823.

Figure 1-8 illustrates time-dependent profiles of the total energy variation (the total energy of the molecule and the pulse) relative to the ground state energy (ΔE), the number of electrons absorbed by the CAP (N_e^{abs}), and the number of electrons in KS MOs with positive energy (N_e^{MO+}). In the simulation without a CAP (represented by the green line), ΔE peaks around 17 fs during the pulse application, then decreases and stabilizes after pulse. Panel d illustrates that the N_e^{MO+} curve mirrors the evolution of ΔE. This is caused by the varying delay in dipole moment propagation compared to the pulse field, which leads to both destructive and constructive superpositions, as demonstrated in (Panel a). This phenomenon is commonly observed in High Harmonic Generation when intense infrared laser pulses are applied in non-linear optics. The electrons in the continuum are reflected back to the molecule, releasing XUV-pulses due to the phase shift between the cycles of the laser pulse and the dipole momentum of the molecule[81]. However, in our case, this phase shift might be artificially produced due to the indefinite retention

of electrons in the continuum states. As evidence, the phase shift disappears when the CAP is introduced into the simulation (as seen in the blue and red lines). With the addition of either a spatial or energy CAP, NBEs are quickly absorbed, causing the number of electrons to no longer remain conserved (as depicted by the blue and red lines in Panel d). This also leads to a reduction in ΔE due to the decrease in the fraction of electron density within the simulation, consequently resulting in a decrease in the total energy of the system. Despite their different approaches, both types of CAP yield similar results. After a sharp increase around 15 fs, when the external electric field is at its peak intensity, a relatively steady situation is achieved. This result shows that the introduction of CAP is crucial for the accurate description of ionization and dynamics during XUV-pulse interactions, and either type of CAP effectively removes the NBEs.

A challenge worth noting, as discovered during the testing phase, is the difficulty associated with ionizing a single electron from small molecules such as N_2. As we increase the strength of the pulse, phase shifting prevents the ionization of a single electron, leading to non-linear pulse behavior. Conversely, by increasing the strength of the CAP, electron dynamics becomes unstable during the simulation, even with more sophisticated electron propagation algorithms and smaller time steps. This phenomenon is observable in both CAP models. We noticed that the instability during simulation is strongly dependent on the rate of electron absorption relative to the total number of electrons in the system. As demonstrated in Chapter Five, for instance, one can easily remove an electron from a large system interacting with an XUV-pulse without any instability in the simulation. This could be because for large systems, a weak pulse is sufficient for ionizing one electron and the molecular orbitals or the electronic spectrum ($\tilde{\varepsilon}_i$) do not change significantly compared to the ground state. This problem could be critical when investigating multiphotonic simulations and simulating non-linear optical properties for small molecular systems. However, for those operating in the linear region (one photon absorption) or interested in the probability of ionization, this issue is not essential. Indeed, DFT calculations provide the probability of ionization or absorption of photons. Hence, ionizing one electron is not necessary. In such cases, a weak field is sufficient.

To evaluate the effects of CAP in the simulation of swift charged particles, we exposed the N_2 molecule to a 70 KeV proton. The same methodologies and parameters as the XUV section are applied, with the exception of the CAP strength, which is reduced to ensure the stability of the

dynamics. Specifically, the threshold distance $R_\circ$ of the spatial CAP is reduced from 15.0 to 12.5 Å, and the energy scale γ_0 of the energy CAP is decreased from 0.25 to 0.05 Ha. Initially, the proton is positioned 30 Å away from the N_2 and it impacted the bond between the nitrogen atoms after 816 as.

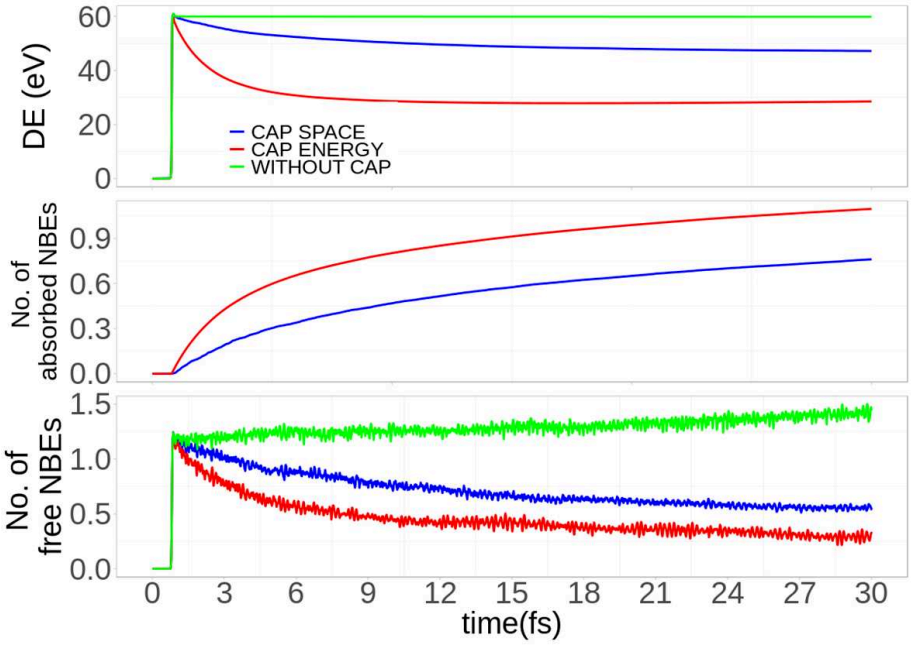

Figure 1-9: illustrates the results of the interaction of a 70 KeV proton with nitrogen. The panels and color codes are similar to those in Figure 1-8, with the exception of the energy variation, which in this case corresponds to energy deposition (variation of molecular energy only). This figure is reproduced with permission from Springer Nature, License Number 5595280327823.

When a proton interacts with the electron cloud of the system over a few attoseconds, it deposits approximately 60 eV of its energy, as seen in Figure 1-9. This energy deposition triggers multiple transitions and ionization channels. In the absence of a CAP, the deposited energy is completely conserved after the collision as expected. However, we observe a gradual increase in the number of NBEs from 1.25 to 1.5 electrons due to various internal electronic transitions occurring within the electron cloud, such as auto-ionization processes. Here, auto-ionization arises not just from highly excited state resonances but also from artificial high ionization transitions. Introducing a CAP reduces the deposited energy as it absorbs the NBEs, which leads to a decrease in the total energy. As shown in Figure 1-9 (panel c), when a large number of electrons are excited in a very short time, the CAP requires more time to absorb these electrons. Unfortunately, the rate of electron absorption impacts the stability of the simulation, thus limiting the strength of the CAP.

Nevertheless, NBEs in higher molecular orbitals have shorter lifetimes than those in lower levels, which helps reduce the risk of high lying resonance states (which decay faster).

In conclusion, while a CAP is not essential for calculating the energy deposited by charged particles due to their short interaction time compared to a laser pulse, it is crucial for simulating ultrafast processes' dynamics.

1.4 Conclusion

We have introduced in this Chapter the general background and the theoretical concepts used in this thesis. We began by outlining the main elements related to radiation damage for both photons and charged particles. In particular, we focused on the physical stage in terms of energy deposition, initial ionization, penetration distance, and electronic relaxation induced by both types of ionizing radiation. Additionally, we discussed the common experimental and theoretical approaches used to reveal the physical stage.

In the second part, we introduced the main mathematical equations involved in RT-TD-ADFT, a workhorse in our work. We introduced local basis sets and auxiliary density fitting approximations. We also described the kinds of perturbations used to explicitly simulate the interaction with ionizing radiation. We discussed the effects of XC functionals, basis sets, and complex absorbing potentials for removing non-bonded electrons in simulations. Despite several issues arising due to XC functional and basis sets, RT-TD-ADFT remains reliable for revealing the physical stage and is unique in its ability to study large molecular systems. We demonstrated that removing NBEs in the simulation is essential for the interaction of XUV-pulses and electron relaxation, but it does not have a significant effect on energy deposition by charged particles. The following Chapters will focus on results.

REFRENCES:

(1) Omar, K. A.; Hasnaoui, K.; De La Lande, A. First-Principles Simulations of Biological Molecules Subjected to Ionizing Radiation. *Annu. Rev. Phys. Chem.* **2021**, *72* (1), 445–465. https://doi.org/10.1146/annurev-physchem-101419-013639.

(2) Omar, K. A.; Korsaye, F. A.; Tandiana, R.; Tolu, D.; Deviers, J.; Wu, X.; Parise, A.; Alvarez-Ibarra, A.; Moncada, F.; Pedroza-Montero, J. N.; Mejía-Rodriguez, D.; Van-Oanh, N.-T.; Cailliez, F.; Clavaguéra, C.; Hasnaoui, K.; De La Lande, A. Current Status of DeMon2k for the Investigation of the Early Stages of Matter Irradiation by Time-Dependent DFT Approaches. *Eur. Phys. J. Spec. Top.* **2023**. https://doi.org/10.1140/epjs/s11734-023-00905-6.

(3) Spitz, D. R.; Azzam, E. I.; Jian Li, J.; Gius, D. Metabolic Oxidation/Reduction Reactions and Cellular Responses to Ionizing Radiation: A Unifying Concept in Stress Response Biology. *Cancer and Metastasis Reviews* **2004**, *23* (3), 311–322. https://doi.org/10.1023/B:CANC.0000031769.14728.bc.

(4) Lomax, M. E.; Folkes, L. K.; O'Neill, P. Biological Consequences of Radiation-Induced DNA Damage: Relevance to Radiotherapy. *Clinical Oncology* **2013**, *25* (10), 578–585. https://doi.org/10.1016/j.clon.2013.06.007.

(5) Platzman, RL. The Physical and Chemical Basis of Mechanisms in Radiation Biology. *Radiation Biology and Medicine. Selected Reviews in the Life Sciences* **1958**, 15–72.

(6) Azzam, E. I.; Jay-Gerin, J.-P.; Pain, D. Ionizing Radiation-Induced Metabolic Oxidative Stress and Prolonged Cell Injury. *Cancer Letters* **2012**, *327* (1–2), 48–60. https://doi.org/10.1016/j.canlet.2011.12.012.

(7) Sansone, G.; Pfeifer, T.; Simeonidis, K.; Kuleff, A. I. Electron Correlation in Real Time. *ChemPhysChem* **2012**, *13* (3), 661–680. https://doi.org/10.1002/cphc.201100528.

(8) Howell, R. W. Auger Processes in the 21st Century. *International Journal of Radiation Biology* **2008**, *84* (12), 959–975. https://doi.org/10.1080/09553000802395527.

(9) Savolainen, J.; Uhlig, F.; Ahmed, S.; Hamm, P.; Jungwirth, P. Direct Observation of the Collapse of the Delocalized Excess Electron in Water. *Nature Chem* **2014**, *6* (8), 697–701. https://doi.org/10.1038/nchem.1995.

(10) Alizadeh, E.; Orlando, T. M.; Sanche, L. Biomolecular Damage Induced by Ionizing Radiation: The Direct and Indirect Effects of Low-Energy Electrons on DNA. *Annu. Rev. Phys. Chem.* **2015**, *66* (1), 379–398. https://doi.org/10.1146/annurev-physchem-040513-103605.

(11) Rodriguez Goyes, J. C.; Gomez, N. J.; Restrepo, V. J.; Gaviria, S.; Hincapie, A.; Arias, E. Biologic Effect of Ionizing Radiation. *Interventional-Cardiology* **2018**, *10* (5). https://doi.org/10.4172/Interventional-Cardiology.1000618.

(12) Ward, J. F. DNA Damage Produced by Ionizing Radiation in Mammalian Cells: Identities, Mechanisms of Formation, and Reparability. In *Progress in Nucleic Acid Research and Molecular Biology*; Elsevier, 1988; Vol. 35, pp 95–125. https://doi.org/10.1016/S0079-6603(08)60611-X.

(13) Gaigeot, M.-P.; Lopez-Tarifa, P.; Martin, F.; Alcami, M.; Vuilleumier, R.; Tavernelli, I.; Hervé du Penhoat, M.-A.; Politis, M.-F. Theoretical Investigation of the Ultrafast Dissociation of Ionised Biomolecules Immersed in Water: Direct and Indirect Effects. *Mutation Research/Reviews in Mutation Research* **2010**, *704* (1–3), 45–53. https://doi.org/10.1016/j.mrrev.2010.01.004.

(14) Becker, D.; Adhikary, A.; Sevilla, M. D. Mechanisms of Radiation-Induced DNA Damage: Direct Effects. In *Recent Trends in Radiation Chemistry*; WORLD SCIENTIFIC, 2010; pp 509–542. https://doi.org/10.1142/9789814282093_0018.

(15) Sonntag, C. von. Radiation-Induced DNA Damage: Indirect Effects. In *Recent Trends in Radiation Chemistry*; WORLD SCIENTIFIC, 2010; pp 543–562. https://doi.org/10.1142/9789814282093_0019.

(16) Vere, T. L.; Becker, D.; Sevilla, M. D. Yields of · OH in Gamma-Irradiated DNA as a Function of DNA Hydration: Hole Transfer in Competition with · OH Formation. *Radiation Research* **1996**, *145* (6), 673. https://doi.org/10.2307/3579357.

(17) Ma, J.; Denisov, S. A.; Adhikary, A.; Mostafavi, M. Ultrafast Processes Occurring in Radiolysis of Highly Concentrated Solutions of Nucleosides/Tides. *IJMS* **2019**, *20* (19), 4963. https://doi.org/10.3390/ijms20194963.

(18) Heslop-Harrison, J. S. (Pat); Schwarzacher, T. Nucleosomes and Centromeric DNA Packaging. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110* (50), 19974–19975. https://doi.org/10.1073/pnas.1319945110.

(19) Hubbell, J. H. Review of Photon Interaction Cross Section Data in the Medical and Biological Context. *Phys. Med. Biol.* **1999**, *44* (1), R1–R22. https://doi.org/10.1088/0031-9155/44/1/001.

(20) Grinstead, J. H.; Barker, P. F. Coherent Rayleigh Scattering. *Phys. Rev. Lett.* **2000**, *85* (6), 1222–1225. https://doi.org/10.1103/PhysRevLett.85.1222.

(21) Klassen, S. The Photoelectric Effect: Reconstructing the Story for the Physics Classroom. *Sci & Educ* **2011**, *20* (7–8), 719–731. https://doi.org/10.1007/s11191-009-9214-6.

(22) Seltzer, S. XCOM-Photon Cross Sections Database, NIST Standard Reference Database 8, 1987. https://doi.org/10.18434/T48G6X.

(23) Qiao, C.-K.; Wei, J.-W.; Chen, L. An Overview of the Compton Scattering Calculation. *Crystals* **2021**, *11* (5), 525. https://doi.org/10.3390/cryst11050525.

(24) Krajewska, K.; Kamiński, J. Z. Recoil Effects in Multiphoton Electron-Positron Pair Creation. *Phys. Rev. A* **2010**, *82* (1), 013420. https://doi.org/10.1103/PhysRevA.82.013420.

(25) Correa, A. A. Calculating Electronic Stopping Power in Materials from First Principles. *Computational Materials Science* **2018**, *150*, 291–303. https://doi.org/10.1016/j.commatsci.2018.03.064.

(26) Seltzer, S. Stopping-Powers and Range Tables for Electrons, Protons, and Helium Ions, NIST Standard Reference Database 124, 1993. https://doi.org/10.18434/T4NC7P.

(27) Groom, D. E.; Klein, S. R. Passage of Particles through Matter. *Eur. Phys. J. C* **2000**, *15* (1–4), 163–173. https://doi.org/10.1007/BF02683419.

(28) Yokoya, A.; Ito, T. Photon-Induced Auger Effect in Biological Systems: A Review. *International Journal of Radiation Biology* **2017**, *93* (8), 743–756. https://doi.org/10.1080/09553002.2017.1312670.

(29) Pirovano, G.; Wilson, T. C.; Reiner, T. Auger: The Future of Precision Medicine. *Nuclear Medicine and Biology* **2021**, *96–97*, 50–53. https://doi.org/10.1016/j.nucmedbio.2021.03.002.

(30) Fourie, H.; Nair, S.; Miles, X.; Rossouw, D.; Beukes, P.; Newman, R. T.; Zeevaart, J. R.; Vandevoorde, C.; Slabbert, J. Estimating the Relative Biological Effectiveness of Auger Electron Emitter 123I in Human Lymphocytes. *Front. Phys.* **2020**, *8*, 567732. https://doi.org/10.3389/fphy.2020.567732.

(31) Ku, A.; Facca, V. J.; Cai, Z.; Reilly, R. M. Auger Electrons for Cancer Therapy – a Review. *EJNMMI radiopharm. chem.* **2019**, *4* (1), 27. https://doi.org/10.1186/s41181-019-0075-2.

(32) Codling, K.; Madden, R. P.; Ederer, D. L. Resonances in the Photo-Ionization Continuum of Ne I (20-150 EV). *Phys. Rev.* **1967**, *155* (1), 26–37. https://doi.org/10.1103/PhysRev.155.26.

(33) Miroshnichenko, A. E.; Flach, S.; Kivshar, Y. S. Fano Resonances in Nanoscale Structures. *Rev. Mod. Phys.* **2010**, *82* (3), 2257–2298. https://doi.org/10.1103/RevModPhys.82.2257.

(34) Baig, M. A. Measurement of Photoionization Cross-Section for the Excited States of Atoms: A Review. *Atoms* **2022**, *10* (2), 39. https://doi.org/10.3390/atoms10020039.

(35) Jahnke, T.; Hergenhahn, U.; Winter, B.; Dörner, R.; Frühling, U.; Demekhin, P. V.; Gokhberg, K.; Cederbaum, L. S.; Ehresmann, A.; Knie, A.; Dreuw, A. Interatomic and Intermolecular Coulombic Decay. *Chem. Rev.* **2020**, *120* (20), 11295–11369. https://doi.org/10.1021/acs.chemrev.0c00106.

(36) Harbach, P. H. P.; Schneider, M.; Faraji, S.; Dreuw, A. Intermolecular Coulombic Decay in Biology: The Initial Electron Detachment from $FADH^-$ in DNA Photolyases. *J. Phys. Chem. Lett.* **2013**, *4* (6), 943–949. https://doi.org/10.1021/jz400104h.

(37) Stoychev, S. D.; Kuleff, A. I.; Cederbaum, L. S. Intermolecular Coulombic Decay in Small Biochemically Relevant Hydrogen-Bonded Systems. *J. Am. Chem. Soc.* **2011**, *133* (17), 6817–6824. https://doi.org/10.1021/ja200963y.

(38) Calegari, F.; Trabattoni, A.; Palacios, A.; Ayuso, D.; Castrovilli, M. C.; Greenwood, J. B.; Decleva, P.; Martín, F.; Nisoli, M. Charge Migration Induced by Attosecond Pulses in Bio-Relevant Molecules. *J. Phys. B: At. Mol. Opt. Phys.* **2016**, *49* (14), 142001. https://doi.org/10.1088/0953-4075/49/14/142001.

(39) Wörner, H. J.; Arrell, C. A.; Banerji, N.; Cannizzo, A.; Chergui, M.; Das, A. K.; Hamm, P.; Keller, U.; Kraus, P. M.; Liberatore, E.; Lopez-Tarifa, P.; Lucchini, M.; Meuwly, M.; Milne, C.; Moser, J.-E.; Rothlisberger, U.; Smolentsev, G.; Teuscher, J.; Van Bokhoven, J. A.; Wenger, O. Charge Migration and Charge Transfer in Molecular Systems. *Structural Dynamics* **2017**, *4* (6), 061508. https://doi.org/10.1063/1.4996505.

(40) Kraus, P. M.; Mignolet, B.; Baykusheva, D.; Rupenyan, A.; Horný, L.; Penka, E. F.; Grassi, G.; Tolstikhin, O. I.; Schneider, J.; Jensen, F.; Madsen, L. B.; Bandrauk, A. D.; Remacle, F.; Wörner, H. J. Measurement and Laser Control of Attosecond Charge Migration in Ionized Iodoacetylene. *Science* **2015**, *350* (6262), 790–795. https://doi.org/10.1126/science.aab2160.

(41) Hennig, H.; Breidbach, J.; Cederbaum, L. S. Charge Transfer Driven by Electron Correlation: A Non-Dyson Propagator Approach. *The Journal of Chemical Physics* **2005**, *122* (13), 134104. https://doi.org/10.1063/1.1869473.

(42) Breidbach, J.; Cederbaum, L. S. Migration of Holes: Formalism, Mechanisms, and Illustrative Applications. *The Journal of Chemical Physics* **2003**, *118* (9), 3983–3996. https://doi.org/10.1063/1.1540618.

(43) Weinkauf, R.; Schanen, P.; Metsala, A.; Schlag, E. W.; Bürgle, M.; Kessler, H. Highly Efficient Charge Transfer in Peptide Cations in the Gas Phase: Threshold Effects and Mechanism. *J. Phys. Chem.* **1996**, *100* (47), 18567–18585. https://doi.org/10.1021/jp960926m.

(44) Cederbaum, L. S.; Zobeley, J. Ultrafast Charge Migration by Electron Correlation. *Chemical Physics Letters* **1999**, *307* (3–4), 205–210. https://doi.org/10.1016/S0009-2614(99)00508-4.

(45) Calegari, F.; Ayuso, D.; Trabattoni, A.; Belshaw, L.; De Camillis, S.; Anumula, S.; Frassetto, F.; Poletto, L.; Palacios, A.; Decleva, P.; Greenwood, J. B.; Martín, F.; Nisoli, M. Ultrafast Electron Dynamics in Phenylalanine Initiated by Attosecond Pulses. *Science* **2014**, *346* (6207), 336–339. https://doi.org/10.1126/science.1254061.

(46) Ikeura-Sekiguchi, H.; Sekiguchi, T. Attosecond Electron Delocalization in the Conduction Band through the Phosphate Backbone of Genomic DNA. *Phys. Rev. Lett.* **2007**, *99* (22), 228102. https://doi.org/10.1103/PhysRevLett.99.228102.

(47) Shimizu, M.; Hayakawa, T.; Kaneda, M.; Tsuchida, H.; Itoh, A. Stopping Cross-Sections of Liquid Water for 0.3–2.0 MeV Protons. *Vacuum* **2010**, *84* (8), 1002–1004. https://doi.org/10.1016/j.vacuum.2009.11.019.

(48) Siiskonen, T.; Kettunen, H.; Peräjärvi, K.; Javanainen, A.; Rossi, M.; Trzaska, W. H.; Turunen, J.; Virtanen, A. Energy Loss Measurement of Protons in Liquid Water. *Physics in Medicine & Biology* **2011**, *56* (8), 2367. https://doi.org/10.1088/0031-9155/56/8/003.

(49) Carrier, J.-F.; Archambault, L.; Beaulieu, L.; Roy, R. Validation of GEANT4 , an Object-Oriented Monte Carlo Toolkit, for Simulations in Medical Physics. *Med. Phys.* **2004**, *31* (3), 484–492. https://doi.org/10.1118/1.1644532.

(50) Vagena, E.; Androulakaki, E. G.; Kokkoris, M.; Patronis, N.; Stamati, M. E. A Comparative Study of Stopping Power Calculations Implemented in Monte Carlo Codes and Compilations with Experimental Data. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* **2020**, *467*, 44–52. https://doi.org/10.1016/j.nimb.2020.02.003.

(51) Friedland, W.; Dingfelder, M.; Kundrát, P.; Jacob, P. Track Structures, DNA Targets and Radiation Effects in the Biophysical Monte Carlo Simulation Code PARTRAC. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **2011**, *711* (1–2), 28–40. https://doi.org/10.1016/j.mrfmmm.2011.01.003.

(52) Ziegler, J. F.; Ziegler, M. D.; Biersack, J. P. SRIM – The Stopping and Range of Ions in Matter (2010). *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* **2010**, *268* (11), 1818–1823. https://doi.org/10.1016/j.nimb.2010.02.091.

(53) Ziegler, J. F. Comments on ICRU Report No. 49: Stopping Powers and Ranges for Protons and Alpha Particles. *Radiation Research* **1999**, *152* (2), 219. https://doi.org/10.2307/3580097.

(54) Janni, J. F. Energy Loss, Range, Path Length, Time-of-Flight, Straggling, Multiple Scattering, and Nuclear Interaction Probability: In Two Parts. Part 1. For 63 Compounds Part 2. For Elements $1 \leqslant Z \leqslant 92$. *Atomic Data and Nuclear Data Tables* **1982**, *27* (2), 147–339. https://doi.org/10.1016/0092-640X(82)90004-3.

(55) Bethe, H. Bremsformel for Elektronen relativistischer Geschwindigkeit. *Z. Physik* **1932**, *76* (5–6), 293–299. https://doi.org/10.1007/BF01342532.

(56) Bloch, F. Zur Bremsung rasch bewegter Teilchen beim Durchgang durch Materie. *Ann. Phys.* **1933**, *408* (3), 285–320. https://doi.org/10.1002/andp.19334080303.

(57) Ziegler, J. F. Stopping of Energetic Light Ions in Elemental Matter. *Journal of Applied Physics* **1999**, *85* (3), 1249–1272. https://doi.org/10.1063/1.369844.

(58) Dingfelder, M.; Inokuti, M.; Paretzke, H. G. Inelastic-Collision Cross Sections of Liquid Water for Interactions of Energetic Protons. *Radiation Physics and Chemistry* **2000**, *59* (3), 255–275. https://doi.org/10.1016/S0969-806X(00)00263-2.

(59) Barkas, W. H.; Dyer, J. N.; Heckman, H. H. Resolution of the $\ensuremath\Sigma^\ensuremath-$-Mass Anomaly. *Phys. Rev. Lett.* **1963**, *11* (1), 26–28. https://doi.org/10.1103/PhysRevLett.11.26.

(60) M.A, W. H. B.; B.Sc, R. K. XXXIX. On the α Particles of Radium, and Their Loss of Range in Passing through Various Atoms and Molecules. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **1905**, *10* (57), 318–340. https://doi.org/10.1080/14786440509463378.

(61) Lindhard, J.; Scharff, M.; Schiøtt, H. E. *Range Concepts and Heavy Ion Ranges*; Munksgaard Copenhagen, 1963; Vol. 33.

(62) Lindhard, J.; Winther, A.; others. Stopping Power of Electron Gas and Equipartition Rule. **1964**.

(63) Faussurier, G.; Blancard, C.; Cossé, P.; Renaudin, P. Equation of State, Transport Coefficients, and Stopping Power of Dense Plasmas from the Average-Atom Model Self-Consistent Approach for Astrophysical and Laboratory Plasmas. *Physics of Plasmas* **2010**, *17* (5), 052707. https://doi.org/10.1063/1.3420276.

(64) Uehara, S.; Toburen, L. H.; Nikjoo, H. Development of a Monte Carlo Track Structure Code for Low-Energy Protons in Water. *International Journal of Radiation Biology* **2001**, *77* (2), 139–154. https://doi.org/10.1080/09553000010012536.

(65) Francis, Z.; Incerti, S.; Karamitros, M.; Tran, H. N.; Villagrasa, C. Stopping Power and Ranges of Electrons, Protons and Alpha Particles in Liquid Water Using the Geant4-DNA Package. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* **2011**, *269* (20), 2307–2311. https://doi.org/10.1016/j.nimb.2011.02.031.

(66) Heredia-Avalos, S.; Garcia-Molina, R.; Fernández-Varea, J. M.; Abril, I. Calculated Energy Loss of Swift He, Li, B, and N Ions in Si O 2 , Al 2 O 3 , and Zr O 2. *Phys. Rev. A* **2005**, *72* (5), 052902. https://doi.org/10.1103/PhysRevA.72.052902.

(67) Kyriakou, I.; Sakata, D.; Tran, H. N.; Perrot, Y.; Shin, W.-G.; Lampe, N.; Zein, S.; Bordage, M. C.; Guatelli, S.; Villagrasa, C.; Emfietzoglou, D.; Incerti, S. Review of the Geant4-DNA Simulation Toolkit for Radiobiological Applications at the Cellular and DNA Level. *Cancers* **2021**, *14* (1), 35. https://doi.org/10.3390/cancers14010035.

(68) Plante, I. A Review of Simulation Codes and Approaches for Radiation Chemistry. *Phys. Med. Biol.* **2021**, *66* (3), 03TR02. https://doi.org/10.1088/1361-6560/abbd19.

(69) Francis, Z.; El Bitar, Z.; Incerti, S.; Bernal, M. A.; Karamitros, M.; Tran, H. N. Calculation of Lineal Energies for Water and DNA Bases Using the Rudd Model Cross Sections Integrated within the Geant4-DNA Processes. *Journal of Applied Physics* **2017**, *122* (1), 014701. https://doi.org/10.1063/1.4990293.

(70) Yost, D. C.; Kanai, Y. Electronic Excitation Dynamics in DNA under Proton and α-Particle Irradiation. *J. Am. Chem. Soc.* **2019**, *141* (13), 5241–5251. https://doi.org/10.1021/jacs.8b12148.

(71) Alvarez-Ibarra, A.; Parise, A.; Hasnaoui, K.; de la Lande, A. The Physical Stage of Radiolysis of Solvated DNA by High-Energy-Transfer Particles: Insights from New First Principles Simulations. *Phys. Chem. Chem. Phys.* **2020**, *22* (15), 7747–7758. https://doi.org/10.1039/D0CP00165A.

(72) Maliyov, I. Ionic Irradiation of Materials : Real-Time Dynamics of Electronic Excitations. Theses, Université Paris-Saclay, 2019. https://theses.hal.science/tel-02495699.

(73) Maliyov, I.; Crocombette, J.-P.; Bruneval, F. Electronic Stopping Power from Time-Dependent Density-Functional Theory in Gaussian Basis. *Eur. Phys. J. B* **2018**, *91* (8), 172. https://doi.org/10.1140/epjb/e2018-90289-y.

(74) Schleife, A.; Kanai, Y.; Correa, A. A. Accurate Atomistic First-Principles Calculations of Electronic Stopping. *Phys. Rev. B* **2015**, *91* (1), 014306. https://doi.org/10.1103/PhysRevB.91.014306.

(75) Gu, B.; Cunningham, B.; Muñoz Santiburcio, D.; Da Pieve, F.; Artacho, E.; Kohanoff, J. Efficient *Ab Initio* Calculation of Electronic Stopping in Disordered Systems via Geometry Pre-Sampling: Application to Liquid Water. *J. Chem. Phys.* **2020**, *153* (3), 034113. https://doi.org/10.1063/5.0014276.

(76) Palacios, A.; Martín, F. The Quantum Chemistry of Attosecond Molecular Science. *WIREs Comput Mol Sci* **2020**, *10* (1). https://doi.org/10.1002/wcms.1430.

(77) Nisoli, M.; Decleva, P.; Calegari, F.; Palacios, A.; Martín, F. Attosecond Electron Dynamics in Molecules. *Chem. Rev.* **2017**, *117* (16), 10760–10825. https://doi.org/10.1021/acs.chemrev.6b00453.

(78) Sato, S. A. First-Principles Calculations for Attosecond Electron Dynamics in Solids. *Computational Materials Science* **2021**, *194*, 110274. https://doi.org/10.1016/j.commatsci.2020.110274.

(79) Belloni, J.; Monard, H.; Gobert, F.; Larbre, J.-P.; Demarque, A.; De Waele, V.; Lampre, I.; Marignier, J.-L.; Mostafavi, M.; Bourdon, J. C.; Bernard, M.; Borie, H.; Garvey, T.; Jacquemard, B.; Leblond, B.; Lepercq, P.; Omeich, M.; Roch, M.; Rodier, J.; Roux, R. ELYSE—A Picosecond Electron Accelerator for Pulse Radiolysis Research. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **2005**, *539* (3), 527–539. https://doi.org/10.1016/j.nima.2004.11.006.

(80) Krause, P.; Sonk, J. A.; Schlegel, H. B. Strong Field Ionization Rates Simulated with Time-Dependent Configuration Interaction and an Absorbing Potential. *The Journal of Chemical Physics* **2014**, *140* (17), 174113. https://doi.org/10.1063/1.4874156.

(81) Corkum, P. B. Plasma Perspective on Strong Field Multiphoton Ionization. *Phys. Rev. Lett.* **1993**, *71* (13), 1994–1997. https://doi.org/10.1103/PhysRevLett.71.1994.

(82) Hervé, M.; Boyer, A.; Brédy, R.; Allouche, A.-R.; Compagnon, I.; Lépine, F. On-the-Fly Investigation of XUV Excited Large Molecular Ions Using a High Harmonic Generation Light Source. *Sci Rep* **2022**, *12* (1), 13191. https://doi.org/10.1038/s41598-022-17416-4.

(83) Pickup, B. T. On the Theory of Fast Photoionization Processes. *Chemical Physics* **1977**, *19* (2), 193–208. https://doi.org/10.1016/0301-0104(77)85131-8.

(84) Hennig, H.; Breidbach, J.; Cederbaum, L. S. Electron Correlation as the Driving Force for Charge Transfer: Charge Migration Following Ionization in *N* -Methyl Acetamide. *J. Phys. Chem. A* **2005**, *109* (3), 409–414. https://doi.org/10.1021/jp046232s.

(85) Lara-Astiaso, M.; Ayuso, D.; Tavernelli, I.; Decleva, P.; Palacios, A.; Martín, F. Decoherence, Control and Attosecond Probing of XUV-Induced Charge Migration in Biomolecules. A Theoretical Outlook. *Faraday Discuss.* **2016**, *194*, 41–59. https://doi.org/10.1039/C6FD00074F.

(86) Krause, P.; Klamroth, T.; Saalfrank, P. Time-Dependent Configuration-Interaction Calculations of Laser-Pulse-Driven Many-Electron Dynamics: Controlled Dipole Switching in Lithium Cyanide. *The Journal of Chemical Physics* **2005**, *123* (7), 074105. https://doi.org/10.1063/1.1999636.

(87) Huber, C.; Klamroth, T. Explicitly Time-Dependent Coupled Cluster Singles Doubles Calculations of Laser-Driven Many-Electron Dynamics. *The Journal of Chemical Physics* **2011**, *134* (5), 054113. https://doi.org/10.1063/1.3530807.

(88) Bruner, A.; Hernandez, S.; Mauger, F.; Abanador, P. M.; LaMaster, D. J.; Gaarde, M. B.; Schafer, K. J.; Lopata, K. Attosecond Charge Migration with TDDFT: Accurate Dynamics from a Well-Defined Initial State. *J. Phys. Chem. Lett.* **2017**, *8* (17), 3991–3996. https://doi.org/10.1021/acs.jpclett.7b01652.

(89) Thomas, L. H. The Calculation of Atomic Fields. *Math. Proc. Camb. Phil. Soc.* **1927**, *23* (5), 542–548. https://doi.org/10.1017/S0305004100011683.

(90) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136* (3B), B864–B871. https://doi.org/10.1103/PhysRev.136.B864.

(91) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140* (4A), A1133–A1138. https://doi.org/10.1103/PhysRev.140.A1133.

(92) Casida, M. E.; Salahub, D. R. Asymptotic Correction Approach to Improving Approximate Exchange–Correlation Potentials: Time-Dependent Density-Functional Theory Calculations of Molecular Excitation Spectra. *The Journal of Chemical Physics* **2000**, *113* (20), 8918–8935. https://doi.org/10.1063/1.1319649.

(93) Li, M.; Kobayashi, R.; Amos, R. D.; Ford, M. J.; Reimers, J. R. Density Functionals with Asymptotic-Potential Corrections Are Required for the Simulation of Spectroscopic Properties of Materials. *Chem. Sci.* **2022**, *13* (5), 1492–1503. https://doi.org/10.1039/D1SC03738B.

(94) Carmona-Espíndola, J.; Gázquez, J. L.; Vela, A.; Trickey, S. B. Generalized Gradient Approximation Exchange Energy Functional with Correct Asymptotic Behavior of the Corresponding Potential. *J. Chem. Phys.* **2015**, *142* (5), 054105. https://doi.org/10.1063/1.4906606.

(95) Grimme, S. Density Functional Theory with London Dispersion Corrections. *WIREs Comput Mol Sci* **2011**, *1* (2), 211–228. https://doi.org/10.1002/wcms.30.

(96) Perdew, J. P.; Parr, R. G.; Levy, M.; Balduz, J. L. Density-Functional Theory for Fractional Particle Number: Derivative Discontinuities of the Energy. *Phys. Rev. Lett.* **1982**, *49* (23), 1691–1694. https://doi.org/10.1103/PhysRevLett.49.1691.

(97) Mori-Sánchez, P.; Cohen, A. J. The Derivative Discontinuity of the Exchange–Correlation Functional. *Phys. Chem. Chem. Phys.* **2014**, *16* (28), 14378–14387. https://doi.org/10.1039/C4CP01170H.

(98) Zhan, C.-G.; Nichols, J. A.; Dixon, D. A. Ionization Potential, Electron Affinity, Electronegativity, Hardness, and Electron Excitation Energy: Molecular Properties from Density Functional Theory Orbital Energies. *J. Phys. Chem. A* **2003**, *107* (20), 4184–4195. https://doi.org/10.1021/jp0225774.

(99) Yang, W.; Cohen, A. J.; Mori-Sánchez, P. Derivative Discontinuity, Bandgap and Lowest Unoccupied Molecular Orbital in Density Functional Theory. *The Journal of Chemical Physics* **2012**, *136* (20), 204111. https://doi.org/10.1063/1.3702391.

(100) Zhang, Y.; Yang, W. A Challenge for Density Functionals: Self-Interaction Error Increases for Systems with a Noninteger Number of Electrons. *The Journal of Chemical Physics* **1998**, *109* (7), 2604–2608. https://doi.org/10.1063/1.476859.

(101) Dutoi, A. D.; Head-Gordon, M. Self-Interaction Error of Local Density Functionals for Alkali–Halide Dissociation. *Chemical Physics Letters* **2006**, *422* (1–3), 230–233. https://doi.org/10.1016/j.cplett.2006.02.025.

(102) Cheng, X.; Zhang, Y.; Jónsson, E.; Jónsson, H.; Weber, P. M. Charge Localization in a Diamine Cation Provides a Test of Energy Functionals and Self-Interaction Correction. *Nat Commun* **2016**, *7* (1), 11013. https://doi.org/10.1038/ncomms11013.

(103) Patchkovskii, S.; Ziegler, T. Improving “Difficult” Reaction Barriers with Self-Interaction Corrected Density Functional Theory. *The Journal of Chemical Physics* **2002**, *116* (18), 7806–7813. https://doi.org/10.1063/1.1468640.

(104) Mishra, P.; Yamamoto, Y.; Johnson, J. K.; Jackson, K. A.; Zope, R. R.; Baruah, T. Study of Self-Interaction-Errors in Barrier Heights Using Locally Scaled and Perdew–Zunger Self-Interaction Methods. *J. Chem. Phys.* **2022**, *156* (1), 014306. https://doi.org/10.1063/5.0070893.

(105) Vargas, J.; Ufondu, P.; Baruah, T.; Yamamoto, Y.; Jackson, K. A.; Zope, R. R. Importance of Self-Interaction-Error Removal in Density Functional Calculations on Water Cluster Anions. *Phys. Chem. Chem. Phys.* **2020**, *22* (7), 3789–3799. https://doi.org/10.1039/C9CP06106A.

(106) Bryenton, K. R.; Adeleke, A. A.; Dale, S. G.; Johnson, E. R. Delocalization Error: The Greatest Outstanding Challenge in Density-functional Theory. *WIREs Comput Mol Sci* **2023**, *13* (2). https://doi.org/10.1002/wcms.1631.

(107) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865–3868. https://doi.org/10.1103/PhysRevLett.77.3865.

(108) Perdew, J. P. Generalized Gradient Approximations for Exchange and Correlation: A Look Backward and Forward. *Physica B: Condensed Matter* **1991**, *172* (1–2), 1–6. https://doi.org/10.1016/0921-4526(91)90409-8.

(109) Lonsdale, D. R.; Goerigk, L. The One-Electron Self-Interaction Error in 74 Density Functional Approximations: A Case Study on Hydrogenic Mono- and Dinuclear Systems. *Phys. Chem. Chem. Phys.* **2020**, *22* (28), 15805–15830. https://doi.org/10.1039/D0CP01275K.

(110) Borlido, P.; Aull, T.; Huran, A. W.; Tran, F.; Marques, M. A. L.; Botti, S. Large-Scale Benchmark of Exchange–Correlation Functionals for the Determination of Electronic Band Gaps of Solids. *J. Chem. Theory Comput.* **2019**, *15* (9), 5069–5079. https://doi.org/10.1021/acs.jctc.9b00322.

(111) Saßnick, H.-D.; Cocchi, C. Electronic Structure of Cesium-Based Photocathode Materials from Density Functional Theory: Performance of PBE, SCAN, and HSE06 Functionals. *Electron. Struct.* **2021**, *3* (2), 027001. https://doi.org/10.1088/2516-1075/abfb08.

(112) Cencek, W.; Szalewicz, K. On Asymptotic Behavior of Density Functional Theory. *The Journal of Chemical Physics* **2013**, *139* (2), 024104. https://doi.org/10.1063/1.4811833.

(113) Zhang, D.; Yang, X.; Zheng, X.; Yang, W. Accurate Density Functional Prediction of Molecular Electron Affinity with the Scaling Corrected Kohn–Sham Frontier Orbital Energies. *Molecular Physics* **2018**, *116* (7–8), 927–934. https://doi.org/10.1080/00268976.2017.1382738.

(114) Zhang, G.-X.; Reilly, A. M.; Tkatchenko, A.; Scheffler, M. Performance of Various Density-Functional Approximations for Cohesive Properties of 64 Bulk Solids. *New J. Phys.* **2018**, *20* (6), 063020. https://doi.org/10.1088/1367-2630/aac7f0.

(115) Weymuth, T.; Couzijn, E. P. A.; Chen, P.; Reiher, M. New Benchmark Set of Transition-Metal Coordination Reactions for the Assessment of Density Functionals. *J. Chem. Theory Comput.* **2014**, *10* (8), 3092–3103. https://doi.org/10.1021/ct500248h.

(116) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. Climbing the Density Functional Ladder: Nonempirical Meta–Generalized Gradient Approximation Designed for Molecules and Solids. *Phys. Rev. Lett.* **2003**, *91* (14), 146401. https://doi.org/10.1103/PhysRevLett.91.146401.

(117) Adamo, C.; Barone, V. Toward Reliable Density Functional Methods without Adjustable Parameters: The PBE0 Model. *The Journal of Chemical Physics* **1999**, *110* (13), 6158–6170. https://doi.org/10.1063/1.478522.

(118) Perdew, J. P.; Ernzerhof, M.; Burke, K. Rationale for Mixing Exact Exchange with Density Functional Approximations. *The Journal of Chemical Physics* **1996**, *105* (22), 9982–9985. https://doi.org/10.1063/1.472933.

(119) Tsai, C.-W.; Su, Y.-C.; Li, G.-D.; Chai, J.-D. Assessment of Density Functional Methods with Correct Asymptotic Behavior. *Phys. Chem. Chem. Phys.* **2013**, *15* (21), 8352. https://doi.org/10.1039/c3cp50441g.

(120) Shao, Y.; Mei, Y.; Sundholm, D.; Kaila, V. R. I. Benchmarking the Performance of Time-Dependent Density Functional Theory Methods on Biochromophores. *J. Chem. Theory Comput.* **2020**, *16* (1), 587–600. https://doi.org/10.1021/acs.jctc.9b00823.

(121) Yanai, T.; Tew, D. P.; Handy, N. C. A New Hybrid Exchange–Correlation Functional Using the Coulomb-Attenuating Method (CAM-B3LYP). *Chemical Physics Letters* **2004**, *393* (1–3), 51–57. https://doi.org/10.1016/j.cplett.2004.06.011.

(122) Köster, A.; Geudtner, G.; Calaminici, P.; Casida, M.; Dominguez, V.; Flores-Moreno, R.; Gamboa, G.; Goursot, A.; Heine, T.; Ipatov, A.; others. DeMon2k, Version 6, The DeMon Developers, Cinvestav, Mexico City (2018). *Google Scholar There is no corresponding record for this reference*.

(123) Obara, S.; Saika, A. Efficient Recursive Computation of Molecular Integrals over Cartesian Gaussian Functions. *The Journal of Chemical Physics* **1986**, *84* (7), 3963–3974. https://doi.org/10.1063/1.450106.

(124) Gill, P. M. W. Molecular Integrals Over Gaussian Basis Functions. In *Advances in Quantum Chemistry*; Elsevier, 1994; Vol. 25, pp 141–205. https://doi.org/10.1016/S0065-3276(08)60019-2.

(125) Calaminici, P.; Alvarez-Ibarra, A.; Cruz-Olvera, D.; Dominguez-Soria, V.; Flores-Moreno, R.; Gamboa, G.; Geudtner, G.; Goursot, A.; Mejía-Rodríguez, D.; Salahub, D.; others. Auxiliary Density Functional Theory: From Molecules to Nanostructures. *Handbook of computational chemistry* **2016**, 1.

(126) Mejía-Rodríguez, D.; Trickey, S. B. Variational Properties of Auxiliary Density Functionals. *Theor Chem Acc* **2021**, *140* (4), 37. https://doi.org/10.1007/s00214-021-02731-2.

(127) Köster, A. M. Hermite Gaussian Auxiliary Functions for the Variational Fitting of the Coulomb Potential in Density Functional Methods. *The Journal of Chemical Physics* **2003**, *118* (22), 9943–9951. https://doi.org/10.1063/1.1571519.

(128) Mejía-Rodríguez, D.; Köster, A. M. Robust and Efficient Variational Fitting of Fock Exchange. *The Journal of Chemical Physics* **2014**, *141* (12), 124114. https://doi.org/10.1063/1.4896199.

(129) Runge, E.; Gross, E. K. U. Density-Functional Theory for Time-Dependent Systems. *Phys. Rev. Lett.* **1984**, *52* (12), 997–1000. https://doi.org/10.1103/PhysRevLett.52.997.

(130) Li, X.; Smith, S. M.; Markevitch, A. N.; Romanov, D. A.; Levis, R. J.; Schlegel, H. B. A Time-Dependent Hartree–Fock Approach for Studying the Electronic Optical Response of Molecules in Intense Fields. *Phys. Chem. Chem. Phys.* **2005**, *7* (2), 233–239. https://doi.org/10.1039/B415849K.

(131) Lopata, K.; Govind, N. Modeling Fast Electron Dynamics with Real-Time Time-Dependent Density Functional Theory: Application to Small Molecules and Chromophores. *J. Chem. Theory Comput.* **2011**, *7* (5), 1344–1355. https://doi.org/10.1021/ct200137z.

(132) Castro, A.; Marques, M. A. L.; Rubio, A. Propagators for the Time-Dependent Kohn–Sham Equations. *The Journal of Chemical Physics* **2004**, *121* (8), 3425–3433. https://doi.org/10.1063/1.1774980.

(133) Gómez Pueyo, A.; Marques, M. A. L.; Rubio, A.; Castro, A. Propagators for the Time-Dependent Kohn–Sham Equations: Multistep, Runge–Kutta, Exponential Runge–Kutta, and Commutator Free Magnus Methods. *J. Chem. Theory Comput.* **2018**, *14* (6), 3040–3052. https://doi.org/10.1021/acs.jctc.8b00197.

(134) Magnus, W. On the Exponential Solution of Differential Equations for a Linear Operator. *Comm. Pure Appl. Math.* **1954**, *7* (4), 649–673. https://doi.org/10.1002/cpa.3160070404.

(135) Cheng, C.-L.; Evans, J. S.; Van Voorhis, T. Simulating Molecular Conductance Using Real-Time Density Functional Theory. *Phys. Rev. B* **2006**, *74* (15), 155112. https://doi.org/10.1103/PhysRevB.74.155112.

(136) Gómez Pueyo, A.; Marques, M. A. L.; Rubio, A.; Castro, A. Propagators for the Time-Dependent Kohn–Sham Equations: Multistep, Runge–Kutta, Exponential Runge–Kutta, and

Commutator Free Magnus Methods. *J. Chem. Theory Comput.* **2018**, *14* (6), 3040–3052. https://doi.org/10.1021/acs.jctc.8b00197.

(137) Wu, X.; Alvarez-Ibarra, A.; Salahub, D. R.; De La Lande, A. Retardation in Electron Dynamics Simulations Based on Time-Dependent Density Functional Theory. *Eur. Phys. J. D* **2018**, *72* (12), 206. https://doi.org/10.1140/epjd/e2018-90219-3.

(138) Fuks, J. I.; Maitra, N. T. Challenging Adiabatic Time-Dependent Density Functional Theory with a Hubbard Dimer: The Case of Time-Resolved Long-Range Charge Transfer. *Phys. Chem. Chem. Phys.* **2014**, *16* (28), 14504–14513. https://doi.org/10.1039/C4CP00118D.

(139) Delgado, J.; Lara-Astiaso, M.; González-Vázquez, J.; Decleva, P.; Palacios, A.; Martín, F. Molecular Fragmentation as a Way to Reveal Early Electron Dynamics Induced by Attosecond Pulses. *Faraday Discuss.* **2021**, *228*, 349–377. https://doi.org/10.1039/D0FD00121J.

(140) Maitra, N. T.; Burke, K.; Woodward, C. Memory in Time-Dependent Density Functional Theory. *Phys. Rev. Lett.* **2002**, *89* (2), 023002. https://doi.org/10.1103/PhysRevLett.89.023002.

(141) Fuks, J. I.; Lacombe, L.; Nielsen, S. E. B.; Maitra, N. T. Exploring Non-Adiabatic Approximations to the Exchange–Correlation Functional of TDDFT. *Phys. Chem. Chem. Phys.* **2018**, *20* (41), 26145–26160. https://doi.org/10.1039/C8CP03957G.

(142) Lacombe, L.; Maitra, N. T. Non-Adiabatic Approximations in Time-Dependent Density Functional Theory: Progress and Prospects. *npj Comput Mater* **2023**, *9* (1), 124. https://doi.org/10.1038/s41524-023-01061-0.

(143) Yost, D. C.; Yao, Y.; Kanai, Y. Examining Real-Time TDDFT Non-Equilibrium Simulations for the Calculation of Electronic Stopping Power. **2018**. https://doi.org/10.48550/ARXIV.1805.01377.

(144) Ali, E.; Madjet, M. E.-A.; De, R.; Frauenheim, T.; Chakraborty, H. S. Ultrafast Nonadiabatic Electron Dynamics in Photoexcited C60: A Comparative Study among DFT Exchange-Correlation Functionals. **2023**. https://doi.org/10.48550/ARXIV.2306.16386.

(145) Riss, U. V.; Meyer, H.-D. Reflection-Free Complex Absorbing Potentials. *J. Phys. B: At. Mol. Opt. Phys.* **1995**, *28* (8), 1475–1493. https://doi.org/10.1088/0953-4075/28/8/012.

(146) Cireasa, R.; Boguslavskiy, A. E.; Pons, B.; Wong, M. C. H.; Descamps, D.; Petit, S.; Ruf, H.; Thiré, N.; Ferré, A.; Suarez, J.; Higuet, J.; Schmidt, B. E.; Alharbi, A. F.; Légaré, F.; Blanchet, V.; Fabre, B.; Patchkovskii, S.; Smirnova, O.; Mairesse, Y.; Bhardwaj, V. R. Probing Molecular Chirality on a Sub-Femtosecond Timescale. *Nature Phys* **2015**, *11* (8), 654–658. https://doi.org/10.1038/nphys3369.

(147) Shore, B. W. Solving the Radial Schrödinger Equation by Using Cubic-Spline Basis Functions. *The Journal of Chemical Physics* **1973**, *58* (9), 3855–3866. https://doi.org/10.1063/1.1679740.

(148) Labeye, M.; Zapata, F.; Coccia, E.; Véniard, V.; Toulouse, J.; Caillat, J.; Taïeb, R.; Luppi, E. Optimal Basis Set for Electron Dynamics in Strong Laser Fields: The Case of Molecular Ion H_2^+. *J. Chem. Theory Comput.* **2018**, *14* (11), 5846–5858. https://doi.org/10.1021/acs.jctc.8b00656.

(149) Kaufmann, K.; Baumeister, W.; Jungen, M. Universal Gaussian Basis Sets for an Optimum Representation of Rydberg and Continuum Wavefunctions. *J. Phys. B: At. Mol. Opt. Phys.* **1989**, *22* (14), 2223–2240. https://doi.org/10.1088/0953-4075/22/14/007.

(150) Nestmann, B. M.; Peyerimhoff, S. D. Optimized Gaussian Basis Sets for Representation of Continuum Wavefunctions. *J. Phys. B: At. Mol. Opt. Phys.* **1990**, *23* (22), L773–L777. https://doi.org/10.1088/0953-4075/23/22/007.

(151) Faure, A.; Gorfinkiel, J. D.; Morgan, L. A.; Tennyson, J. GTOBAS: Fitting Continuum Functions with Gaussian-Type Orbitals. *Computer Physics Communications* **2002**, *144* (2), 224–241. https://doi.org/10.1016/S0010-4655(02)00141-8.

(152) Tennyson, J.; Brown, D. B.; Munro, J. J.; Rozum, I.; Varambhia, H. N.; Vinci, N. Quantemol-N: An Expert System for Performing Electron Molecule Collision Calculations Using the R-Matrix Method. *J. Phys.: Conf. Ser.* **2007**, *86*, 012001. https://doi.org/10.1088/1742-6596/86/1/012001.

(153) Laporta, V.; Celiberto, R.; Tennyson, J. Dissociative Electron Attachment and Electron-Impact Resonant Dissociation of Vibrationally Excited O 2 Molecules. *Phys. Rev. A* **2015**, *91* (1), 012701. https://doi.org/10.1103/PhysRevA.91.012701.

(154) Tandiana, R.; Omar, K. A.; Luppi, E.; Cailliez, F.; Van-Oanh, N.-T.; Clavaguéra, C.; De La Lande, A. Use of Gaussian-Type Functions for Describing Fast Ion-Matter Irradiation with Time-Dependent Density Functional Theory. *J. Chem. Theory Comput.* **2023**, *19* (21), 7740–7752. https://doi.org/10.1021/acs.jctc.3c00656.

(155) Parise, A.; Alvarez-Ibarra, A.; Wu, X.; Zhao, X.; Pilmé, J.; De La Lande, A. Quantum Chemical Topology of the Electron Localization Function in the Field of Attosecond Electron Dynamics. *J. Phys. Chem. Lett.* **2018**, *9* (4), 844–850. https://doi.org/10.1021/acs.jpclett.7b03379.

(156) Lopata, K.; Govind, N. Near and Above Ionization Electronic Excitations with Non-Hermitian Real-Time Time-Dependent Density Functional Theory. *J. Chem. Theory Comput.* **2013**, *9* (11), 4939–4946. https://doi.org/10.1021/ct400569s.

(157) Tolu, D.; Guillaumont, D.; De La Lande, A. Irradiation of Plutonium Tributyl Phosphate Complexes by Ionizing Alpha Particles: A Computational Study. *J. Phys. Chem. A* **2023**, acs.jpca.3c02117. https://doi.org/10.1021/acs.jpca.3c02117.

(158) Krause, P.; Sonk, J. A.; Schlegel, H. B. Strong Field Ionization Rates Simulated with Time-Dependent Configuration Interaction and an Absorbing Potential. *The Journal of Chemical Physics* **2014**, *140* (17), 174113. https://doi.org/10.1063/1.4874156.

CHAPTER TWO

ON SOME RELATIVISTIC EFFECTS IN FAST ION-MOLECULE COLLISIONS

Contents

2 Relativistic effects in fast ion-molecule collisions

One of the primary features of charged particles that makes them suitable for radiotherapy is their ability to deposit a significant amount of energy at the end of their trajectory, referred to as the Bragg peak. Beyond the Bragg peak, energy deposition diminishes as the kinetic energy of the charged particles increases. When energetic particles are used, they lose their kinetic energy, transferring it to the target as the particles penetrate up to the Bragg peak. At this point, the charged particles release the majority of their energy. This characteristic of charged particles enables targeting of deep-seated tumor cells, while protecting healthy cells in the path of the charged particles. However, the energy deposition curve increases again at high kinetic energy due to the relativistic effects of charged particles. Relativity of charged particles plays a crucial role in radiotherapy, specifically in proton therapy. This is one of the reasons that limits the maximum kinetic energy of a proton to approximately 250 MeV[1,2]. As a result, the healthy tissue at the beginning of the charged particle penetration is no longer protected.

Relativity effects can manifest in different situations depending on the speed of the particles of the system of interest. Either the electrons of the target material can approach the speed of light, or the irradiating particles can. The finite propagation of electric fields mediating electrostatic interactions may also have to be considered on the attosecond time scale.

The core electrons of a target, particularly for heavy atoms, are potentially subjected to relativity effects. The latter further impact the binding of electrons in upper electronic shells. For atoms with an atomic number greater than 54, the reduction of binding energy due to relativity is significant[3]. As a result, the calculated stopping power of relativistic heavy atoms is expected to be less than that of non-relativistic ones[3–5]. Fei Mao and colleagues investigated the contribution of *f*-electron excitations to the electronic stopping power of platinum for a proton using the RT-TD-DFT framework[6]. In their work, the core electrons were replaced by a relativistic pseudopotential model. They demonstrated the importance of including electron relativity in the stopping power calculation for Pt. However, as we are mainly interested in biomolecules, which typically have an atomic number of less than 20, the effect of target electron relativity is expected to be negligible.

Another source of relativistic effect which is the main focus of this Chapter is that arising from the ionizing particle's velocity. Both experimental[7–9] and theoretical[10] evidences show that energy deposition, or alternatively the stopping power, increases with kinetic energy in the relativistic region. However, in the high relativistic region, the stopping power is lower compared to what is predicted by the Bethe-Bloch theory. This reduction is proportional to the density of the target[11] and is more significant for light atoms[12]. This phenomenon is known as the density effect. Due to relativity, electric fields become concentrated and expand perpendicular to the direction of particle's propagation track. This leads to the polarization of electrons in the medium. Consequently, a portion of the long-distance electric field is screened. The density effect becomes significant when the kinetic energy of the particles approaches their rest mass. For instance, in the case of a proton, this value is 938.27 MeV. The threshold for the increase in the stopping power curve, known as minimum ionization loss, begins to rise gradually from about 300 MeV (for protons). The minimum ionization depends solely on the target material and it decreases with increasing atomic number[13].

Wu and colleagues proposed to include such relativistic effects of charged particles in RT-TD-ADFT[14]. They proposed to use the Liénard-Wiechert potential (LWP) instead of the classical Coulomb potential in the Kohn-Sham external potential. They found that the use of the LWP is necessary for calculating energy deposition, for example, for protons above 1 MeV. This preliminary work called for deeper analyses of such relativistic effects, which is the motivation of the present Chapter.

RT-TD-DFT is a promising model for calculating energy deposition or stopping power, as it can accurately describe the physical processes incorporated in charged particle interactions. However, in practice, stopping power, a quantity that is comparable to experimental data, is underestimated by RT-TD-DFT at intermediate energies (corresponding to a maximum proton energy of around 2 MeV) when compared to a variable data for stopping power[15]. Correa and coworkers reported the stopping power for Ni in the velocity range between 0.5-20 au (corresponding to proton energies between 0.006-9.92 MeV)[16]. They found a dramatic underestimation of the stopping power at high velocity. To date, higher energy particles have not been explored in the context of RT-TD-DFT.

The source of stopping power underestimation in RT-TD-DFT calculations is not clear until now. However, some technical problems may be responsible for example, the quality of the basis set. Maliyov *et al*[17,18] have shown that the calculation of stopping power is substantially sensitive to the quality of the basis set. Especially at intermediate velocity, the underestimation of stopping power highly depends on the quality of the basis set for core and deep valence electrons. Of course, the size of the simulated system could be another factor. Additionally, the effects of relativistic phenomena, which have not been explored by RT-TD-DFT, may also play a role in these discrepancies.

In this chapter, we generally investigate the impact of relativity on the electric field of charged particles. Moreover, we explore the energy deposition of protons in guanine-cytosine nucleobase paired with water molecules over a wide energy range, from 0.01 MeV to 635 MeV. This is accomplished by incorporating relativistic effects for the charged particles in the RT-TD-ADFT simulation. In Section 2.2, we discuss the mechanisms involved in energy deposition. Section 2.3 is dedicated to exploring the effects of relativity on the electric field of charged particles. Section 2.4 focuses on RT-TD-ADFT simulations for the energy deposition of protons in a biomolecule. This includes detailing the methodology and system setup, testing the basis set, evaluating the energy deposition of protons in the range from 0.01 to 635 MeV, and discussing the effects of relativity on proton interactions. We conclude the chapter with a summary of our findings and perspectives.

2.1 Interaction of charged particles depending on their energy

The description of the interaction of charged particles with electrons resting on molecules is a substantial challenge for RT-TD-DFT simulations as the mechanisms at play strongly depend on the projectile's kinetic energy. For the sake of simplicity, we may define four regimes depending on the irradiated material and the irradiating ion. Even though it is delicate to define strict limits between these regimes, they help to set the ideas. Taking the example of protons[12] these regions would be:

- Very-low energy region (< 0.01 MeV). Energy deposition via elastic collision in nuclear vibrational modes is substantial here[19,20]. This kind of excitation that does not involve electronic excitations is not of interest to us in the context of this thesis and will not be considered any further.
- Low-energy region (from 0.01 to 0.2 MeV). This energy domain is characterized by the intricate non-adiabatic excitation and ionization electron dynamics. Atom nuclei have no time to respond due to their large mass during the energy deposition process. In general, valence electrons have a significant role in energy transition[21]. The most notable feature of this energy region is the electron polarization during impact. Notably, when the proton has a lower velocity than electrons, the latter move and accumulate around the projectile.
- Intermediate-energy region (from 0.2 to 1 MeV). There, core electrons more substantially contribute to energy transfer[17]. While electron polarization remains, it is significantly reduced.
- High-energy region (above 1 MeV). There, relativistic effects arising from the high velocity of the ion play an ever-greater role in the electron-ion interaction. Additionally, electron density polarization re-appears due to the relativistic effects.

I will now describe in more detail each of these last three regimes for which electronic excitations are the main energy deposition channel.

2.1.1 Low-energy region (from 0.01 to 0.2 MeV)

Energy deposition is highly dependent on the complex interaction between the charge and the electron cloud. Theories based on first-order approximations, such as the Bethe-Bloch theory[12,22] and the Lindhard model[23–25] , cannot adequately describe this regime. Indeed, according to the first-order approximation, the electrons on the target material do not move significantly during collision ("electron rest approximation"). However, if the charge is slow enough, electrons have time to move significantly relative to the ion, creating electrostatic induction (polarization). If the projectile's charge is positive, the electrons move in its direction. Conversely, if it is negative, the electrons are repelled.

- The augmentation of electron density along the ion track and the deficiency of electron density on the surrounding fragments create typical charge oscillations that are called by some authors plasma oscillations[26]. The projectile has the ability to excite this plasma. This typical excitation is crucial for conductor and semi-conductor materials[27].
- Polarization-induced electron accumulation around the proton may screen the proton's electric field. As a consequence, the interaction of electrons with the screened charge decreases, especially for long-range interactions, hence energy deposition. The effect of screening increases as more electrons accumulate around the proton. This effect is known as charge neutralization. In general, this effect is significant below 1 MeV for protons.
- For positive charge particles like protons, electrons that are faster than the projectile move toward the charged particle during collision. As a result, the distance between these electrons and the ion decreases compared to other electrons. Consequently, energy deposition increases because Coulomb interaction decreases with distance between an electron and a charged particle. This is known as Barkas's effect[28]. This effect becomes more pronounced as ion energy decreases due to an increase in the number of electrons that are faster than the ion. However, it can keep on even in the high-energy region (~100 MeV)[12] due to the fact that the velocities of some core electrons (depending on their energy) are always larger than the velocity of the projectile. In the low-energy region, Barkas effect is in competition with the screening effect.

These three kinds of mechanisms efficiently involve the energy deposition of charged particles in the low-energy region. That is why it is one of the most difficult regions for simulation, even for experiments. However, compared to other regions, this region of energies is technically easier for RT-TD-DFT simulations.

Our group reported the irradiation of biomolecules with fast particles using RT-TD-ADFT[29]. It was found that a positive charge pulls in the surrounding electron density and partially reduces the fragments of molecules that are about to be hit. The slower the projectile, the stronger the polarization. RT-TD-ADFT simulations are thus able to reproduce the Barkas effect. This phenomenon has been qualified as a mechanism of "tidal" ("ebb-and-flow") ionization.

2.1.2 Intermediate-energy region (from 0.2 to 1 MeV)

The aforementioned three mechanisms are still at play, but as the ion's kinetic energy is higher, they are less pronounced. Core electron excitations, on the other hand, become significant[21].

In general, when a projectile goes through a molecule, it is decelerated by the Coulomb interaction with the electrons, and as a result, the energy will be released with the equivalent of ionization or excitation of target electrons. The frequency created by charged particles has broadband energy. In addition, the frequencies increase with the energy of charged particles (we will come back to this in the next sub-section), so that the probability of these frequencies that synchronize with the binding energy of core electrons augments. The participation of core electrons in the contribution of stopping power expects to be increasing with the energy of the projectile. However, the relativistic effect has not yet been effectively incorporated into the energy deposition contribution.

2.1.3 High-energy region (above 1 MeV)

Because of the very high velocity of the ion, the electric field it generates is subjected to some relativistic effects. The distances between the ion and the electrons reduce (the so-called distance-contraction-effect) and the collision time increases (time-dilatation-effect). As a consequence, energy deposition increases with the speed of the projectile.

In addition, the electric field generated by the ion is substantially concentrated in the perpendicular direction of particle propagation. This leads to the polarization of electrons and to the screening of the projectile's electric field for long-distance electrons. In the following sub-section, I go into deeper detail about the mechanisms at play in the relativistic area.

2.2 Relativistic effect

2.2.1 Liénard-Wiechert potential

The interaction of the charged particle with the target electrons is achieved by the introduction of electrostatic potential in Kohn-Sham external potential. For non-relativistic ions, this potential is the classical Coulomb potential (ϕ_{CP}) between the charged particle and the target electrons:

$$\phi_{CP}(r,t) = \frac{q_{proj}}{r} \tag{2.1}$$

where q_{proj} is the charge of the projectile. r is the distance between the projectile and the electron. On the other hand, as the speed of the charge increases and approaches the speed of light, the relativistic effect can be introduced in two ways:

One way is to change the velocity of the projectile to a relativistic velocity. Alternatively, we can modify the ion's potential, that is we can replace the non-relativistic Coulomb potential with the Liénard-Wiechert potential (ϕ_{LWP})[30] (eq. 2.2). This is the method we have chosen.

$$\phi_{LWP}(r,t) = \frac{q_{proj}}{r\left[1 - \frac{v_{proj}^2 sin^2\theta}{c^2}\right]^{1/2}} \tag{2.2}$$

where θ is the angle between the propagation direction and the distance vector between an electron and the particle, and r is the distance as illustrated in Figure 2-1. $\gamma = \left(1 - v_{proj}^2 sin^2\theta/c^2\right)^{-1/2}$ is the angle-dependent Lorentz factor.

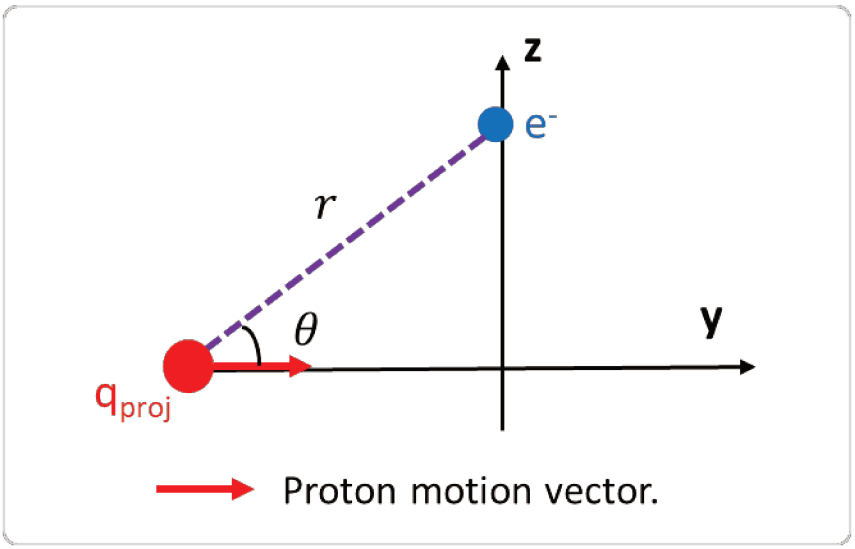

Figure 2-1: A schematic illustration of the angle between the propagation direction of a projectile and an electron.

The Liénard-Wiechert potential reduces to the classical Coulomb potential for ions traveling at speeds much lower than the speed of light because $\gamma \approx 1$. When $v_{proj}^2 \to c^2$, $\gamma \to 1$ if $\theta = 0$ (for

electrons positioned on the particle trajectory) and $\gamma \to +\infty$ if $\theta = \pi/2$ (for electrons positioned perpendicular to the particle trajectory). This means that as the ion's velocity approaches the speed of light, the potential generated by the projectile is constant in the direction of propagation and expands perpendicular to the direction of propagation. As shown in Table 2-1, we calculated ϕ_{LWP} for a proton $q_{proj} = 1\ (au)$ with 0.01 and 635 MeV of kinetic energy, depending on the θ value and assuming a distance proton-electron distance $r = 1\ bohr$.

Table 2-1: Liénard-Wiechert potential for proton depending on the proton energy and Liénard-Wiechert angle with $r = 1\ bohr$.

	$\phi_{LWP}\ (au)$	
θ	0.01 MeV	635 MeV
0°	1.00	1.00
22.5°	1.00	1.05
45°	1.00	1.21
77.5°	1.00	1.60
90°	1.00	1.67

γ decreases when the kinetic energy increases, which leads to a shorter distance between charged particle and an electron. Consequently ϕ_{LWP} increases, so that the augmentation of the potential charged particle by the relativistic effect is expected to increase energy deposition compared to the non-relativistic case.

2.2.2 The mechanisms of interaction

Relativistic effects change the mechanism of irradiation. To better appreciate this, we study the electric field produced by the charged particle. The electric field can be divided into a transverse and a longitudinal component, denoted E_z and E_y, respectively. E_z accelerates electrons perpendicular to the direction of the projectile, and E_y accelerates the target electrons parallel to the proton propagation motion, as illustrated in Figure 2-2.

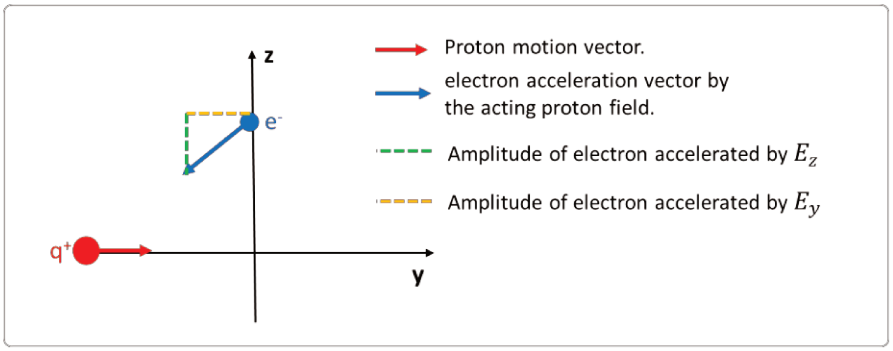

Figure 2-2: Schematic illustration of the action of transverse and longitudinal electric fields on an electron.

The transverse and longitudinal electric field strengths can be expressed as follows:

$$E_z(t) = \frac{q_{proj} b}{(b^2 + (\gamma' v t)^2)^{3/2}}$$
$$E_y(t) = \frac{-q_{proj} \gamma v t}{(b^2 + (\gamma' v t)^2)^{3/2}} \quad (2.3)$$

b is the impact parameter. v and t are the velocity of charged particle and time. Here, γ' is the relativistic factor, $\gamma' = (1 - v^2/c^2)^{-1/2}$.

As an illustration, we represent E_y and E_z as a function of time or of the proton's position in Figure 2-3 (b and c). We consider kinetic energies of 0.01, 0.06, 2, and 635 MeV. The atomic unit of electric field is e/a_0, with e and a_0 the charge of the electron and the Bohr radius respectively. We take the Fourier transform of E_y and E_z, leading to the E_y^ω and E_z^ω functions. We plot on panel d the relative intensity, which is defined as ($\left|E_{y/z}^\omega\right|^2 / max\left|E_{y/z}^\omega\right|^2$).

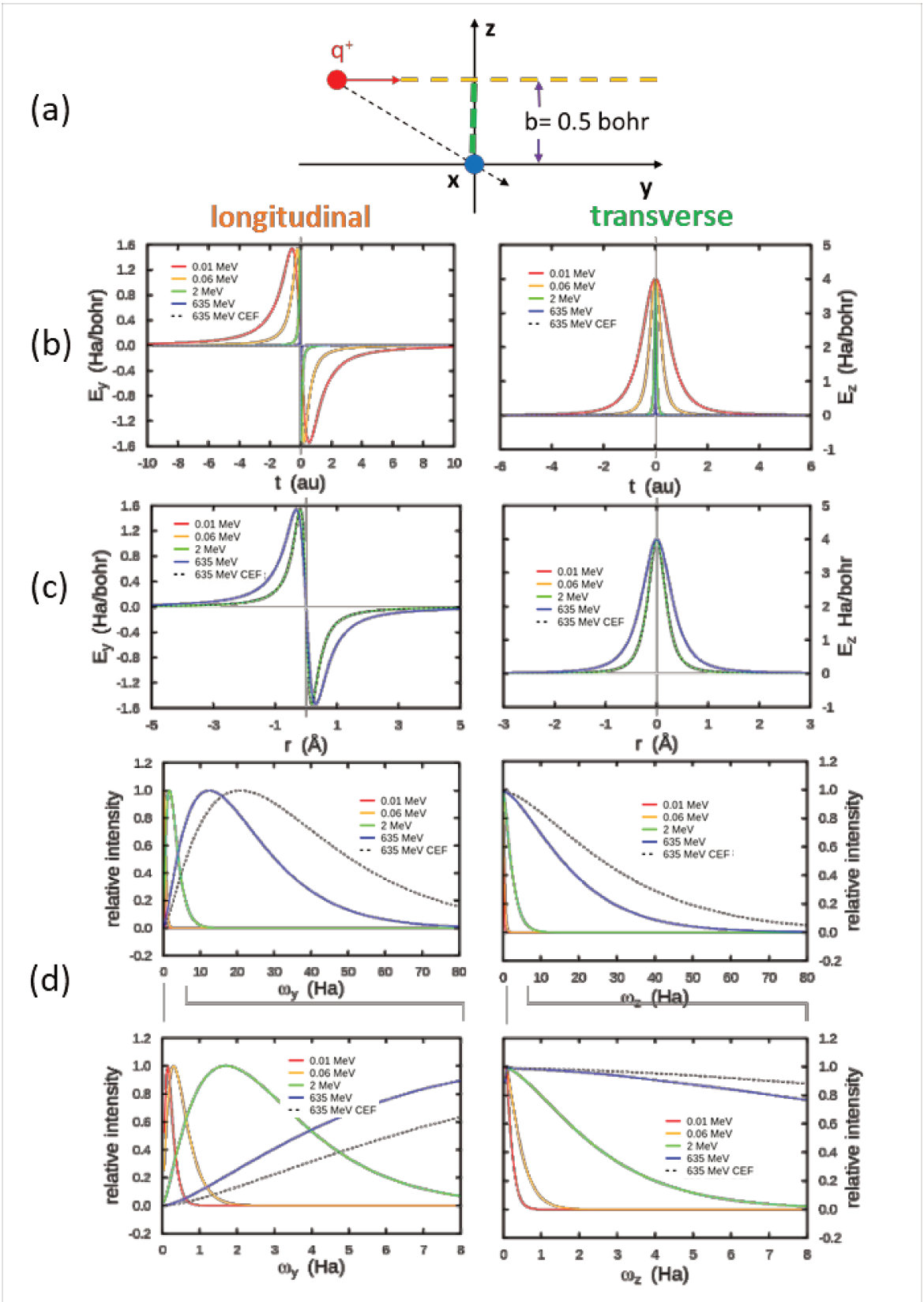

Figure 2-3: (a) Schematic demonstration of the transverse (left) and longitudinal (right) electric fields. (b) and (c): electric field strength as a function of time and distance. (d) Spectrum diagrams of the corresponding electric field. The gray dotted line corresponds to the electric field calculated from a Coulomb electric field (CEF) at 635 MeV.

As seen on Figure 2-3, b and c, E_y is anti-symmetric around the collision event. The force exerted on the electron before and after collision is the same, but of opposite directions. Therefore, E_y is not effective in transferring energy to the electron. This statement is valid under the conditions of electron rest and binary collision, where the electrons do not move during the collision and the charged particle interacts independently with each electron of the target molecules one by one, neglecting both screening and the correlation between electrons. However, these conditions tend to break down, particularly in the low-energy region.

The transverse electric field E_z is the main component responsible for energy deposition. It is symmetric around the collision time and leads to increase energy deposition before and after collision. The longitudinal force acting on the electron leads to an acceleration of the electron toward the ion. The electron continually gains energy until the force decreases as the ion goes away.

As evidenced on panel c, E_z starts to increase from -3 Å for 635 MeV, while it starts to increase at -1.5 Å for 2 MeV. In fact, there is a tiny difference between E_z at 2 MeV and 0.06 MeV, but we cannot see it in the figure. This corresponds to a relativistic effect. Relativity contracts the distance between the electron and the charged particle, and the interaction between them increases. The dotted black line corresponds to the calculation of the electric field strength at 635 MeV using a non-relativistic electric field. We will return to this point later.

Panel d shows the spectrum of energy produced by the charged particle. As seen, higher-energy projectile contains more high-frequency. As a result, the participation of core and inner valence electrons in energy deposition is expected to be augmented by charged particle energy. But it is not all, the frequency of the charged particle also highly depends on the impact parameter, as illustrated in Figure 2-4. The frequency increases significantly by reducing the impact parameter. In a close collision, core electron excitation is possible even at 0.01 MeV. This finding emphasizes the importance of a good description of core electrons in simulating charged particle interactions at high energies and close collisions, which is a difficult task for RT-TD-DFT simulations with non-local basis sets.

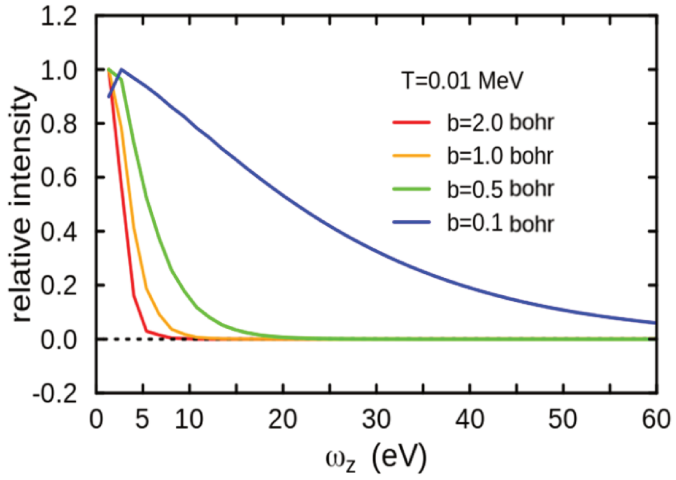

Figure 2-4: Spectrum diagrams of the transversal electric field (T) of 0.01 MeV proton according to the impact parameters (b).

The dotted black line in Figure 2-3 refers to E_y, E_z, and the relative intensity of a 635 MeV proton, calculated using a classical (non-relativistic) electric field (CEF).

As depicted in Figure 2-3(b), When comparing the dotted black line with the solid blue line (results derived from a relativistic electric field) for a proton at 635 MeV, no significant difference is observed for E_y and E_z within the time domain at the scale used in the figure. However, when the time range is reduced to -0.05 to +0.05 au, as shown in Figure 2-5, it is apparent that E_y and E_z from the relativistic electric field (REF) extend in time compared to those from the CEF.

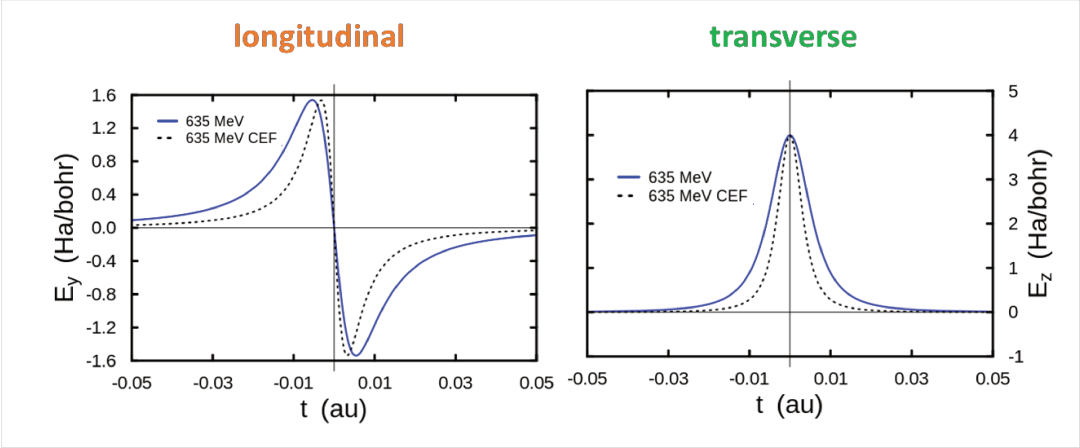

Figure 2-5: The transverse and longitudinal electric field strengths for a proton at 635 MeV, calculated using the relativistic electric field (solid blue line) and classical electric field (dotted black line).

In Figure 2-3(c), the effects of relativity are clearly visible in the spatial domain. The dotted black line perfectly coincides with the red line (proton at 0.01 MeV) and the orange line (proton at 0.06 MeV), where the effect of relativity is negligible at these kinetic energies of the proton. However,

the electric field strength of the blue line starts at a larger distance compared to the others. This is due to the contraction of distance caused by relativity and the increase in the strength of charged particles.

Furthermore, as illustrated in Figure 2-3(d), the relativistic effect significantly influences the frequencies generated by the charged particles. A red-shift is observed for a proton at 635 MeV resulting from the reference (blue line). This can be attributed to the fact that relativistic effects extend the interaction time length, leading to the generation of less energetic frequencies compared to the CEF. In this figure, we display the relative intensity of the Fourier-transformed E_y and E_z. For further details, we plotted the intensity as a function of frequency for a proton at 635 MeV with the reference and CEF in Figure 2-6. The results show that the intensity of the spectrum resulting from the reference significantly increases compared to that of the non-relativistic one. This can be attributed to the contraction of distance, which leads to an increase in the interaction with charged particles. Consequently, the relativistic effect reduces the frequencies of the charged particle due to the dilation of the interaction time. On the other hand, it increases the intensity of the frequency due to the contraction of the interaction distance.

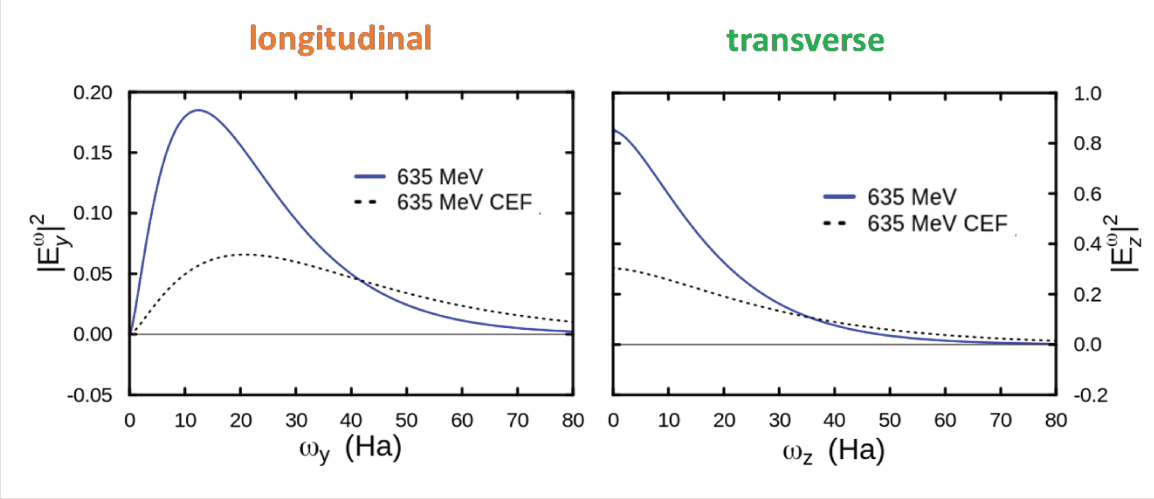

Figure 2-6: display the spectra for a 635 MeV proton calculated from the relativistic (solid blue line) and classical (dotted black line) electric fields. The plots represent the intensity ($\left|E_{y/z}^{\omega}\right|^2$) versus frequencies.

Finally, the analysis we have just reported, demonstrates that the energy transition from a charged particle essentially depends on the transversal interaction. In addition, by the relativistic effect, the transversal interaction expands in terms of interaction and spatiality. The interaction starts at a longer distance than for non-relativistic charged particles. We also demonstrate that

the relativistic effect leads to a red-shift and increase in the intensity of charged particle spectra. Moreover, our findings indicate that the frequencies generated by a charged particle heavily depend on the particle's energy and the distance of collision.

2.3 Assessment of RT-TD-ADFT simulations for charged particle interaction

In this section we assess the capability of first-principle simulations based on RT-TD-ADFT to capture the various energy deposition mechanisms. We first detail the molecular system we will be considering and the computational details, including some tests to select an adequate basis set. Then we will consider irradiation of our molecular system by proton with kinetic energies in the range of 0.01 to 635 MeV. The effect of relativity on the ion-electron interaction will be examined in detail.

2.3.1 Methodology and system setup

We consider a guanine-cytosine nucleobase with a pair of water molecules (Figure 2-7). We have optimized the supramolecular system geometry with deMon2k (version 6.0.2)[31]. To reduce the computational cost, we have used variational density fitting[32,33]. The classical electronic repulsion and XC contributions to the Kohn-Sham potential are computed with auxiliary densities. We chose the GGA (generalized-gradient-approximation) exchange correlation functional PBE (Perdew–Burke–Ernzerhof)[34]. We used the DZVP[35] basis sets (double zeta valence polarization) and GEN-A2* auxiliary sets[36] to expand the electronic Kohn-Sham molecular orbitals and the fitted density, respectively. We used the D3 empirical dispersion correction[37]. The two-dimensional molecular system was approximately placed in the Oxz plane, as shown in Figure 2-7.

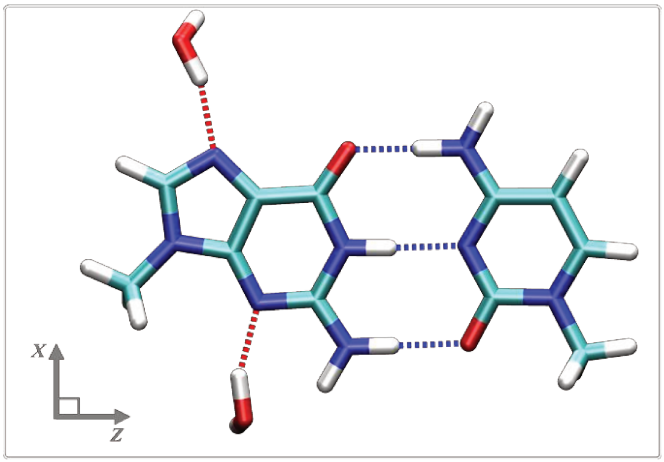

Figure 2-7: supramolecular system considered in the Chapter: a guanine (left), cytosine (right) and two water molecules. The dashed lines make visible the intermolecular hydrogen bonds.

To evaluate the amount of electronic energy deposition upon collision, we have used RT-TD-ADFT as implemented in a house version of deMon2k (version 4.4.5)[38]. Unless otherwise stated, for RT-TD-ADFT simulations, we have chosen the PBE functional for correlation and Correct Asymptotic Potential functional (CAP)[39] for exchange. We make here the adiabatic approximation. The propagation time step is limited by the physical processes involved in the dynamics of electrons. In our simulations, the time step is dependent on the kinetic energy of the projectile and ranges from 1 as at 0.01 MeV to 8.07×10^{-4} as at 635 MeV. In a preliminary phase, we have tested various propagation parameters to propagate the electron density in time. We have concluded that the efficiency and computational cost of time propagation is adequate with the second order Magnus propagator[40] coupled to a predictor-corrector algorithm[41]. We use a Taylor expansion[42] to evaluate the exponential of the Kohn-Sham matrix comprising 45 terms (see Chapter 1).

2.3.2 Choice of basis set

The quality of the basis set significantly impacts energy deposition. We have considered three basis sets in our tests: TZVP, TZVP-FIP2, and TZVP-AOC.

TZVP (Triple zeta valence polarization) is a basis set optimized for DFT calculations. TZVP-FIP2 stands for TZVP augmented with second-order field-induced polarization functions[43]. This basis set was optimized to improve the description of molecules' polarization responses. It was used for example to calculate polarizabilities and hyperpolarizabilities. We have selected this basis set here because we expect that the polarization of the electrons cloud by fast ions might play a

critical role in energy deposition. Finally, we have considered the TZVP-AOC (standard TZVP with additional fitting Gaussian-type orbitals for describing the continuum). These AOC basis set functions are added to help reproducing the continuum wave functions. They have been generated using the GTOBAS program from Faure *et al*[44]. The GTOBAS program accepts in input three parameters: the maximum angular quantum number (set to 2 here), the spatial extent of the fitting domain (3 Bohr) and the maximum energy of the continuum wave functions (5 Ry), respectively. A study from the ICP research group to which I contributed, have shown the capability of these basis sets to accurately simulate energy deposition by swift ions[45].

To assess the performance of the three basis sets to capture energy deposition, we have considered 56 single, independent, trajectories of the proton. Each of these trajectories crosses the average molecular plane perpendicularly, at different points, highlighted by the green beads on Figure 2-8. The set of trajectories include collisions at the centroid of the covalent bonds, on hydrogen bonds, on lone pairs or in regions of very low electron densities. The kinetic energy of the proton was set to 0.07 or 0.30 MeV. The proton was initially placed at y = 30 Å. For these simulations, we have chosen the PBE XC functional.

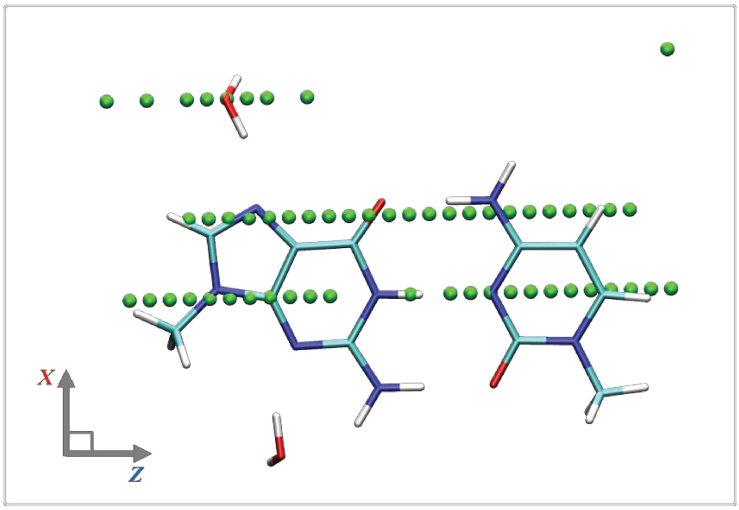

Figure 2-8: Supramolecular system (guanine, cytosine, and a pair of water molecules) is shown in licorice representation. Each green ball represents the passage of a single proton trajectory in the y-axis view.

The energy deposition for each trajectory is calculated from the difference between the ADFT energy at the end of the simulation and the ground state energy. The deposited energies summed over the 56 trajectories (ΔE_{56}), the averages ($\overline{\Delta E}$) and the standard deviations σ are collected in Table 2-2.

Table 2-2: Energy variation of the supramolecular system upon 56 collisions by a proton of kinetic energy of 0.07 and 0.30 MeV. N_{AO} is the number of atomic orbitals.

Basis set	N_{AO}	0.07 MeV			0.30 MeV		
		ΔE_{56} (eV)	$\overline{\Delta E}$ (eV)	σ (eV)	ΔE_{56} (eV)	$\overline{\Delta E}$ (eV)	σ (eV)
TZVP	581	1801	32.1	± 17.7	955	17.1	± 8.9
TZVP-FIP2	1080	1955	34.9	± 17.2	1014	18.1	± 8.7
TZVP-AOC	1618	2327	41.6	± 19.7	1217	21.7	± 12.5

ΔE_{56} and $\overline{\Delta E}$ decrease for a proton kinetic energy from 0.07 MeV to 0.30 MeV. This is in line with the expected position of the Bragg peak. Each of the three basis sets correctly describes this trend. However, the ΔE_{56} value significantly depends on the basis set. The larger the basis set the larger the ΔE_{56}. Taking by convention TZVP-AOC as the reference, the energy deposition calculated with TZVP is underestimated by 24% and 21.5% at 0.07 MeV and 0.30 MeV, respectively. With TZVP-FIP2, the underestimation is of 16% and 16.7%. As observed in the table, the standard deviations are notably large compared to the mean values. This outcome is expected due to significant variations in the spatial distribution of electron density within isolated systems. The energy deposition strongly depends on the local electron density surrounding the projectile's trajectory.

As a compromise between accuracy and computational cost, we have chosen the TZVP-FIP2 for our simulations. Actually, we are more interested in this work to investigate trends in deposited energy than the absolute value of deposited energy. We thus hope the TZVP-FIP2 is accurate enough to draw meaningful conclusions.

2.3.3 Energy deposition on the 0.01 – 635 MeV kinetic energy range

In this section we scan the proton kinetic energy in the 0.01 - 635 MeV range and examine the capability of RT-TD-ADFT to capture the physics at play that leads to energy deposition. For the sake of simplicity, we consider a single trajectory, still perpendicular to the bead as shown in Figure 2-9. The projectile is initially positioned 30 Å from the system and the simulation is run until the proton has travelled 60 Å. We have considered either a Coulomb potential or a Liénard-Wiechert potential to describe the electron-fast ion interaction.

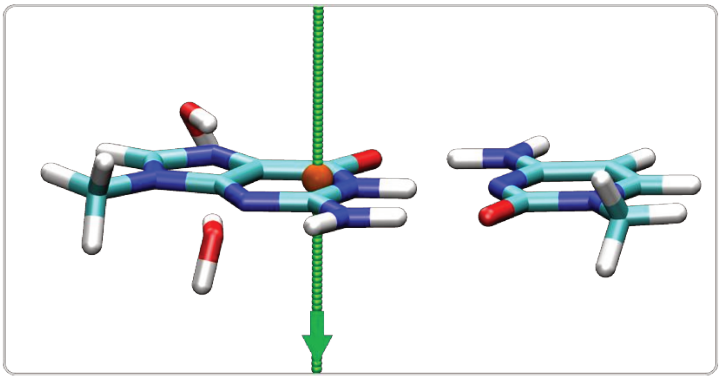

Figure 2-9: The supramolecular system (guanine, cytosine, and pair water molecules) is shown in licorice representation. The green line represents the position and direction of the proton's trajectory. The orange bead defines a reference position (y-axis=0) used in later analyses.

The energy deposition curve as function of the proton's kinetic energy is shown in Figure 2-10. The red line (Coulomb potential) and the blue line (Liénard-Wiechert potential) are superimposable from 0.01 MeV to around 2 MeV, where they start to diverge. We will successively examine three energy regions in more detail in the next three sub-sections.

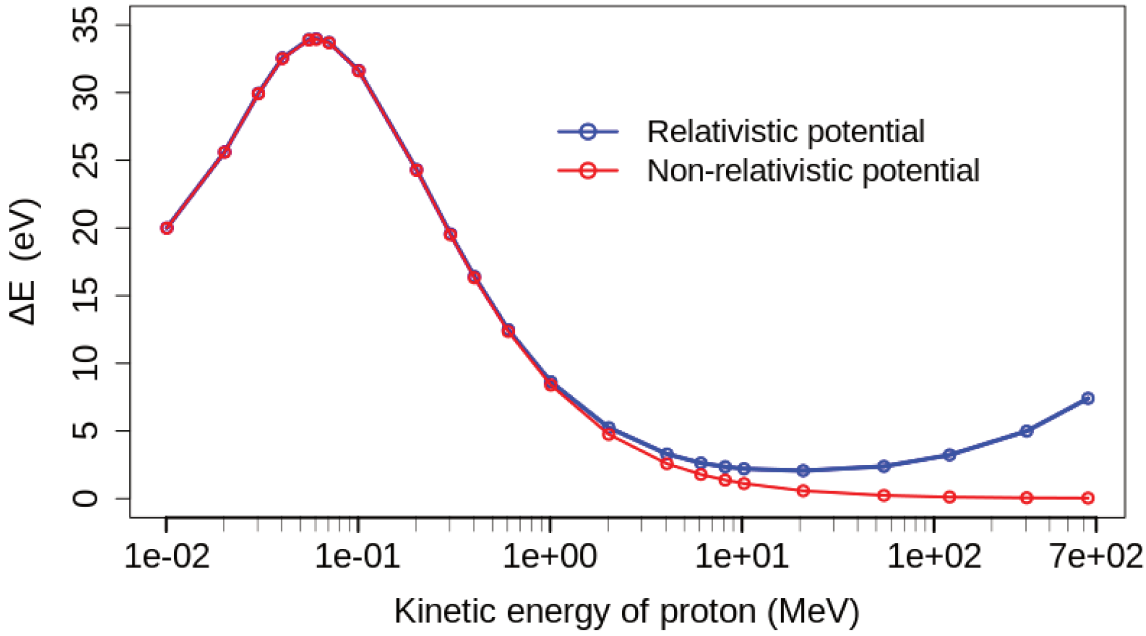

Figure 2-10: energy deposition curve as function of the proton's kinetic energy. The x-axis uses a logarithmic scale.

2.3.3.1 Low-energy region (0.01 - 0.2 MeV)

This region is one of the important parts of the energy deposition curve where charged particles lose the majority of their energy. As seen on Figure 2-10, the proton deposits 20 eV into the

supramolecular system at 0.01 MeV. Then, the deposited energy increases until it reaches a maximum of 34 eV at 0.055 MeV. After that, energy deposition declines. In this region, relativity has a negligible effect.

The non-monotonic shape of the energy deposition curve over this range of kinetic energy results from complex interactions with the electron cloud. In this region the electrons are much faster than the ions. Three mechanisms are usually invoked, namely plasma oscillations, fast-ion induced polarization and Barkas effect.

To evidence the electron dynamics during collision, we have calculated the electron density at the center of the molecular system for this trajectory. This point is evidenced by the orange bead on Figure 2-9. We report in Figure 2-11, the variation of electron density with the ground state at this fixed point is space during the simulation. Thus, a positive value means an accumulation of electron density with respect to the ground state.

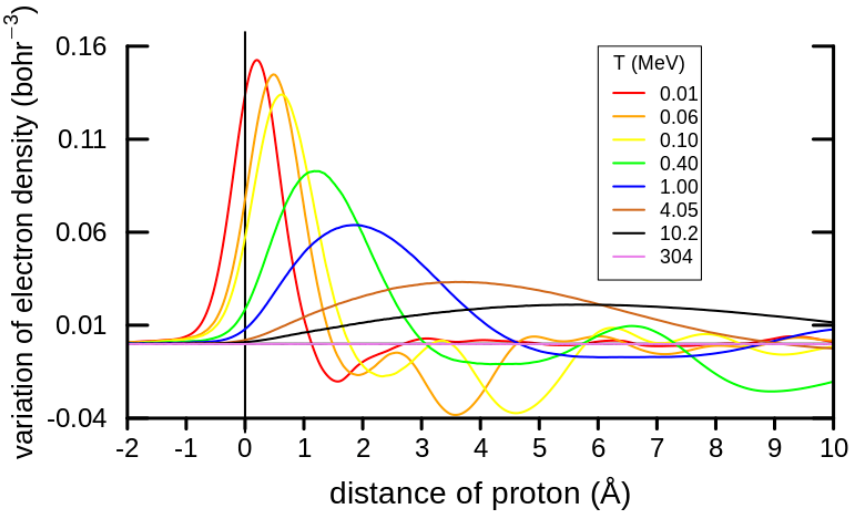

Figure 2-11: electron density variation in one fixed point on the trajectory position at a center of plane geometry (as shown as a red ball in (Figure 2-9 as a function of the proton's distance from this point. Proton kinetic energy range between 0.01 and 635 MeV. The simulations are conducted with LWP.

The electron density already increases before the arrival of the proton, electrons are thus not at rest during collision. Electron accumulation around the proton is mainly responsible for electric field screening. The variation of electron density at the proton superimposed on the point (the distance of proton =0 in the figure) decreases as the proton's energies increase. Therefore, it can

be expected that the effect of screening charged particles becomes more significant as the energy of the proton decreases. Generally, this effect is important for protons with energy below 1 MeV. However, in our system, the threshold extends to 4 MeV. Of course, this effect competes with Barkas's effect.

After a collision, the variation in electron density decreases and exhibits oscillations, which are attributed to charge migration between fragments. Charge migration, characterized by its oscillatory nature, is responsible for charge delocalization.

These three types of mechanisms efficiently involve energy deposition of charged particles in the low-energy region, which is why it is one of the most difficult regions for simulation by the models dependent on first-order perturbation theory.

2.3.3.2 Intermediate-energy region (0.2-1 MeV)

We name this region of energy as intermediate because the participation of core electrons starts to be significant in the contribution of stopping power. As shown in Figure 2-10, energy deposition of the proton decreases from 20 to 8.6 eV with proton energies from 0.2 to 1 MeV. This diminution of energy deposition relates to the interaction time between the charged particle and the target electrons, which reduces as a function of the charged particle energies. The three mechanisms detected above remain significant. However, the relativistic effect has not yet been effectively incorporated into energy deposition contribution.

2.3.3.3 High-energy region (> 1 MeV)

The Liénard-Wiechert potential allows in principle to account for relativistic effects in the ion-electron interaction. As shown in Figure 2-10, the energy loss calculated using the non-relativistic potential (red line) decreases smoothly as a function of proton energy, approaching zero at 635 MeV. It is a fact that the interaction time of the proton with the target electrons decreases with increasing proton energy. With a relativistic potential, the energy deposition decreases from 1 to 55 MeV, but more slowly than with a non-relativistic potential. Above 55 MeV the energy deposition increases further. The competition between the interaction time and the expansion of

the particle potential determines when the energy deposition curve starts to rise in the high energy region. In Figure 2-12, we quantify the relativistic potential effect as follow:

$$\%\Delta E_R = \frac{\Delta E_r - \Delta E_{nr}}{\Delta E_r} \times 100 \tag{2.4}$$

where ΔE_r and ΔE_{nr} is the energy deposition with relativistic and non-relativistic potential respectively.

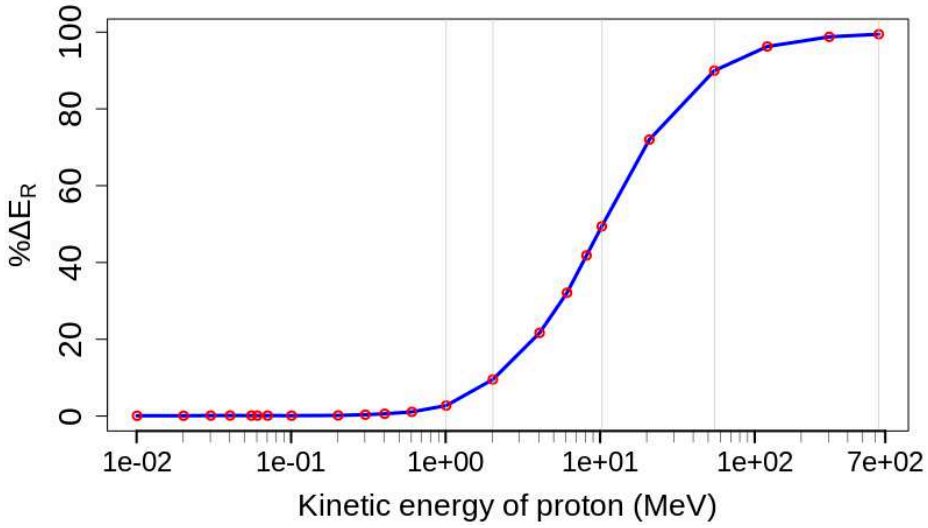

Figure 2-12: relativity effect on energy deposition. The x-axis (kinetic energy of a proton) has a logarithmic scale. The vertical lines correspond to remarkable values at 1, 2, 10.2, 55 and 635 MeV.

As already said, the effect of relativity is negligible below 1 MeV. It increases between 2 and 55 MeV and largely dominates above at 635 MeV. This is related to the characterization of the relativistic effect as shown in Section 2.2. An important feature of the energy transfer process at high relativistic velocities is the polarization of the target material. When a charged particle goes to high energy, the electric field becomes strongly concentrated in the transverse direction. Thus, when it passes into a dense material, the atom in the target material becomes strongly polarized. As a result, the ion's transverse electric field is shielded for electrons over a long distance by polarization. This effect becomes more and more important as the target density and proton energy increase. In our result, the relativistic effect between 2 and 55 MeV is more significant than that between 55 and 635 MeV. This may be related to the polarization effect at high relativistic energies.

To summarize we have pointed out different mechanisms that contribute to energy deposition as a function of the proton kinetic energy We have shown the necessity of including relativistic external potential above 1 MeV in RT-TD-ADFT simulations.

2.3.4 Relativistic effect on proton interaction cut-off

The size of the simulation is a technical issue in RT-TD-DFT. Due to simulation cost constraints, the finite microscopic simulation in space and time is used to simulate the interaction of charged particles with molecular systems. The electric field of the charged particle is able to propagate infinitely in universe; however, its strength depends on the distance. The strength of the electric field decreases with increasing distance. The constraint of simulation in space is limited by the interaction cut-off of charged particle with the target. That means the distance at which a charged particle can deposit energy into a target. As shown in Section 2.2, the interaction distance between a charged particle and an electron has been affected by relativity.

In this sub-section, we investigate how relativity affects energy deposition in our supramolecular system through proton interactions cut-off in the parallel and perpendicular directions of its propagation. It's important to note that in our simulations, the transverse and longitudinal electric fields are incorporated into both parallel and perpendicular interactions. These fields are not separable within the RT-TD-DFT simulations.

2.3.4.1 Parallel interaction cut-off

The parallel interaction cut-off refers to the distance at which a projectile begins to deposit energy into the target along the direction of its propagation. This type of interaction is important for optimizing the length of simulation sizes. The passage of a proton in our system is on the y-axis, and it arrives at the center of a supramolecular system at y = 0. In Figure 2-13, we analyse the energy variation in the supramolecular system, that is the variation of electronic energy with respect to the ground state as a function of the proton distance in the y-axis for the same trajectroy as above. Proton energies of 0.01 MeV (red line), 0.06 MeV (yellow line), 2 MeV

(orange line), 10.2 MeV (green line), 55 MeV (blue line) , and 635 MeV (purple line) are considered.

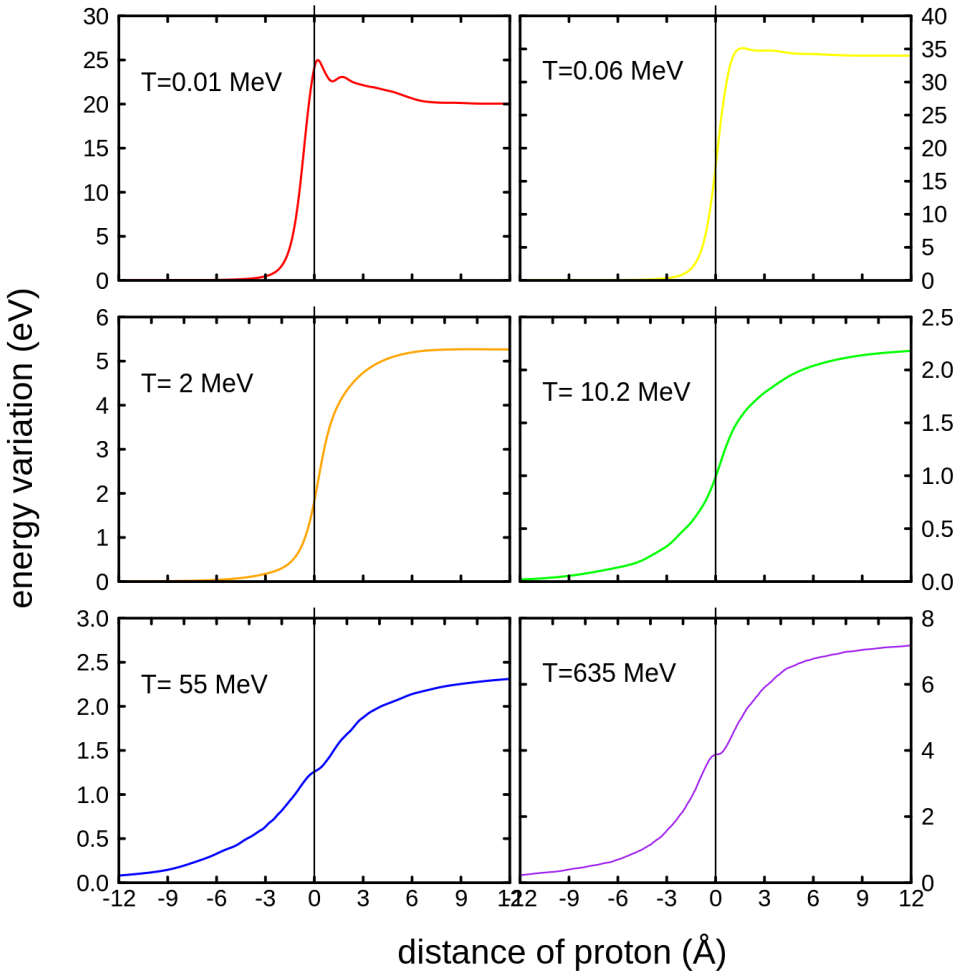

Figure 2-13: Variation of energy with respect to the ground state according to the distance of the proton from the supramolecular (y-axis = 0). T is the kinetic energy of proton.

For a proton a 0.01 MeV kinetic energy, energy deposition starts 3 Å before reaching the center of the supramolecular system, and it reaches a maximum at y = 0. Energy deposition decreases from 25 eV to 20 eV when the proton just leaves the center of the molecule. For the proton of 0.06 MeV (around the Bragg peak), energy deposition starts at the same distance, but the position of maximum energy deposition shifts. Energy deposition reduces by about 1 eV after collision. The behavior of energy deposition changes in the case of the proton at 2 MeV. Energy deposition begins at -6 Å and equilibrates at 6 Å without decrease after collision. The distance for energy deposition cut-off increases to about ±11 Å in the case of a 10.2 MeV proton. The essential difference in energy deposition is observed for the proton at 55 and 635 MeV. The proton

interacts with the supramolecular system at a very long distance. Energy deposition begins before -12 Å and continues anti-symmetrically after collision. The significant difference observed here is the fluctuation of energy deposition when the proton reaches a y = 0. This phenomenon is noted for the proton above 20 MeV and it increases with the proton's kinetic energy.

Furthermore, instead of the energy variation, the energy deposition per each time step ($\Delta E/\Delta t$) (energy variation between simulation time steps) with respect to the distance of the proton is shown in Figure 2-14.

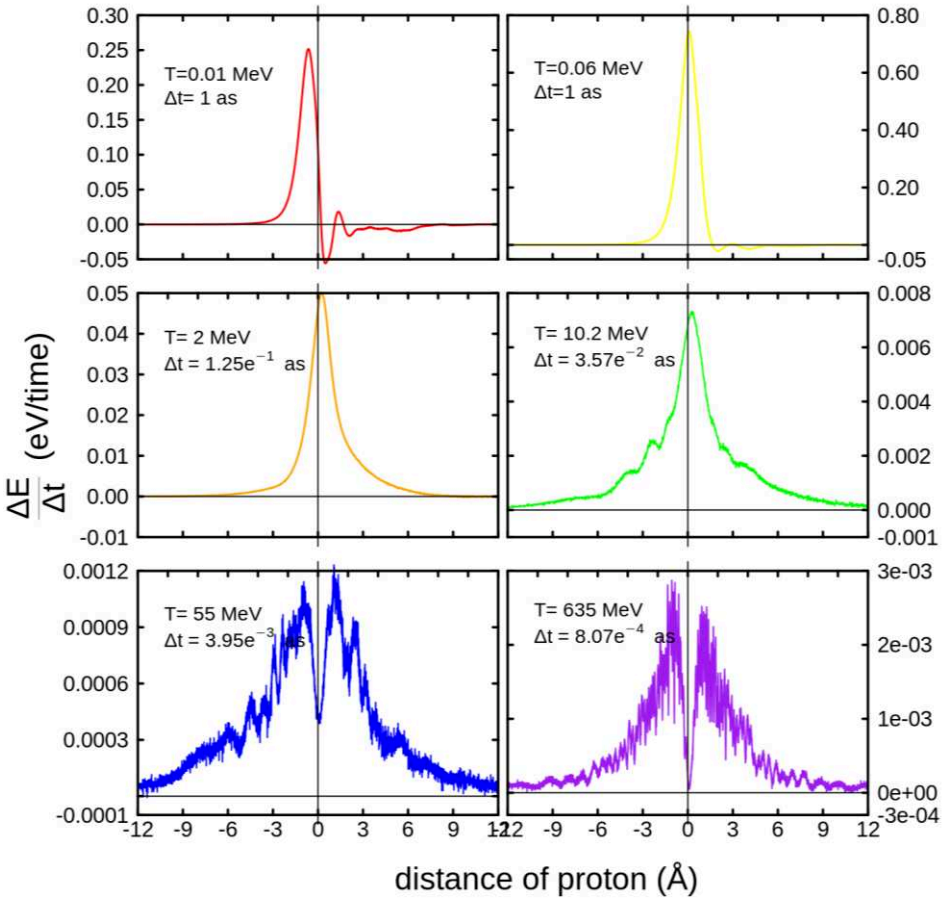

Figure 2-14: Energy deposition per simulation time step over the distance of the proton from the supramolecular system (y-axis = 0). T and dt are the kinetic energy of proton and simulation time steps, respectively.

In Figure 2-14, the time step (Δt) at 635 MeV is very small because the velocity of the proton is very high (2.4 Å/as). Several processes may be engaged in the proton's interaction with the electrons even at this distance. As a result, if the time steps in this system are greater than the physical process time, the electron dynamics becomes unstable. Therefore, we used 8.07×10^{-4} as time steps to ensure the simulation stability.

$\Delta E/\Delta t$ increases for 0.01 MeV before the proton strikes the supramolecular center and then sharply declines. When the proton just departs from the center, it becomes negative. It means that during the collision, some electrons undergo de-excitation. Because the supramolecular system is two-dimensional, such a de-excitation and shift of the $\Delta E/\Delta t$ maximum position are unexpected. The maximum of $\Delta E/\Delta t$ should be located at a center of the molecule and symmetric around this position. At 0.06 MeV, the maximum of $\Delta E/\Delta t$ is located at the center, although it is not symmetric. The curve drops faster after collision. In the case of 2 MeV and 10.2 MeV protons, a small $\Delta E/\Delta t$ shift towards after y=0 is noted. However, a significant qualitative difference is noticeable for 55 MeV and 635 MeV protons. As the kinetic energy of the proton increases, $\Delta E/\Delta t$ approaches zero at the molecule's center. Around y=0, $\Delta E/\Delta t$ sharply and symmetrically augments to maximum and trends to long-distance (more than 12 Å).

We return to Figure 2-3 in Section 2.2 to explain the difference in behavior of proton interaction by changing its energy. The contribution of the longitudinal electric field E_y is anti-symmetric around the collision point. Therefore, it is expected that the energy does not transfer by E_y. On the other hand, E_y has a prominent role to polarize electrons for low ion kinetic-energy as the electron's velocities are faster than that of the ion. In addition, we can see the contribution of E_y in energy deposition of low-energy particle because electrons have time to move before collision, so that the system's energy increases. In contrast, the same amount of energy that was increased by E_y is reduced after collision. This is called adiabatic energy transfer. The contribution of E_y to energy deposition increases with decreasing ion velocity as the velocity of number of electrons are faster than the charged particle.

Returning to our result displayed on Figure 2-14, the E_y contribution is clearly visible for the proton at 0.01 MeV and decreases at 0.06 MeV. It is not visible for the proton from 2 MeV to 635 MeV. When calculating the energy deposition and the stopping power for a complete collision (taking into account the energy deposition in the range before and after the collision), this energy deposition value is automatically discarded. Sometimes this value is extracted or given as a statistical error bar in the stopping power calculation by finding the adiabatic energy surface from a projectile particle simulation at zero energy[26]. This method is very time consuming and may be inaccurate because, as shown in Figure 2-3 (d) the frequency produced by a charged particle is

highly dependent on the energy of the particle. As the energy of the particle decreases, the frequency decreases, resulting in different excitations of the electrons.

The transverse electric field E_z is the main component responsible for energy deposition by charged particles. It is perfectly symmetric and leads to increasing energy deposition before and after collision. If an electron is faster than the particle, it will move toward the projectile, reducing the distance (impact parameter) between the electron and the projectile and increasing energy deposition (Barkas effect). Of course, this process is in competition with the decrease of energy deposition by screening the electric field via electron polarization. As a consequence, it is expected that energy deposition is more important after collision compared to before collision. Returning to our result in Figure 2-14, energy deposition after collision declines more slowly than before collision at 2 MeV and 10.2 MeV. It is a fact that the effect of polarization is not essential at this energy, but the Barkas effect remains important because a number of electrons are even faster in the deep valence and the core shells. This behavior disappears at 55 MeV and 635 MeV because the movement of electrons is negligible compared to the velocity of projectile (Barkas effect = 0).

The interesting question in Figure 2-14, is the significant reduction of energy deposition at 55 MeV and 635 MeV in the molecule's center (y=0). As shown in Figure 2-3 (b), E_z is maximum in this position and energy deposition is expected to be maximum. As explained before, the electric field of charged particles concentrates in the perpendicular direction, which strongly polarizes the electrons of the target; consequently, the portion of the electric field is shielded (density effect). This effect increases with the particle's energy at the relativistic region. So, we see this sharp decreasing energy deposition at y=0. Thanks to the RT-TD-ADFT simulation, this effect can be seen at the microscopic level.

In addition, Figure 2-14 indicates that increasing the proton's energies, particularly at relativistic velocity, causes energy deposition to begin at a greater distance from the center of the molecule. The cut-off interaction distance between the charged particle and the electrons of the target is more than 12 Å before and after collision at 635 MeV. This is due to the relativistic effect. Because the relativity contracts the distance between an electron and a charged particle, the interaction between them increases, as seen in Figure 2-3 (c).

2.3.4.2 *Perpendicular interaction cut-off*

The perpendicular interaction cut-off refers to the distance at which a projectile continues to deposit energy into the target perpendicular to the direction of its propagation. This interaction is crucial for optimizing the width of simulation sizes. To deal with the influence of perpendicular interaction on energy deposition depending on the proton's energies, we performed a new trajectory that was located on the center of supramolecular system, as shown by the green ball in Figure 2-15. By assuming the full supramolecule as a large system, we calculated the energy deposition (ΔE_l) for proton energies ranging from 0.01-635 MeV. In order to evaluate the effect of perpendicular interaction, we assume that the contribution of energy deposition of molecules around the guanine comes from the perpendicular interaction. Then, another simulation has been done with only guanine for the same trajectory (small system). The minimum distance between the proton and the near fragment around guanine is 3.9 Å. Their energy deposition is referenced as (ΔE_s).

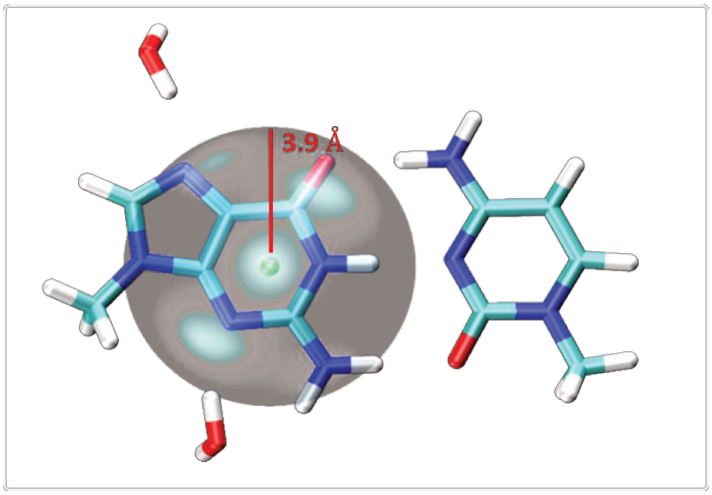

Figure 2-15: The supramolecular system (guanine, cytosine, and pairs of water molecules) is shown in licorice representation. The green ball represents the proton's position at the molecule's center (y=0). The blue circle corresponds to the radius of 3.9 Å around the proton.

The effect of perpendicular interaction on energy deposition (ΔE_w) is expressed as a percentage in Figure 2-16, and it is calculated as follows:

$$\%\Delta E_W = \frac{\Delta E_l - \Delta E_s}{\Delta E_l} 100 \qquad (2.5)$$

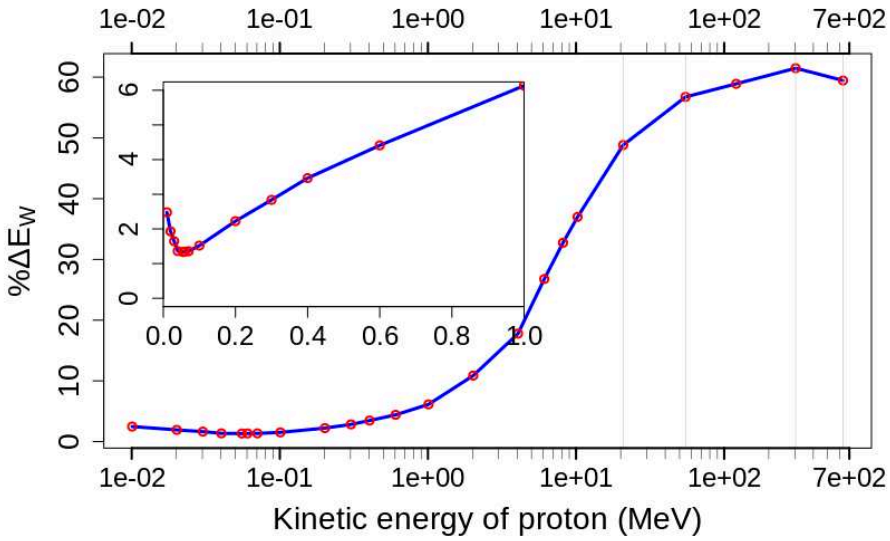

Figure 2-16: percentage of energy deposition difference between supramolecular system (large system) and guanine system (small system) as a function of the proton energies. The x-axis (kinetic energy of a proton) has a logarithmic scale. The vertical lines correspond to the results obtained for proton kinetic energies of 20.8, 55, 304, and 635 MeV. The inset represents the zoom of the figure from 0.01 to 1 MeV.

The minimum effect of perpendicular interaction is 1.3% at Bragg's peak (0.055 MeV), and it augments to 2.5% at 0.01 MeV. Moreover, $\%\Delta E_w$ significantly raises from 1.3% to 49% at 0.06 MeV and 20.8 MeV, successively. However, this augmentation slows down between 55 and 304 MeV, where it is 57% and 61%, respectively, and drops to 59% at 635 MeV. In general, $\%\Delta E_w$ increases according the proton energy. The different responses of $\%\Delta E_w$ to proton energies are related to the mechanisms involved in the processes of energy deposition in each energy region.

The main difference between large and small systems is that the small system omits energy deposition produced by long-distance collision with cytosine and the two water molecules. The augmentations of $\%\Delta E_w$ below Bragg's peak may be related to the contribution of energy deposition for long-distance and short distance collision. Energy deposition at the Bragg's peak is huge, that is why the long-distance contribution is expected to be small compared to the total energy deposition. By consequence, the $\%\Delta E_w$ at the Bragg peak may be minimal with respect to lower energy, despite the polarization effect, which is greater at lower energy.

In the other side, the augmentation of $\%\Delta E_w$ after Bragg's peak is related to the polarization of the target electrons and relativistic effect. The accumulation of electrons around the proton leads

to the screening of the proton's electric field for long-distance electrons; as a result, the contribution of the long-distance energy deposition decreases. This effect declines in accordance with the proton energies; that is why, $\%\Delta E_w$ augments with proton energies.

The accumulation of electrons around the proton becomes insignificant after about 4 MeV, as shown previously in Figure 2-11. The relativistic effect, on the other hand, plays a role in the increase of $\%\Delta E_w$ from 4 MeV to 20 MeV. As explained in Section 2.2 (Table 2-1), the Liénard-Wiechert potential is expended perpendicular to the axis of the proton propagation with increasing relativistic velocities. In particular, this effect leads to increase the long-distance energy deposition contribution because the interaction of a long-distance electron with a proton crucially depends on the perpendicular interaction. That is why, $\%\Delta E_w$ significantly augments. Whereas at high-relativistic, the augmentation of $\%\Delta E_w$ goes slowly. It is related to the local-electron polarization by relativistic effect (density effect), which particularly screens the long-distance interaction. It is obvious that $\%\Delta E_w$ marginally reduces at 635 MeV compared to 304 MeV.

Finally, the result demonstrates that the relativistic effect increases not only the energy deposition but also its spatial contribution. The contraction of distance by relativity increases the interaction of charged particle with electrons in molecular systems at longer distances.

This result is crucial to indicate the initial position of the charged particle in the RT-TD-ADFT simulation which must be far enough from the target to avoid the perturbation of the electron density at the beginning of the simulation. In addition, this result could be helpful for optimizing simulation size in the calculation of energy deposition or stopping power for condensed phase. In general, a cluster or a simulation box of target molecules is used for calculating stopping power. The result of stopping power highly depends on the size of the system (simulation size). If it is small, a part of energy deposition discards in the calculation that comes from long-distance contribution. According to our result, by increasing the projectile's energies a larger size of simulation is required, otherwise, the result of stopping power at high energies may be underestimated. Therefore, it is advisable to optimize the simulation's size at higher energies.

2.4 Conclusion

In this chapter, we have studied charged particle interactions, with a particular focus on the impact of relativity in the charged particle-electron interaction in RT-TD-ADFT calculations for energy deposition. From the analysis of the electric field generated by the charged particle, we have established that the energy transfer from a charged particle essentially depends on the transverse interaction. In addition, by the relativistic effect, the transversal interaction expands in terms of time interaction and spatial contribution. Dilatation of time and contraction of distance lead to a red-shift and an increase in the intensity of the charged particle spectra. In addition, we have shown that the interaction of charged particles starts at a longer distance than non-relativistic charged particles. Moreover, our findings also indicate that the frequencies generated by a charged particle are substantially dependent on the particle's energy and the collision distance (impact parameters).

From our RT-TD-ADFT simulations, we have discussed distinct mechanisms that contribute to energy deposition at different proton kinetic energies. Furthermore, we have underlined the necessity of including a relativistic external potential above 1 MeV in RT-TD-ADFT simulations for accurate energy deposition calculations. Neglecting relativity in these simulations can significantly impact the results, especially as proton energies increase. In addition, the relativity affects the spatial distribution of energy deposition.

Our study has thus provided valuable insights into the interactions and energy transitions of charged particles, with a specific focus on their kinetic energy. Importantly, it emphasizes the inclusion of relativistic effects of charged particles in RT-TD-ADFT calculations.

2.5 Perspective

The main challenges in calculating stopping power with RT-TD-DFT, which make it non-useful for applications such as track structure simulation or range penetration of charged particles, are related to trajectory sampling, energy range, and transportability of trajectories. The calculation of stopping power requires a high number of trajectories to ensure all possible impact

parameters, particularly for biomolecules, and this requires substantial computational resources. As mentioned in introduction, the range of the stopping power curve calculated by RT-TD-DFT is limited to energies below 2 MeV, whereas, for applications like radiotherapy, energies below 350 MeV are required. Additionally, the trajectory results of RT-TD-DFT are not transportable, meaning that any small change in system geometry requires the simulation to be repeated, thus limiting the applicability of RT-TD-DFT data for track structure simulations.

The range of energy in the stopping power curve can be extended by incorporating the relativistic effects of charged particles. To overcome the issue of transportability of trajectories, machine learning could be a useful tool. By training a model using RT-TD-DFT trajectory data, the trajectories can be recalculated in different situations without the need for new RT-TD-DFT calculations. Our preliminary results are encouraging in this direction.

REFERENCES:

(1) Mohan, R.; Grosshans, D. Proton Therapy – Present and Future. *Advanced Drug Delivery Reviews* **2017**, *109*, 26–44. https://doi.org/10.1016/j.addr.2016.11.006.

(2) Van Abbema, J. K.; Van Goethem, M.-J.; Mulder, J.; Biegun, A. K.; Greuter, M. J. W.; Van Der Schaaf, A.; Brandenburg, S.; Van Der Graaf, E. R. High Accuracy Proton Relative Stopping Power Measurement. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* **2018**, *436*, 99–106. https://doi.org/10.1016/j.nimb.2018.09.015.

(3) Montanari, C. C.; Archubi, C. D.; Mitnik, D. M.; Miraglia, J. E. Energy Loss of Protons in Au, Pb, and Bi Using Relativistic Wave Functions. *Phys. Rev. A* **2009**, *79* (3), 032903. https://doi.org/10.1103/PhysRevA.79.032903.

(4) Montanari, C. C.; Miranda, P. A.; Alves, E.; Mendez, A. M. P.; Mitnik, D. M.; Miraglia, J. E.; Correa, R.; Wachter, J.; Aguilera, M.; Catarino, N.; Da Silva, R. C. Stopping Power of Hydrogen in Hafnium and the Importance of Relativistic 4 f Electrons. *Phys. Rev. A* **2020**, *101* (6), 062701. https://doi.org/10.1103/PhysRevA.101.062701.

(5) Montanari, C. C.; Mitnik, D. M.; Archubi, C. D.; Miraglia, J. E. Energy Loss of Protons in W Using Fully Relativistic Calculations and Mean Excitation Energies of W, Au, Pb, and Bi. *Phys. Rev. A* **2009**, *80* (1), 012901. https://doi.org/10.1103/PhysRevA.80.012901.

(6) Li, S.-M.; Mao, F.; Zhao, X.-D.; Jin, W.-Q.; Zuo, W.-Q.; Li, B.-S.; Wang, F.; Zhang, F.-S. Contribution of f -Electron Excitation to Electronic Stopping Power of Platinum for Protons. *Phys. Rev. B* **2022**, *106* (1), 014103. https://doi.org/10.1103/PhysRevB.106.014103.

(7) Collaboration, C. Measurement of the Muon Stopping Power in Lead Tungstate. *J. Inst.* **2010**, *5* (03), P03007–P03007. https://doi.org/10.1088/1748-0221/5/03/P03007.

(8) Whittemore, W. L.; Street, J. C. The Density Effect for Cosmic-Ray Mesons. *Phys. Rev.* **1949**, *76* (12), 1786–1791. https://doi.org/10.1103/PhysRev.76.1786.

(9) Hereford, F. L. The Specific Primary Ionization and Energy Loss of Fast Electrons in Matter. *Phys. Rev.* **1948**, *74* (5), 574–579. https://doi.org/10.1103/PhysRev.74.574.

(10) Sternheimer, R. M. The Density Effect for the Ionization Loss in Various Materials. *Phys. Rev.* **1952**, *88* (4), 851–859. https://doi.org/10.1103/PhysRev.88.851.

(11) Sternheimer, R. M.; Berger, M. J.; Seltzer, S. M. Density Effect for the Ionization Loss of Charged Particles in Various Substances. *Atomic Data and Nuclear Data Tables* **1984**, *30* (2), 261–271. https://doi.org/10.1016/0092-640X(84)90002-0.

(12) Ziegler, J. F. Stopping of Energetic Light Ions in Elemental Matter. *Journal of Applied Physics* **1999**, *85* (3), 1249–1272. https://doi.org/10.1063/1.369844.

(13) Groom, D. E.; Klein, S. R. Passage of Particles through Matter. *Eur. Phys. J. C* **2000**, *15* (1–4), 163–173. https://doi.org/10.1007/BF02683419.

(14) Wu, X.; Alvarez-Ibarra, A.; Salahub, D. R.; De La Lande, A. Retardation in Electron Dynamics Simulations Based on Time-Dependent Density Functional Theory. *Eur. Phys. J. D* **2018**, *72* (12), 206. https://doi.org/10.1140/epjd/e2018-90219-3.

(15) Gu, B.; Cunningham, B.; Muñoz Santiburcio, D.; Da Pieve, F.; Artacho, E.; Kohanoff, J. Efficient *Ab Initio* Calculation of Electronic Stopping in Disordered Systems via Geometry Pre-Sampling: Application to Liquid Water. *J. Chem. Phys.* **2020**, *153* (3), 034113. https://doi.org/10.1063/5.0014276.

(16) Caro, M.; Correa, A. A.; Artacho, E.; Caro, A. Stopping Power beyond the Adiabatic Approximation. *Sci Rep* **2017**, *7* (1), 2618. https://doi.org/10.1038/s41598-017-02780-3.

(17) Maliyov, I. Ionic Irradiation of Materials : Real-Time Dynamics of Electronic Excitations. Theses, Université Paris-Saclay, 2019. https://theses.hal.science/tel-02495699.

(18) Maliyov, I.; Crocombette, J.-P.; Bruneval, F. Electronic Stopping Power from Time-Dependent Density-Functional Theory in Gaussian Basis. *Eur. Phys. J. B* **2018**, *91* (8), 172. https://doi.org/10.1140/epjb/e2018-90289-y.

(19) Cabrera-Trujillo, R.; El Hachimi, A. G. Calculation of the Electronic, Nuclear, Rotational, and Vibrational Stopping Cross Sections for H Atoms Irradiation on H_2, N_2 and O_2 Gas Targets at Low Collision Energies. *J. Phys. B: At. Mol. Opt. Phys.* **2020**, *53* (13), 135203. https://doi.org/10.1088/1361-6455/ab8834.

(20) Martínez-Flores, C.; Domínguez-Gutiérrez, F. J.; Cabrera-Trujillo, R. Rotational and Vibrational Effects on the Energy Loss of Hydrogen Colliding on Glycine at Low Irradiation Energies. *Radiation Physics and Chemistry* **2020**, *166*, 108513. https://doi.org/10.1016/j.radphyschem.2019.108513.

(21) Yost, D. C.; Yao, Y.; Kanai, Y. Examining Real-Time TDDFT Non-Equilibrium Simulations for the Calculation of Electronic Stopping Power. **2018**. https://doi.org/10.48550/ARXIV.1805.01377.

(22) Bethe, H. Bremsformel for Elektronen relativistischer Geschwindigkeit. *Z. Physik* **1932**, *76* (5–6), 293–299. https://doi.org/10.1007/BF01342532.

(23) Lindhard, J. ON THE PROPERTIES OF A GAS OF CHARGED PARTICLES. *Kgl. Danske Videnskab. Selskab Mat.-fys. Medd.* **1954**, *Vol:*

(24) Montanari, C. C.; Miraglia, J. E. Chapter Seven - The Dielectric Formalism for Inelastic Processes in High-Energy Ion–Matter Collisions. In *Theory of Heavy Ion Collision Physics in Hadron Therapy*; Belkić, D., Ed.; Advances in Quantum Chemistry; Academic Press, 2013; Vol. 65, pp 165–201. https://doi.org/10.1016/B978-0-12-396455-7.00007-8.

(25) Dielectric Stopping Theory. In *Particle Penetration and Radiation Effects*; Cardona, M., Fulde, P., von Klitzing, K., Queisser, H.-J., Merlin, R., Störmer, H., Series Eds.; Springer Series in Solid-State Sciences; Springer Berlin Heidelberg: Berlin, Heidelberg, 2006; Vol. 151, pp 141–180. https://doi.org/10.1007/3-540-31718-X_5.

(26) Correa, A. A. Calculating Electronic Stopping Power in Materials from First Principles. *Computational Materials Science* **2018**, *150*, 291–303. https://doi.org/10.1016/j.commatsci.2018.03.064.

(27) Schleife, A.; Kanai, Y.; Correa, A. A. Accurate Atomistic First-Principles Calculations of Electronic Stopping. *Phys. Rev. B* **2015**, *91* (1), 014306. https://doi.org/10.1103/PhysRevB.91.014306.

(28) Barkas, W. H.; Dyer, J. N.; Heckman, H. H. Resolution of the $\ensuremath\Sigma^\ensuremath-$-Mass Anomaly. *Phys. Rev. Lett.* **1963**, *11* (1), 26–28. https://doi.org/10.1103/PhysRevLett.11.26.

(29) Alvarez-Ibarra, A.; Parise, A.; Hasnaoui, K.; de la Lande, A. The Physical Stage of Radiolysis of Solvated DNA by High-Energy-Transfer Particles: Insights from New First Principles Simulations. *Phys. Chem. Chem. Phys.* **2020**, *22* (15), 7747–7758. https://doi.org/10.1039/D0CP00165A.

(30) Dikshit, B. Space–Time Diagram Approach in Derivation of Lienard–Wiechert Potential for a Moving Point Charge. *Can. J. Phys.* **2013**, *91* (7), 519–521. https://doi.org/10.1139/cjp-2013-0073.

(31) Köster, A.; Geudtner, G.; Calaminici, P.; Casida, M.; Dominguez, V.; Flores-Moreno, R.; Gamboa, G.; Goursot, A.; Heine, T.; Ipatov, A.; others. DeMon2k, Version 6, The DeMon Developers, Cinvestav, Mexico City (2018). *Google Scholar There is no corresponding record for this reference*.

(32) Dunlap, B. I.; Rösch, N.; Trickey, S. B. Variational Fitting Methods for Electronic Structure Calculations. *Molecular Physics* **2010**, *108* (21–23), 3167–3180. https://doi.org/10.1080/00268976.2010.518982.

(33) Calaminici, P.; Alvarez-Ibarra, A.; Cruz-Olvera, D.; Dominguez-Soria, V.; Flores-Moreno, R.; Gamboa, G.; Geudtner, G.; Goursot, A.; Mejía-Rodríguez, D.; Salahub, D.; others. Auxiliary Density Functional Theory: From Molecules to Nanostructures. *Handbook of computational chemistry* **2016**, 1.

(34) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865–3868. https://doi.org/10.1103/PhysRevLett.77.3865.

(35) Godbout, N.; Salahub, D. R.; Andzelm, J.; Wimmer, E. Optimization of Gaussian-Type Basis Sets for Local Spin Density Functional Calculations. Part I. Boron through Neon, Optimization Technique and Validation. *Can. J. Chem.* **1992**, *70* (2), 560–571. https://doi.org/10.1139/v92-079.

(36) Calaminici, P.; Flores–Moreno, R.; Köster, A. M. A Density Functional Study of Structures and Vibrations of Ta3O and Ta3O−. *Computing Letters* **2005**, *1* (4), 164–171. https://doi.org/10.1163/157404005776611420.

(37) Wu, Q.; Yang, W. Empirical Correction to Density Functional Theory for van Der Waals Interactions. *The Journal of Chemical Physics* **2002**, *116* (2), 515–524. https://doi.org/10.1063/1.1424928.

(38) Wu, X.; Teuler, J.-M.; Cailliez, F.; Clavaguéra, C.; Salahub, D. R.; de la Lande, A. Simulating Electron Dynamics in Polarizable Environments. *Journal of Chemical Theory and Computation* **2017**, *13* (9), 3985–4002. https://doi.org/10.1021/acs.jctc.7b00251.

(39) Carmona-Espíndola, J.; Gázquez, J. L.; Vela, A.; Trickey, S. B. Generalized Gradient Approximation Exchange Energy Functional with Correct Asymptotic Behavior of the Corresponding Potential. *J. Chem. Phys.* **2015**, *142* (5), 054105. https://doi.org/10.1063/1.4906606.

(40) Gómez Pueyo, A.; Marques, M. A. L.; Rubio, A.; Castro, A. Propagators for the Time-Dependent Kohn–Sham Equations: Multistep, Runge–Kutta, Exponential Runge–Kutta, and Commutator Free Magnus Methods. *J. Chem. Theory Comput.* **2018**, *14* (6), 3040–3052. https://doi.org/10.1021/acs.jctc.8b00197.

(41) Cheng, C.-L.; Evans, J. S.; Van Voorhis, T. Simulating Molecular Conductance Using Real-Time Density Functional Theory. *Phys. Rev. B* **2006**, *74* (15), 155112. https://doi.org/10.1103/PhysRevB.74.155112.

(42) Castro, A.; Marques, M. A. L.; Rubio, A. Propagators for the Time-Dependent Kohn–Sham Equations. *The Journal of Chemical Physics* **2004**, *121* (8), 3425–3433. https://doi.org/10.1063/1.1774980.

(43) Calaminici, P.; Jug, K.; Köster, A. M. Density Functional Calculations of Molecular Polarizabilities and Hyperpolarizabilities. *The Journal of Chemical Physics* **1998**, *109* (18), 7756–7763. https://doi.org/10.1063/1.477421.

(44) Faure, A.; Gorfinkiel, J. D.; Morgan, L. A.; Tennyson, J. GTOBAS: Fitting Continuum Functions with Gaussian-Type Orbitals. *Computer Physics Communications* **2002**, *144* (2), 224–241. https://doi.org/10.1016/S0010-4655(02)00141-8.

(45) Tandiana, R.; Omar, K. A.; Luppi, E.; Cailliez, F.; Van-Oanh, N.-T.; Clavaguéra, C.; De La Lande, A. Use of Gaussian-Type Functions for Describing Fast Ion-Matter Irradiation with Time-Dependent Density Functional Theory. *J. Chem. Theory Comput.* **2023**, *19* (21), 7740–7752. https://doi.org/10.1021/acs.jctc.3c00656.

CHAPTER THREE

EFFECT OF HYDROGEN BONDS ON ENERGY DEPOSITION

Contents

3 Effect of hydrogen bonds on energy deposition

Hydrogen bonding is an essential non-covalent interaction for molecular biology. It is responsible for the structural organization of macrobiomolecules such as proteins, DNA, or carbohydrates, and is essential for the functional properties of biomolecules[1]. Hydrogen bonding is weak compared to covalent chemical bonding, but it can piled-up the biomolecules and alter slightly their electronic structure. These modifications may be impacting the electronic stopping power (S_e) or energy deposition in biological systems[2]. This effect is crucial for the semi-empirical models based on Bragg's additive rule. According to this rule[3], the stopping power of an atom does not depend on the chemical nature and physical state of the system.

In this work, we investigate the effect of hydrogen bonding on energy deposition over a wide range of proton energies from 0.01 to 630 MeV in biological molecules. We consider a Watson-Crick guanine-cytosine nucleobase-pair with two water molecules hydrogen-bonded to the guanine nitrogen atoms, which are essential molecules of DNA structure.

This chapter starts with an introduction to Bragg's additive rule, and we discuss the effect of chemical and physical bonds on this rule. Section 3.2 details the methodology and strategies used to estimate the effect of hydrogen bonds on energy deposition in our system. In Section 3.3, we demonstrate the results of the effects of hydrogen bonds, obtained from 667 RT-TD-ADFT trajectories for protons with energy ranging from 0.03 to 0.3 MeV. Section 3.4 refers to the results obtained in a wide range of proton energy (0.01-635 MeV) based on four trajectories. The chapter concludes with two sections; Section 3.5 is a conclusion, and Section 3.6 corresponds to the discussion and perspectives about the fluctuations in the effects of hydrogen bonds on energy deposition at high proton energies.

3.1 Bragg's additive rule

The Bragg's additive rule (BAR) for calculating stopping power is essential for semi-empirical models such as SRIM[4] and PSTAR[5]. In 1905, Bragg and Kleeman[3] proposed that the stopping power of a compound ($S_{compound}$)(e.g. molecules, aggregates) can be estimated by a linear combination of the stopping power of its individual atoms. Meanwhile, it does not depend on the chemical nature and physical state of the compound. Here the term of the stopping power, means that *the stopping power is normalized by the density of the target material*.

$$s_{compound}(T_p) = \sum_x r_x \, S_x(T_p) \tag{3.1}$$

Where, r_x is the ratio of the number density of the atom x to the total density of all atoms in the target. S_x corresponds to the stopping power of atom x as a function of kinetic energy of projectile T_p.

In reality, the electronic structure is altered by any change in atomic bonds. Consequently, atomic excitation and ionization processes will be changed. So that, the brut BAR needs to add two types of corrections, chemical bond and phase state.

Chemical bond correction (CBC): this kind of correction has been extensively investigated by experimentalists[2]. They showed that the accuracy of BAR is reasonable for well high energetic projectile. Whereas, in the lower energy region near and below the maximum S_e, the overestimation of BAR may be up to 50%. It is expected because chemical and ionic bonds modify only the energy of valence electrons without significant alteration of core electrons. Moreover, at the lower energies the stopping is dominated by the valence electrons, but at high energies, it is controlled by the core electrons. Since the valence electron ratio is higher in light atoms than in heavy atoms, the light atoms have a larger bonding effect on stopping power. In general, CBC can be incorporated into BAR by scaling the stopping curve with available experimental data for this compound, or by employing Bethe-Bloch formula at high energies[6], for instance in SRIM.

Kohanoff *et al.* investigated Bragg's additivity rule for vapor water from hydrogen and oxygen molecules using RT-TD-DFT simulation[7]. They found a good agreement between RT-TD-DFT

simulation with available experimental data and SRIM simulation. However at very low energies, the position of Bragg peak in RT-TDDFT is red-shift by 15 keV, and the peak value of S_e is underestimated by 4%. The authors assumed that this divergence between RT-TD-DFT and experimental data comes from the quality of the atomic basis sets.

Phase state correction (PSC): when the physical state changes, not only the intermolecular force changes, but the intermolecular distances between molecules are also significantly modified. Two factors influence S_e. First, moving from gas phase to condensed phase, intermolecular forces are established, for instance crystalline bonds in solid, hydrogen bonds, and dispersion forces. These forces are weaker compared to chemical bonds, but they can modify the energy of valence electrons. The same effect of chemical bond will be expected but, weaker. Second factor, the intermolecular distances significantly reduce and aggregate the molecules. So that, the environment of molecules hardly alters.

The effect of physical state on S_e has been marginally investigated due to the high uncertainty within the experimental stopping power for liquid and solid phase. For instance, due to the high incertitude in experimental S_e for liquid water (which is an essential molecule in radiation chemistry), the majority of Monte Carlo codes use the stopping cross section of liquid water calculated from Lindhard's and Bethe-Bloch theory[8]. Despite this, we discuss here some experimental data for water to clarify this effect. For 0.3-2 MeV He^{2+} ions in vapor and ice water Matteson *et al.* found that the S_e of vapor water is greater than ice water by 12-4% respect to energy of He^{2+} ions[9]. Thwaites reported the S_e for He^{2+} ions in vapor and liquid water for energy range between 1.75-4.75 MeV[10]. He found also that the S_e in vapor is greater by 4.3-1.03% as a function of kinetic energy of He^{2+} ions. In other side, Palmer and Akhavan-Rezayat investigated also the S_e for 0.5-8 MeV of He^{2+} ions (equivalent 0.125-2 MeV proton energies) in vapor and liquid water[11]. They found large fluctuation of phase effect on S_e. The stopping power of vapor water is significantly greater below 1.5 MeV and above 2.5 MeV. It is about 14% at 0.5 and 8 MeV. In the case of proton, the S_e of protons are reported in vapor water by (Phillips[12]) and ice D_2O by (Wenzel and Whaling[13]) for energies 0.18-541 keV. The phase effect for vapor is 10% at 541 keV and tends to 14% at 18 keV. In general, experimental data show that the phase effect is

more important at low energy compared to high energy such as observed in the case of chemical bonds.

In summary, both covalent and non-covalent interactions affect the validity of BAR. One of them may be more significant than the other, depending on the projectile and the target. However, since the chemical bond is stronger than the physical bond, we cannot say that it is more crucial. For example, the difference between stopping He^{2+} ions in vapor water and those calculated by pure BAR from oxygen and hydrogen molecules is only 1%. But, this difference between vapor and ice water raises to 9% at Bragg peak[4].

All of this encourages us to investigate how hydrogen bonds influence the interaction of a charged particle with biological molecules.

3.2 Methodology and strategies

This section includes the methodology and the strategies used to determine the effect of hydrogen bonds on energy deposition in our system.

3.2.1 Methodology

We consider the same simple system as in Chapter Two, namely a guanine-cytosine nucleobase with a pair of water molecules which is a two-dimensional molecular system in the Oxz plane (as shown in Figure 3-1).

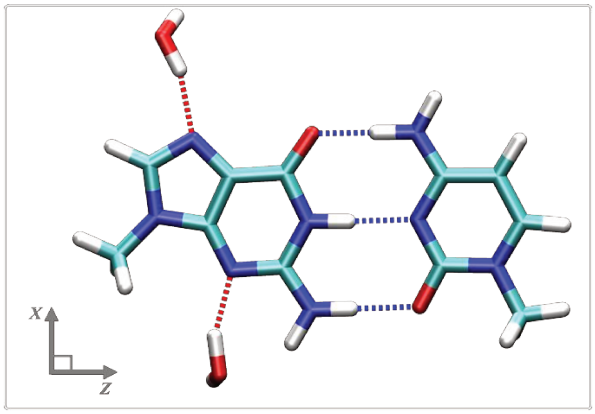

Figure 3-1: the supramolecular system considered in the Chapter: guanine (left), cytosine (right), and two water molecules. The dashed lines make visible the intermolecular hydrogen bonds.

We employed RT-TD-ADFT, as available in our customized version of deMon2k (version 4.4.5)[14], to assess the quantity of electronic energy deposition during a collision. The PBE functional was our choice for correlation, while we selected the Correct Asymptotic Potential functional (CAP)[15] for exchange. The simulations were performed with TZVP-FIP2 (stands for TZVP augmented with second-order field-induced polarization functions)[16] basis set and GEN-A2* auxiliary basis set[17]. We propagated the simulations with the combination of the second-order Magnus propagator[18] and a predictor-corrector algorithm[19]. The Kohn-Sham matrix's exponential was evaluated using a Taylor expansion[20], comprising 45 terms. In our simulations, the interaction of charged particles was described using the Liénard-Wiechert potential. Additionally, we abstained from employing a complex absorbing potential to remove electrons in the continuum state. Our focus in this chapter was primarily on calculating energy deposition rather than ionization states. As depicted in Figure 3-2, the energy deposition remains entirely conserved after the collision.

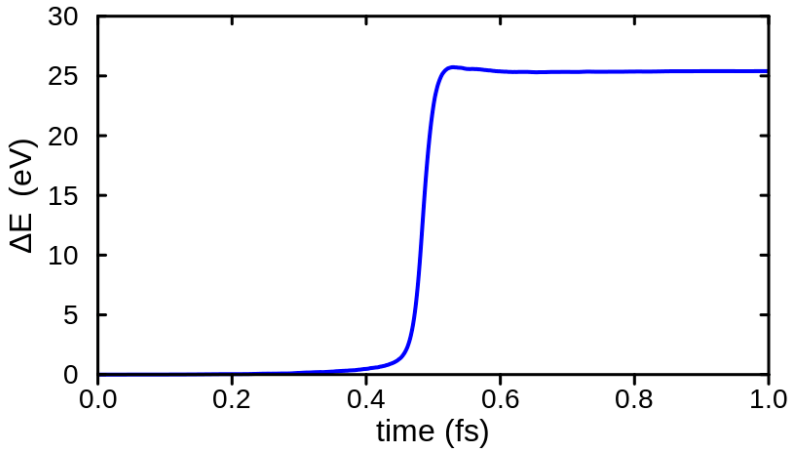

Figure 3-2: Energy variation relative to the ground state energy over time for a 0.2 MeV proton.

3.2.2 Strategies for determining the energy deposition of hydrogen bonds

We attempt to illustrate an ideal experimental concept to estimate stopping power in both condensed and gaseous phases, before outlining our approach to assessing the effect of the hydrogen bonds in our system. Experimentally, when an ion beam interacts with a target, the ions are able to sample the target atoms or molecules in all directions. It means that all impact parameters between ions and target atoms take place. Thus, at the microscopic level the number of trajectories that impact an atom in the gas and condensed phases are not the same. The

number of impacting trajectories increases and is limited by the interatomic distance. In Figure 3-3, we show the simplified form of the interaction of ions beam with atoms in gaseous and condensed phases. Considering that the density is 1 atom/cm^3 in condensed phase and 0.5 atom/cm^3 in gas phase, the number of trajectories for each atom doubles in the gaseous phase when these two systems are irradiated by the same ions and intensity beam. For instance, suppose each atom in the condensed phase collides with one trajectory; in the gaseous phase, this atom collides with two. We cannot directly compare the stopping power of these systems to investigate the effect of phase on stopping power. They have to be normalized by density number. According to the normalization, the two systems have the same atomic geometry; however, the trajectories superimpose in the gas phase. In the gas phase, the stopping power is equal to the sum of two trajectories for the first atom without the second atom and two trajectories for the second atom without the first atom. This superposition of trajectories in the gaseous phase is not an artifact because they are a part of measurement before being normalized.

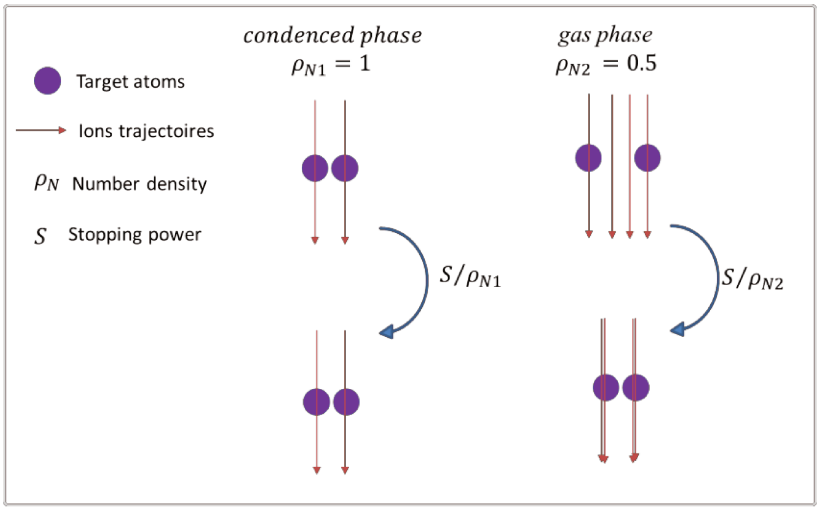

Figure 3-3: the general scheme of an ion beam's interaction with a condensed and gaseous target, and normalization of their trajectories by the density at the microscopic level.

Thereby, we prepared our supramolecular system and its isolated fragments. Our supramolecular system consists of a guanine, a cytosine, and two water molecules as mentioned before. They are connected by five hydrogen bonds. In order to prevent the geometric difference between the supramolecular system and its fragments, the geometry of each isolated molecule is defined from the supramolecular system with the non-targeted molecules “removed”. They are described as "ghost atoms” (an atom having a basis set but no electrons). The supramolecular system basis set

was employed even for the irradiation simulations on the individual molecules. This technique increases the computational time (about the same cost in the supramolecular system and isolated molecule), but it could avoid basis-set-superposition errors. In our results, we calculated energy deposition rather than stopping power to avoid the complexity and artifacts that come from the number of trajectories, and the choice of the length of the simulation. As shown in Chapter Two, the length of the simulation depends on the energies of the proton. However, the ratio of energy deposition in two systems reflects the ratio of stopping power because the two systems have exactly the same geometry and distance interaction with the projectile.

3.3 Effect of hydrogen bonding at the low-energy region

In this section, we calculate the effect of hydrogen bonds on energy deposition by sampling our systems with a swift proton for the energy range between 0.03 and 0.3 MeV. We begin by showing the result of the average hydrogen bond energy deposition over 667 trajectories for each proton kinetic energy, and then we analyze the effect of hydrogen bonds depending on the trajectories.

3.3.1 Effect of hydrogen bonding on energy deposition at the low-energy region

In order to calculate hydrogen bond energy deposition for low-energy proton, we consider a planar grid of 29x23 equidistant points generating a set of 667 inelastic single-trajectories of the proton (Figure 3-4). Each proton trajectory is perpendicular to the generating plane, and runs through the plane at the level of the bead. The same grid of trajectories was exactly performed for each isolated molecule. Finally, we run 5x667 RT-TD-DFT simulations for each kinetic energy of the proton. In this way, supramolecular and isolated systems are comparable to evaluate the contribution of hydrogen-bond energy deposition.

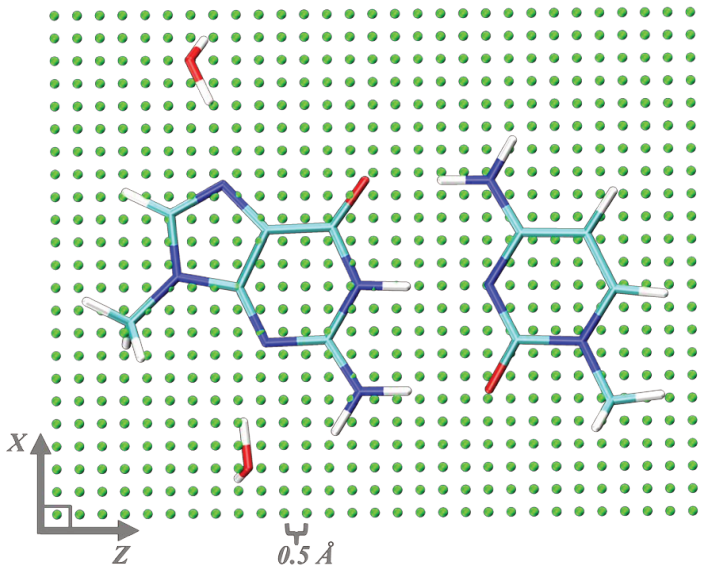

Figure 3-4: the supramolecular system (guanine, cytosine, and a pair of water molecules) is shown in licorice representation. Each green ball represents the passage of a single proton trajectory in the y-axis view.

Upon collision with an atom following a propagation line i, the electron cloud is excited and gains energy (ΔE_i). Within the proton kinetic energy range considered in this work (30 keV-300 keV), excitation of nuclear degrees of freedom can be safely neglected over electronic excitations. In each kinetic energy, the proton moves form y=30 Å to y=-30 Å. ΔE_i is thus the electronic energy variation of the supramolecular system before and after collision by the projectile. ΔE_i consists of the summation of energy deposition of all isolated fragments ΔE_i^x and hydrogen bond contribution ΔE_i^H.

$$\Delta E_i = \sum_x \Delta E_i^x + \Delta E_i^H \qquad (3.2)$$

Before showing the contribution of hydrogen bonds, the result of average energy deposition over 667 trajectories $\overline{\Delta E}$ for supramolecular (blue line) and isolated fragments (red line) is presented in Figure 3-5 as a function of the kinetic energy of the proton.

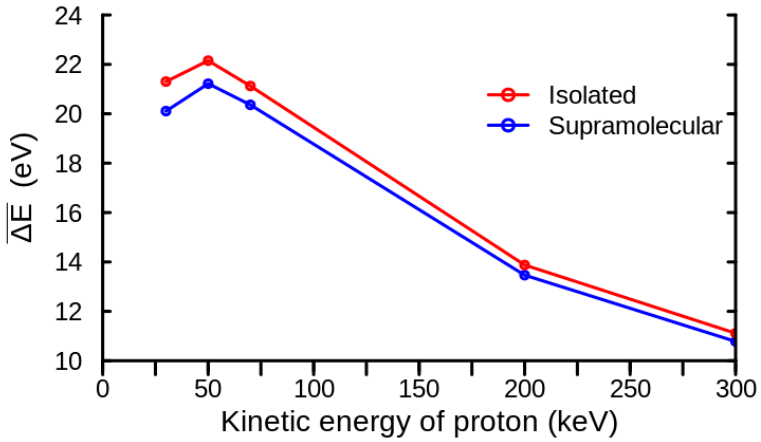

Figure 3-5: Average energy deposition curve for isolated (red) and supramolecular (blue) systems over all trajectories (667).

The average energy deposition for the two systems increases with increasing kinetic energy until Bragg's peak (50 keV) is reached. After that, they start to decrease. The important point here is that, as experimentally expected for other systems, the deposited energy within supramolecular system is lower than that of the isolated system in this range of energies.

Furthermore, the effect of hydrogen bonds ΔE^H can be calculated as follows:

$$\Delta E^H = \sum_i \Delta E_i^H = \sum_i \Delta E_i - \sum_i \sum_X \Delta E_i^X \tag{3.3}$$

And the percentage of hydrogen-bond contribution:

$$\%\Delta E^H = \frac{\sum_i \Delta E_i - \sum_i \sum_X \Delta E_i^X}{\sum_i \Delta E_i} 100 \tag{3.4}$$

collects values of hydrogen bonds effect for various kinetic energies of the projectile as shown in Table 3-1.

Table 3-1: deposited energy of the supramolecular system ΔE(eV), isolated molecular fragments ΔE^x(eV) and hydrogen bonds ΔE^H (eV) as a function of the kinetic energy of the proton E_{kin} (keV).

E_{kin}(keV)	ΔE(eV)	ΔE^x(eV)	ΔE^H(eV)	$\%\Delta E^H$
30	13412	14209	-796	-5.9%
50	14153	14775	-622	-4.4%
70	13581	14090	-509	-3.7%
200	8982	9256	-274	-3.1%
300	7190	7412	-222	-3.1%

The quantities of ΔE and ΔE^x are similarly corresponding to the kinetic energy of the projectile. The Bragg's peak appears at 50 keV. However, the value of ΔE^x is always greater than ΔE, so that the values of ΔE^H are negative. The effect of hydrogen bonds decreases with the kinetic energy of the proton. It is about (-6%) at 30 keV and (-3%) at 300 keV. In addition, the Bragg's peak disappears in the case of ΔE^H. This means that the contribution of ΔE^H does not follow the energy deposition curve. Overall, this hydrogen bond effect is not negligible, and could significantly impact the accuracy of charge particle penetration in radiotherapy (when the stopping power accuracy of less than ± 2% is desired)[21].

3.3.2 Effect of hydrogen bonding on energy deposition depending on the trajectories position

In microscopic Monte Carlo track structure simulations, energy deposition of a single charged particle can be calculated according to the stopping cross-section of the medium. Then, the passage and track shape of charged particles are investigated. So, the passage of charged particles depends on the local energy deposition for each collision. Despite this, the stopping cross-section of water has been used until now to determine the track and damage of biomolecules, on the assumption that water is the most abundant in cells, and that the stopping cross-section of water is close to one of biological molecules[22]. However, some data have shown that the stopping cross-section of certain biomolecules is significantly different from water which may impact the track structure[23,24]. Anyway, our question is: does the effect of hydrogen bonds rely on the position of trajectories? If it is the case, it could be an additional factor affecting the track structure.

In RT-TD-DFT simulations (microscopic system), we can explicitly determine the energy deposition for each trajectory. It is an opportunity to investigate the response of the deposited energy to the position of the trajectory. For this purpose, we have divided our system into three categories of trajectories: *intra*, *inter* and *far* which correspond to all trajectories that are located on the molecular fragments, between molecular fragments, and far away from molecular fragments, successively (see Figure 3-6 for the color code).

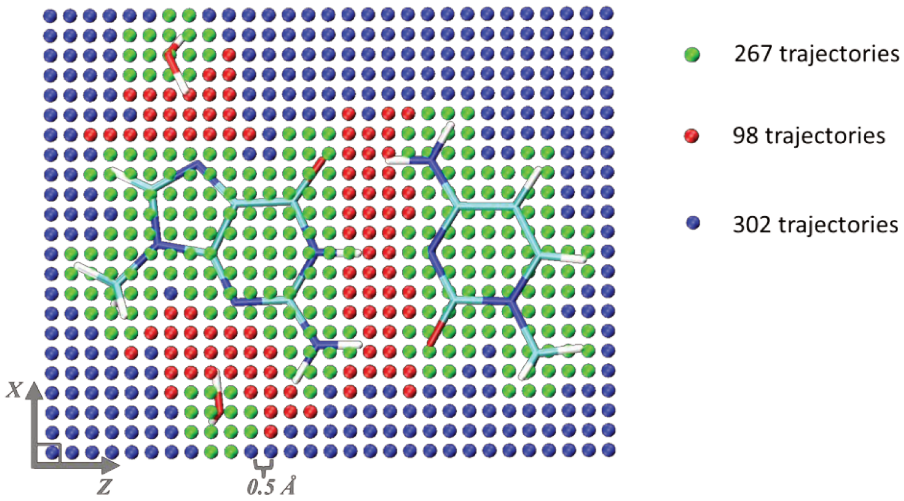

Figure 3-6: the supramolecular system is shown in licorice representation. Each ball represents the passage of a single proton trajectory in the y-axis view. The green, red, and blue balls correspond to the groups of trajectories which are located on the fragments (Intra), between fragments (Inter), and around fragments (far), successively.

The separation of trajectories has been based on electron density criteria of molecular fragments within the supramolecular system at point (r_i). r_i refers to the position of each trajectory at a bead of molecules. We separated *intra* from *inter* and *far* trajectories from the Hirshfeld partition weight is <= 0.995. Intra and *far* trajectories have been separated where the electron density is >= 0.023 for *intra*, and <=0.0027 for *far* trajectories. ΔE^H of each group of trajectories has been calculated by Eq. 3.4, but instead of the total trajectories, *intra*, *inter*, and *far* groups of trajectories have been used. The result of total energy deposition in the supramolecular system, hydrogen-bond energy deposition, and its percentage for each group is provided in Table 3-2.

Table 3-2: Energy deposition within the supramolecular system ΔE (eV), and hydrogen bonds $\Delta E^H(eV)$ for different groups of trajectories (intra, inter and far) as a function of the kinetic energy of the proton.

	intra			inter			far		
Ekin(keV)	ΔE	ΔE^H	$\%\Delta E^H$	ΔE	ΔE^H	$\%\Delta E^H$	ΔE	ΔE^H	$\%\Delta E^H$
30	9425	-177	-1.9	2163	-532	-24.6	1824	-88	-4.8
50	10297	-152	-1.5	2157	-383	-17.8	1699	-87	-5.1
70	9983	-135	-1.4	2027	-292	-14.4	1572	-82	-5.2
200	6474	-87	-1.3	1369	-109	-7.9	1138	-79	-6.9
300	5066	-74	-1.5	1121	-75	-6.7	994	-72	-7.3

As expected, the response of energy deposition to hydrogen bonds is highly dependent on trajectory position. First, we start with the energy deposition in the supramolecular system. ΔE of *intra* group is dominant. For example: at 50 keV, it corresponds to 73% *intra*, 15% *inter*, and 12% *far* groups where the percentage of trajectories are 40%, 15%, and 45%, respectively. That is expected because the electron density in the *far* group is considerably low. Furthermore, the red shift of Bragg's peak is observed for both the *inter* and *far* groups, with the *far* group showing a significant red shift. As shown in the Chapter Two, the screening of the electric field of charged particles by the local accumulation of electron density is one important factor responsible for the reduction of energy deposition at low velocity. This effect is negligible for the *far* trajectories because of the low electron density in this region.

Second, we analyze hydrogen-bonds energy deposition. Overall, ΔE^H is decreasing when increasing the energies of the proton, whereas their quantities rely heavily on the kind of trajectories. $\%\Delta E^H$ shows the most pronounced differences. In the *intra* group, $\%\Delta E^H$ is relatively small and less sensitive to the energies of protons. It slowly tends from -1.5% to -1.9% with decreasing kinetic energy. However, hydrogen bonds have a significant effect on the *inter* group, ranging from -6.7% to -24.6%. Surprisingly, the effect in the *far* group is completely reversed when compared to the other groups. The effect of hydrogen bonds augments with the kinetic energy of the proton, from -4.8% to -7.3%.

To summarize, the effect of hydrogen bonds strongly depends on the impact position. The most significant effect occurs when particles directly pass through the target's hydrogen bonds. That means, even if the hydrogen bonds or in general phase effects are not significant for stopping power, they could be crucial at the trajectories level.

To understand the reasons for the change in energy deposition in our system due to the formation of hydrogen bonds, we should first explain two mechanisms. As pointed out previously, changing the physical state or forming hydrogen bonds can affect two properties: first, the energy of valence electrons, and second, intermolecular distances (aggregation).

The first mechanism stated that when molecules change phases from gas to liquid or solid states, their valence electrons are stabilized by the formation of physical bonds. Consequently, excitation and ionization of these electrons become more difficult. As a result, the energy transition from the projectile to the target is less efficient in the condensed phase than in the gaseous phase. Because this modification only affects the valence electrons, it is more significant at low energies and negligible at high energies. This mechanism could be incorporated into the semi-empirical models, for example, by changing the mean ionization energy term in the Bathe-Bloch equation.

The second mechanism holds that as the phase changes, the molecules come closer together and the intermolecular distances decrease. As a result, the local polarization of electron density in the condensed phase screens more of the charged particle's electric field. For instance, in the condensed phase when a charged particle collides with an atom, the electric field of the charge particle passes through the electron density of the impacted atom to the neighbor atom, as shown in Figure 3-3. Nonetheless, the effect of an electric field on a neighboring atom is not shielded by the electron density of the impacted atom in the gas phase. Polarization and electron density accumulation are more important at low energy; thus, energy deposition within the condensed phase is lower, and the difference decreases as particle energies increase. This mechanism may be implicitly included in semi-empirical models for low-energy particles by introducing an empirical effective charge for ion particles (neutralize ion charge by electron attachment). However, the effect of target or physical phases are not incorporated in the empirical effective charge, which only depends on the projectile's charge.

By forming hydrogens bonds our supramolecular system is stabilized by 0.124 eV. This energy is not significant compared to the energy deposition. However, we cannot say that it does not modify significantly the ionization and excitation energies. It is hard to separate the ionization and excitation energy without screening effect in the RT-TD-DFT simulation because these two effects are automatically included.

In the *intra* group, hydrogen bonds have little effect. This is probable due to the fact that these trajectories are relatively far from the zones modified by the hydrogen bonds. Moreover, the neighboring molecules are far away from the impacting molecule (low energy deposition), and the interaction of charged particles decreases with distance, according to the Coulomb law. As a result, the screening effect is not essential because the energy deposition of the environment is originally small.

In the case of the *inter* group, the two mechanisms could be entirely involved in the hydrogen bond energy deposition. First, the proton directly impacts the valence electrons that are implied in hydrogen bonds. Second, the surrounding molecules are near the impact zone. Consequently, the energy deposition within fragments is important. In addition, the screening effect in the supramolecular system becomes important due to local polarization. As a result, trajectories at such positions are extremely sensitive to hydrogen bonding as they approach low kinetic energies.

In the case of the *far* group, both mechanisms fail to explain this behavior of the group of trajectories. It is a fact that the electron density is low around this kind of trajectory, electron accumulation and screening of the proton electric field are insignificant. On the other side, the Barkas' effect can efficiently contribute to energy deposition because the electrons of the target drag towards the proton in this energy region. Supramolecular and isolated systems may respond differently to Barkas' effect. Consequently, different behaviors of the hydrogen-bond effect are possible.

3.4 Effect of hydrogen bonding on energy deposition for protons within the energy range of 0.01-630 MeV

The effect of the physical phase is generally thought to be that it disappears at high proton energies. In SRIM, for instance, the stopping cross section of liquid water exactly coincides with that of gaseous water at 1 MeV. Unfortunately, due to computational costs, it is difficult to study the effect of higher proton energies on all trajectories in our system. Instead of that, we chose four trajectories; 2-1-1 from *intra*, *inter*, and *far* groups, respectively, as illustrated in Figure 3-7, hoping that these four trajectories could be representative of the whole system (667 trajectories). In the selection, approaching the value of each trajectory to the average value of its group trajectories over 0.03-0.3 MeV (the range in the whole system) and the ratio of trajectories in each group are considered. Except for the *far* group, only one trajectory was chosen because energy deposition is significantly lower in comparison to the other groups.

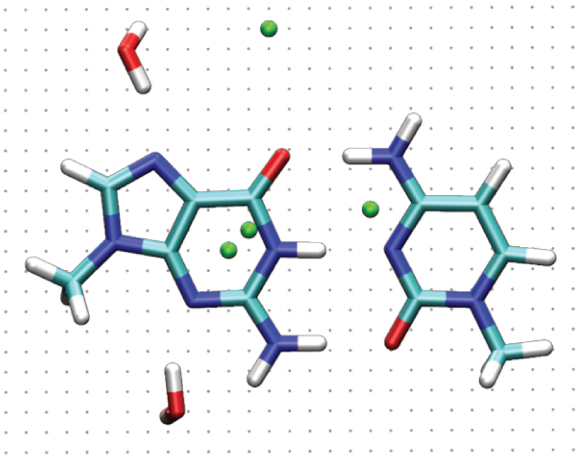

Figure 3-7: The green balls represent the passage of a single proton trajectory in the y-axis view.

The RT-TD-DFT simulations for these four trajectories were run from 0.01 to 630 MeV. In general, proton radiotherapy is performed with proton energies ranging from 50 to 350 MeV. It is therefore important to reveal these extremely high energies.

The result of average energy deposition $\overline{\Delta E_4}$ for the supramolecular and isolated systems over the 4 trajectories are presented in Figure 3-8 as a function of the kinetic energy of the proton. Both systems exhibit a relatively similar response of $\overline{\Delta E_4}$ to proton kinetic energy. From low energies, $\overline{\Delta E_4}$ increases with kinetic energy until it reaches the Bragg's peak at 0.05 MeV, after

which it decreases, and from 20.8 MeV, it gradually increases due to the relativistic effect. In comparison with the whole system, the location of Bragg's peak exactly coincides. The value of energy deposition, on the other hand, is slightly lower in the four-trajectory system. For instance, $\overline{\Delta E_4}$= 18 eV, and for whole system (667 trajectories) $\overline{\Delta E}$= 21.2 eV. This indicates that our assumption for a four-trajectory system is reasonably successful. Comparing the difference between the isolated and supramolecular systems, a significant difference is observed in low energies and extremely high energies.

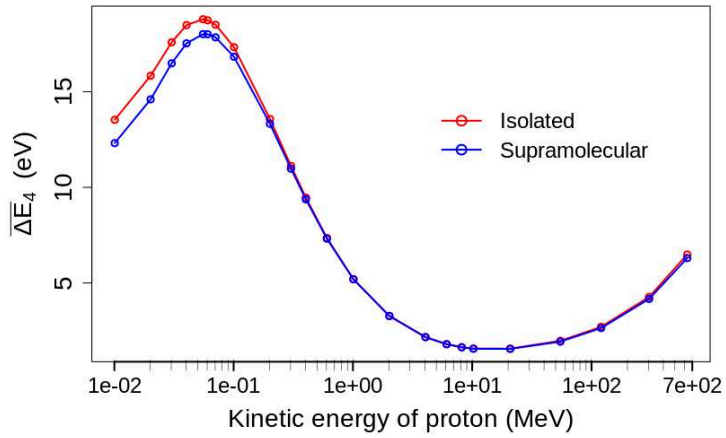

Figure 3-8: Average energy deposition curve for isolated (red line) and supramolecular (blue line) systems over four trajectories. The x-axis (kinetic energy of a proton) has a logarithmic scale.

Furthermore, for the four-trajectory system, we calculated the percentage of energy deposition of hydrogen-bond $\%\Delta E_4^H$. The result is shown in Figure 3-9.

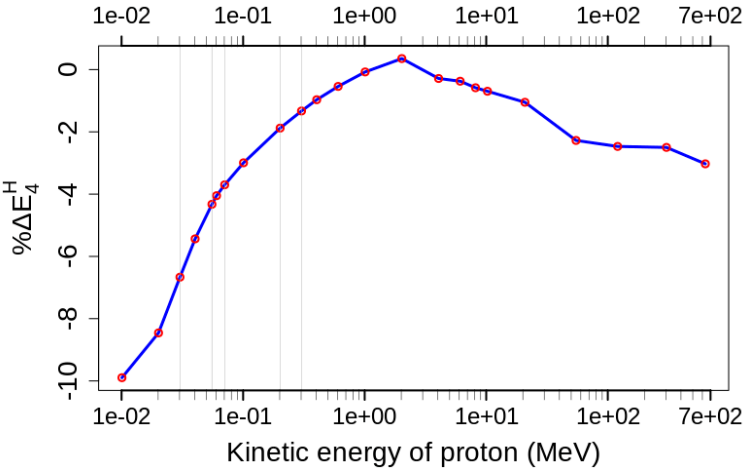

Figure 3-9: The percentage of hydrogen-bond energy deposition over the average of four trajectories. The proton's kinetic energy axis scales logarithmically. The vertical line highlights the consistent kinetic energy of the proton in the whole system at 0.03, 0.05, 0.07, 0.2, 0.3 MeV.

First, we attempt to compare the four-trajectory system to the whole system, in order to ensure the validity of the four-trajectory system result. In the whole system energy range, the effect of hydrogen bonds likewise decreases with proton energies in both systems. The $\%\Delta E_4^H$ is -6.6%, -4.3%, and -2% at 0.03, 0.05, and 0.2 MeV, respectively, whereas, for the whole system, it is -5.9%, -4.4%, and -3.1% for the same proton energies. Even though the values are not exactly the same, they are close enough to trust the conclusion of the four-trajectory system. In the low energies, the effect of hydrogen bonds raises from -6.6% to -10% at 0.03-0.01 MeV, respectively. Besides that, the $\%\Delta E_4^H$ decreases as a function of proton energies up to 2 MeV where the effect of hydrogen bonds completely vanishes. That is expected at high energies, as the participation of the valence electrons diminishes in the contribution of energy deposition. That is why the effects of phase state or of chemical bond are considered nil above about 1-2 MeV for proton in the majority of semi-empirical models. Our result is also in agreement in this region. However, surprisingly, the effect gradually increases after 2 MeV. It reaches -2.5% at 303 MeV and -3% at 635 MeV. This percentage is not negligible because energy deposition is small in comparison with the low-velocity region. In fact, in terms of penetration in proton radiotherapy, it exhibits the same error as maximum energy deposition. Stopping power normalizes the energy deposition by particle penetration. Any percentage of error in stopping power, whether at high or low velocity, has the same effect on Bragg's peak position.

This finding is completely unexpected in the literature. This fluctuation in the hydrogen bond effect could be due to a relativistic effect. Because the contribution of valence electrons to energy deposition is negligible in the high-energy region, valence electron modification has no effect. The effect of screening the electric field, however, is crucial. As demonstrated in Chapter Two, the relativistic effect can cause the electric field of a charged particle to shrink and concentrate perpendicular to the axis of propagation. This concentration increases energy deposition in the local position while decreasing the long-distance contribution of energy deposition. The local electron density surrounding the charged particle provides long-distance electric field shielding. This effect becomes more significant as the electric field shrinks (kinetic energy increases). When comparing the condensed phase to the gaseous phase, the local electron density is more important, that is why energy deposition in the condensed phase is lower than that in the gaseous phase at high energies, and this difference increases with charged particle energies.

3.5 Discussion and perspectives on fluctuations in hydrogen bond energy deposition at high energy

It is possible to accurately calculate our systems using semi-empirical models in order to compare this result with a theoretical calculation. We attempt to investigate water instead of it because we are especially interested in how the phase effect behaves in a relativistic region. We calculated the stopping power of proton in the vapor and liquid water from the PSTAR database[5] which is included in the ICRU Reports 49 data. ICRU data[25] is considered as a standard date for calculating the stopping power of proton, α-particle, and Lithium in several semi-empirical models, for instance, SRIM. At low energies, fitting formulas are used based on experimental stopping power data. At high energies, stopping powers are evaluated using the correct Bethe-Bloch formula that includes all corrections: mean ionization, shell, Barkas, Bloch, and density correction. The result of the percentage of hydrogen-bonds effect (phase effect) over 0.01-635MeV is represented in Figure 3-10. In the calculation the mean ionization energies for vapor and liquid water are 71.6 and 75 eV, respectively.

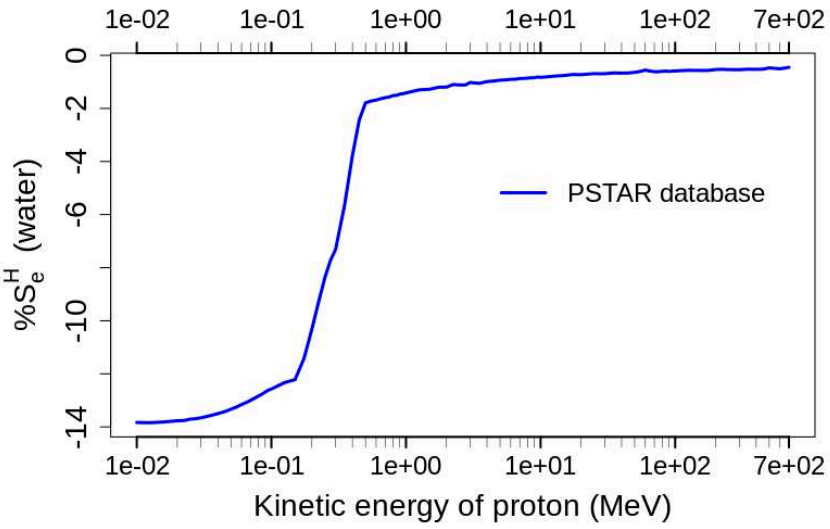

Figure 3-10: Percentage of hydrogen-bond effect on the stopping power of vapor and liquid water as a function of proton energy calculated using the PSTAR database. The x-axis (kinetic energy of a proton) has a logarithmic scale.

The result shows that the effect of hydrogen bond decreases as a function of proton energies. The maximum effect is about 14% for 0.01 MeV. Then, it slowly diminishes to 13% around 0.08

MeV and drops to 2% at 0.5 MeV. It may be because in the ICRU data, the Bethe-Bloch formula is used above 0.5 MeV for proton. Our region of interest is above 1 MeV. According to the PSTAR calculation, the effect of hydrogen-bond gradually decreases toward 635 MeV. No fluctuation is observed in this region. In fact, the density correction in Bethe-Bloch formula is calculated according to Sternheimer's method[26] which is responsible for the correction of the screening of the electric field by polarization electron density in the relativistic region. According to Sternheimer, this effect is significant when the energy of the proton is the same as or larger than the energy of the rest mass of the proton, for instance this correction includes 1% at 1000 MeV. So, in our region (1-635 MeV), any influence of the screening of the electric field by relativistic effect is not incorporated in the Bethe-Bloch formula. The small difference observed in this region between liquid and vapor water is due to the difference in the mean ionization energies.

The conclusion is also difficult to verify according to experiment due to high uncertainty and a lack of experimental data for biomolecules and water in this region. The unique experimental data reported by Palmer and Akhavan-Rezay[11] represents this fluctuation of stopping power for α-particles in vapor and liquid water. They discovered that from 2 MeV to 8 MeV (corresponding to 0.5-2 MeV for proton), the effect of phase once again increases with the kinetic energy of the projectile from $0.0_{\pm 3}$ % to $13_{\pm 3}$ %. This behavior is easily dismissed because it is not supported by any experimental data or a theoretical explanation. The authors were unable to find a satisfactory explanation. Of course, this fluctuation does not correspond to the fluctuation in the energy range of proton in our system. Thwaites[10] performed new measurements after Palmer and Akhavan-Rezay to confirm this fluctuation of the phase effect. He assessed the stopping of α-particles in the vapor and liquid water at 1.75-4.75 MeV (0.437-1.063 MeV for proton). He found that the phase effect decreases with energy of α-particle from $3.3_{\pm 3}$% to $0_{\pm 10}$%. The major fluctuations in the stopping power ratio were not observed in this energy region by Thwaites. If we compare Thwaites' result to our system with protons, we see that they agree well in this energy region.

In addition, in the case of proton in water, Shimizu and coworker[27] have recently measured the stopping power of proton in liquid water at 0.3-2 MeV. They found that the stopping power of liquid water is lower about $15_{\pm 5}$ to $10_{\pm 5}$% in comparison with the vapor phase. This proves that the effect of phase is essential after 2 MeV. Siiskonen and coworker[28], extended the stopping power of proton for liquid water in the energy range 4.7-15.2 MeV. They found a good agreement

between their result (uncertainty of 4.6% in experiment) and, ICRU Report 49 and Janni's tabulation[29] which are based on Bethe's formula. Finally, we cannot conclude about our result for high-energies because the uncertainty in the experimental data is larger than our result.

Overall, this behavior of the phase effect has not been observed before. To make the conclusion more robust, we must increase the number of trajectories or, in the best case, change the system to a simpler one, such as water, which can be more easily sampled. This is one of the directions we plan to explore in future work.

3.6 Conclusion

In this chapter, we have investigated the effect of hydrogen bonds in the guanine-cytosine nucleobase paired with water molecules on the energy deposition of protons with energies ranging from 0.01 to 635 MeV. We have demonstrated that our strategy for evaluating this effect is robust compared to existing data for water.

During the sampling of our system with 667 trajectories, we have shown that the effect of the hydrogen bond is more pronounced at low energy, being about -6% at 0.03 MeV and tending to -3% at 0.3 MeV. Additionally, our results indicate that the effect of the hydrogen bond strongly depends on the trajectories and is more significant where the proton impacts the hydrogen bonds. Divergent behavior is observed for trajectories, for instance where the trajectories are far from the molecules, the effect of hydrogen bonds on energy deposition increases with the kinetic energy of the proton.

From 4 trajectories, we have demonstrated that the effect of hydrogen bonds on energy deposition fluctuates at 2 MeV, rapidly increases below 2 MeV, and gradually increases above 2 MeV.

Overall, our results reveal that the effect of hydrogen bonds on energy deposition is non-negligible in biomolecules and should be taken into account for both low- and high-energy protons.

REFERENCES

(1) Jeffrey, G. A.; Saenger, W. *Hydrogen Bonding in Biological Structures*; Springer Berlin Heidelberg, 2012.

(2) Thwaites, D. I. Bragg's Rule of Stopping Power Additivity: A Compilation and Summary of Results. *Radiation Research* **1983**, *95* (3), 495. https://doi.org/10.2307/3576096.

(3) M.A, W. H. B.; B.Sc, R. K. XXXIX. On the α Particles of Radium, and Their Loss of Range in Passing through Various Atoms and Molecules. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **1905**, *10* (57), 318–340. https://doi.org/10.1080/14786440509463378.

(4) Ziegler, J. F.; Ziegler, M. D.; Biersack, J. P. SRIM – The Stopping and Range of Ions in Matter (2010). *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* **2010**, *268* (11), 1818–1823. https://doi.org/10.1016/j.nimb.2010.02.091.

(5) Seltzer, S. Stopping-Powers and Range Tables for Electrons, Protons, and Helium Ions, NIST Standard Reference Database 124, 1993. https://doi.org/10.18434/T4NC7P.

(6) Sigmund, P.; Schinner, A. Electronic Stopping in Oxides beyond Bragg Additivity. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* **2018**, *415*, 110–116. https://doi.org/10.1016/j.nimb.2017.11.023.

(7) Gu, B.; Muñoz-Santiburcio, D.; Pieve, F. D.; Cleri, F.; Artacho, E.; Kohanoff, J. Bragg's Additivity Rule and Core and Bond Model Studied by Real-Time TDDFT Electronic Stopping Simulations: The Case of Water Vapor. *Radiation Physics and Chemistry* **2022**, *193*, 109961. https://doi.org/10.1016/j.radphyschem.2022.109961.

(8) Uehara, S.; Toburen, L. H.; Nikjoo, H. Development of a Monte Carlo Track Structure Code for Low-Energy Protons in Water. *International Journal of Radiation Biology* **2001**, *77* (2), 139–154. https://doi.org/10.1080/09553000010012536.

(9) Matteson, S.; Powers, D.; Chau, E. K. L. Physical-State Effect in the Stopping Cross Section of $\mathrmH_2$O Ice and Vapor for 0.3-2.0-MeV $\ensuremath\alpha$ Particles. *Phys. Rev. A* **1977**, *15* (3), 856–864. https://doi.org/10.1103/PhysRevA.15.856.

(10) Thwaites, D. I. Stopping Cross-Sections of Liquid Water and Water Vapour for Alpha Particles within the Energy Region 0.3 to 5.5 MeV. *Physics in Medicine & Biology* **1981**, *26* (1), 71. https://doi.org/10.1088/0031-9155/26/1/008.

(11) Palmer, R. B. J.; Akhavan-Rezayat, A. . Range-Energy Relations and Stopping Power of Water, Water Vapour and Tissue Equivalent Liquid for α Particles over the Energy Range 05 to 8 MeV; Commission of the European Communities (CEC), 1978; pp 739–748.

(12) Phillips, J. A. The Energy Loss of Low Energy Protons in Some Gases. *Phys. Rev.* **1953**, *90* (4), 532–537. https://doi.org/10.1103/PhysRev.90.532.

(13) Wenzel, W. A.; Whaling, W. The Stopping Cross Section of $\mathrmD_2$O Ice. *Phys. Rev.* **1952**, *87* (3), 499–503. https://doi.org/10.1103/PhysRev.87.499.

(14) Wu, X.; Teuler, J.-M.; Cailliez, F.; Clavaguéra, C.; Salahub, D. R.; de la Lande, A. Simulating Electron Dynamics in Polarizable Environments. *Journal of Chemical Theory and Computation* **2017**, *13* (9), 3985–4002. https://doi.org/10.1021/acs.jctc.7b00251.

(15) Carmona-Espíndola, J.; Gázquez, J. L.; Vela, A.; Trickey, S. B. Generalized Gradient Approximation Exchange Energy Functional with Correct Asymptotic Behavior of the Corresponding Potential. *J. Chem. Phys.* **2015**, *142* (5), 054105. https://doi.org/10.1063/1.4906606.

(16) Calaminici, P.; Jug, K.; Köster, A. M. Density Functional Calculations of Molecular Polarizabilities and Hyperpolarizabilities. *The Journal of Chemical Physics* **1998**, *109* (18), 7756–7763. https://doi.org/10.1063/1.477421.

(17) Calaminici, P.; Flores–Moreno, R.; Köster, A. M. A Density Functional Study of Structures and Vibrations of Ta3O and Ta3O−. *Computing Letters* **2005**, *1* (4), 164–171. https://doi.org/10.1163/157404005776611420.

(18) Gómez Pueyo, A.; Marques, M. A. L.; Rubio, A.; Castro, A. Propagators for the Time-Dependent Kohn–Sham Equations: Multistep, Runge–Kutta, Exponential Runge–Kutta, and Commutator Free Magnus Methods. *J. Chem. Theory Comput.* **2018**, *14* (6), 3040–3052. https://doi.org/10.1021/acs.jctc.8b00197.

(19) Cheng, C.-L.; Evans, J. S.; Van Voorhis, T. Simulating Molecular Conductance Using Real-Time Density Functional Theory. *Phys. Rev. B* **2006**, *74* (15), 155112. https://doi.org/10.1103/PhysRevB.74.155112.

(20) Castro, A.; Marques, M. A. L.; Rubio, A. Propagators for the Time-Dependent Kohn–Sham Equations. *The Journal of Chemical Physics* **2004**, *121* (8), 3425–3433. https://doi.org/10.1063/1.1774980.

(21) Carrier, J.-F.; Archambault, L.; Beaulieu, L.; Roy, R. Validation of GEANT4 , an Object-Oriented Monte Carlo Toolkit, for Simulations in Medical Physics. *Med. Phys.* **2004**, *31* (3), 484–492. https://doi.org/10.1118/1.1644532.

(22) Friedland, W.; Dingfelder, M.; Kundrát, P.; Jacob, P. Track Structures, DNA Targets and Radiation Effects in the Biophysical Monte Carlo Simulation Code PARTRAC. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **2011**, *711* (1–2), 28–40. https://doi.org/10.1016/j.mrfmmm.2011.01.003.

(23) Francis, Z.; El Bitar, Z.; Incerti, S.; Bernal, M. A.; Karamitros, M.; Tran, H. N. Calculation of Lineal Energies for Water and DNA Bases Using the Rudd Model Cross Sections Integrated within the Geant4-DNA Processes. *Journal of Applied Physics* **2017**, *122* (1), 014701. https://doi.org/10.1063/1.4990293.

(24) Zein, S. A.; Bordage, M.-C.; Francis, Z.; Macetti, G.; Genoni, A.; Dal Cappello, C.; Shin, W.-G.; Incerti, S. Electron Transport in DNA Bases: An Extension of the Geant4-DNA Monte Carlo Toolkit. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* **2021**, *488*, 70–82. https://doi.org/10.1016/j.nimb.2020.11.021.

(25) Berger, M. J.; Inokuti, M.; Andersen, H. H.; Bichsel, H.; Powers, D.; Seltzer, S. M.; Thwaites, D.; Watt, D. E. 8. Stopping-Power and Range Tables. *Reports of the International Commission on Radiation Units and Measurements* **1993**, *os-25* (2), 76–81. https://doi.org/10.1093/jicru_os25.2.76.

(26) Sternheimer, R. M.; Berger, M. J.; Seltzer, S. M. Density Effect for the Ionization Loss of Charged Particles in Various Substances. *Atomic Data and Nuclear Data Tables* **1984**, *30* (2), 261–271. https://doi.org/10.1016/0092-640X(84)90002-0.

(27) Shimizu, M.; Hayakawa, T.; Kaneda, M.; Tsuchida, H.; Itoh, A. Stopping Cross-Sections of Liquid Water for 0.3–2.0 MeV Protons. *Vacuum* **2010**, *84* (8), 1002–1004. https://doi.org/10.1016/j.vacuum.2009.11.019.

(28) Siiskonen, T.; Kettunen, H.; Peräjärvi, K.; Javanainen, A.; Rossi, M.; Trzaska, W. H.; Turunen, J.; Virtanen, A. Energy Loss Measurement of Protons in Liquid Water. *Physics in Medicine & Biology* **2011**, *56* (8), 2367. https://doi.org/10.1088/0031-9155/56/8/003.

(29) Janni, J. F. Energy Loss, Range, Path Length, Time-of-Flight, Straggling, Multiple Scattering, and Nuclear Interaction Probability: In Two Parts. Part 1. For 63 Compounds Part 2. For Elements $1 \leqslant Z \leqslant 92$. *Atomic Data and Nuclear Data Tables* **1982**, *27* (2), 147–339. https://doi.org/10.1016/0092-640X(82)90004-3.

CHAPTER FOUR

PHYSICAL STAGE OF A NUCLEOPROTEIN COMPLEX SUBJECTED TO FAST TIONS

Contents

4 Physical stage of a nucleoprotein complex subjected to fast ions

Nucleoproteins are involved in many biological processes such as replication, transcription, protection and repairing of genome[1–3]. One example of a nucleoprotein complex is the nucleosome which is the basic structural unit of eukaryotic chromatin. A nucleosome consists in a core of histone proteins around which DNA is wrapped. The histones serve as a scaffold for the DNA and help regulate gene expression by controlling access to the genetic material[4]. Another example of a nucleoprotein complex is the ribosome. Ribosomes are the cellular machines that synthesize proteins, and they consist of both protein and RNA molecules. The protein component of the ribosome helps to stabilize the structure and catalyze chemical reactions, while the RNA component (ribosomal RNA) provides the functional machinery for protein synthesis[5].

Understanding radiation-induced damage to nucleoprotein complexes is essential because DNA is rarely naked in a cell, but rather functions in a crowed environment interacting with proteins, water and other metabolites. The interrogation that motivates this work is the interplay between DNA binding and radiation damages. According to some authors, DNA association to proteins has a significant protective effect in complexes when compared to free DNA targets[6–9]. In this chapter, we investigate the energy deposition and charge migration in a protein/DNA complexes exposed to α-particle (He^{2+}) by electron dynamics simulations.

Our stating point is provided by the work of John E. McGeehan and coworkers[10]. They investigated the irradiation of the bacterial protein C.Esp1396I/DNA complex. They used modern X-ray macromolecular crystallography at 100K to map radio-induced damages at different doses. They measured the spatial deflection and accumulation of electron density during data collection of X-ray diffraction and indicated the location of damages in the protein/DNA complex. They found that some residues are more sensitive to damages than others. For instance, the majority of damages take place on residues with carboxylic acid and sulfonated groups in the protein such as glutamic or aspartic acid or methionine residues (Figure 4-1). In the nucleic chain, thymine and adenine are more sensitive to damage than cytosine and guanine. In addition, they found that damages are more likely to be located on protein groups rather than on DNA.

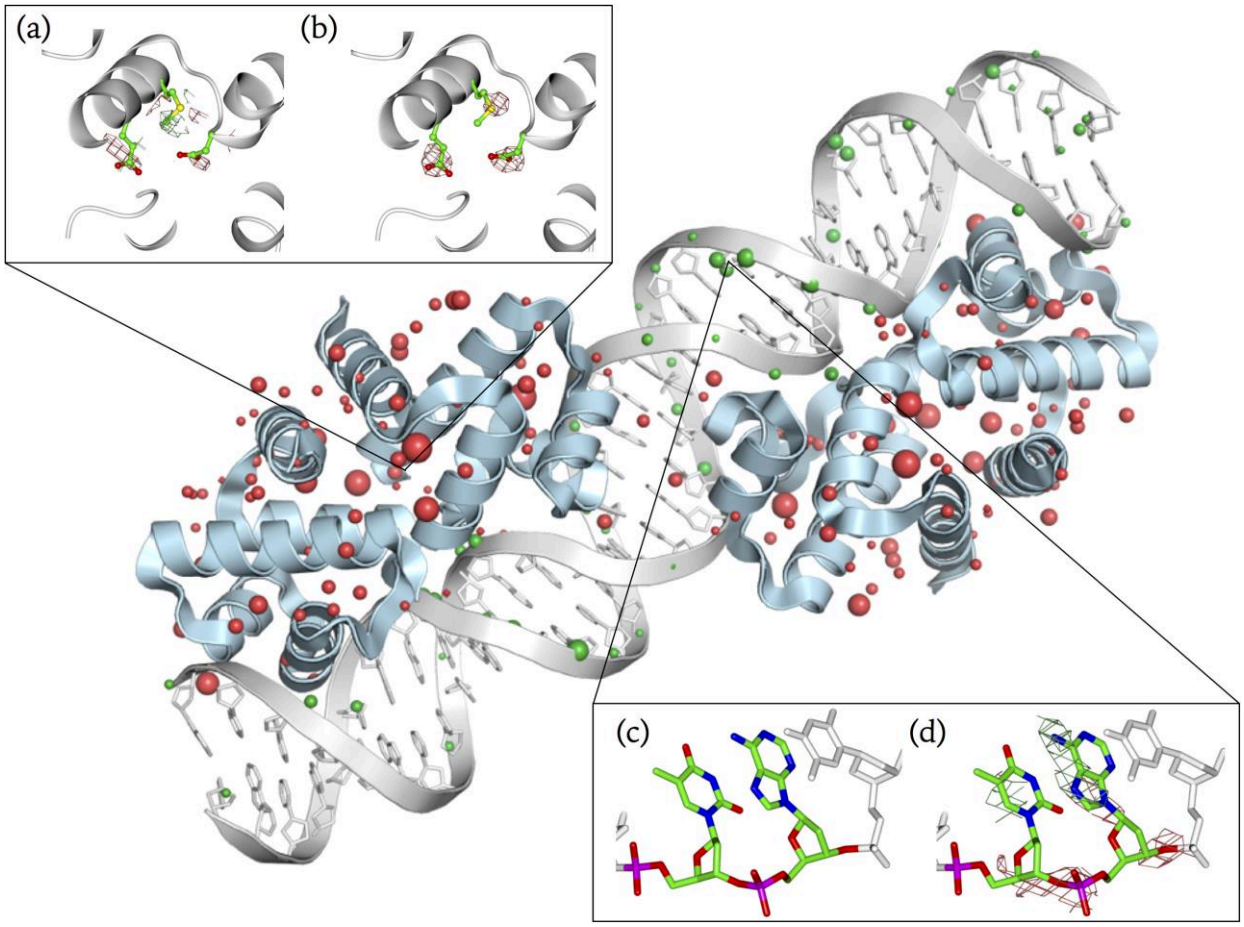

Figure 4-1: Representation of specific damage distribution throughout the C.Esp1396I complex for structures derived from 44 MJ/kg or (MGy) dose dataset. Specific damage sites are represented as spheres, with radii proportional to electron density loss (electrons per Å^3). Spheres closer (resp. further) than 2 Å to (resp. from) DNA strands are coloured in blue (resp. red). (Inserts) Damage sites in C.Esp1396I protein and DNA: (a)-(b) Protein chain with GLU-54, MET-57, and ASP-64 residues and (c)-(d) DNA chain with T24 and A25, shown at (a), (c) low-dose (6.2 MJ/kg) and (b), (d) high-dose (44.6 MJ/kg) exposures. The green and red contour colors signify electron density gain and loss, respectively. The figure is reproduced from references [10,11].

In Figure 4-1(a-b), a distinct loss of electron density is noticeable around acidic residue side-chain carboxyl groups and methylthio side-chain groups in the protein chain at both low (panel a) and high doses (panel b). This suggests radiation-induced decarboxylation and Met CH_3-S bond cleavage. For DNA chains (Figure 4-1(c-d)), electron density changes are negligible at low doses (panel c). However, at high doses (panel d), there is a reduction in electron density for T24 and A25 moieties, possibly indicating sugar-phosphate C-O bond cleavage between T24 and A25 nucleotides. An increase in electron density is observed with dose near T24 and A25 bases, which is consistent with low energy electron (LEE) attachment to nucleobases[12]. Figure 4-1 (center) depicts the site of electron density change in the C.Esp1396I/DNA complex at 44 MJ/kg. The

magnitude and color of the spheres correspond to the variation in electron density and their proximity to the DNA chain. As observed, the alterations in electron density, which are associated with radiation damage, predominantly occur on the protein chains.

In another work of John E. McGeehan and coworkers, investigating irradiation of protein/RNA complexes, similar conclusions were drawn. On the other hand, it was found that when protein residues are bound directly to RNA, less radiation damages are induced[13]. Various damage mechanisms are at play and require clarification[11]. One example is the proposed mechanism for radiation-induced decarboxylation of acidic amino acids, which involves the migration of holes to the sidechain carboxyl group[14,15]. The decarboxylation process is influenced by the interplay among hole migration, hole recombination, and engagement with solvent species generated through radiolysis. It has been suggested that when an acidic residue is bound to RNA or DNA, the average solvent-accessible area is reduced, leading to an increase in hole-recombination and a decrease in decarboxylation[13]. However, According to a study conducted by Allan et al, which focused on the diffusion of radiolytic species at low temperatures, it was found that the highly oxidizing hydroxyl radical is immobile at an experimental temperature of 100 K[16]. As a result, it is not clear whether decarboxylation mechanism is dependent on the solvent-accessible area of the carboxyl group at this temperature[17,18]. Instead, it is more likely that it is related to the creation of holes and their migration.

In this chapter, our objective is to investigate the physical stage that involves energy deposition and charge migration in protein/DNA complexes irradiated by α-particles, with a particular focus on the acidic amino acid's residues within these complexes. Our protein/DNA complex system is a large molecular system, which could complicate the simulation if fully treated at the DFT (Density Functional Theory) level and the subsequent analysis of results. To address this, we employ a combination of Quantum Mechanics/Molecular Mechanics (QM/MM) approaches in our simulation, and we search for analytical tools to facilitate the analysis of such large systems.

This chapter encompasses four sections. Section 4.1 covers the preparation of the QM/MM systems and provides the methodological details of the simulations. Section 4.2 investigates the impact of electrostatic induction of MM-region on energy deposition and post-irradiation charge migrations. A part of this work has been published in the book "Multiscale Dynamics Simulations:

Nano- and Nano-bio Systems in Complex Environments" in the Chapter "Electron and Molecular Dynamics Simulations with Polarizable Embedding"[19]. In Section 4.3, we study charge migration in a protein/DNA complex. In Section 4.4, we introduce a new analytical technique for studying charge migration in large systems and propose protocols to identify correlations among inter-fragments charge variations. Finally, it concludes with two sections: a general conclusion and a discussion on charge migration in the carboxylate group.

4.1 System preparation

4.1.1 Molecular dynamic simulations

The system was prepared by classical molecular dynamic (MD) simulations with NAMD 2.11[20] . The crystal structure of bacterial protein C.Esp1396I/DNA complex (PDB code : 3clc)[21] was solvated with the TIP3P water[22] model in a 130x130x130 $Å^3$ periodic box as illustrated in (Figure 4-2(a)). The system was neutralized by adding 382 sodium ions. The Amber FF14SB (Force Field 2014 with Side-Chain Backbone)[23] and DNA.OL15 (DNA Optimized Linear 2015)[24] force fields were used for the protein and DNA parts, respectively. We selected the Amber default protonation state, which corresponds to physiological conditions. This signifies that for protein part: the protein terminals are in the form of zwitterions, the amino acid side chains with acidic groups, such as aspartic acid (ASP) and glutamic acid (GLU), are typically deprotonated, while those with basic groups, such as lysine (LYS), arginine (ARG), and histidine (HIS), are generally protonated. For DNA, the phosphate groups in the DNA backbone are mono-deprotonated.

After geometry optimization, the system was equilibrated within the NPT ensemble (constant number of particles, pressure and temperature) at a temperature of 100 K for a duration of 15 ns. Equilibration was carried out without imposing any restraints on atomic positions and utilized a time step of 1 fs. This low temperature was chosen to maintain the crystal structure used in the experimental study, while allowing for relaxation of bond lengths and angles. The final structure was used as the starting point for QM/MM simulations.

4.1.2 Preparing QM/MM system

The protein/DNA complex we aim to investigate is a large molecular system that cannot be fully treated at the DFT level. To address this issue, we used a combination of quantum and molecular mechanism approaches (QM/MM for Quantum Mechanics/Molecular Mechanics)[25]. The atoms pertaining to the so-called QM region are described using RT-TD-DFT, while the surrounding atoms are represented using classical molecular mechanics force fields (MM region). The choice of the force field may have a significant impact on the physical processes of radiation, such as energy deposition and charge migrations. deMon2k affords two categories of force fields: polarizable and non-polarizable[17,18]. Polarizable force fields accounts for the polarization of atoms caused by the electron cloud distortion in response to an external electric field (electrostatic induction). Polarizable FF allows for a more accurate description of molecular interactions and properties, particularly for systems involving charged or polar species[26–28]. Unlike non-polarizable force fields, that assume constant atomic charges, polarizable force fields allow the charges to vary depending on the molecular environment. This provides a more realistic representation of the electrostatic interactions between molecules, and can lead to more accurate predictions of molecular properties.

Preparing QM/MM systems directly from MM simulations is challenging due to the absence of periodic boundary conditions and the different format of force field parameters used in deMon2k. To overcome this, we utilized the QIB package[29] to extract QM/MM and QM regions and convert AMBER parameters to deMon2k format, as depicted in (Figure 4-2(b)). QM/MM systems were

then prepared using the Amber polarizable force field ff02 parameter[30] and the Amber non-polarizable force field ff14SB parameter[31].

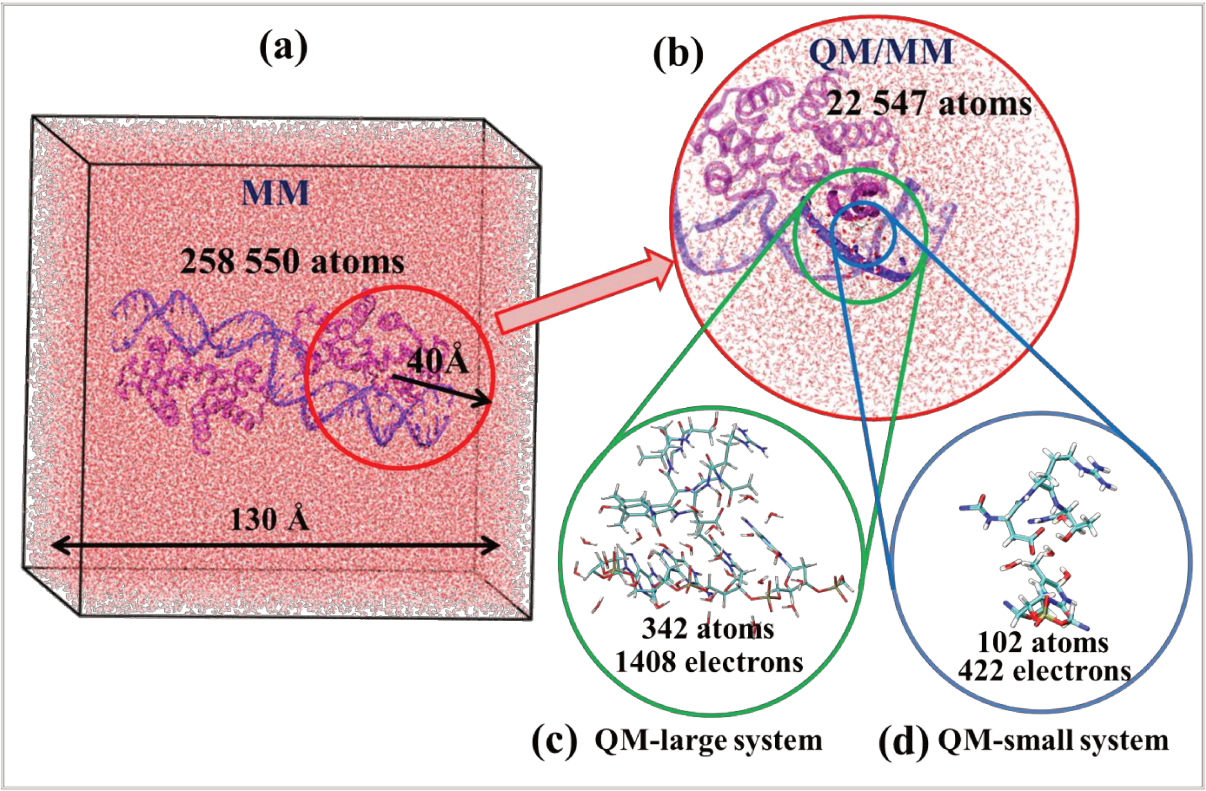

Figure 4-2: (a) represents the solvated protein/DNA complex in a cubic box, (b) shows the spherical pruning for QM/MM simulation in deMon2k with a radius of 40 Å, (c) and (d) show the QM region for large and small systems respectively.

We have created two QM/MM systems with varying QM sizes. The large one has QM region containing 342 atoms and 1 408 electrons. This system includes seven amino acids (residue number): LEU32-ASP33-ARG34-THR35-TYR36-ILE37-SER38, four nucleotides (residue number): A369-G370-T371-C372, and 30 water molecules (Figure 4-2(c)). The second one is smaller, with a QM region comprising only 102 atoms and 422 electrons. This system contains three amino acids (ASP33-ARG34-THR35), one nucleotide (T371), and four water molecules. (Figure 4-2(d)). In both systems, the QM region is centered on the ASP33 residue.

A key challenge with the QM/MM method pertains to the treatment of the boundary between the QM and MM regions. A technique known as the link atom approach is commonly employed to address this issue. In this approach, the bond that connects the QM and MM regions is described at the MM level, and a hydrogen atom (in our case) is introduced to cap the open

valences on the QM side. This link atom is subjected to quantum mechanical treatment, while the remaining part of the molecule on the MM side is treated classically.

It is important to note that the choice and position of the link atoms can influence the results. In our large system, we have four bonds at the QM/MM frontier: between the α-carbon and carbonyl group of ASN31 and GLY39 in the protein part, and between C5' of the phosphate group and C4' sugar moiety of A373, as well as between O3' of the phosphate and C3' sugar moiety of A369. We use the names of the atoms, similar to a PDB file, as illustrated in Figure 4-3 for nucleotides and amino acids. In our small system, we also have four link atoms. Two of them are located in the protein part between the α-carbon and carbonyl group of LEU32, and between the α-carbon and amine group of ILE37. The others are found in the DNA part between C5' of the phosphate group and C4' sugar moiety of T371, and also between O3' of the phosphate and C3' sugar moiety of C372. One important parameter is the distance ratio, which is the ratio of the bond length between the broken bonded atom and the replacing atom (in this case, H) to the bond length between the broken bonded atom and its original atom. This ratio can be optimized to minimize the effects of the XC functional and the types of atoms. Therefore, each link atom was optimized using the PBE exchange and correlation functional[32].

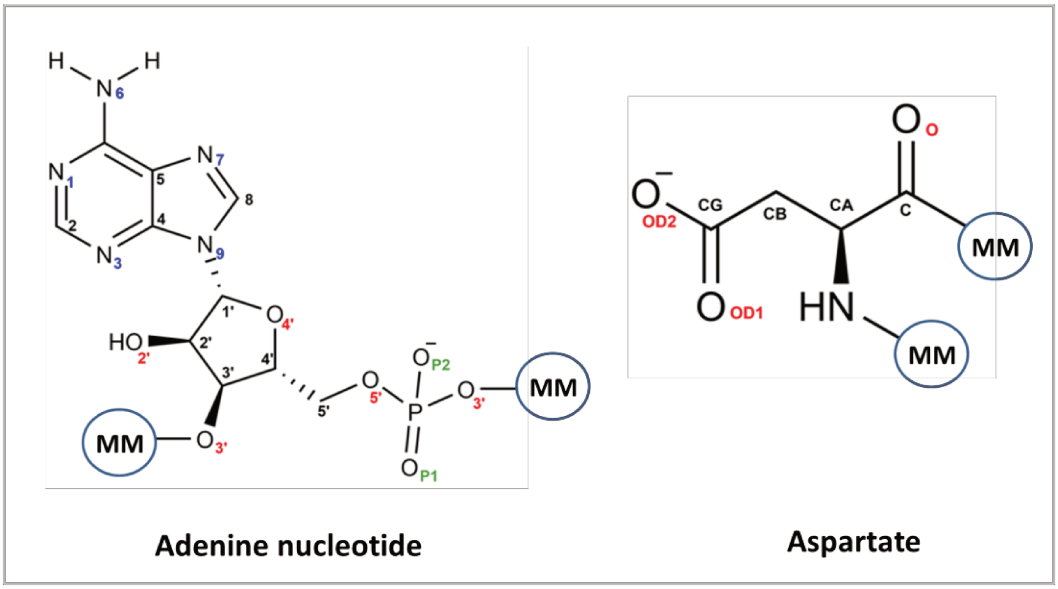

Figure 4-3: Structure of a nucleotide and an amino acid, with the names of their atoms labeled according to PDB file.

4.1.3 Electron dynamic simulation

We carried out electron dynamic simulations (ED) using pure RT-TD-ADFT under the nuclear fixed approximation to study the effect of polarizable and non-polarizable environments on α-particle irradiation of different QM regions. Time propagation was operated by the second order Magnus propagator[33] coupled to a predictor/corrector algorithm[34] with a time step of 1 as, for a total of 10 fs. On this time scale, it is assumed that nuclear motion exerts negligible effects on electron dynamics, particularly at a temperature of 100 K. We use a Taylor expansion[35] to evaluate the exponential of the Kohn-Sham matrix with 55 terms. We used the PBE[36] exchange correlation functional with DZVP-GGA basis set[37] adapted for generalized gradient approximation exchange-correlation functional and the GEN-A2* auxiliary basis set[38]. The kinetic energy of the α-particle in all simulations was set to 250 keV (a kinetic energy close to the Bragg's peak). We conducted population analyses on-the-fly using the Hirshfeld partition[39] at intervals of 10 as. Auxiliary density was employed for this analysis, which is a reliable method for population analysis and reduces computational costs significantly[40]. We would like to mention that we did not employ Complex Absorbing Potentials to remove non-bonded electrons, which might influence the electron dynamics and population analysis

4.2 Effect of electrostatic embedding on physical stage

The focus of this section is to investigate the impact of the electrostatic induction of MM on the energy deposition and the consequent charge migrations within the large and small systems. It's important to note that the term 'charge migration' in this chapter does not solely correspond to the pure charge migration process explained in Chapter One. Here, 'charge migration' encompasses all ultra-fast processes responsible for charge transfer within the molecule. This includes phenomena such as charge migration, auto-ionization, and Interatomic Coulomb decay (ICD), as all these processes are inseparably included in the RT-TD-DFT simulations.

4.2.1 Effect of electrostatic embedding in large system

We start by considering the large system, i.e. with the largest QM region. The α-particle was initially positioned 40 Å away from the center of the QM region. As illustrated in Figure 4-4, we arbitrarily defined as propagation line, one that crossed the bond center between the α-carbon

and the carbonyl group of aspartate (ASP33), the bond center between the side chain carboxylate group and the β-carbon of aspartate (ASP33), as well as the bond center between the nucleobase and the sugar moiety of thymine (T371). Defining an instant of collision is difficult due to the charged particle's interaction with the entire electronic system. We define collision positions being referred to (red beads) corresponds to the nearest fragments along the trajectory of the α-particle.

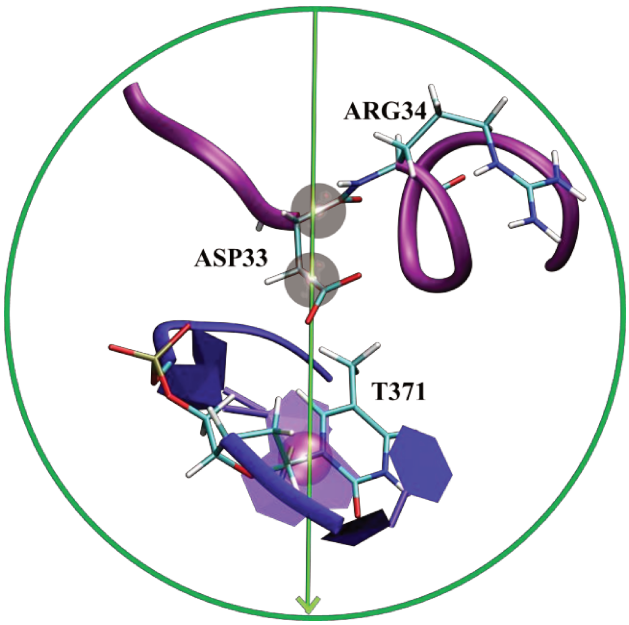

Figure 4-4: QM region defined in the large system. The MM atoms and QM water molecules are not shown for clarity. The peptide chain is represented in violet, the nucleic chain in blue. The green line refers to the travelling path of the α-particle throughout the QM region. The locations of main collision of α-particle with the system are symbolized by the red beads.

a. Effect on energy deposition

Figure *4-5* shows the variation of the Kohn-Sham energy (which represents only the DFT energy during collision in the context of ADFT, excluding projectile energy interaction and MM energy) with respect to the Self-Consistent Field (SCF) energy (ground state) in both polarizable and non-polarizable MM regions. The curves show a sudden increase of 431 eV in the range of 1 to 1.5 femtosecond after the start of the ED simulation, followed by a plateau until the end of the simulation. This rapid increase corresponds to the excitation of the electron cloud during collision. After the α-particle has traversed the QM region, no more energy is transferred and the energy remains constant.

As illustrated in the figure, the impact of electrostatic induction from the MM region on energy deposition is negligible. This can be attributed to several factors. Firstly, energy deposition is a local property; the interaction between a charged particle and electrons diminishes with increasing distance between them. In our simulation, the α-particle passes through the center of the QM region, and the size of the QM region is sufficiently large to limit the efficient incorporation of bordering QM atoms in energy deposition. This is because the atoms at the border of the QM region are distant from the site of collision[41]. Second, the process of energy transfer is extremely rapid, taking only a few attoseconds. As a result, the electron cloud of the QM-region does not undergo significant changes during the collision, especially for the electron cloud of border atoms located further away from the collision site. Third, the influence of the simulation cluster size on energy deposition is anticipated to be minimal at the Bragg peak, particularly as a result of the perpendicular interaction of charged particles, as shown in Chapter Two. This may also be the reason why introducing electrostatic induction within the MM region does not affect energy deposition.

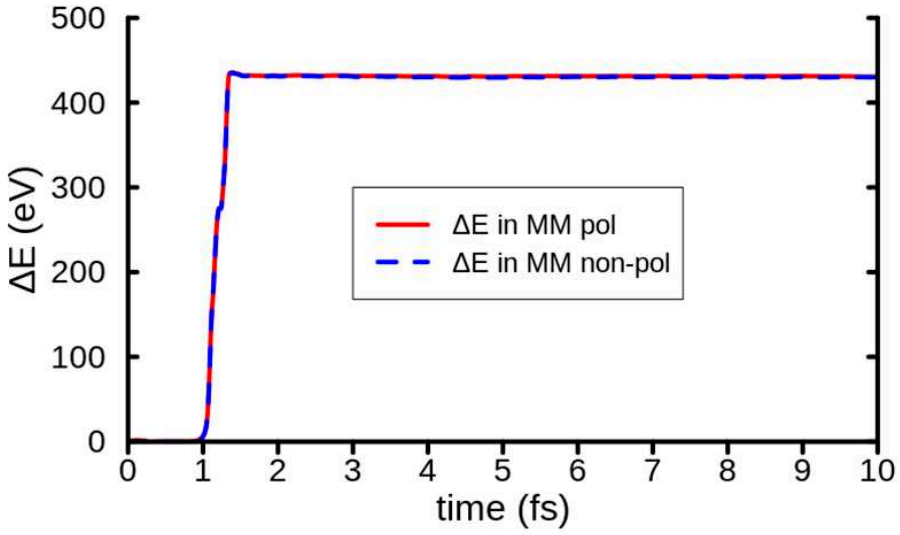

Figure 4-5: Energy variation in the larger QM/MM system during 10 fs upon collision with a 0.25 MeV He^{2+} in polarizable force field (red sold line) and non-polarizable force field (dashed blue line) environments.

b. Charge migration

To streamline the analysis of the impact of MM's electrostatic induction on charge migration, we separate the QM region into protein and DNA sub-ensembles. It is important to note that the simulation was carried out for the entire system. In this sub-section we only primary focus on

general trends. Deeper analyses of charge migrations will be conducted in Section 4.3 and Section 4.4 using dedicated analytical tools.

Figure 4-6 depicts the atomic charge variations within the protein relative to the ground state and following the polarizable QM/MM simulation (top-right panel). At the beginning of the simulation, charges are equal to zero because the projectile is positioned 40 Å away from the QM region and it does not cause any perturbation in the electron cloud. At the impact time about 1.5 fs, the variation in charge increases on the impacted fragments of ASP-frag (Aspartate without carboxylate group in tan color) and COO (carboxylate group of ASP in red color), which exhibit variations of +0.8 and +1, respectively. To better highlight the effect of electrostatic induction on charge migrations, we have plotted the difference between a polarizable and a non-polarizable simulation (right-bottom panel).

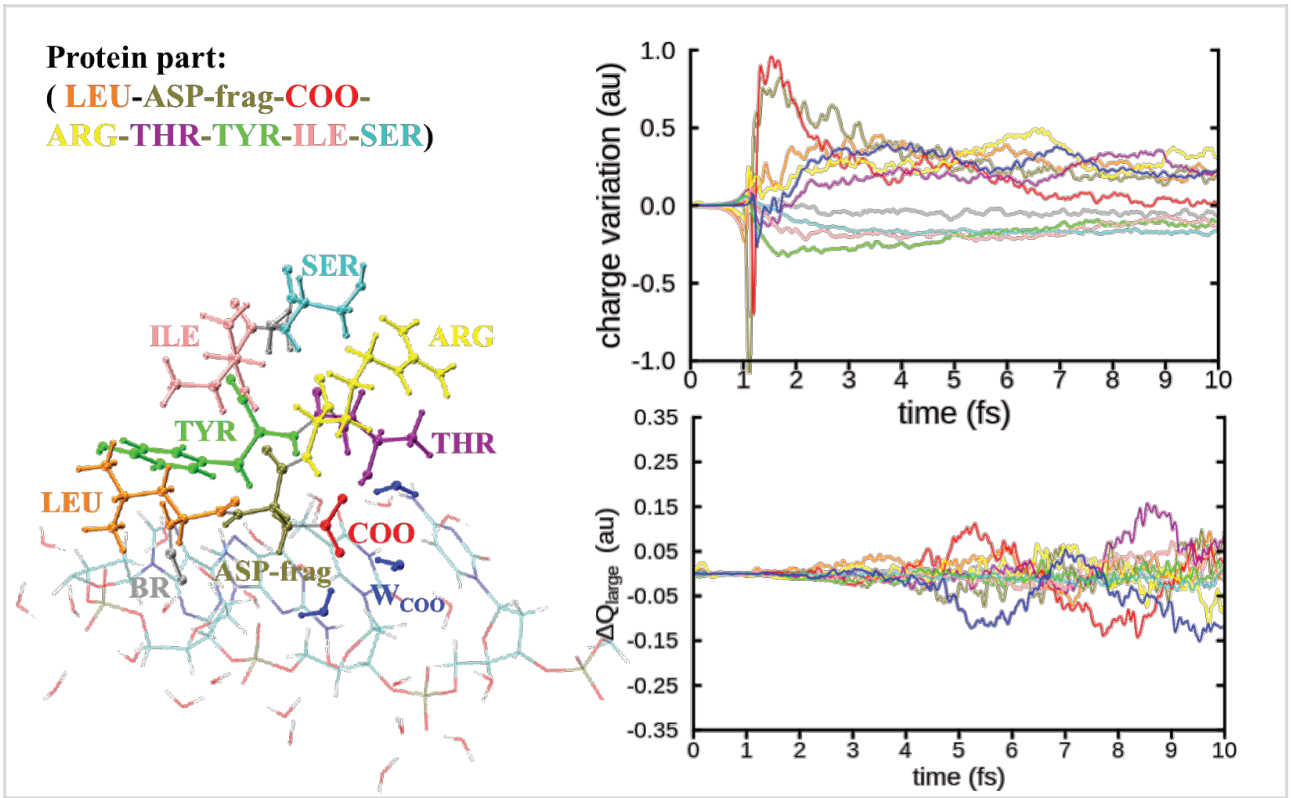

Figure 4-6: charge variation on amino acid residues within the large system under polarizable environments, while the right-bottom plot displays the differences in charge variation between polarizable and non-polarizable environments over a 10 fs. For ease of visualization, the figure presents the analyses of the protein part. The color code of fragments in protein part represents in the left.

As seen in the figure, the difference in charge variation between polarizable and non-polarizable FF simulations (ΔQ_i) is negligible (±0.05 au) at the beginning of the simulation, up to 4.5 fs. After

that point, some differences become more apparent. The majority of the differences are observed for COO (red line), W_{COO} (a water molecule, blue line), and THR (purple line), both of them form hydrogen bonds with COO. The maximum differences in these fragments amount to ± 0.15 (au), whereas the maximum differences in other fragments are less than ± 0.1 au. This could be attributed to the electron cloud or charge on COO, which is substantially perturbed (+1 au). As a result, the electron cloud on COO is more influenced by the induced dipole moments on the atoms. Since W_{COO} and COO are connected through hydrogen bonds, the induction of the electron cloud on COO affects them as well. This explanation is difficult to confirm because the ASP-frag is also significantly ionized (+0.8 au), yet the ΔQ remains relatively small.

Charge variations for the DNA part are shown in Figure 4-7. Like for the protein part, the strongest variations are seen on the impacted fragments. In DNA part, the α-particle traverses the bond between the thymine (T) base and thymine moiety (dT). Consequently, the charge at the onset of the collision is primarily situated on the T (cyan line) and dT (dark-red line), with values of +1 atomic unit (au) and +0.2 au, respectively (Figure 4-7 (right-top)).

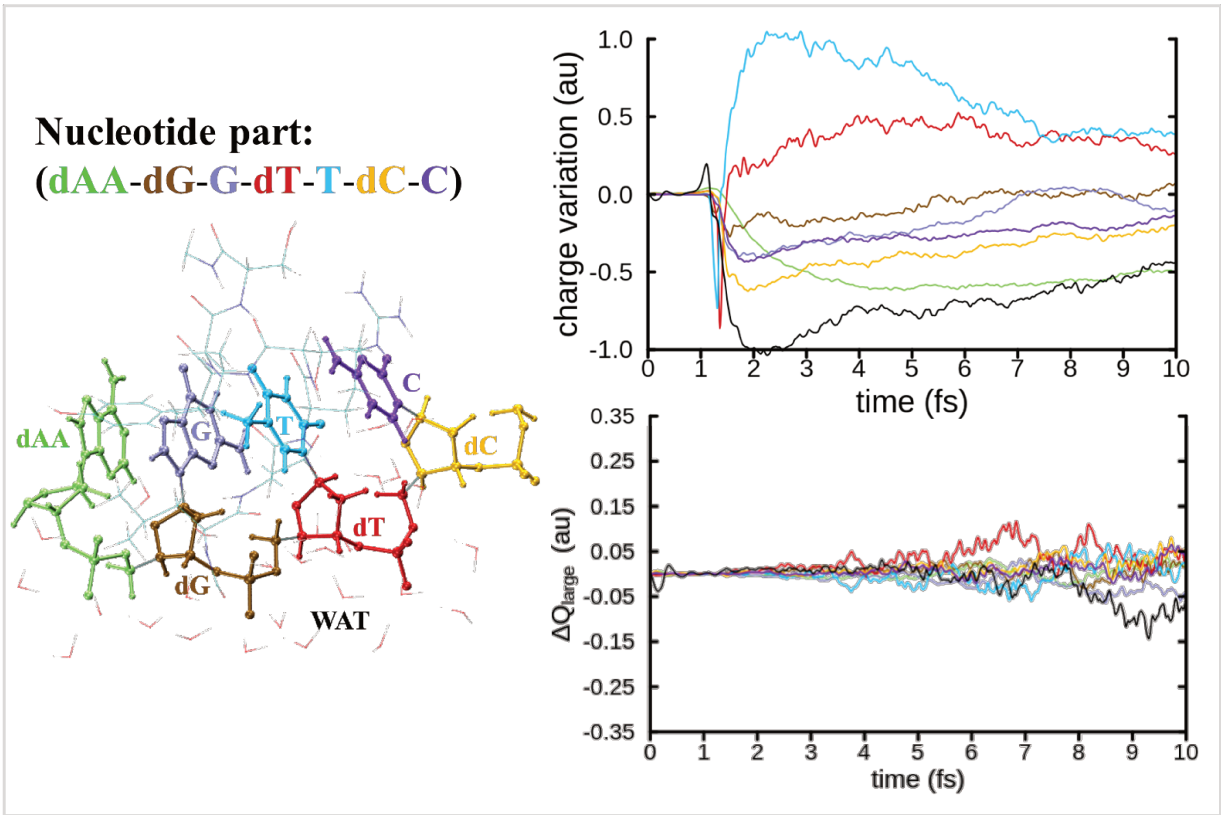

Figure 4-7: displays the charge variation analysis of the DNA component within the large system. The right-top plot represents the charge variation of fragments within the DNA part of the system under polarizable environment, while the right-bottom plot illustrates the differences in charge variation between polarizable and non-polarizable environments. The color code of fragments in DNA part is represented on the left. All water molecules, except for those that form a hydrogen bond with COO, are referred to as WAT.

In the DNA part, the impact of the environment's electrostatic induction on charge migration (Figure 4-7 (right-bottom)) is initially negligible, just like in the protein part. However, differences appear after 4.5 fs. Most of the differences are observed for the impacted fragments (T and dT) as well as the micro-solvated water shell where the maximum variations are ± 0.07, ± 0.12, and ± 0.14 (au), respectively. For the remaining fragments, the differences are below ± 0.05 au. As explained in the protein part, the perturbation of electron density on the impacted fragments and the potential hydrogen bonding may both be significant factors. In our system, the phosphate groups of DNA moieties form six hydrogen bonds with QM water molecules, except for the cytosine moiety (dC) which forms hydrogen bonds with MM water molecules. The DNA bases also form hydrogen bonds with the MM second DNA strand. However, a significant difference is observed for dT and T, which confirms that not only are the hydrogen bonds relevant, but also

the initial charge on the fragments. Furthermore, the magnitude of deviation in the DNA part is generally smaller in comparison to the protein part.

c. Partial conclusion

To summarize, we found that electrostatic induction in the MM environment has a negligible effect on energy deposition upon irradiation by a 250 keV He^{2+} ion. As a corollary, the effect of induction on charge migrations is negligible at the beginning of collision, but becomes noticeable 3 fs after collision. In general, the difference in charge migration between polarizable and non-polarizable models is mainly observed for the impacted fragments and those fragments that form hydrogen bonds with them.

The fact that electrostatic induction of MM has a small effect, may be related to the size of the QM-region, which is relatively large. It would be interesting to see how this effect changes with the size of the QM-region. Therefore, in the next sub-section, we investigate the smaller QM region.

4.2.2 Effect of electrostatic embedding in small system

We have replicated the simulations for a smaller QM region, as depicted in left of Figure 4-11. The irradiation conditions (nature of the ions, their initial positions, and their propagation vectors) are the same as before. The α-particle reaches the QM region at 1.035 fs and exits at 1.376 fs respectively for the large and small QM region.

Figure 4-8 depicts the ground state electron density encountered at the passage of an α-particle from static calculations. We identify three peaks that correspond to electron density when the position of α-particle locates: the bond between the α-carbon and the carbonyl group of aspartate, the bond between the side chain carboxylate group and the β-carbon of aspartate, and, finally, the bond between the nucleobase and the sugar moiety of thymine, respectively (Figure 4-4). The remarkable superposition of the two curves indicate that the reduction of the QM region's size should not impact significantly the amount of deposited energy. The smaller system should enable us to investigate the effect of electrostatic induction.

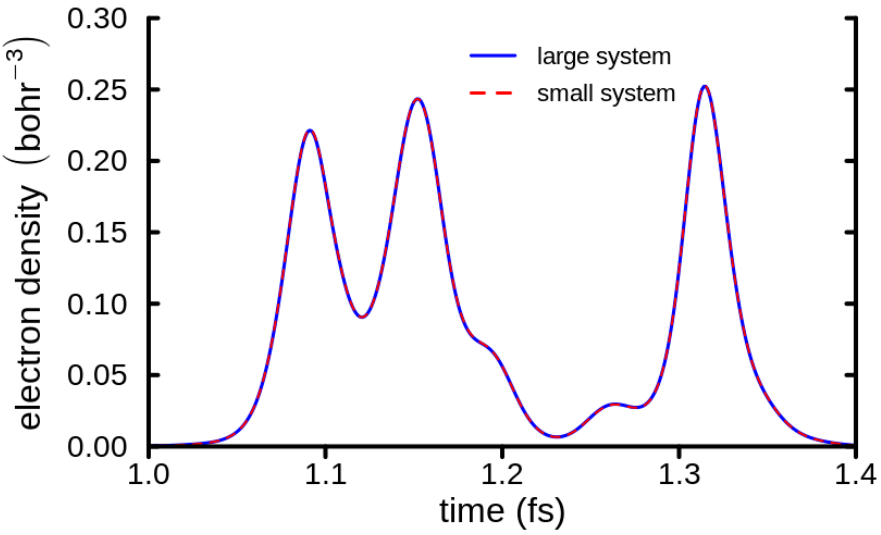

Figure 4-8: Electron density of the ground state for the large system (represented by the solid blue line) and the small system (represented by the red dashed line) at the passage of an α-particle.

I. Effect on energy deposition

We show energy deposition for the small QM region system during 10 fs in (Figure 4-9). 427 eV are deposited by the α-particle. Despite the substantial reduction in QM region's size, the difference in energy deposition between the large and small systems is only 4 eV. This proves the local properties of energy deposition. As seen in Figure 4-8, the electron density along the path of the α-particle is the same in both the large and small systems. The energy deposition depends particularly on the electron density near the path.

This very small difference could be attributed to long-distance interactions between the α-particle and the fragments not included anymore in the QM region, which substantially decreases with distance.

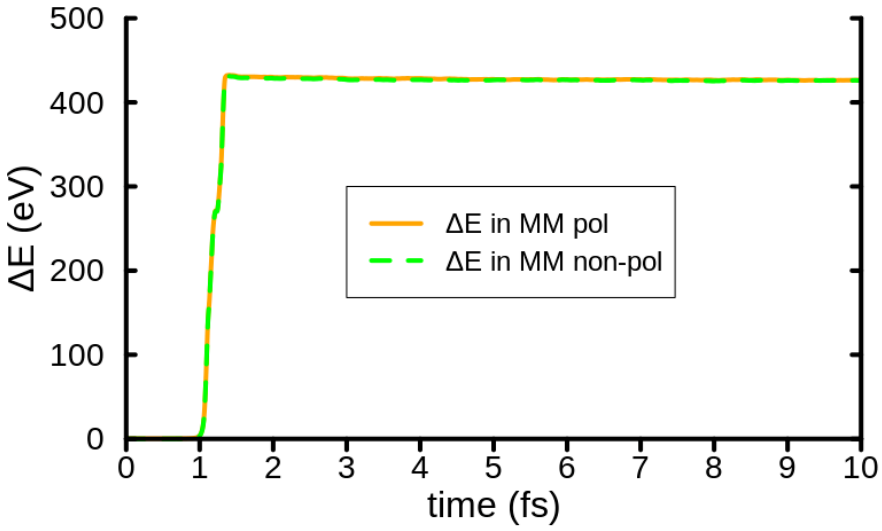

*Figure 4-9: Energy variation in **small system** over 10 fs upon collision with a 0.25 MeV He2+ in polarizable force field (red sold line) and non-polarizable force field (dashed blue line) environments.*

As already discussed, electrostatic induction has almost no effect on energy deposition (Figure 4-9). These results confirm the local nature of energy deposition of charged particles. Although the size of the QM region substantially decreases, its effect on the energy deposition is small. Therefore, the effect of electrostatic induction of environment is expected to be insignificant.

I. Difference of charge variation between large and small system

We have plotted the difference in charge variation for residues that are common between the small and large systems in a polarizable environment, as shown in Figure 4-10. Despite the small difference in energy deposition compared to the large system, a significant difference in charge variation is observed between the large and small systems after the collision and over time. This is particularly evident in the case of T (cyan) and ARG (yellow), where the difference exceeds 0.8 au. This difference can be attributed to the substantial reduction in initial ionization in the small system. For instance, the charge of COO decreased from +1 to +0.9 au, while ASP-frag dropped from +0.8 to +0.45 au. Furthermore, T decreased from +1 to -0.2 and dT dropped from +0.2 to -0.5. In addition, the change in initial ionization could affect the electron dynamics within the system. The smallest difference is noted for COO, which is around ± 0.2 au. This could be due to the location of COO in the center of the QM region, making it less affected by the reduction in size. These results underscore the importance of the QM size, which can significantly influence

both the initial ionization and electron dynamic response to ionizing radiation in biological systems, even when the same amount of energy is deposited in the molecules.

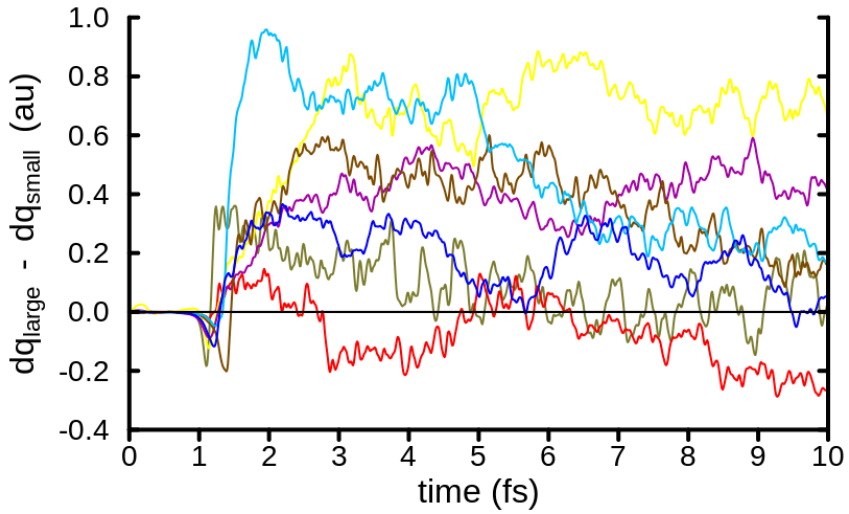

Figure 4-10: difference of charge variation between large and small systems. The color code is the same as Figure 4-11.

a. Effect on charge migration

Figure 4-11 displays the charge variation profile and its dependence on the electrostatic induction of the environment in small system. We compare the charge fluctuations between the MM and MMpol approaches of the environment in bottom-right panel of the figure. The difference in charge variation between the polarizable and non-polarizable embedding is noticeable at 2 fs after collision in our small system, whereas in the large system, it is observable at 3 fs. Additionally, the deviation magnitude in the small system is significantly larger than the large system, reaching up to ±2.5 au, while in the large system, it was a maximum of ±0.15 au. The maximum values of ΔQ_{small} for ASP-frag, COO, ARG, THR, W_{COO}, dT, and T are ±0.34, ±0.29, ±0.35, ±0.31, ±0.35, ±0.19, and ±0.13 au, respectively.

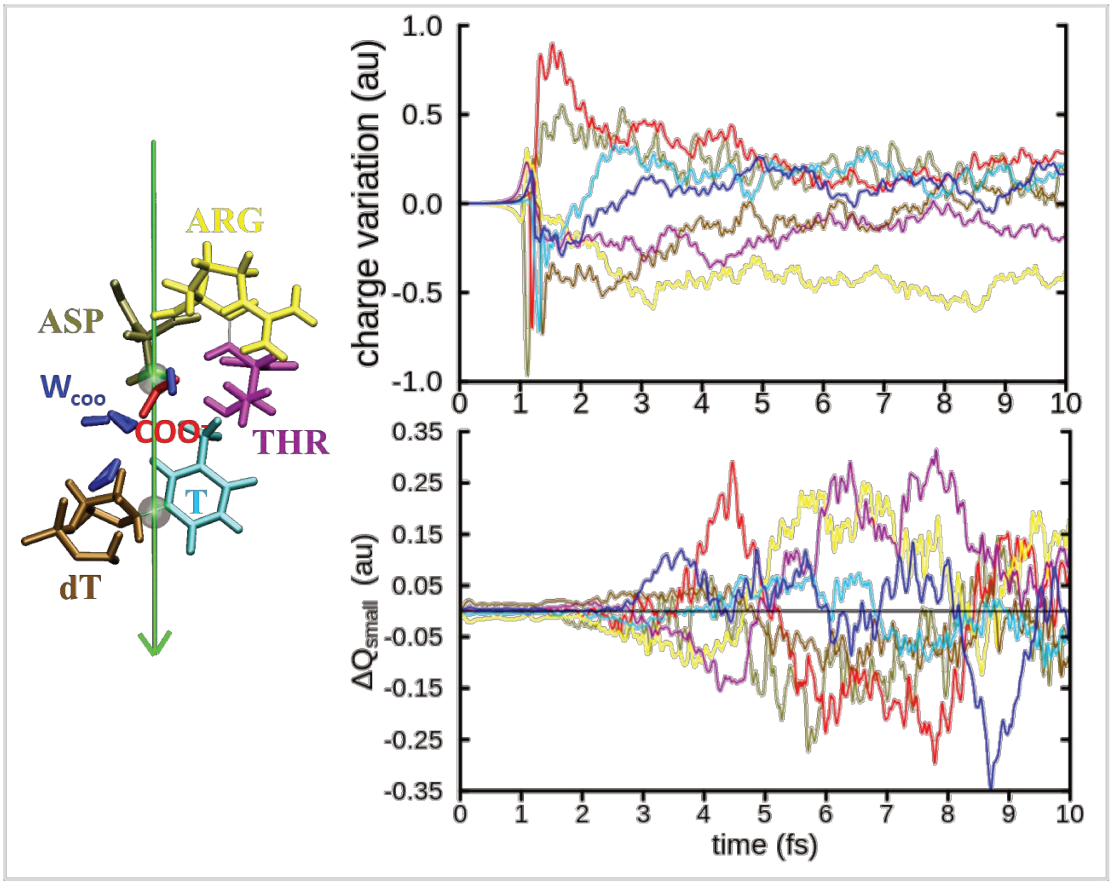

Figure 4-11: Charge variation with respect to the ground state for the small QM region system (right-top) using polarizable force field for environment. The lower-right panel illustrates the difference in charge variation calculated in polarizable and non-polarizable MM environments. left: QM region irradiated by the α-particle travelling along the green line; the green bead symbolizes the collision site.

It is challenging to establish which of the MM and MM polarizable (MMpol) approaches are more suitable for describing charge migration. By reducing the size of the QM-region, these processes undergo substantial changes, that are nor compensated by the inclusion of induction in the MM region. The best strategy seems to enlarge as much as possible the size of the QM region if one is interested in charge migrations following irradiation by ions.

b. Partial conclusion

To conclude, incorporating electrostatic induction into the computational protocol has no significant effect on energy deposition by α-particles. The MMpol approach, which induces dipoles in the MM region, does not affect the initial ionization, but it can affect charge migrations. A reduction in the size of the QM-region leads to a marginal decrease in energy deposition, subject to the condition that there is no significant modification in the electron density and fragments surrounding the projectile passage. Furthermore, the impact of electrostatic induction

in the environment remains insignificant with regard to energy deposition. The size of QM region can significantly affect the initial ionization and charge variation dynamics of fragments. While the effect of dipole induction from the environment on initial ionization is negligible, it plays an essential role in the process of charge migration. Moreover, as the size of the QM region decreases, the amplitude of deviation in charge variation of fragments, calculated within MMpol and MM environments, increases.

4.3 Comparison of two projectile's trajectories

In this section, we compare the irradiation of the QM region for two perpendicular trajectories. The first one is the one already discussed in Section 4.2. The second trajectory is depicted in Figure 4-13. In the "second trajectory", the α-particle traverses between the protein and DNA parts, predominantly impacting the bond between the carboxylate group and the β-carbon of the Aspartate. We only consider the large QM region within a polarizable QM/MM scheme.

4.3.1 Charge migration in first trajectory

I. Charge migration

To facilitate the comparison between the two trajectories, we reproduce in Figure 4-12, the upper-right panels of Figure 4-6 and Figure 4-7 showing charge variations in protein and DNA moieties.

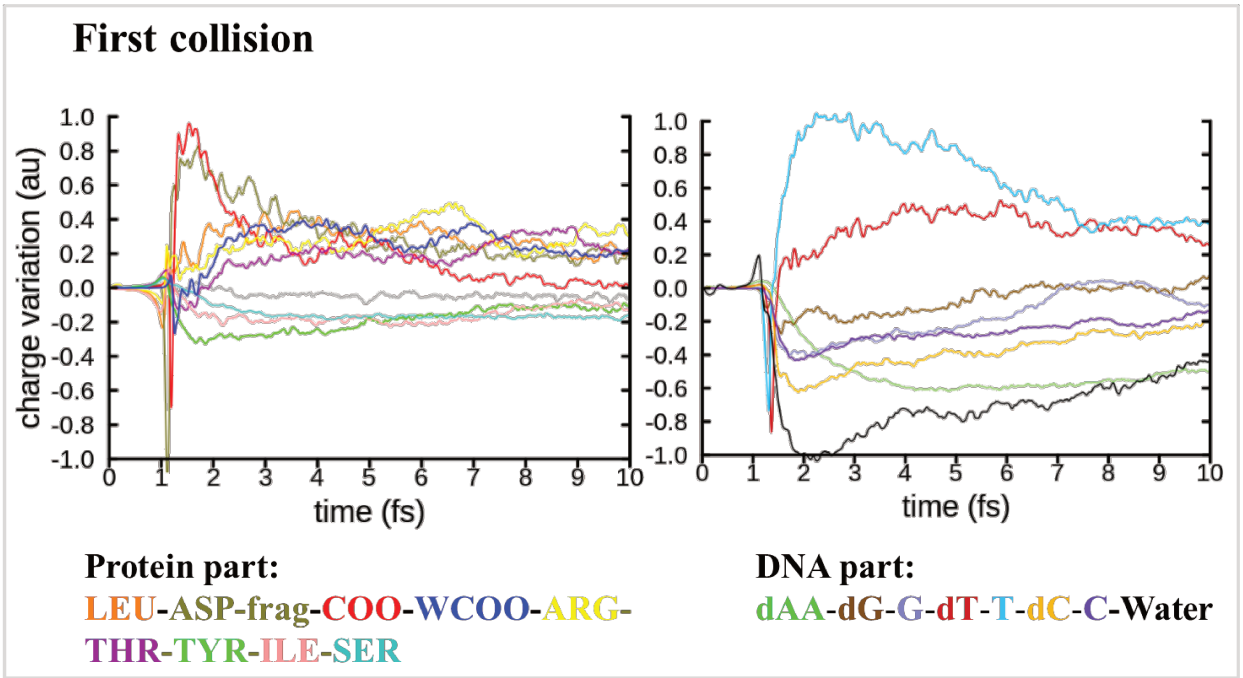

Figure 4-12: displays the charge variation of fragments in the large system for the first trajectory under a polarizable environment over a duration of 10 fs. To aid interpretation, the charge variation results are presented separately for the protein part (left) and the nucleotide part (right). The bottom of the figures provides the color code and acronyms corresponding to the fragments.

In general, the charge on the impacted fragments {ASP-frag (Aspartate without carboxylate group) , COO (carboxylate group of ASP), T (Thymine base), and dT (Thymine moiety)} initially decreases sharply and then increases, indicating an ebb-and-flow ionization mechanism[42]. Before collision, the incoming positively charged α-particle polarizes the electron cloud, attracting electrons to it. This causes a decrease in the charge on the ASP-frag, COO, T, and dT. After collision, electrons flow back to the surrounding molecular moieties. Due to the strong excitation of the electron cloud, electrons are emitted, leaving positive charges on the impacting fragments. The initial sharp decrease in charge at the beginning of the collision assists us in identifying the impacted fragments, such as ASP-frag, COO, T, and dT. However, it is important to note that the amplitude of the sharp declining charge on these impacted fragments is not directly related to the amplitude of their initial ionization, for instance, at the beginning of the collision, the charges on COO and ASP-frag are -0.7 au and -1.05 au respectively, while their initial ionizations are +1 au and +0.8 au, respectively.

As we begin analyzing the charge variation on protein residues, we observe an initial concentration of positive charge at the site of impact on the COO and ASP-frag, which is evident

as the red and tan lines peak at nearly +1 au and +0.8 au, respectively. However, after 1fs, the hole begins to migrate and is no longer localized at the collision site. Instead, the charge becomes delocalized across surrounding fragments. In particular, on longer timescales, the charge variation on the COO fragments tend to decrease in absolute value (the charge in the ground state). This is likely indicative of electron-hole recombinations. As explaining in beginning the mechanism of decarboxylation depend on the hole migration and recombination. Based on this result, it appears that the hole recombination occurs very rapidly.

Concerning charge migrations within DNA, we similarly observe that the positive charge initially localizes at the collision site on T and dT during the onset of the impact. As time progresses, the charge on T becomes more dispersed, but interestingly, the charge on dT increases.

Furthermore, the electron density accumulates on the SER, TYR, and ILE residues in protein, micro-solvated water molecules, and dAA (Adenine nucleotide) in the DNA. The accumulation of electron density on fragments could be a result of electron migration and the presence of secondary electrons.

Establishing a correlation between charge migration across fragments presents a significant challenge. Simple plots of charge variation struggle to elucidate this due to the complexity introduced by the ionization of multiple fragments and the intricate relations of charge migration. Therefore, in the next section, we aim to develop a methodology that simplifies the analysis of charge migration and facilitates correlation establishment.

a. Partial conclusion

In conclusion, our investigation revealed that the charge initially localized on the impacted fragment in our system upon collision with an α-particle. However, over time, the charge undergoes complex migration processes and becomes delocalized across surrounding fragments. Notably, the hole migration in the carboxylate group of aspartate is rapid, which is crucial for radiation-induced decarboxylation processes.

4.3.2 Second trajectory

It is known that the process of energy deposition is significantly influenced by the trajectory of the projectile, and any variation in this energy deposition could subsequently alter the pattern of

charge migration. We highlighted the importance of charge migration in the carboxylate group of aspartate. In this sub-section, we further explore the charge migrations within our large system, specifically focusing on the response of the carboxylate group of aspartate to changes in the trajectory of the α-particle. To this end, we carried out an additional trajectory simulation of α-particles within the large system, a situation we have termed the 'second trajectory'.

Starting from a distance of 40 Å away from the center of the QM region (COO^- group), the α-particle traversed between the protein and DNA parts, predominantly impacting the bond between the carboxylate group and the β-carbon of the Aspartate, as shown in Figure 4-13.

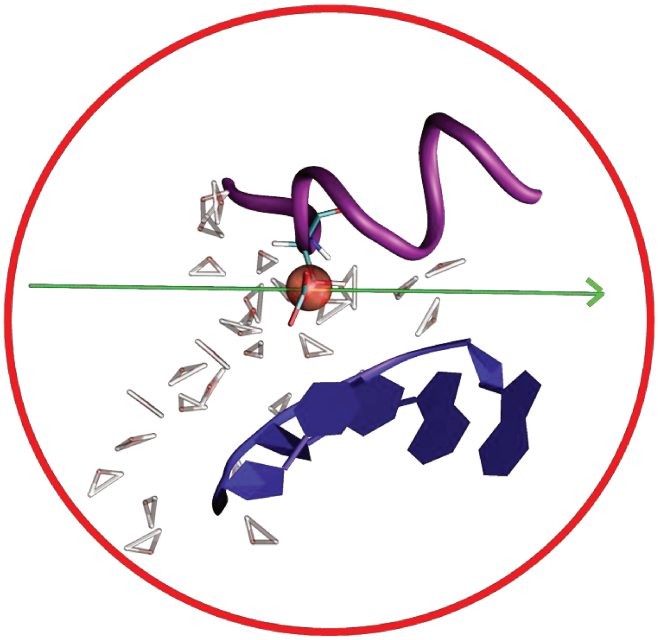

Figure 4-13: quantum mechanical (QM) region of the large system, wherein the amino acid chain is represented by the violet color, the nucleotide chain by the blue color, and the water molecules by the red and white colors. The green line in the figure indicates the path of the α-particle throughout the QM region during the second trajectory. Moreover, the locations of the main collisions between the α-particle and the target are denoted by the red bead representation

I. Energy deposition

In the second trajectory, the energy deposition of the α-particle amounts 172 eV. That is lower than first trajectory. This reduction relates to the decreasing electron density along the track compared to the first trajectory. As illustrated in Figure 4-14, there is a significant reduction of the ground state electron density along the track of second trajectory. In the case of the second

trajectory, we observe a single prominent electron density peak, indicating the position of the α-particle between the carboxylate group and the β-carbon of aspartate. Additionally, a smaller peak at 1.27 fs corresponds to the electron density in proximity to TYR. The energy deposition of charged particles is closely tied to the electron density surrounding the projectile, which explains why the energy deposition is lower in the second trajectory.

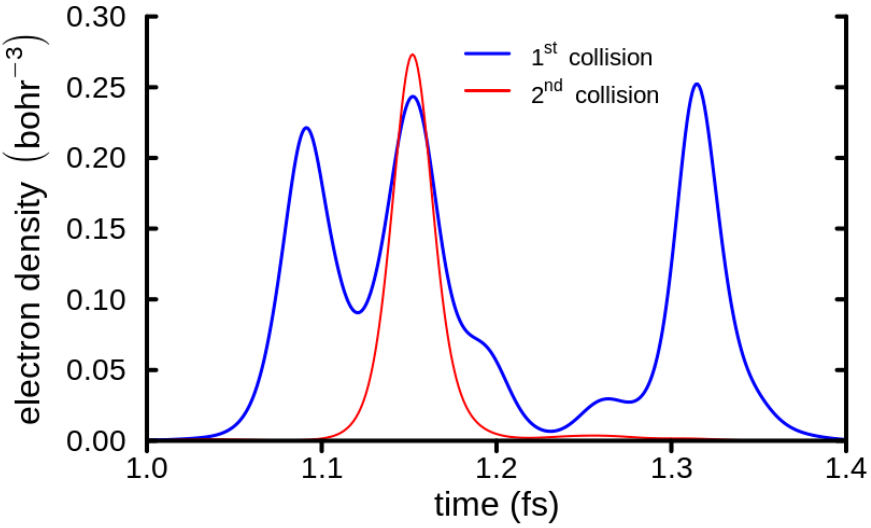

Figure 4-14: Electron density of the ground state at the passage of an α-particle through the large system, with the first trajectory represented by the blue line and the second trajectory by the red line.

a. Charge migration

The figure depicts the charge variations for second trajectory. Similar to the first trajectory, the ionization appears on the impacted fragments (ASP-frag and COO) and delocalizes over time. however, the initial charge on ASP-frag and COO in the second trajectory decreases to +0.55 au and +0.7 au, respectively, compared to +0.8 au and +1 au in the first trajectory.

Focusing on the COO, the charge increases upon impact and rapidly spreads across surrounding fragments. In comparison to the first trajectory, the charge achieves its absolute value (charge of ground state) faster. In the second trajectory, hole recombination on the COO fragment starts to be noticeable approximately 4 fs after collision, whereas it takes about 8 fs in the first trajectory. The observed differences could be attributed to various factors. Altering the trajectory could potentially modify the nature of ionization and excitation, thereby influencing the charge migration process. The amount of initial ionization, which has reduced to 0.7 au, might also be

another factor. Additionally, the charge on fragments surrounding the COO fragment might be affected, with a higher probability in the first trajectory. For example, the initial charge on ASP-frag, T, and dT in the second trajectory is +0.55, +0.1, and -0.1 au, respectively, compared to +0.8, +1, and +0.2 au in the first trajectory. Consequently, a reduction in electrons around the COO fragment leads to a decline in charge recombination, which is more likely in the first trajectory. In any case, these findings corroborate the ultra-fast charge recombination of the side chain carboxylate group upon exposure to charged particles.

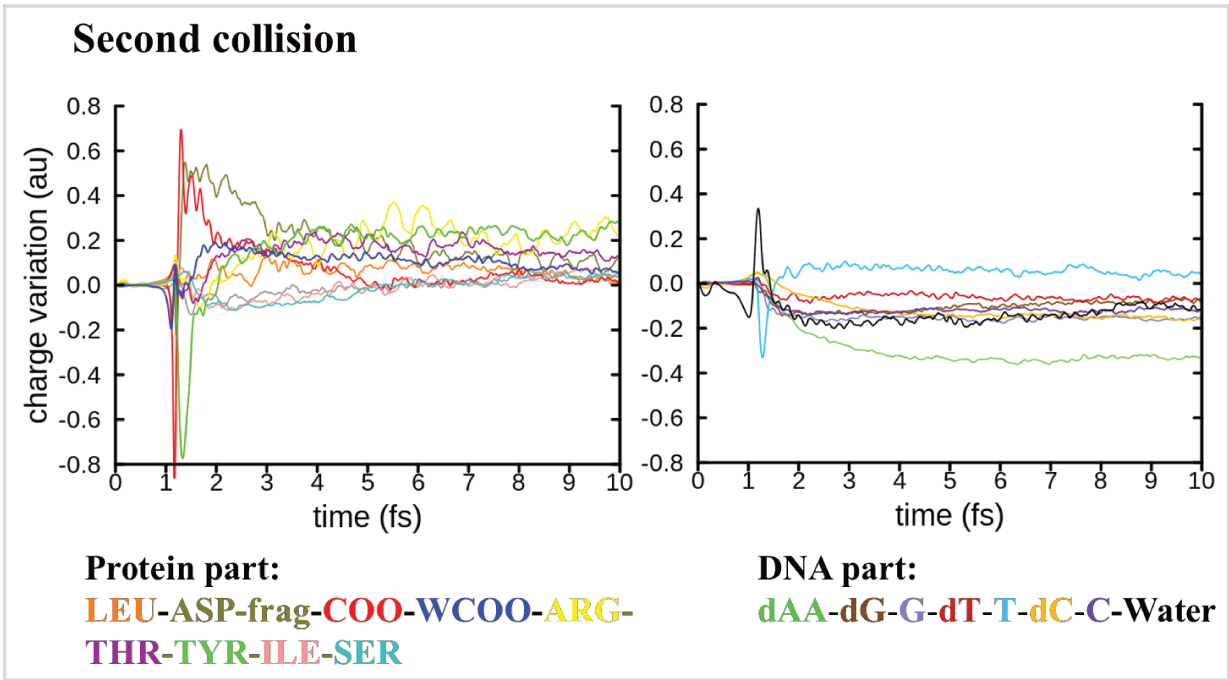

Figure 4-15: illustrates the charge variation of fragments in the large system during the second trajectory under a polarizable environment for a duration of 10 fs. The charge variation results are divided into two parts for easier interpretation: the protein part (left) and the nucleotide part (right). The color code and acronym of fragments are provided at the bottom of the figures.

Generally, the initial ionization process and charge migration deviate from those observed in the first trajectory, particularly in the DNA component. This is due to the altered location of energy deposition and the changes in the nature of ionization and excitation.

b. Partial conclusion

Overall, the results indicate that upon irradiation of our protein/DNA complex with α-particle, the initial positive charge localizes at the site of impact but rapidly becomes delocalized across

neighboring fragments. Over extended periods, especially concerning the carboxylate group of Aspartate, the charge variation tends to diminish in absolute value. This likely signifies electron-hole recombination, a crucial process in radiation-induced decarboxylation mechanisms. Additionally, the direction of the α-particle does not alter the hole-recombination mechanism of the carboxylate group. Instead, it only influences the speed of hole-recombination, as demonstrated by the two different trajectories.

4.4 Analysis of the correlation of charge migration among fragments

The augmentation of system sizes afforded by RT-TD-ADFT simulations leads to a huge complexity of the data to be analyzed. For instance, when charged particles impact multiple molecules, several residues may become ionized. It is challenging to identify the correlation of charge migration between residues, as shown in Figure 4-12 and Figure 4-15. Therefore, we have been motivated to search for a methodology to better analyze charge migration within large systems. In this section, we propose two methodologies to evidence such correlations. The first one is the correlation matrix that provides instantaneous correlation between two fragments over a time series. The second one is the cross-correlation analysis, that reveals correlations between two fragments by introducing a time delay in the series.

4.4.1 Correlation matrix

In this sub-section, we mathematically define the correlation matrix[43] and analyze the charge migration correlations during the first trajectory. Additionally, we calculate the correlations during the second trajectory.

I. Mathematical definition

The correlation matrix is a time-series analysis tool initially employed to measure the degree of similarity between two time-domain signals[44]. The correlation coefficient R_{xy} between two variables X and Y is:

$$R_{xy}(t) = \frac{cov(X(t), Y(t))}{(\sigma_{X(t)} * \sigma_{Y(t)})}$$

where $cov(X(t), Y(t))$ is the covariance between the two variables and σ_X and σ_y are the standard deviations respectively. The resulting correlation coefficient R_{xy} is a value between -1 and 1. A value of 1 indicates perfect positive correlation, a value of -1 indicates perfect negative correlation (anti-correlation). A null value indicates an absence of correlation between the two variables. In our case, X and Y are charge variations on two given residues overtime.

The computation of the correlation coefficient was carried out for the charge variation of all residues involved in large QM region system, using the *R package for statistical computing* and the *RStudio* graphical interface (version 3.4.4)[45].

a. Correlations in the time range of 1.75-6 fs

We have calculated the correlation matrix for all fragments within the 1.75-6 fs time range (Figure 4-16). This specific time window is chosen because it avoids the pre-collision time and the actual moment of irradiation that is characterized by the ebb-and-flow mechanism. Here we wish to focus on post-irradiation charge migration. After 6 fs, there appears to be no significant change in the charge variation on fragments, as seen in Figure 4-12.

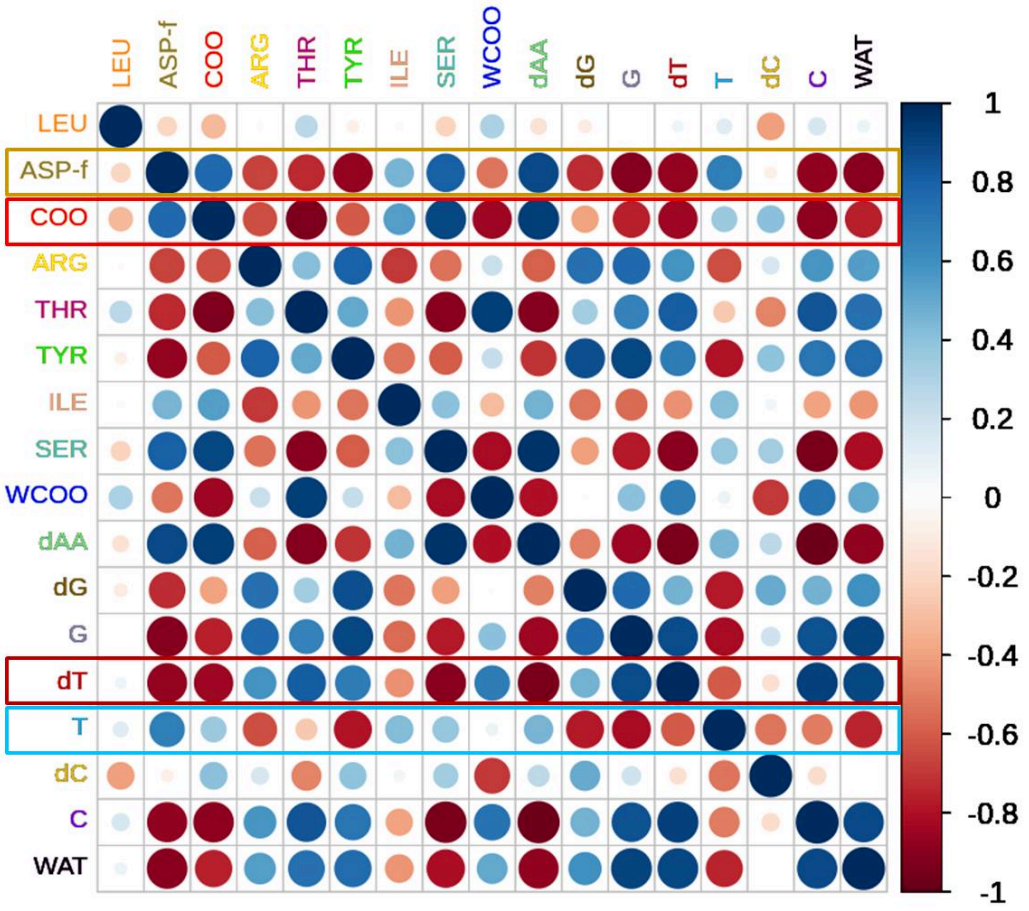

Figure 4-16: correlation matrix for charge variation between residues in the large system during the first trajectory, which was calculated over the time range of 1.75-6 fs. The size and color of the circles scale in accordance with the correlation coefficient between two fragments.

The correlation matrix offers a simplified visualization of the data, enabling new insights into the relationships between fragments. However, it's important to note that the correlation coefficient solely indicates similarity in charge variation between fragments. For a more detailed examination of charge recombination and electron emission, it's necessary to complement the analysis with the individual fragment charge information and how these charges change over the given time interval. These additional details are provided in Table 4-1 . Additionally, since the value of the correlation coefficient is dependent on the data range, it generally detects some minor

correlations. However, a significant correlation is typically considered to be larger than 0.8 upon our experiences.

Table 4-1: charge variation in fragments at 1.75 fs and its evolution over the time range of 1.75-6 fs. Red arrows represent an increase in charge over time, while blue arrows signify a decrease in charge during time.

Residue Name	LEU	ASP-frag	COO	ARG	THR	TYR	ILE	SER	WCOO	dAA	dG	G	dT	T	dC	C	WAT
Δq (au) at 1.75 fs	0.20	0.79	0.79	0.11	-0.01	-0.29	-0.17	-0.05	0.05	-0.18	-0.20	-0.40	0.11	0.78	-0.24	-0.42	-0.92
Δq vector	↗	↘	↘	↗	↗	↗	↘	↘	↗	↘	↗	↗	↗	↘	↗	↗	↗

By examining the correlation coefficient and charge information in Table 4-1, we can observe that the decrease in charge on the COO carboxylate group, indicated by the red box in Figure 4-16, is not only correlated with molecules moieties that are in contact such as THR and W_{COO}, but it also exhibits a inverse correlation (< -0.8) with distant fragments dT, C. which may indicate electron-hole recombination: decreasing the charge on the COO fragment while increasing the charge on the THR, dT or C fragments. On the other hand, a strong positive correlation is observed between COO and SER, as well as dAA, as the charge on these fragments decreased over time. This implies that an increase in electron density on COO coincides with a correlated electron density on SER and dAA, an unexpected phenomenon. This could be associated with the intricate processes of ionization and excitation during the interaction of charged particles. In these interactions, multiple electrons at different orbital levels within a fragment could be ionized, leading to various paths of charge migration. For instance, one atom or orbital might accept electrons while another atom or orbital within the same fragment emits electrons. Another hypothesis, originates from other fragments correlating with COO, such as THR, W_{COO}, dT, and C. The correlation function seeks linear relationships within the scattered data across two fragments, which is dependent on the oscillation (frequency) of charge variation between them. For instance, the electron emitted from THR to dAA establishes a characteristic oscillation of charge variation between them. If THR has a correlation with COO, then COO may appear to correlate with dAA, even though in reality, the electron is migrating from THR to dAA. As depicted in Figure 4-16, SER and dAA demonstrate a strong anti-correlation with THR, W_{COO}, dT, and C, which, in turn, show a strong correlation with COO.

For the sake of illustration we have represented in Figure 4-17 the QM region, coloring each residue according to the correlation with the COO terminal group with the help of the Visual Molecular Dynamics (VMD) software package[46].

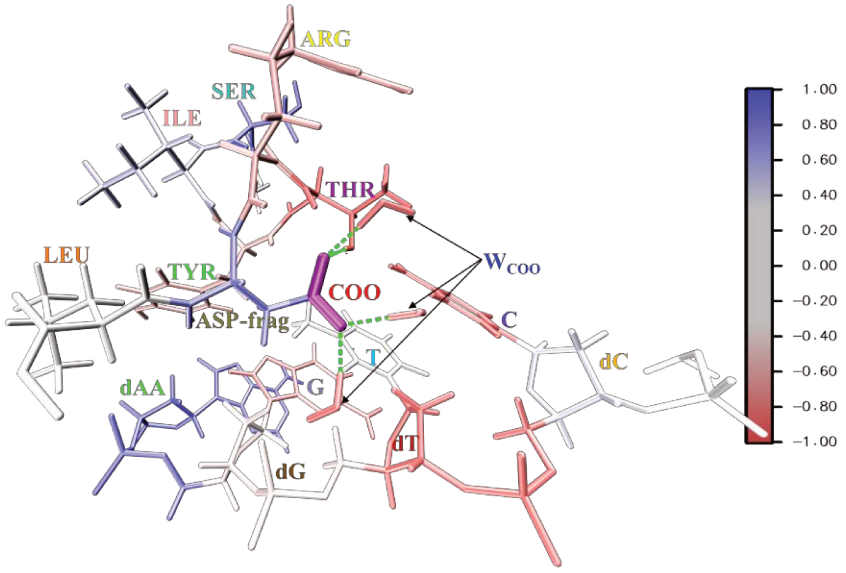

Figure 4-17: Licorice representation of the QM-region for the large system. The colors of the residues and fragments are scaled according to the values of the correlation coefficients with the COO fragments (represents in purple color). The residue labels correspond to the names and colors us used in this chapter. The green dash-lines represent hydrogen bonds with COO.

Considering now correlations involving the ASP-frag (highlighted in the tan box in Figure 4-16), a decrease in charge is observed to correlate with TYR, G, dT, C, and WAT, as per the correlation coefficient. Notably, these fragments are spatially distant from the ASP-frag and do not share a direct connection with it through either a hydrogen or chemical bond. In the case of thymine (cyan box in Figure 4-16), the decreasing charge is found to be correlated with neighboring fragments, dG and G, as well as with the distant fragment, TYR. In the instance of thymine moiety dT alteration, despite it being an impacted fragment, the initial charge fluctuation on it is relatively minimal (0.11). However, this variation increases progressively over time. The increasing charge on dT shows a strong inverse correlation with both ASP-frag and COO, which are associated with the hole-recombination mechanism, and SER and dAA, which are related to the process of electron emission.

Note that the correlation coefficient depends not only on the amplitude and vector of charge variation between fragments but also on the oscillation of charge over time. If there is a

correlation of charge migration between two fragments, the charge decrease on one fragment corresponds with an increase on the other, creating a typical oscillation. This oscillation serves as evidence of charge migration from one fragment to another. For example, in Figure 4-12 displaying the charge variation over time, a strong correlation could be anticipated between T and dT. However, the correlation coefficient between them is insignificant, as the oscillations of their charges do not coincide. Similarly, this holds true for ASP-frag and W_{COO} as well. This demonstrates the power of the correlation function in recognizing the correlation of charge migration between two fragments.

Bearing in mind that the process of charge migration between fragments could change very quickly, especially for complex systems, the correlation matrix however provides a global picture of the correlation between them over a given time range. This means that if the correlation coefficient between two fragments is weak, the charge migration between them is not entirely absent. It may occur in a very short time, but it is negligible within the chosen time range. Also, the correlation between fragments could change by altering the time range.

b. Correlations in the time range of 6-10 fs

The selection of the time range is crucial for the correlation function since the amplitude of charge migration vary over time. In Figure 4-18, we calculated the correlation coefficient between fragments for the same system as mentioned above, within a time range of 6-10 fs.

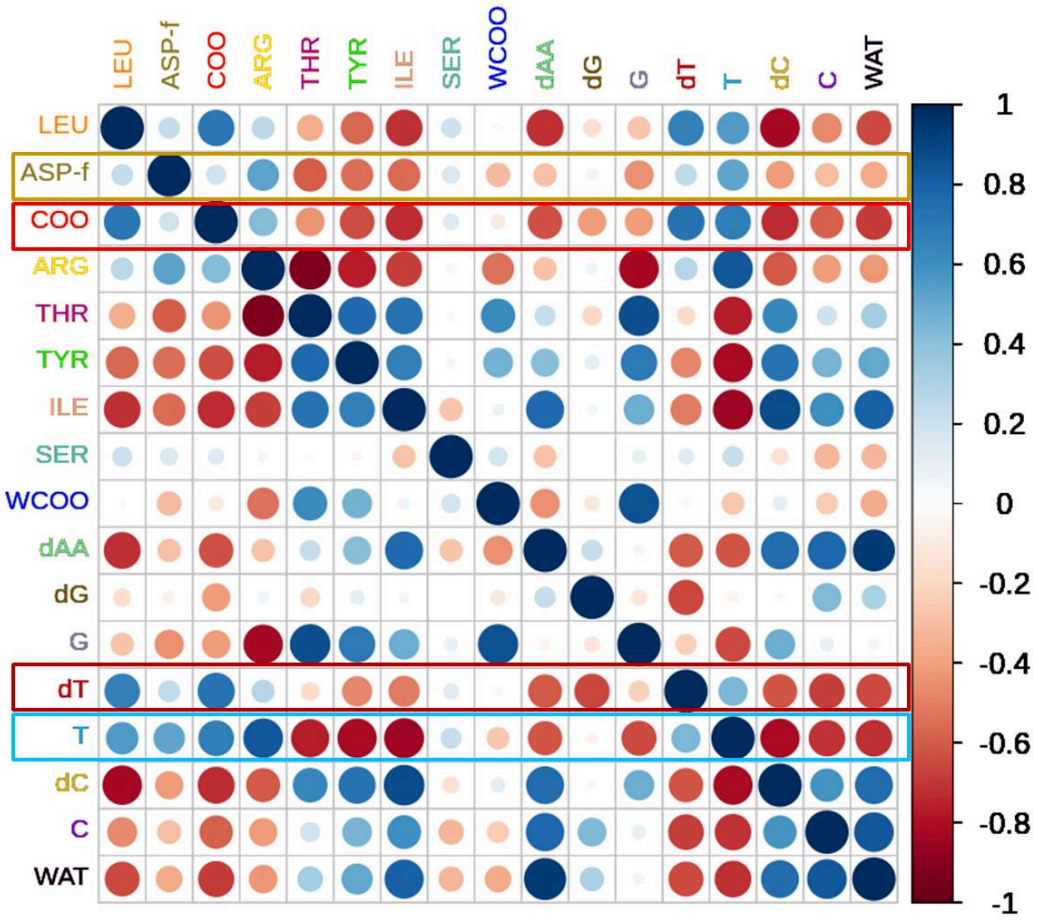

Figure 4-18: Correlation matrix of charge variation between residues in the large system during the first trajectory. The correlation coefficients were calculated over a time range of 6-10 fs. The size and color of the circles represent the magnitude of the correlation coefficient between two fragments.

During the time interval of 6-10 fs, significant modifications in the correlation between fragments are seen. The correlation coefficient between ASP-frag and other fragments is found to be less than 0.6, indicating the absence of correlation according to our criteria. This observation is supported by Figure 4-12, which demonstrates that the charge variation on ASP-frag remains stationary after 6 fs. For COO, the correlation pattern has changed compared to the time range of 1.75-10 fs. A relatively strong anti-correlation (-0.72) was observed between COO and both ILE and dC. However, the correlation coefficient with other fragments is less than 0.7, notably with THR and W_{COO}, which exhibited a strong correlation in earlier time range. In the case of dT, the strongest correlation observed is -0.66 with both dG and C. Despite the small correlation, as evident in Figure 4-12, this correlation of dT is as expected. They appear to exhibit the same oscillation in charge variation with dT. In the case of T, correlation coefficients greater than 0.8

are observed with TYR, ILE, and dC. These results show that the correlation of charge migration between fragments changes over time.

c. Correlation in the second trajectory

To investigate the relationship of charge migration between fragments in the second trajectory, we calculated a correlation matrix for the time range of 1.75-6 fs, as illustrated in Figure 4-19.

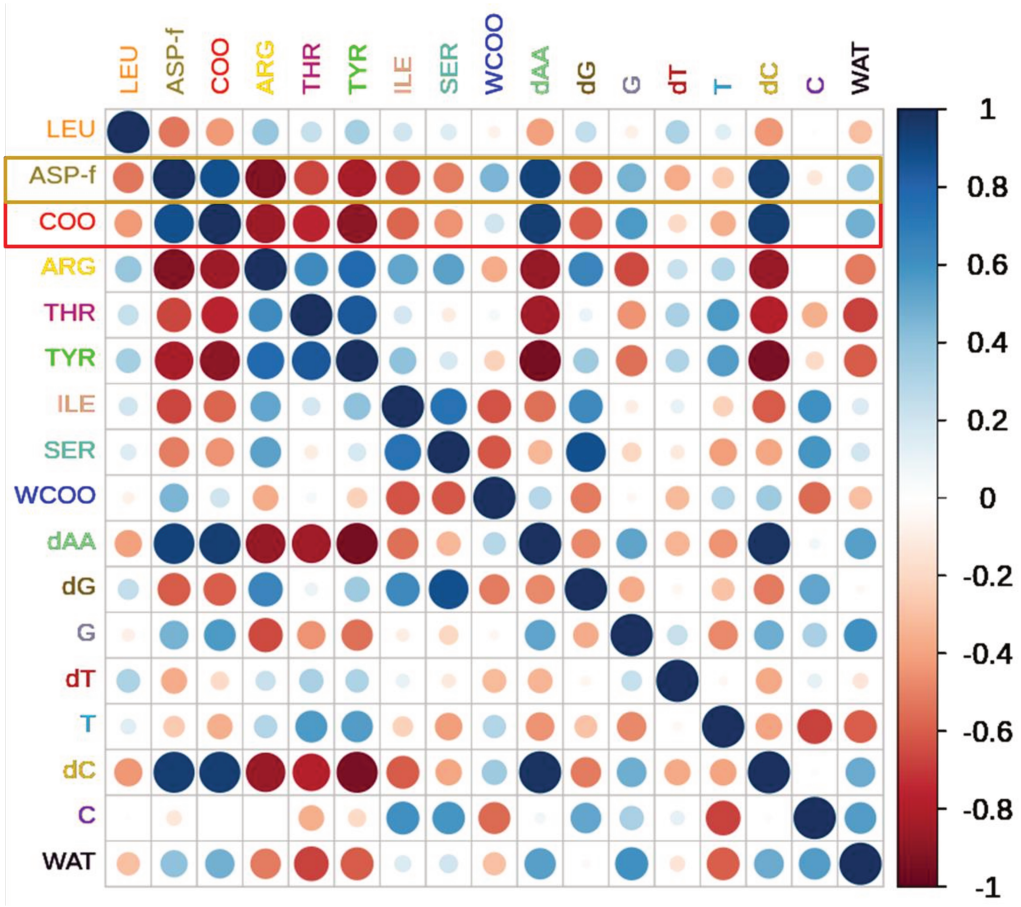

Figure 4-19: Correlation matrix of charge variation between residues in the large system during the second trajectory. The correlation coefficients were calculated over a time range of 1.5-6 fs. The size and color of the circles represent the magnitude of the correlation coefficient between two fragments.

In the second trajectory, the patterns between fragments change significantly compared to the first trajectory, especially for the principal impacted fragments. This may be due to alterations in the initial ionization and changes in the nature of ionization and excitation, as discussed in the previous section.

Within this time range, the reduction of charge variation on COO exhibits a strong anti-correlation (> 0.8) with ARG and TYR. This is different from the first trajectory, where THR, W_{COO}, dT, and C

were the correlating fragments. However, the week correlation of THR and negligible correlation of W_{COO} are observed with COO which have hydrogen bond with it. As illustrated in Figure 4-15, showing charge variation over time, there appears to be a correlation in charge migration between COO, W_{COO}, and THR, especially at the onset of the collision. Therefore, we calculated the correlation between these fragments in the time range of 1.5-2.5 fs. We observed a strong correlation of -0.94 with THR and -0.81 with W_{COO}. This suggests that the correlation of charge migration between these fragments is fleeting; after 2.5 fs, it changes entirely.

The same correlation is observed in the case of ASP-frag, which correlates with ARG and TYR. However, in the first trajectory, the correlating fragments were TYR, G, dT, C, and WAT. We also observe a strong positive correlation for dAA and dC with both ASP-frag and COO, which corresponds to a decrease in charge on these fragments—an unlikely event. As explained in the context of the first trajectory system, this may relate to the pattern of charge oscillation on those fragments which have a strong correlation with dAA and dC, and in turn, with ASP-frag and COO. As evidenced in the first trajectory, and similarly in the second trajectory, dAA and dC show a strong anti-correlation with those fragments which also strongly correlate with ASP-frag and COO, such as ARG and TYR. These results reveal the complexity of the charge migration process, which depends substantially on the nature of ionization and the time duration.

d. Partial conclusion

We have considered correlation matrix to analyze and identify the correlation of charge migration between fragments in RT-TD-ADFT, particularly in the context of large molecular systems. In addition to simplifying the analysis, the correlation matrix effectively succeeds in revealing the correlation between two fragments charge fluctuations, a task that is challenging when using a simple plot of charge variation over time. Another significant feature of the correlation function is its dependency on the data range, which proves useful in constructing a correlation map between two fragments over different time periods.

4.4.2 Cross-correlation analysis

The delay in charge migration between two fragments may occur in a large system. When the hole migrates from distant to target fragments through intermediate fragments, it may require

time, resulting in a delayed charge migration between them. Therefore, in this subsection, we examine the delay of charge migration between fragments in our first trajectory system.

I. Mathematical definition

One useful analytical method for determining the delays between two signals is the cross-correlation function. This measure is used to evaluate the similarity or correlation between two signals or time series based on the time-lag applied to one of them. In other words, it helps determine how much one signal or series resembles another when shifted in time. This technique is widely employed in signal processing, image processing[47,48], and time series analysis[44].

For continuous-time signals, $X(t)$and $Y(t)$, the cross-correlation function is defined as:

$$R_{xy}(t+\tau) = \frac{cov(X(t+\tau), Y(t))}{(\sigma_{X(t+\tau)} * \ \sigma_{Y(t)})}$$

Here, τ, representing time lags, that can take the values of 0, ±1, ±2, ±3, and so forth. A negative τ value indicates a correlation between the x-variable at a time before t and the y-variable at time t. Conversely, a positive τ value suggests a correlation between the x-variable at a time after t and the y-variable at time t. Where $R_{xy}(t+\tau)$ is a correlation coefficient at time-lag τ.

a. Cross-correlation with COO

In Figure 4-20, we report the cross-correlation function of charge migration between the COO fragment and all other fragments (considering the first trajectory). The data chosen for this calculation ranged from 1.75 to 6 fs. The time delay calculated was ± 1 fs. Given that the data of charge variation for COO shifts and decreases over time, the negative delay corresponds to the time needed for the hole to migrate from COO to the target fragments. A positive delay suggests the inverse, which is not meaningful in this context.

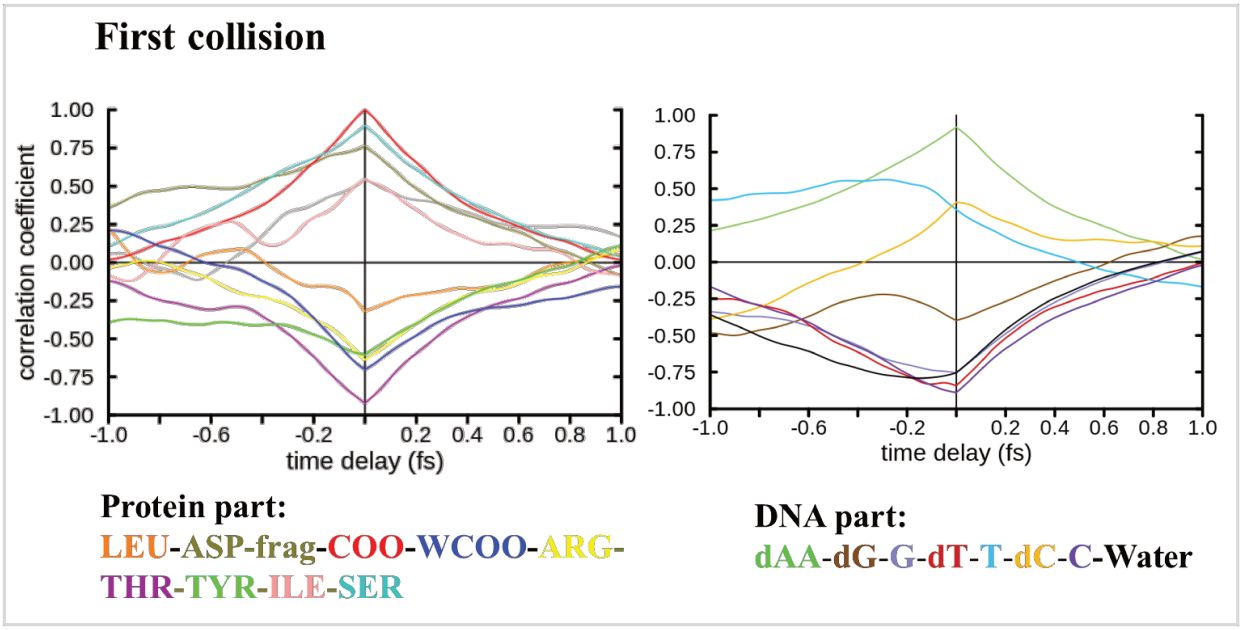

Figure 4-20: Cross-correlation of charge variation for the COO fragment with other fragments in the first trajectory system, over a time range of 1.75-6 fs. The x-axis represents the correlation coefficient for each time delay between the COO data and that of other fragments. Color coding corresponds to various residues and fragments.

According to the figures, the correlation between COO and other fragments diminishes with increasing delay from zero, except for T and WAT. This suggests the absence of delay in the charge migration process. For T and WAT, the correlation increases in the negative delay, reaching a maximum correlation coefficient of 0.56 and -0.79 respectively. However, the coefficient for T is not significant. As for WAT, it rises from -0.75 to -0.79 at -0.14 fs, which is not a significant jump. This suggests that the status of charge migration correlation for WAT does not change significantly over time delay. This calculation has been examined for other fragments and the results are consistent. The fact is, the process of charge migration is ultra-fast. This generally refers to the movement of an electron from one energy state (or molecular orbital) to another, often resulting from electron correlation and intra- and intermolecular energy transfer[49]. Essentially, when an electron decreases in one molecular orbital of a fragment, it simultaneously increases in the molecular orbitals of another fragment without passing through the intermediate fragments. Therefore, the delay in charge migration within our system is not significant.

b. Partial conclusion

In summary, the cross-correlation function is indeed a powerful tool for identifying correlations between two data sets over a delayed period of time. However, in our specific instance, the delay

of charge migration between fragments is not significant due to the extremely rapid nature of this charge migration. Nevertheless, this tool could hold potential interest for investigations concerning the electron transfer process. These processes generally occur over a longer time scale, and intermediate fragments play a key role in the transition process.

4.5 Addendum: perspectives for the modeling of decarboxylation

In this chapter, we placed specific emphasis on the irradiation of the carboxylic group of aspartic acid. This is due to the fact that the decarboxylation of acidic amino acids plays a significant role in the process of radiation-induced damage to proteins. However, the mechanisms underlying this process, particularly at low temperatures (100 K), remain unclear. Decarboxylation is prompted when a hole directly locates on the side chain carboxylic group, or when it reacts with highly oxidative radiolytic species such as hydroxyl radicals ($OH^{\cdot}$) through an indirect mechanism.

At low temperatures, the diffusion of radiolytic species is negligible, as confirmed by Allan and colleagues[16]. This suggests that decarboxylation is unlikely to occur through indirect damage under these conditions. Furthermore, both theoretical[50] and experimental[51] evidence indicate that the local temperature of macromolecular system does not increase by more than 10 K due to X-ray absorption, even at high doses. This assertion supports the conclusion that radiolytic species are not activated by X-ray absorption.

The mechanism of hole generation on the carboxylic group can predominantly occur in two ways. Firstly, the carboxylic group can be ionized directly by absorbing X-rays or secondary electrons. Secondly, the creation of a hole on the carboxylic group could result from hole migration or transition from other residues.

According to the cross-section, radiation damage to acidic amino acids is expected to be less severe than for several other residues, such as Arginine. However, this does not seem to be the case. Furthermore, even though the decarboxylation process is chemically more favorable than other fragmentations, all acidic residues should exhibit similar radiation responses within the protein, according to their respective cross sections. This implies that decarboxylation of acidic

amino acids should not depend on their position within the protein or their respective environments. However, as demonstrated in the introduction, radiation-induced decarboxylation in acidic residues substantially depends on the amino acid's environment within the protein/DNA complex. Therefore, in consideration of initial ionization, hole transition is expected to play a significant role in the radiation damage of acidic amino acids.

In terminology, the processes of hole migration and transfer are essentially different. The main distinction lies in that hole migration refers to the rapid redistribution of charge within the system due to electron dynamics, assuming that the coupling dynamics and movement of atomic nuclei in the system are insignificant. On the other hand, in the hole transfer process, the coupling dynamics and movement of atomic nuclei within the system play a crucial role in charge redistribution. At 100 K, macromolecular X-ray diffraction experiments[52] demonstrate that the atomic movement and alterations in protein globule size during the experiment are related to the fragment product induced by the X-ray. This implies that fundamental movement of atomic nuclei occurs after decarboxylation, although the subsequent change in atomic position is not drastic at atomic level. It is crucial to note, however, that while we cannot definitively claim that charge migration is the dominant process, it could indeed be significant. Minor nuclear movements might exert a substantial influence on the system's electron density redistribution, and the absorption of X-rays may activate vibrational modes.

These observations indeed underscore the importance of the charge migration process involved in the decarboxylation of acidic residues. However, we must bear in mind that our results cannot be directly compared with the consequences of X-ray interactions, as the nature of ionization and the dynamics of electron relaxation by X-rays significantly differ, as outlined in Chapter One. For example, X-rays essentially ionize the core electrons via the Photoelectric Effect at low energies or outer shell electrons via Compton Scattering at moderate energies. Conversely, charged particles primarily target valence electrons, especially at the Bragg's peak. A distinct characteristic of charged particles is their high linear energy transfer (LET), which allows them to deposit substantial energy in a localized position along their path far more efficiently than X-rays, which possess low LET. Consequently, hot spots of radicals and excited molecules are produced along the path of the charged particle, leading to an increase in localized temperature at the atomic

level. Therefore, the movement of atomic nuclei and the occurrence of indirect effects are expected to rise during charged particle interactions.

To the best of our knowledge, existing literature does not provide clear studies on the decarboxylation of acidic amino acids in proteins induced by charged particles, particularly at low temperatures, and differentiated from X-ray interactions. However, according to two trajectories specifically targeting the side chain carboxylic group of Aspartic acid, the hole-recombination on the carboxylic group is remarkably fast, taking less than 10 fs due to charge migration. Restrained within fixed nuclei and extremely short timeframes, our results suggest that the decarboxylation process facilitated by the charge migration mechanism is less probable during charged particle interactions. However, this conclusion requires further substantiation, for instance, by increasing the number of trajectories and including nuclear dynamics in our simulations. These considerations may form the basis of our future work.

4.6 Conclusion

In this Chapter we have explored the physical stage of the alpha irradiation of a protein/DNA complex by means of RT-TD-ADFT simulations within a QM/MM framework. Both a polarizable and a non-polarizable force fields have been considered to deal with the environment. We have varied the size of the QM region and we have focused our attention on two specific irradiation conditions (propagation vectors of the α-particle).

We have evaluated the influence of electrostatic induction in the MM environment. Our results indicate that electrostatic induction has a negligible effect on energy deposition. On the other hand, for the system investigated here, The MMpol approach dose not is not affect initial ionization, although it can influence charge migrations. Further, as the size of the QM region diminishes, the differences in charge variation of fragments, as calculated within MMpol and MM environments, tend to increase. The best strategy to follow seems to increase as much as computationally feasible the size of the QM region.

We have examined ultra-fast charge migrations within the protein/DNA system, occurring on the scale of a few femtoseconds. Notably, when a hole is located on the carboxylate group of the aspartate residue, it migrates entirely to the surrounding fragment in less than 10 fs. This likely indicates electron-hole recombinations, a key process in radiation-induced decarboxylation mechanisms. Moreover, the direction of the α-particle does not seem to affect the hole-recombination mechanism of the carboxylate group. Rather, it only influences the speed of hole-recombination, as evidenced by the two different trajectories.

In the last section we have explored the use of correlation matrices to analyze charge migration correlations among fragments. The correlation matrix is a powerful analytical tool that simplifies data representation and provides valuable insights into relationships among different fragments in large molecular systems. We tested the cross-correlation function to identify any delay in charge migration between fragments. However, in our particular situation, the delay in charge migration proved to be insignificant, owing to the swift nature of this process.

REFERENCES

(1) Hogan, D. J.; Riordan, D. P.; Gerber, A. P.; Herschlag, D.; Brown, P. O. Diverse RNA-Binding Proteins Interact with Functionally Related Sets of RNAs, Suggesting an Extensive Regulatory System. *PLoS Biol* **2008**, *6* (10), e255. https://doi.org/10.1371/journal.pbio.0060255.

(2) Echols, H. Nucleoprotein Structures Initiating DNA Replication, Transcription, and Site-Specific Recombination. *The Journal of biological chemistry* **1990**, *265* (25), 14697—14700. https://doi.org/10.1016/s0021-9258(18)77163-9.

(3) Cox, M. M.; Goodman, M. F.; Kreuzer, K. N.; Sherratt, D. J.; Sandler, S. J.; Marians, K. J. The Importance of Repairing Stalled Replication Forks. *Nature* **2000**, *404* (6773), 37–41. https://doi.org/10.1038/35003501.

(4) Luger, K.; Richmond, T. J. The Histone Tails of the Nucleosome. *Current Opinion in Genetics & Development* **1998**, *8* (2), 140–146. https://doi.org/10.1016/S0959-437X(98)80134-2.

(5) Noller, H. F.; Lancaster, L.; Mohan, S.; Zhou, J. Ribosome Structural Dynamics in Translocation: Yet Another Functional Role for Ribosomal RNA. *Quarterly Reviews of Biophysics* **2017**, *50*, e12. https://doi.org/10.1017/S0033583517000117.

(6) Davídková, M.; Štísová, V.; Goffinont, S.; Gillard, N.; Castaing, B.; Spotheim-Maurizot, M. Modification of DNA Radiolysis by DNA-Binding Proteins: Structural Aspects. *Radiation Protection Dosimetry* **2006**, *122* (1–4), 100–105. https://doi.org/10.1093/rpd/ncl442.

(7) Gillard, N.; Begusova, M.; Castaing, B.; Spotheim-Maurizot, M. Radiation Affects Binding of Fpg Repair Protein to an Abasic Site Containing DNA. *Radiation Research* **2004**, *162* (5), 566–571. https://doi.org/10.1667/RR3247.

(8) Spotheim-Maurizot, M.; Davídková, M. Radiation Damage to DNA in DNA–Protein Complexes. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **2011**, *711* (1–2), 41–48. https://doi.org/10.1016/j.mrfmmm.2011.02.003.

(9) Eon, S.; Culard, F.; Sy, D.; Charlier, M.; Spotheim-Maurizot, M. Radiation Disrupts Protein–DNA Complexes through Damage to the Protein. The *Lac* Repressor–Operator System. *Radiation Research* **2001**, *156* (1), 110–117. https://doi.org/10.1667/0033-7587(2001)156[0110:RDPDCT]2.0.CO;2.

(10) Bury, C.; Garman, E. F.; Ginn, H. M.; Ravelli, R. B. G.; Carmichael, I.; Kneale, G.; McGeehan, J. E. Radiation Damage to Nucleoprotein Complexes in Macromolecular Crystallography. *Journal of Synchrotron Radiation* **2015**, *22* (2), 213–224. https://doi.org/10.1107/S1600577514026289.

(11) Bury, C. S.; Carmichael, I.; McGeehan, J. E.; Garman, E. F. Radiation Damage within Nucleoprotein Complexes Studied by Macromolecular X-Ray Crystallography. *Radiation Physics and Chemistry* **2016**, *128*, 118–125. https://doi.org/10.1016/j.radphyschem.2016.05.023.

(12) Pan, X.; Sanche, L. Dissociative Electron Attachment to DNA Basic Constituents: The Phosphate Group. *Chemical Physics Letters* **2006**, *421* (4), 404–408. https://doi.org/10.1016/j.cplett.2006.01.099.

(13) Bury, C. S.; McGeehan, J. E.; Antson, A. A.; Carmichael, I.; Gerstel, M.; Shevtsov, M. B.; Garman, E. F. RNA Protects a Nucleoprotein Complex against Radiation Damage. *Acta Crystallogr D Struct Biol* **2016**, *72* (5), 648–657. https://doi.org/10.1107/S2059798316003351.

(14) Sevilla, M. D.; D'Arcy, J. B.; Morehouse, K. M. An Electron Spin Resonance Study of .Gamma.-Irradiated Frozen Aqueous Solutions Containing Dipeptides. Mechanisms of Radical Reaction. *J. Phys. Chem.* **1979**, *83* (22), 2887–2892. https://doi.org/10.1021/j100485a016.

(15) Burmeister, W. P. Structural Changes in a Cryo-Cooled Protein Crystal Owing to Radiation Damage. *Acta Crystallogr D Biol Crystallogr* **2000**, *56* (3), 328–341. https://doi.org/10.1107/S0907444999016261.

(16) Allan, E. G.; Kander, M. C.; Carmichael, I.; Garman, E. F. To Scavenge or Not to Scavenge, That Is STILL the Question. *Journal of Synchrotron Radiation* **2013**, *20* (1), 23–36. https://doi.org/10.1107/S0909049512046237.

(17) Gerstel, M.; Deane, C. M.; Garman, E. F. Identifying and Quantifying Radiation Damage at the Atomic Level. *J Synchrotron Rad* **2015**, *22* (2), 201–212. https://doi.org/10.1107/S1600577515002131.

(18) Fioravanti, E.; Vellieux, F. M. D.; Amara, P.; Madern, D.; Weik, M. Specific Radiation Damage to Acidic Residues and Its Relation to Their Chemical and Structural Environment. *J Synchrotron Rad* **2007**, *14* (1), 84–91. https://doi.org/10.1107/S0909049506038623.

(19) Alvarez-Ibarra, A.; Omar, K. A.; Hasnaoui, K.; De La Lande, A. Chapter 4. Electron and Molecular Dynamics Simulations with Polarizable Embedding. In *Theoretical and Computational Chemistry Series*; Salahub, D. R., Wei, D., Eds.; Royal Society of Chemistry: Cambridge, 2021; pp 117–143. https://doi.org/10.1039/9781839164668-00117.

(20) Phillips, J. C.; Hardy, D. J.; Maia, J. D. C.; Stone, J. E.; Ribeiro, J. V.; Bernardi, R. C.; Buch, R.; Fiorin, G.; Hénin, J.; Jiang, W.; McGreevy, R.; Melo, M. C. R.; Radak, B. K.; Skeel, R. D.; Singharoy, A.; Wang, Y.; Roux, B.; Aksimentiev, A.; Luthey-Schulten, Z.; Kalé, L. V.; Schulten, K.; Chipot, C.; Tajkhorshid, E. Scalable Molecular Dynamics on CPU and GPU Architectures with NAMD. *The Journal of Chemical Physics* **2020**, *153* (4), 044130. https://doi.org/10.1063/5.0014475.

(21) McGeehan, J. E.; Streeter, S. D.; Thresh, S.-J.; Ball, N.; Ravelli, R. B. G.; Kneale, G. G. Structural Analysis of the Genetic Switch That Regulates the Expression of Restriction-Modification Genes. *Nucleic Acids Research* **2008**, *36* (14), 4778–4787. https://doi.org/10.1093/nar/gkn448.

(22) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *The Journal of Chemical Physics* **1983**, *79* (2), 926–935. https://doi.org/10.1063/1.445869.

(23) Tian, C.; Kasavajhala, K.; Belfon, K. A. A.; Raguette, L.; Huang, H.; Migues, A. N.; Bickel, J.; Wang, Y.; Pincay, J.; Wu, Q.; Simmerling, C. Ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution. *Journal of Chemical Theory and Computation* **2020**, *16* (1), 528–552. https://doi.org/10.1021/acs.jctc.9b00591.

(24) Zgarbová, M.; Šponer, J.; Otyepka, M.; Cheatham, T. E. I.; Galindo-Murillo, R.; Jurečka, P. Refinement of the Sugar–Phosphate Backbone Torsion Beta for AMBER Force Fields Improves the Description of Z- and B-DNA. *Journal of Chemical Theory and Computation* **2015**, *11* (12), 5723–5736. https://doi.org/10.1021/acs.jctc.5b00716.

(25) Warshel, A.; Levitt, M. Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme. *Journal of Molecular Biology* **1976**, *103* (2), 227–249. https://doi.org/10.1016/0022-2836(76)90311-9.

(26) Baker, C. M. Polarizable Force Fields for Molecular Dynamics Simulations of Biomolecules. *WIREs Comput Mol Sci* **2015**, *5* (2), 241–254. https://doi.org/10.1002/wcms.1215.

(27) Warshel, A.; Kato, M.; Pisliakov, A. V. Polarizable Force Fields: History, Test Cases, and Prospects. *J. Chem. Theory Comput.* **2007**, *3* (6), 2034–2045. https://doi.org/10.1021/ct700127w.

(28) Ren, P.; Ponder, J. W. Polarizable Atomic Multipole Water Model for Molecular Mechanics Simulation. *J. Phys. Chem. B* **2003**, *107* (24), 5933–5947. https://doi.org/10.1021/jp027815+.

(29) Samaniego-Rojas, J. D.; Hernández-Segura, L.-I.; López-Sosa, L.; Delgado-Venegas, R. I.; Gomez, B.; Lambry, J.-C.; De La Lande, A.; Mineva, T.; Alejandre, J.; Zúñiga-Gutiérrez, B. A.; Flores-Moreno, R.; Calaminici, P.; Geudtner, G.; Köster, A. M. Chapter 1. QM/MM with Auxiliary DFT in DeMon2k. In *Theoretical and Computational Chemistry Series*; Salahub, D. R., Wei, D., Eds.; Royal Society of Chemistry: Cambridge, 2021; pp 1–54. https://doi.org/10.1039/9781839164668-00001.

(30) Wang, Z.-X.; Zhang, W.; Wu, C.; Lei, H.; Cieplak, P.; Duan, Y. Strike a Balance: Optimization of Backbone Torsion Parameters of AMBER Polarizable Force Field for Simulations of Proteins and Peptides. *Journal of Computational Chemistry* **2006**, *27* (6), 781–790. https://doi.org/10.1002/jcc.20386.

(31) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99SB. *J. Chem. Theory Comput.* **2015**, *11* (8), 3696–3713. https://doi.org/10.1021/acs.jctc.5b00255.

(32) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865–3868. https://doi.org/10.1103/PhysRevLett.77.3865.

(33) Gómez Pueyo, A.; Marques, M. A. L.; Rubio, A.; Castro, A. Propagators for the Time-Dependent Kohn–Sham Equations: Multistep, Runge–Kutta, Exponential Runge–Kutta, and Commutator Free Magnus Methods. *J. Chem. Theory Comput.* **2018**, *14* (6), 3040–3052. https://doi.org/10.1021/acs.jctc.8b00197.

(34) Cheng, C.-L.; Evans, J. S.; Van Voorhis, T. Simulating Molecular Conductance Using Real-Time Density Functional Theory. *Phys. Rev. B* **2006**, *74* (15), 155112. https://doi.org/10.1103/PhysRevB.74.155112.

(35) Castro, A.; Marques, M. A. L.; Rubio, A. Propagators for the Time-Dependent Kohn–Sham Equations. *The Journal of Chemical Physics* **2004**, *121* (8), 3425–3433. https://doi.org/10.1063/1.1774980.

(36) Carmona-Espíndola, J.; Gázquez, J. L.; Vela, A.; Trickey, S. B. Generalized Gradient Approximation Exchange Energy Functional with Correct Asymptotic Behavior of the Corresponding Potential. *J. Chem. Phys.* **2015**, *142* (5), 054105. https://doi.org/10.1063/1.4906606.

(37) Calaminici, P.; Janetzko, F.; Köster, A. M.; Mejia-Olvera, R.; Zuniga-Gutierrez, B. Density Functional Theory Optimized Basis Sets for Gradient Corrected Functionals: 3d Transition Metal Systems. *The Journal of Chemical Physics* **2007**, *126* (4), 044108. https://doi.org/10.1063/1.2431643.

(38) Calaminici, P.; Flores–Moreno, R.; Köster, A. M. A Density Functional Study of Structures and Vibrations of Ta3O and Ta3O−. *Computing Letters* **2005**, *1* (4), 164–171. https://doi.org/10.1163/157404005776611420.

(39) Hirshfeld, F. L. Bonded-Atom Fragments for Describing Molecular Charge Densities. *Theoret. Chim. Acta* **1977**, *44* (2), 129–138. https://doi.org/10.1007/BF00549096.

(40) Tandiana, R.; Clavaguéra, C.; Hasnaoui, K.; Pedroza-Montero, J. N.; De La Lande, A. Reliability and Performances of Real-Time Time-Dependent Auxiliary Density Functional Theory. *Theor Chem Acc* **2021**, *140* (9), 126. https://doi.org/10.1007/s00214-021-02819-9.

(41) Wu, X.; Teuler, J.-M.; Cailliez, F.; Clavaguéra, C.; Salahub, D. R.; de la Lande, A. Simulating Electron Dynamics in Polarizable Environments. *Journal of Chemical Theory and Computation* **2017**, *13* (9), 3985–4002. https://doi.org/10.1021/acs.jctc.7b00251.

(42) Alvarez-Ibarra, A.; Parise, A.; Hasnaoui, K.; de la Lande, A. The Physical Stage of Radiolysis of Solvated DNA by High-Energy-Transfer Particles: Insights from New First Principles Simulations. *Phys. Chem. Chem. Phys.* **2020**, *22* (15), 7747–7758. https://doi.org/10.1039/D0CP00165A.

(43) Vinod, H. D. Matrix Algebra Topics in Statistics and Economics Using R. In *Handbook of Statistics*; Elsevier, 2014; Vol. 32, pp 143–176. https://doi.org/10.1016/B978-0-444-63431-3.00004-8.

(44) Derrick, T.R. and Thomas, J.M. Time-Series Analysis: The Cross-Correlation Function. In *Innovative Analyses of Human Movement*; Human Kinetics Publishers: Illinois, 2004; pp 189–205.

(45) TEAM, R. A Language and Environment for Statistical Computing. *http://www.R-project.org* **2009**.

(46) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *Journal of Molecular Graphics* **1996**, *14* (1), 33–38. https://doi.org/10.1016/0263-7855(96)00018-5.

(47) Quaegebeur, N.; Padois, T.; Gauthier, P.-A.; Masson, P. Enhancement of Time-Domain Acoustic Imaging Based on Generalized Cross-Correlation and Spatial Weighting. *Mechanical Systems and Signal Processing* **2016**, *75*, 515–524. https://doi.org/10.1016/j.ymssp.2015.12.012.

(48) Hyde, J. S.; Jesmanowicz, A. Cross-Correlation: An FMRI Signal-Processing Strategy. *NeuroImage* **2012**, *62* (2), 848–851. https://doi.org/10.1016/j.neuroimage.2011.10.064.

(49) Jahnke, T.; Hergenhahn, U.; Winter, B.; Dörner, R.; Frühling, U.; Demekhin, P. V.; Gokhberg, K.; Cederbaum, L. S.; Ehresmann, A.; Knie, A.; Dreuw, A. Interatomic and Intermolecular Coulombic Decay. *Chem. Rev.* **2020**, *120* (20), 11295–11369. https://doi.org/10.1021/acs.chemrev.0c00106.

(50) Mhaisekar, A.; Kazmierczak, M. J.; Banerjee, R. Three-Dimensional Numerical Analysis of Convection and Conduction Cooling of Spherical Biocrystals with Localized Heating from Synchrotron X-Ray Beams. *J Synchrotron Rad* **2005**, *12* (3), 318–328. https://doi.org/10.1107/S0909049505003250.

(51) Snell, E. H.; Bellamy, H. D.; Rosenbaum, G.; Van Der Woerd, M. J. Non-Invasive Measurement of X-Ray Beam Heating on a Surrogate Crystal Sample. *J Synchrotron Rad* **2007**, *14* (1), 109–115. https://doi.org/10.1107/S090904950604605X.

(52) Petrova, T.; Lunin, V. Y.; Ginell, S.; Hazemann, I.; Lazarski, K.; Mitschler, A.; Podjarny, A.; Joachimiak, A. X-Ray-Radiation-Induced Cooperative Atomic Movements in Protein. *Journal of Molecular Biology* **2009**, *387* (5), 1092–1105. https://doi.org/10.1016/j.jmb.2009.02.030.

CHAPTER FIVE

ULTRAFAST RESPONSES OF INSULIN^{+6} AND SUBSTANCE PH^{+} TO XUV-PULSES IN GAS PHASE

Contents

5 Ultrafast responses of Insulin^{+6} and Substance PH^{+} to XUV-pulses in gas phase

This PhD thesis started with studies on nucleobase models (Chapters, 2 and 3), we then moved on to DNA/protein complexes (Chapter 4), and we now focus in Chapter 5 on protein irradiation. Proteins are fundamental in biology due to their diverse and multifaceted functions[1]. While it is impossible to list all their functions here, a few notable examples include their roles as enzymes, structural supports, transport mechanisms, storage vessels, signal transmitters, regulators, facilitators of cell movement, antibodies, receptors and hormones. The function of a protein hinges on several factors: the type and sequence of amino acids, the secondary structure formed by hydrogen bonds between amino acid backbones, the tertiary structure arising from interactions among amino acid side chains, and for some proteins, the quaternary structure resulting from the assembly of multiple protein units. Any minor alterations in protein structure can profoundly impact their essential biological functions.

Ionizing radiation (IoR) can significantly affect protein function through energy deposition and the formation of radical ions[2]. IoR can alter the chemical structure of proteins by inducing changes such as the breaking of chemical and hydrogen bonds. Examples include the decarboxylation of amino acids and the disruption of disulfide bridges[3,4]. These changes can lead to modifications in both secondary and tertiary structures, as seen in instances like chromatin radiation-induced modifications[5,6].

Most studies on radiation damage to biomolecules have focused on damage caused by charged particles or X- or γ-rays. This may be due to their use in radiotherapy and medical imaging. Nowadays, attosecond spectroscopies relying for instance on Extreme Ultraviolet (XUV) rays are prone to play central role in revealing the earliest stages of the mechanisms leading to radiation damages. One characteristic of XUV rays is their strong interaction with valence electrons, leading to depopulation.

In this chapter, we investigate the ultra-fast electronic responses of a protein, specifically insulin in its six-time protonated form, and a peptide, substance P in its singly protonated form, to an XUV-pulse using real-time time-dependent auxiliary density functional theory (RT-TD-ADFT).

Attosecond pump-probe experiments with XUV pulses have played a pivotal role in uncovering the ultra-fast mechanisms taking place in small and relevant biomolecules in the gas phase[7–9]. However, investigations of larger biomolecules have typically been restricted to amino acids[10] or small peptides[11]. At this point, it remains unclear whether the size of the biomolecule or the secondary structure of proteins influences ionization, electronic relaxation, energy redistribution, and the ultra-fast events triggered by XUV.

Insulin and substance P are found in Human. Insulin is a hormone that plays a critical role in the regulation of glucose metabolism. Produced by the beta cells of the islets of Langerhans in the pancreas, insulin facilitates the uptake of glucose from the bloodstream into cells, primarily muscle and adipose tissue cells. This process allows cells to use glucose for energy or store it for future use[12,13]. Insulin is a biopolymer comprised of 51 amino acid residues[14]. In its six-times protonated form, it has 794 atoms. Substance P is an undecapeptide[15], that functions as a neurotransmitter and neuromodulator. It plays a role in pain transmission in the central nervous system and also has various peripheral roles, including the regulation of blood pressure, smooth muscle contraction, and salivary secretion[16–18].

Studying large biomolecules, such as peptides or proteins, poses significant challenges for both experimental and theoretical methods. Recently, our experimental collaborators, led by F. Lépine at ILM in Lyon, have pioneered a technique using XUV-IR pump-probe lasers to probe biomolecules as large as entire proteins[19]. This work is a collaboration with experimentalists from F. Lépine's team and theoreticians from M. Vacher's group at CEISAM Nantes, who are specialists in surface hopping non-adiabatic dynamics.

Our experimental colleagues observed an ultrafast proton transfer occurring within a few fs, as well as structural relaxation taking hundreds of fs, both induced by the XUV-pulse. These phenomena are associated with the local characteristics of the moments immediately following ionization and the H-bond network. Our objectives for this project are twofold. First, we aim to investigate the molecular conformations of insulin and substance P. This will help us understand the interplay between the molecular conformations adopted by the proteins and the ultra-fast events probed experimentally. A second goal is to investigate the mechanism of ultra-fast ionization prompted by the XUV-pulse. Given that the nuclei of molecules remain fixed in RT-TD-

ADFT, and considering the extended duration of the pulse in the experiments, our focus is on the primarily ionization events. Simulations that incorporate the effect of nuclear motion are taken in charge in M. Vacher's group (CEISAM, Nantes). The manuscript detailing this project is currently in preparation. To maintain confidentiality, this chapter will only concentrate on my contributions to this project.

Describing entire large biomolecules, especially insulin, using RT-TD-DFT presents significant challenges. These challenges arise not only due to the size of the molecule but also because of the necessity to explicitly introduce the pulse into the simulation, which demands a comprehensive description of the continuum state. To the best of our knowledge, this is the first time a biomolecule of this size has been studied in the context of XUV-pulse interaction at the RT-TD-DFT level.

Directly introducing the pulse into the simulation is critical for accurately describing ionization processes and ultra-fast electron responses. F. Calegari *et al.*[8,11,20] have demonstrated that the actual nature of ionization and charge migration induced by an XUV-pulse in phenylalanine differs greatly from simulations where a localized hole is simply created in the dynamics. They showed that the XUV-pulse significantly depopulate numerous valence molecular orbitals. As a result, the charge becomes strongly delocalized across the molecule, complicating the task of tracking charge migrations. To observe these phenomena, the authors performed a Fourier transformation of electron density fluctuations for a specific fragment. Their findings revealed that the spectrum of this fragment does not display a singular signal of periodic charge migration. Instead, several signals of charge migration were observed on the fragment. This observation suggests the presence of multiple intricate charge migrations that are not captured by localized hole approximations.

As discussed in Chapter One, the continuum state is inadequately described by a local basis set, which causes the continuum to convert into discrete states. This can significantly impact the ionization process. For small molecules, it is well-established that the standard local basis set is ill-suited to describe the ionization process[21]. One approach to addressing this limitation is to add Atomic Orbitals designed for the Continuum state (AOC) to the standard basis set[22,23]. However, incorporating AOCs greatly increases computational costs, rendering this approach impractical

for large molecular systems. Yet, for larger molecules, it remains uncertain whether a local basis set can reliably describe the ionization process. It is postulated that in large molecular systems, the continuum state is enhanced due to the sharing of Molecular Orbitals (MO) among atoms within the molecule. White *et al.*[24] demonstrated that introducing a ghost atom into the high harmonic generation spectrum of H_2 significantly improved the high-energy spectrum, corresponding to the quality of the continuum state. However, it's still undetermined how the size of a real molecular system affects the continuum. This is crucial for our work, as using AOCs for large biomolecules like insulin is not feasible. Thus, before undertaking RT-TD-ADFT simulations for insulin^{+6} and substance PH^+, we conducted an in-depth investigation of the use of local basis set using small molecules and several peptides.

This chapter is organized into five sections. Section 5.1 focuses on benchmarking the use of local basis sets in XUV simulations for both small molecules and relatively large peptides. Section 5.2 delves into the structural study of insulin^{+6} and substance PH^+ using molecular dynamics simulations. In Section 5.3, we explore the primary ionization events in insulin^{+6} and substance PH^+ triggered by XUV-pulse, utilizing RT-TD-ADFT simulations. The chapter concludes with two sections: conclusions and perspectives.

5.1 Local basis set for XUV simulation

Extreme UV (XUV) rays, with an energy range of about 10-130 eV, interact strongly with valence electrons, causing excitations and ionization. Accurate simulation of the XUV ionization need a proper description of the continuum. Local basis sets, for example Gaussian or Slater type orbitals, are effective in describing bound states; however, as mentioned in the first chapter, the continuum state is poorly described by standard local basis sets. This leads to the conversion of the continuous state into discrete states and produces artificially strong absorption peaks in the high-energy spectrum.

The strategy to improve the continuum description is to complement the standard basis set with an atomic basis set designed for the continuum. In addition, the quality of the continuum is

expected to improve with increasing molecular size as atoms share molecular orbitals, although it is still unclear how this happens.

This section starts with an introduction to the basic building blocks used in this test and the detailed methodology. First, we investigate the addition of continuum atomic orbitals to a standard basis set and the effect of molecular size on continuum enhancement in small molecules such as N_2. Second, we extend our benchmark on a series of peptides.

5.1.1 Basis set and methodology

a. Dedicated atomic orbitals for the continuum

A strategy for enhancing the representation of continuum states is to augment the standard basis set with Gaussian atomic orbitals specifically fitted for the continuum. Several methodologies have been put forth to achieve this. Notably, Nestmann and Peyerimhoff[25], with subsequent extensions by Faure and coworkors[26], introduced a method to derive Gaussians optimized for the continuum (AOC) in the context of electron-molecule scattering[27,28]. This specific procedure is incorporated into the GTOBAS software[26], allowing users to produce tailored AOCs for any atom they're working with. Atomic Gaussian basis functions are fitted to reproduce the continuum wave functions described as Coulomb or Bessel functions. The user can refine the AOCs for their specific system and the targeted physical process by setting parameters.

The fitting procedure requires the effective nuclear charge of the atom of interest, for nitrogen for instance, it could be 7 or 6. Additionally, the definition of a boundary radius, R is needed, as it defines the finite range of space where the Coulomb functions are calculated. Other parameters are the maximum angular momentum number (L) of the AOC, and the upper bound energy (E), which determines the number of Coulomb functions to be fitted, and consequently, the number of fitted basis functions. We will denote the AOC as Lx_Ex_Rx, where "x" stands for the fitting parameters for each set.

In addition, we have investigated the use of effective core potentials. As XUV rays primarily affect valence electrons, the use of pseudo-potential could be used to reduce the computational cost of large molecular simulations. Foglia and colleagues[29] examined the influence of core electrons on spectroscopy within the range of valence excitation using RT-TD-DFT. Their findings indicate that

core electrons have a minimal effect on the spectral information. For this purpose, we adopted the Relativistic Effective Core Potential | Stuttgart-Dresden (RECP | SD) for the valence electrons and RECP for the effective core electrons[30,31]. In our tests, we refer to this as ECP.

In peptide tests, the molecular systems are relatively large, making it challenging to use AOC with extensive parameters. Therefore, the reference basis set employed to evaluate the performance of the standard basis set and ECP consists of TZVP augmented with AOC, using parameters E=5 Ry, L=2, and R=3 bohr for heavy atoms (C, N, O, S). For all hydrogen atoms, we used DZVP-GGA (the standard DZVP adapted for the generalized gradient approximation)[32]. In our tests, this is referred to as ref-AOC.

b. Simulation details

In this section, all simulations were conducted using RT-TD-ADFT (Real-Time Time-Dependent Auxiliary Density Functional Theory), as implemented in our in-house version of deMon2k (version 6.1.6)[33]. The PBE[34] was chosen for exchange and correlation functionals. For density fitting, we used the automatically generated GEN-A2* auxiliary basis set[35]. Time propagation was facilitated by the second-order Magnus propagator[36], integrated with a predictor/corrector algorithm[37], and a time step of 1 as was used. The exponential of the Kohn-Sham matrix was evaluated using a Taylor expansion[38] with 55 terms.

One advantage of a local basis set is that it allows for direct investigation of the continuum state without any transformation. This is especially beneficial when examining features of the absorption spectrum at high energies, as these features directly reflect the quality of the continuum state. In the case of N_2, we calculated the absorption spectrum in the range of 0-100 eV for each basis set using RT-TD-ADFT. To achieve this, we applied three 'kick' pulses during the RT-TD-ADFT simulation over 40 fs with a time step of 1 as in the xyz-direction, aligning N_2 on the z-axis with a bond distance of 1.090 Å. As discussed in Chapter One, the absorption spectrum can be derived from the Fourier transform (FT) of the time-dependent dipole moment. We set the damping rate to 2.42 fs in the FT to enhance the clarity of the figures when comparing between basis sets. This results in broadening the absorption peak compared to the standard, which is typically set at 4.8 fs. To accurately capture the spectrum of the continuum state, we employed a complex absorption potential (CAP) in energy space to eliminate non-bonded electrons (NBEs) in

the simulation (see Chapter One). This step aimed to prevent the artificial auto-ionization process, which often causes a sharp peak in the high-energy spectrum. For this purpose, we utilized a CAP with parameter; 1 Ha^{-1} energy scale for γ_0 and a 0.05 Ha^{-1} damping strength for ξ. Additionally, a vacuum energy cut-off was approximated at 0.01 Ha for all basis sets.

To explore the impact of the basis set on the specific XUV-pulse used in this research, we tested several basis sets across various peptides. We introduced a squared cosine-shaped pulse in RT-TD-ADFT under the dipole approximation, as detailed in Chapter One. The maximum electric field and the total duration of the pulse are set at 0.004 (Ha/e.bohr) (5.61 x 10^{11} W/cm^2) and 30 fs, respectively, with the pulse's centroid energy at 30 eV. We only applied the pulse in the x-direction. The shape of the pulse is depicted in Figure 5-1.

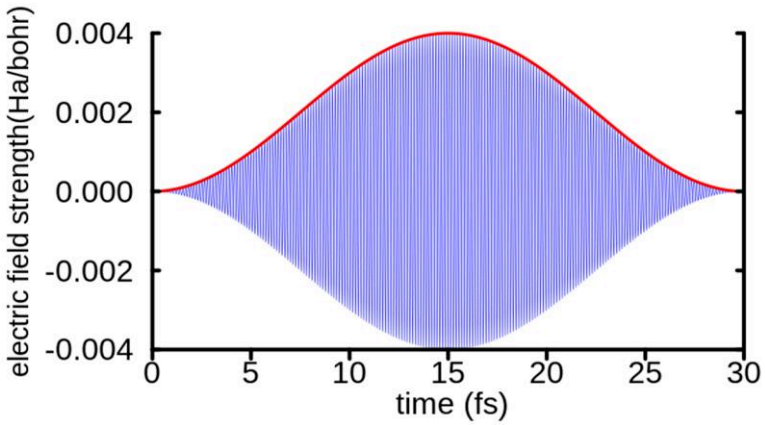

Figure 5-1: Shape of square-cosine pulse with a maximum electric field of 0.004 Ha/e.bohr (5.62 x 10^{11} W/cm^2), a frequency of 1.10245 Ha (30 eV), and a pulse duration of 30 fs.

As demonstrated in Chapter One, the NBEs significantly influence the interaction of XUV-pulses, affecting even the analysis of ionization or charge in the simulation. Therefore, we employed a CAP in energy space to remove electrons in the continuum state. The parameters of CAP were optimized with a 0.1 Ha^{-1} energy scale for γ_0 and a 0.05 Ha^{-1} damping strength for ξ. The vacuum energy cut-off was approximated using the electron affinity for each basis set.

5.1.2 Nitrogen molecule

In order to study the description of the continuum within the XUV energy range, we decided to analyze the absorption cross section of the continuum for N_2. The advantage of this analysis is that the spectrum can be more easily obtained through just three simulations, and it directly provides information about the quality of the continuum across a wide energy range. We don't

need for extensive simulations at each energy point. The downside is that it is not directly comparable with experimental data, such as the photoionization cross-section, since the spectrum contains not only ionization transitions but also excitation transitions. Based on our calculations for HOMO and LUMO energies, the highest excitation is possible up to 28 eV for deep valence electrons. This leads to the appearance of a sharp peak in the spectrum below 28 eV, which is a characteristic feature of excitation transitions. Yet, in our energy range of interest (0-100 eV), excitations of core electrons don't occur, as they are only possible above 300 eV.

Figure 5-2 (gray line) illustrates the experimental measurement[39,40] of the photoionization cross section for N_2, sourced from the Leiden Observatory website[†] at the University of Leiden. To facilitate comparison with our calculations, the original data in cm^{-2} was converted to $bohr^{-2}$. The data indicates sharp peaks between 18-20 eV, corresponding to the auto-ionization of the Rydberg series, followed by a smooth, exponential-like decrease towards 100 eV. Particularly noteworthy is the substantial drop in cross-section at 100 eV compared to, for instance, 25 eV.

To investigate the behavior observed in the photoionization cross-section, we computed the absorption cross-section by calculating the absorption spectrum of isolated N_2 using RT-TD-DFT. Subsequently, we normalized this spectrum by the density number of N_2 gas at 100 K, which stands at 0.0383 $atom/bohr^3$. While this normalization is not absolute, it provides a suitable scale for comparing the spectrum with experimental data. The result of absorption cross section of N_2 calculated with cc-pVDZ[41] is illustrated in Figure 5-2.

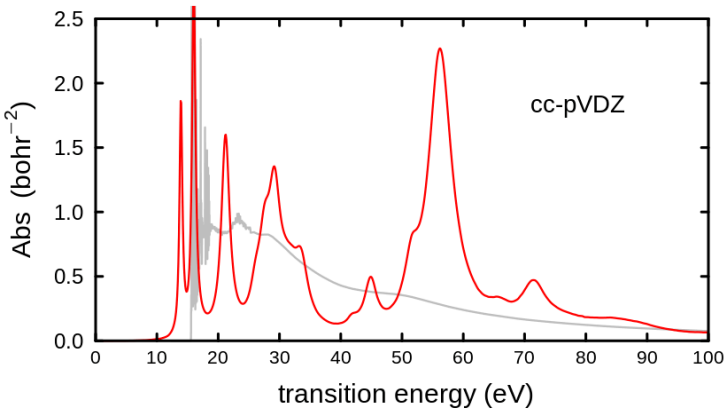

Figure 5-2: Absorption cross section of nitrogen with the cc-pVDZ basis set in red line and experimental ionization cross section of nitrogen in gray line.

[†] https://home.strw.leidenuniv.nl/~ewine/photo/data/photo_data/cross_sections/N2/N2.txt

The figure shows that the spectrum is not continuous; instead, discrete absorption peaks are distributed across the energy range, notably with a sharp, strong peak at 57 eV. This result indicates that photoionization of small molecules cannot be accurately described using a standard basis set optimized to describe bound states. In this subsection, we evaluate the improvement in the continuum state by adding fitted AOC to cc-pVDZ and incorporating ghost atoms. The latter indicates an enhancement in the continuum by expanding the size of the molecular system.

a. Addition of Gaussian type orbital to continuum

To investigate the capability of AOCs in describing the continuum, various sets of AOCs were prepared and added to cc-pVDZ basis set. The following AOC parameters used were: E values of 3 and 6 Ry, maximum L values of 3 and 6, and R values of 3, 6, and 8 bohr. First, we show the number of molecular orbitals (MOs) as a function of their energies in Figure 5-3. We used the ground state MO obtained by a self-Consistent-Field procedure. As evident from the figure, the electronic spectra are very similar below zero, indicating that the AOC only affect, as hoped, the unbounding region. We notice significant differences in the representation of the continuum state. In its standard form, cc-pVDZ basis set gives only 20 MOs that describe the continuum state below 100 eV. This number increases to 77 MOs in L3_E6_R3 within the same energy range. By expanding the boundary radius parameters in AOC, the number of MOs rises from 77 to 145 and 186 for L3_E6_R6 and L3_E6_R8, respectively. A remarkable enhancement is observed when the number of angular momenta increases from 3 to 6, even though the upper energy is decreased to 3 Ry. This corresponds to 429 MOs in L6_E3_R6 within the same energy range.

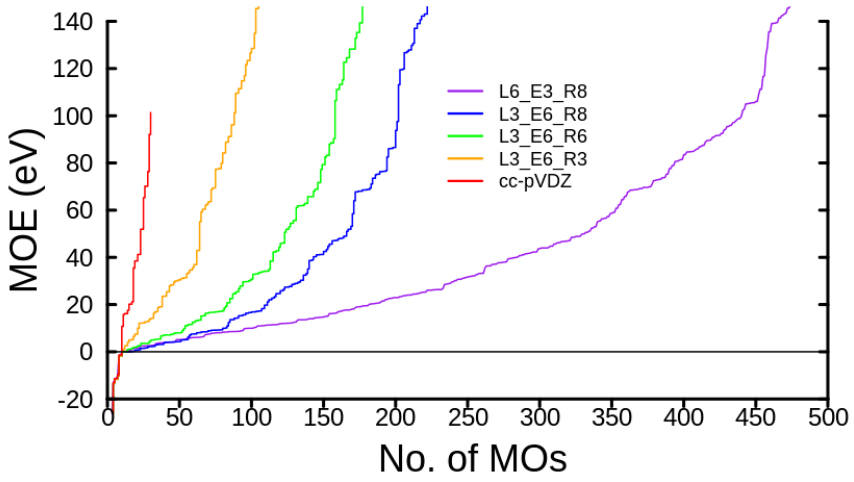

Figure 5-3: Number of molecular orbitals of di-nitrogen with respect to molecular orbital energy for cc-pVDZ and cc-pVDZ basis sets with additional Atomic Orbitals fitting the Continuum (AOC).

A high number of MOs in the continuum state is essential to ensure electronic transitions. However, the oscillation strength of these transitions is also influenced by the spatial contribution of the molecular orbital. This implies that solely relying on the density-of-states does not conclusively indicate that electron transitions to the continuum are described accurately.

To go further, we have calculated the absorption cross section for the continuum using our basis sets, (Figure 5-4). In general, all AOC basis sets enhance the continuum spectrum. They mitigate the sharp peak observed in cc-pVDZ and broaden the peak, which is a hallmark of the continuum. By increasing the density of states, these sharp peaks diminish, and the typical form of the continuum is reproduced. In L3_E6_R3, the discretization of the continuum and the sharp peaks remain significant. However, they decrease with increasing R, and substantial improvement is observed for L6_E3_R8, as anticipated based on the number of MOs in the continuum state.

In comparison with the experimental measurements of the photoionization cross section for N_2, there are some discrepancies, particularly regarding the presence of peaks. However, the exponential-like decline of intensity with photon energy is remarkably reproduced, especially for L6_E3_R8.

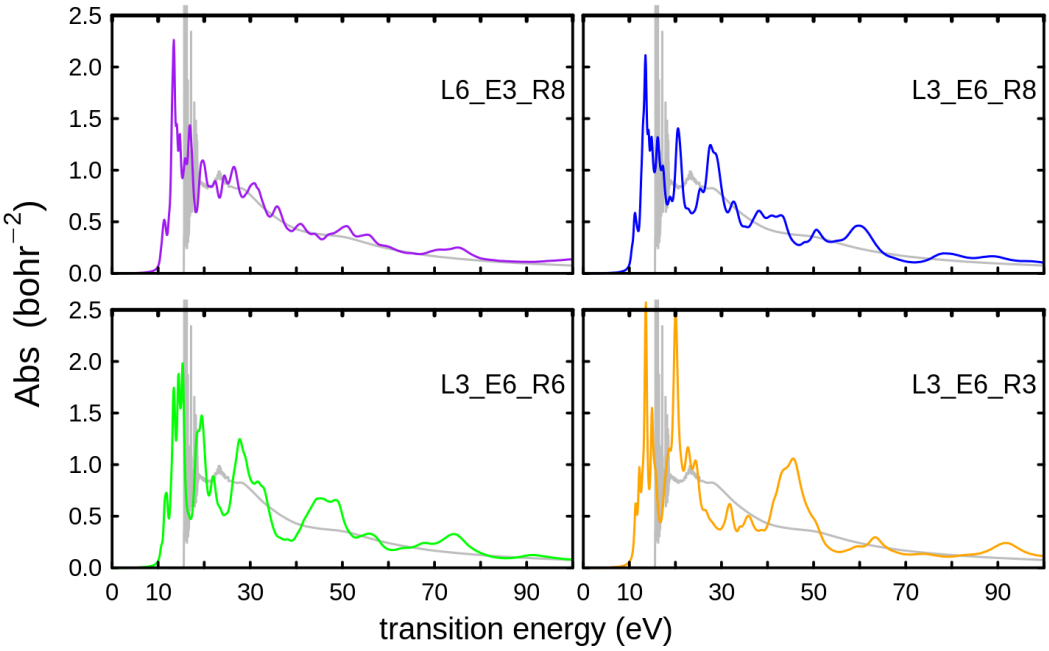

Figure 5-4: Absorption cross section of dinitrogen using the cc-pVDZ basis set with different AOCs, alongside the experimental ionization cross-section depicted by the gray solid line.

This result indicates that the addition of AOC significantly improves the description of the continuum state. By adjusting the parameters in AOC fitting, it is possible to accurately describe the continuum, particularly for small molecules. However, for larger systems, using high parameters in AOC becomes challenging due to a substantial increase in computational costs.

b. Effect of ghost atoms

Increasing the size of a molecular system is expected to enhance the quality of the continuum state through the overlap and shared MOs between atoms. One strategy to clearly evidence this effect is the use of ghost atoms. A ghost atom lacks electrons and only possesses basis functions centered on it, much like a real atom. Any enhancement attributed to a ghost atom is reflective of the enhancement in the corresponding host atom within a molecule. To illustrate this, we prepared an N_2 configuration surrounded by six ghost nitrogen atoms. The N_2 molecule is positioned at the center, 2 Å away from each ghost atom, as depicted in Figure 5-5.

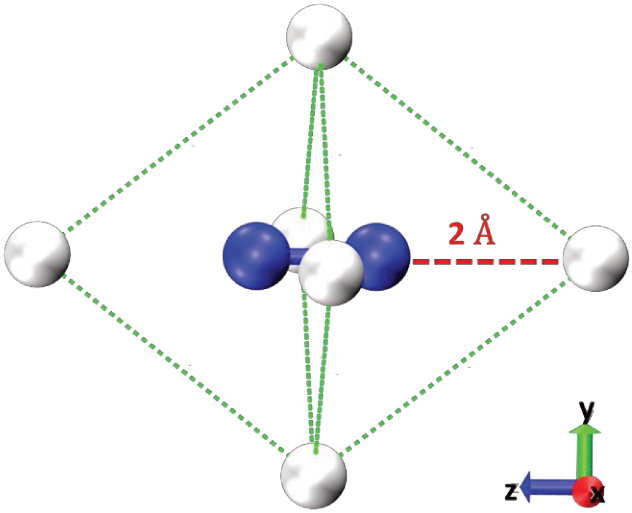

Figure 5-5: Di-nitrogen atoms in blue, accompanied by six ghost nitrogen atoms located in the xyz-directions from the dinitrogen, at a distance of 2Å.

In Figure 5-6, we report the number of MOs as a function of the molecular orbital energy (MOE) for both N_2 and the ghost model. To distinguish ghost model from N_2, the basis set with the ghost atom is generally written as (gs_Lx_Ex_Rx, where) 'x' denotes the maximum parameters in the AOC fitting. Specifically, for this test, we selected those basis sets that inadequately represent the continuum state in the N_2 examination.

The number of molecular orbitals in the continuum significantly increases with the addition of a ghost atom. For instance, it increases from 20 to 78 for cc-pVDZ, from 77 to 264 for L3_E6_L3, and from 145 to 539 for L3_E6_L6. In the case of gs_L3_E6_L6, the density of states is even higher than that of L6_E3_R8. Based on the density-of-states, an improvement in the continuum is expected with the use of the ghost model.

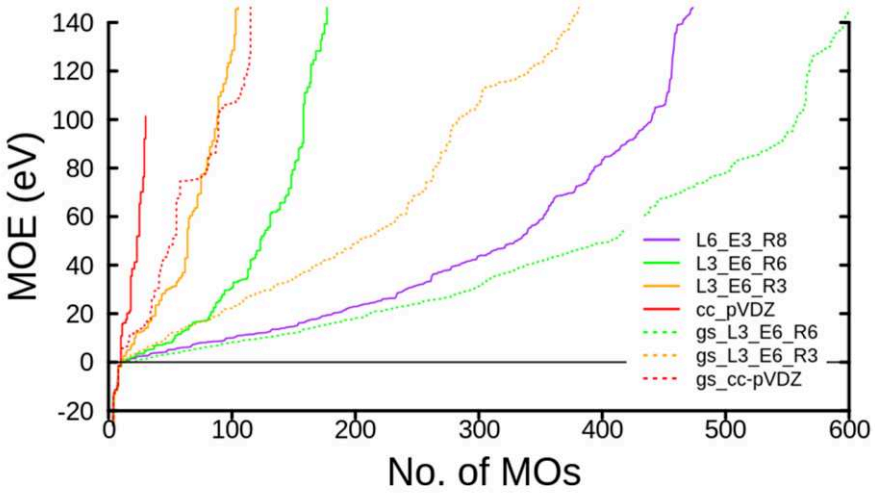

Figure 5-6: Representation of the number of molecular orbitals with respect to molecular orbital energy for di-nitrogen, shown in solid lines, and di-nitrogen with six ghost nitrogen atoms, shown in dotted lines.

We report on Figure 5-7 the absorption cross section. For reference, the number of MOs in the continuum for ECP is 8 for N_2, and 43 for the ghost model in the range below 100 eV.

As observed, the inclusion of ghost atoms notably enhances the spectrum of L3_E6_L6, and also in the case of L3_E6_L3 (as indicated by dashed lines). This enhancement is spread over the entire energy spectrum. The discreteness of the continuum and the sharp peaks are considerably reduced in the spectra, and the shape with exponentially-like decreasing intensity is clearly visible. On the other hand, a significant enhancement is not observed for cc-pVDZ, especially above 50 eV where strong peaks persist. Improvements are generally seen below 50 eV, except at 30 eV where a gap in the spectrum appears. In the case of ECP, despite the remaining discretization of the continuum, a notable enhancement is observed in the range between 20-40 eV and 80-90 eV. The strong peaks broaden, and the gaps in the spectrum are significantly reduced.

Regardless of the inferior density of states in ECP compared to cc-pVDZ, the description of the continuum is better in the range below 60 eV, except around 20 eV. This discrepancy might be related to the spatial contribution of MOs. An increase in the spatial contribution of MOs is anticipated to enhance the likelihood of overlap and sharing of MOs between neighboring atoms. The maximum spatial radius of ECP for N in N_2 is 6.79 Å, whereas it is 6.05 Å for cc-pVDZ. For L3_E6_R3 and L3_E6_R6, the radii are 14.89 Å and 71.44 Å, respectively.

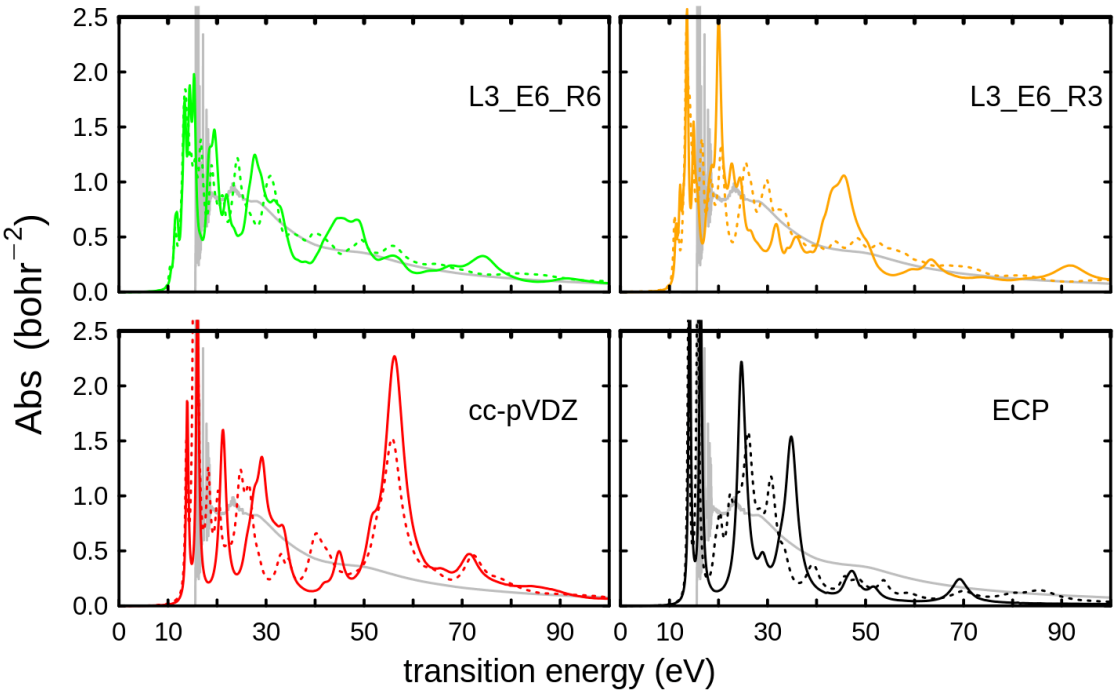

Figure 5-7: Absorption cross-section of dinitrogen shown in solid lines, and dinitrogen with six ghost nitrogen atoms shown in dotted lines, using cc-pVDZ, cc-pVDZ with additional AOCs, and ECP (Relativistic Effective Core Potentials | Stuttgart-Dresden (RECP | SD)). The gray solid line refers to the experimental ionization cross-section.

To clearly see the effect of the addition of ghost atoms, we plotted the absorption cross section for gs_L3_E6_R3 and gs_L3_E6_R6 alongside the best basis set for N2 (L6_E3_R8) in Figure 5-8.

When comparing gs_L3_E6_R3 with L6_E3_R8, the absorption cross-section of gs_L3_E6_R3 shows a high level of agreement with L6_E3_R8, except around 23 eV, where some divergence is observed. Notably, gs_L3_E6_R3 performs even better above 70 eV. For gs_L3_E6_R6, the improvement in the spectrum is even more pronounced; however, it still lags behind in the 25-35 eV range when compared to L6_E3_R8.

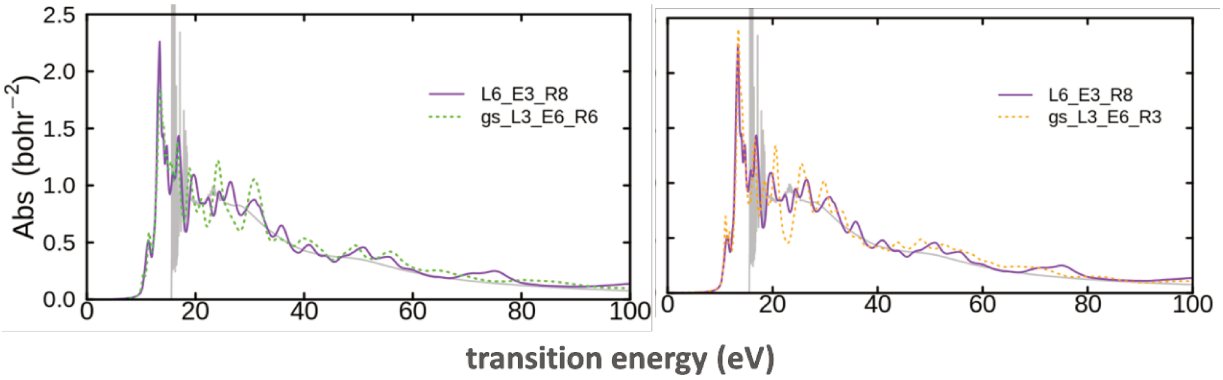

Figure 5-8: Absorption cross section of dinitrogen using cc-pVDZ_L6_E3_R8 as a reference, shown in solid purple line, and dinitrogen with six ghost nitrogen atoms using cc-pVDZ_L3_E6_L6 and cc-pVDZ_L3_E6_L3, shown in dotted orange and green lines, respectively. The experimental ionization cross-section of dinitrogen is represented by the gray line.

This result shows that the presence of ghost atoms, or in other words, the presence of basis functions on surrounding atoms, significantly improves the quality of continuum states. This directly reflects the increase in the size of the molecular system. Even with a poor basis set in certain energy regions, the enhancement is notable. This suggests that for large biomolecules, simulating ionization processes in specific energy regions should be feasible using a standard basis set. This is a significant assumption, and we verify it in the following sub-section.

5.1.3 Peptides

The use of AOC for large biomolecules such as insulin and substance P, represent a great challenge even if low AOC parameters are used. In this sub-section, we investigate the possibility of using standard basis set or ECP for describing interaction of XUV-pulse with the same pulse parameters for insulin^{+6} or substance PH^{+}.

For this purpose, we selected peptides with varying residues and charge states to account for the effects of atom types, residues, and protonation states. The first peptide, P1, is the Met-enkephalin, which contains five residues. The other peptides, P2, P3, P4, and P5, are extracted from insulin^{+6} sequence and the two ends of the peptides are capped with acetyl group (Act) and a protonated N-terminus. The profiles of these peptides are presented in Table 5-1 and their structures are shown in Figure 5-9.

Table 5-1: Peptide Profiles. P1 represents Met-enkephalin, and the last four peptides are extracted from insulin[+6]. 'Act' denotes the acetyl group. Peptides P2 to P5 hold an N-acetylation termination (Act). Peptide P2 consists of four amino acid-residue with a disulfide bridge with an acetylated cysteine.

Peptide	Atom number	Electron number	Charge	Atoms	Sequence
P1	75	304	0	S,O,N,C, and H	Tyr-Gly-Gly-Phe-Met
P2	83	352	+1	S,O,N,C, and H	Cys-Cys-Gly-Glu-Arg
P3	63	242	+1	O,N,C, and H	Gly-Ser-His-Leu
P4	65	248	0	O,N,C, and H	Leu-Tyr-Gln
P5	72	262	+1	O,N,C, and H	Thr-Pro-Lys-Gln

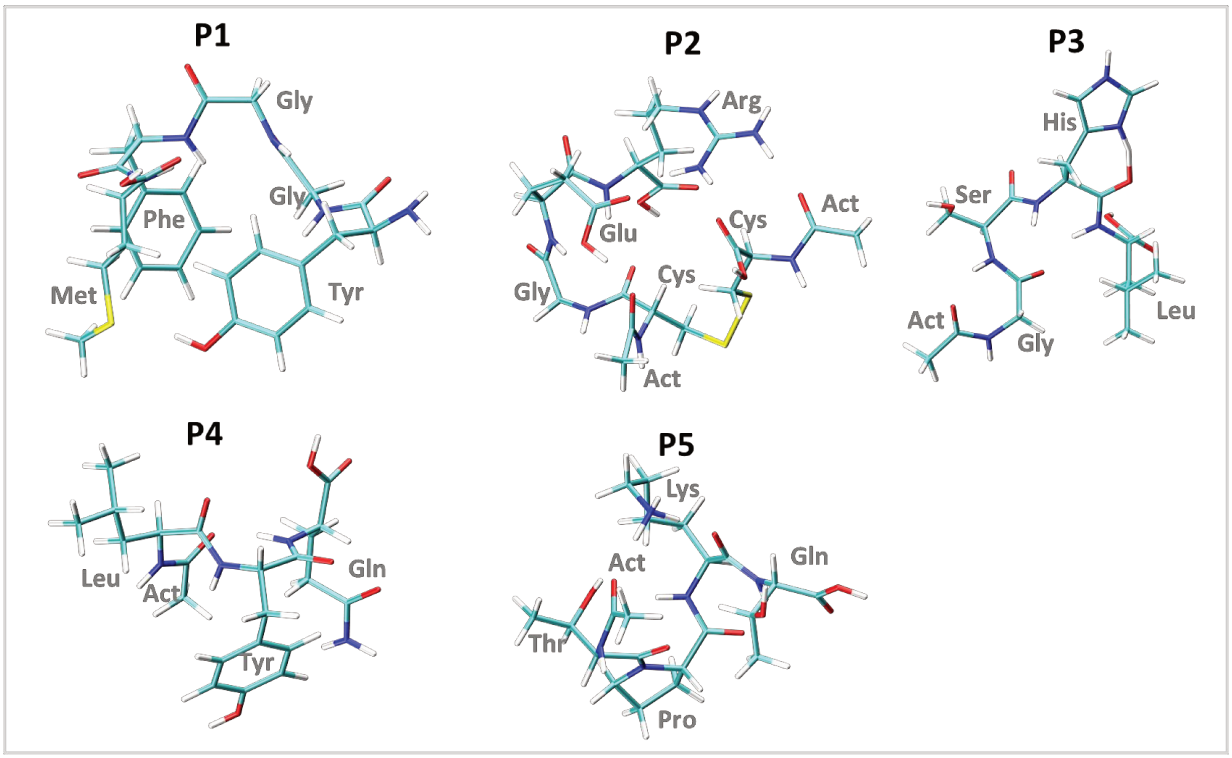

Figure 5-9: Licorice representation of P1, P2, P3, P4, and P5 structures.

For the basis sets, we selected TZVP, DZVP-GGA, and ECP. The basis set chosen as a reference is TZVP augmented with AOC, referred to as ref-AOC and introduced in the basis set sub-section. In the case of the ghost model for N_2, AOC shows good performance in describing the continuum, even with low fitting parameters. That's why we use it as a reference.

In this sub-section, we analyze the variation in energy and the number of holes as functions of time. Additionally, we investigate the distribution of holes across residues.

a. Energy variation and ionization state

One way to assess the quality of a basis set is by observing the evolution of energy and electrons in the simulation over time. To illustrate the effect of the basis set on these variations, we initially focus on peptide P1. The number of MOs are given in Table 5-2. The number of MOs with ref-AOC is 2.3, 3, and 4.8 times larger than that of TZVP, DZVP-GGA, and ECP, respectively.

Table 5-2: Number of molecular orbitals in P1 based on the basis set.

Basis set	No. of MOs
ref-AOC	2368
TZVP	1044
DZVP-GGA	776
ECP	495

In Figure 5-10, for each basis set, we show the energy variation with respect to the ground state (ΔE) (on the left) and the number of non-bonded electrons absorbed by CAP (right panel), which corresponds to the number of holes, during a 50 fs period, where the total duration of the pulse is 30 fs and it is centered at 15 fs.

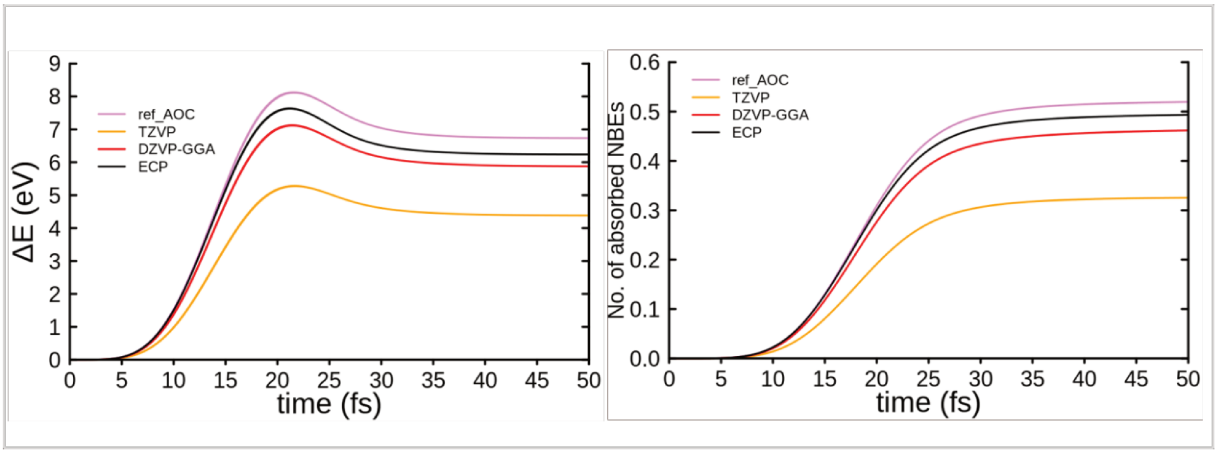

Figure 5-10: Total energy variation (left) and number of non-bonded electrons absorbed by CAP (right) as a function of time for P1.

In the figure of energy variation, ΔE rapidly increases from 5 fs to 20 fs. It then decreases in a dome-like shape until 30 fs (the end of the pulse) and afterward, ΔE remains stable until the end of the simulation. This pattern of energy evolution is observed across all basis sets. However, this dome-like shape doesn't appear in ionization profile. The number of absorbed NBE rises from 5 to 30 fs at a rate corresponding to the electric field strength of the pulse. After 30 fs, the number of holes increases slightly due to a few electrons remaining in non-bonded states; the CAP needs time to absorb them.

The dome-shaped variation in energy arises from the energy of the pulse. Unlike in the case of charged particles, here the energy variation corresponds to the total energy, which comprises both the Kohn-Sham energy and the energy of the pulse. Even at the end of the pulse, we cannot refer to it as 'energy deposition' as in the case of charged particles. This is because the number of electrons is no longer fixed due to CAP. Instead, it corresponds to the partial energy deposition by the pulse as a consequence of the excitation of electrons and the ionic form of the system.

When comparing the basis sets, we observe that the amplitude of ionization mirrors the energy variation. Moreover, the ionization state is strongly influenced by the choice of basis set. At the end of the simulation, 0.52 electrons are absorbed by the CAP using the ref-AOC basis set. A significant difference is noted with TZVP, where the discrepancy from ref-AOC is 0.2. For DZVP-GGA, this gap narrows to 0.05. Interestingly, there is a close agreement between ref-AOC and ECP, with a difference of less than 0.02, despite the fact that each basis set involves a very different number of MOs.

The same simulations have been carried out for the other peptides. Table 5-3 collects the mean error and mean relative error across the five peptides taking ref-AOC as reference.

Table 5-3: Mean error and mean relative error of energy variation and number of holes at 50 fs over P1, P2, P3, P4, and P5 with respect to the ref-AOC basis set.

Basis set	ΔE		No. of NBE	
	ME (eV)	%ME	ME	%ME
TZVP	-2.89	- 40.2	-0.20	-41.8
DZVP-GGA	-1.19	-17.0	-0.07	-14.2
ECP	-0.02	-0.2	-0.01	-1.5

As for peptide P1, ECP shows excellent performance for all peptides; the underestimation is significantly low for both energy variation and the number of absorbed electrons. On the other hand, the underestimation with DZVP-GGA and TZVP is significant, especially for TZVP where it exceeds 40%. These results demonstrate that the energy variation and total number of holes substantially depend on the basis set for the specific pulse we tested. Nevertheless, ECP quantitatively reproduces the energy variation and the total ionization process when compared to a rich basis set like ref-AOC.

b. Charge variation on residues

An important feature that could be impacted by an incorrect description of continuum states is the nature of ionization. Artificial electron transitions, or avoid transitions due to a lack of orbitals in the continuum state, can significantly affect the nature of ionization at the atomic and fragment levels in molecules. This is crucial for our work, as we aim to observe the ionization processes caused by the XUV-pulse at the atomic or fragment level.

To define the charge on the atoms or residues in the molecule, we may follow a Becke[42] or Hirshfeld[43] partition scheme[44], both finally rendering qualitatively similar results. The number of holes (h) on each residue or atom (i) in the molecule is calculated by the difference between the charge at 50 fs and that of the ground state. The charge of an atom or residue is denoted as q.

$$h_i = q_i^{50} - q_i^0 \tag{5.1}$$

The normalized number of holes for a residue or atom in a molecule is calculated as:

$$Nh_i = \frac{h_i}{\sum_i h_i} \tag{5.2}$$

To exhibit the effect of the basis set on hole spread over residues, we analyzed the hole for each residue in P1, as presented in Table 5-4.

Table 5-4: Number of holes and percentage of normalized number of holes for each residue in P1 and total number of hole for P1 as a function of basis set.

Basis set	Thr		Gly		Gly		Phe		Met		total
	h	$\%Nh$	h	$\%Nh$	h	$\%Nh$	h	$\%Nh$	h	$\%Nh$	h
ref-AOC	0.148	28.7	0.047	9.0	0.054	10.4	0.140	27.1	0.128	24.8	0.52
TZVP	0.096	29.7	0.027	8.3	0.034	10.5	0.085	26.1	0.082	25.4	0.32
DZVP-GGA	0.134	29.7	0.043	9.4	0.050	11.0	0.116	25.2	0.114	24.8	0.46
ECP	0.150	30.4	0.043	8.7	0.053	10.9	0.121	24.7	0.124	25.3	0.49

As observed in the table, the distribution of holes across residues strongly depends on the total number of holes. Compared to ref-AOC, the largest discrepancy is noted for TZVP, followed by DZVP-GGA. However, as previously observed for energy variation and the total ionization state, ECP demonstrates exemplary performance at the residue level for ionization state. Quantitatively, the number of holes on residues calculated with ECP is very close to ref-AOC; the most notable difference corresponds to Phe, which is 0.019.

What is particularly interesting in this table is that the normalized number of holes for all basis sets is notably close to ref-AOC. This indicates that, qualitatively, the nature of ionization isn't largely influenced by the basis sets. This feature is crucial because the total ionization can be modulated by the strength of the electric field.

The results here are based on the simulation of the XUV-pulse in one direction (x-direction). In Table 5-5, we computed $\%Nh$ for ECP and DZVP-GGA in response to xyz-direction pulses.

Table 5-5: Percentage of normalized number of holes calculated from the xyz-direction of the pulse and mean percentage of normalized number of holes over all directions.

Basis set	DZVP-GGA				ECP			
	$\%Nh$ (x)	$\%Nh$ (y)	$\%Nh$ (z)	<% Nh>	$\%Nh$ (x)	$\%Nh$ (y)	$\%Nh$ (z)	<% Nh>
Thr	29.2	29.0	27.2	28.5	31.3	31.8	28.7	30.7
Gly	9.9	9.5	8.9	9.4	8.6	8.7	8.8	8.7
Gly	10.9	12.0	10.9	11.3	10.3	10.2	10.5	10.3
Phe	25.0	25.0	25.8	25.3	24.5	25.4	26.2	25.3
Met	25.0	24.5	27.1	25.5	25.3	23.9	25.7	25.0

As shown in the table, some residues exhibit different responses to the orientation of the pulse, but the differences remain small For DZVP-GGA, a difference in $\%Nh$ for pulse directions is observed across all residues. However, in the case of ECP, this difference is more pronounced for polar residues like Thr, Phe, and Met. When comparing the <$\%Nh$> with the x-direction for both DZVP-GGA and ECP, the difference between the two basis sets diminishes for all residues.

All the previous results are at the residue level. The good performance of the basis set might be due to the dilution of error because of the size of the residues. To delve deeper, we visualized P1 by displaying the value of $\%Nh$ for each atom, as calculated with ref-AOC and ECP, in Figure 5-11. In the figure, the value of $\%Nh$ for atoms is represented using a color scale.

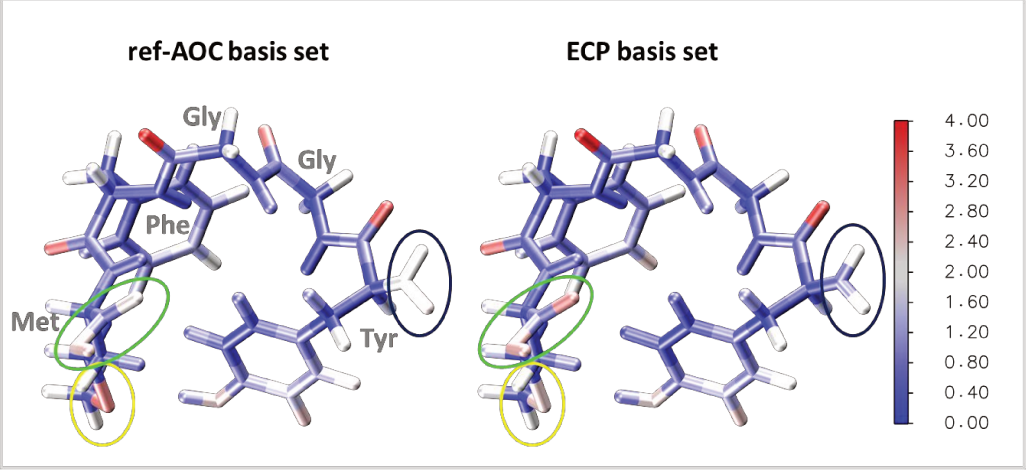

Figure 5-11: VMD atomic representation of P1, with the color of atoms scaled according to the percentage of normalized number of holes, calculated with the ref-AOC and ECP basis sets. The colored circles indicate the zones where ref-AOC and ECP differ the most.

The figure indicates that the $\%Nh$ at the atomic level, as calculated from ref-AOC and ECP, is quite similar. Small differences are observed for O in the C-terminus, N in the N-terminus, and S in Met (circled in figure 5-11). As seen in the pulse orientation test, polar residues are more sensitive to pulse orientation. Thus, these differences are expected to diminish when simulations are completed for all three directions. This result confirms that the nature of ionization is independent on the basis set, even at the atomic level. This performance at the atomic level is observed not only for ECP but also for TZVP and DZVP-GGA.

We have summarized the effects of the basis sets on the number of holes and the normalized number of holes at the residue level for all peptides by calculating the relative error and standard deviation of relative error, as presented in Table 5-6.

Table 5-6: mean relative error and mean standard deviation of relative error of number of holes and normalized number of holes over the residues in P1, P2, P3, P4, and P5 with respect to the ref-AOC basis set.

Basis set	h		Nh	
	%ME	**MSE**	**%ME**	**MSE**
TZVP	-40.4	1.4	-1.4	1.5
DZVP-GGA	-13.2	2.0	-0.7	1.7
ECP	-2.7	1.3	-0.5	1.3

As seen in the case of P1, for all peptides, the error in the number of holes is significant for TZVP and DZVP-GGA, corresponding to underestimations of -40.4% and -13.2%, respectively. However, this error appears to be systematic because the standard deviation of error is small, indicating the data scattering is minor. On the other hand, as seen in the case of P1, the error in ECP is notably smaller compared to the standard basis set, which demonstrates the performance of ECP.

For Nh, the results confirm the ability of all basis sets to describe the nature of ionization. The error and standard deviation of error in Nh are remarkably low. The error in all basis sets is significantly below the confidence zone (± 5%), which proves the robustness of basis sets in correctly describing the nature of ionization. Notably, for ECP, the error is significantly small.

Overall, these results indicate that the simulation of XUV-pulses is feasible for large molecular systems using standard basis sets, or even ECP for those interested in the nature of ionization

within molecules. We demonstrate that, qualitatively, the nature of ionization is independent of the choice of the basis set. Particularly, despite the poor number of MOs in ECP, it performs exceptionally well, both qualitatively and quantitatively, in terms of the nature of ionization and energy variation. The correct description of the nature of ionization is not only at the residue level but at the atomic level as well.

5.1.4 Partial conclusion

In this section, we have explored the simulation of molecule irradiation by XUV-pulses using local basis sets. Initially, we tested the addition of atomic orbital fitting to the continuum (AOC) in a standard basis set to describe the continuum state in N_2. Then, we extended this test by adding ghost atoms around the N_2 molecule. This addition reflects to the participation of basis functions from surrounding atoms in larger molecular systems, enhancing the continuum state. Subsequently, we investigated the use of standard basis sets and ECP for the ionization processes in relatively large peptides when exposed to XUV-pulse.

We concluded that the addition of AOC notably enhances the description of the continuum state. By augmenting the parameters in AOC fitting, one can achieve an accurate depiction of the continuum, especially for smaller molecules. Furthermore, we demonstrated that the quality of continuum states considerably benefits from the inclusion of ghost atoms. Remarkably, the improvement is more pronounced in high energy region for AOC with lower fitting parameters in the ghost model than for those with higher parameters in N_2 alone.

In the testing of peptides, what is particularly interesting is that, despite the significant difference in the number of MOs between basis sets like TZVP, DZVP-GGA and ECP compared to ref-AOC, the results show that the nature of ionization at both residue and atomic levels, induced by the XUV-pulse, is independent on the basis set. However, there is a basis set-dependency when considering energy variations and ionization states in peptides. ECP demonstrates excellent performance in both qualitative and quantitative descriptions of the interaction with the XUV-pulse compared to ref-AOC. This could be attributed to the inclusion of ionization and excitation information in the ECP fitting process, or it may more accurately capture the core-valence correlation changes between neutral and partially ionized atoms[45].

These findings are valuable for the simulation of XUV-pulse interactions using local basis sets. For small molecular systems, our findings advocate the addition of AOC. In larger molecular systems, where adding AOC becomes challenging due to computational costs, our results indicate that standard basis sets or ECP are sufficient to gain qualitative information about the process of molecular ionization.

5.2 Structure investigation of insulin^{+6} and substance PH^{+}

In the XUV-pump/IR-probe experiment developed in Lépine's group at ILM [19], the combination of electrospray ionization and mass spectrometry techniques[46] is employed to evaporate samples in the liquid phase and select irradiated molecules based on their mass-to-charge ratio. Using electrospray ionization, the sample in solution is evaporated and ionized. The solvent gently evaporates around room temperature (approximately 300 K) through the continuous evaporation process. As this occurs, the droplets decrease in size until the surface charge can no longer be stabilized by the liquid flow. This leads to a Coulombic explosion where the droplet disintegrates, releasing analyte molecules as gas-phase ions, such as a protein charged by protonation. Following this, the desired ion is selected for study using a mass spectrometer. This technique ensures that the molecule retains its structure without fragmentation. However, the structural conformation can deviate substantially from the crystalline structure of the molecule. Investigating structural conformations at the atomic level is crucial for both experimental and theoretical approaches. This is especially pertinent when seeking to understand certain ultra-fast processes or protonation states in experiments or when selecting the initial structure in simulations.

In the XUV-pump/IR-probe experiment, Insulin^{+6} and Substance PH^{+} are investigated in the gas phase due to their abundant ionization states and significant mass-to-charge (m/z) ratios.

In this section, I delve into the structural investigation of Insulin^{+6}, through long classical molecular dynamics (MD) simulations, and detail the selection process of its initial structure for XUV simulations. In the subsequent part, the focus will shift to Substance PH^{+}. Since I did not personally

conduct the MD simulations of substance P, my focus will be limited to describing the structures chosen as starting points for the XUV simulations.

5.2.1 Insulin^{+6}

The functionality of proteins is intricately linked to their complex structures, beginning with their primary structure comprising the specific sequence of amino acids connected by peptide bonds. The secondary structure, established through hydrogen bonds between amino acid backbones, gives rise to formations like alpha helices and beta sheets. Tertiary structure emerges from interactions among amino acid side chains, shaping the three-dimensional form of the protein. Additionally, in certain proteins, the quaternary structure is the result of assembling multiple protein units.

Insulin is composed of 51 amino acids across two chains[14]. Chain A includes residues 1-21, while chain B encompasses residues 22-51. The tertiary structure of insulin features three disulfide bonds, with two interchain bridges formed between A7-B7 and A20-B19, and one intrachain bridge connecting residues A6 and A11 as shown in Figure 5-12.

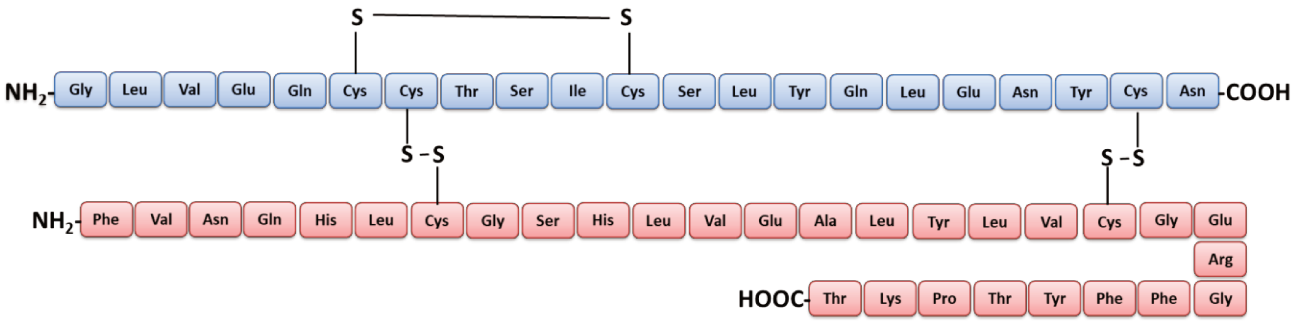

Figure 5-12: Sequence of amino acids in insulin. The chains A and B are represented in blue and red color, respectively. The amino acid codes are for: alanine (Ala), arginine (Arg), asparagine (Asn), cysteine (Cys), glutamic acid (Glu), glutamine (Gln), glycine (Gly), histidine (His), isoleucine (Ile), lysine (Lys), leucine (Leu), phenylalanine (Phe), proline (Pro), serine (Ser), threonine (Thr), tyrosine (Tyr), and valine (Val).

Before conducting the RT-TD-DFT simulation, our goal is to identify a reference structure through MD simulations that will hopefully approximate the experimental structure. To achieve this, we start by solvating insulin crystalline structure in water at 300K, and after adding 6 protons to

insulin, we perform a long MD simulation in vacuum at 300K, in accordance with experimental conditions.

c. Solvated and thermalization of insulin in water

For this study, we chose the crystal structure of human insulin (PDB code: 2hiu.pdb[14]) as obtained by nuclear magnetic resonance spectroscopy (NMR). The insulin crystal structure was solvated in TIP3P water[47] within a 68x68x68 Å periodic box, and the charge of the system was neutralized using sodium counter-ion. The Amber force field FF14SB[48] was used for insulin. Initially, the insulin in water was optimized for 10,000 steps while applying harmonic restraints on the heavy atoms of insulin through a classical molecular dynamics (MD) simulation using the NAMD 2.11 package[49]. Subsequently, the system was heated from 50 K to 300 K over 20 ps with 10 kcal/mol harmonic restraints applied to the heavy atoms of the solute at a constant volume. The system's equilibration then continued in an NPT ensemble at 300 K, gradually reducing the restraint. Finally, an unrestrained MD simulation was performed for 100 ns with a timestep of 1 fs. The Root-Mean-Square Deviation (RMSD) of atomic positions for solvated insulin is illustrated in Figure *5-13*. After equilibration, the final step structure was used as starting point for MD simulations in the gas phase.

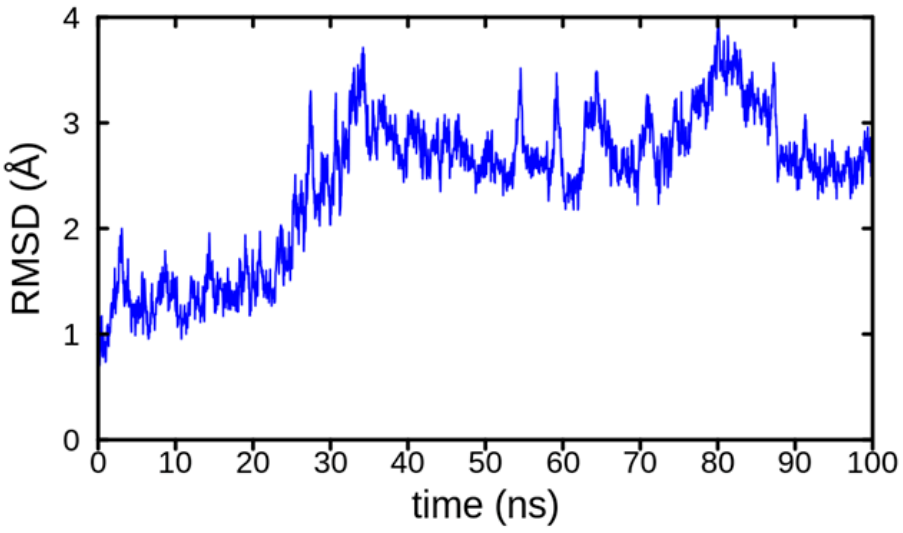

Figure 5-13: RMSD calculation of atomic positions of insulin in water for the NVT ensemble MD simulation.

d. Molecular dynamics simulation of insulin^{+6} in vacuum

I. Protonation state of insulin

Protonating specific amino acids to achieve insulin^{+6} in vacuum is a complex process, particularly as the protonation state of human insulin in gas phase is not well-documented in the literature. Consequently, we decided to protonate those amino acids with the highest probability of proton acceptance. Insulin contains four basic amino acids: Lysine (LYS50), Arginine (ARG43), and two Histidines (HIS26, HIS31). Additionally, it has two N-terminal residues: Glycine (GLY1) and Phenylalanine (PHE22), which are given priority for protonation. Through this process, we achieved insulin with a net charge of +6 in vacuum. The positions of these charged amino acids are represented in Figure *5-14*.

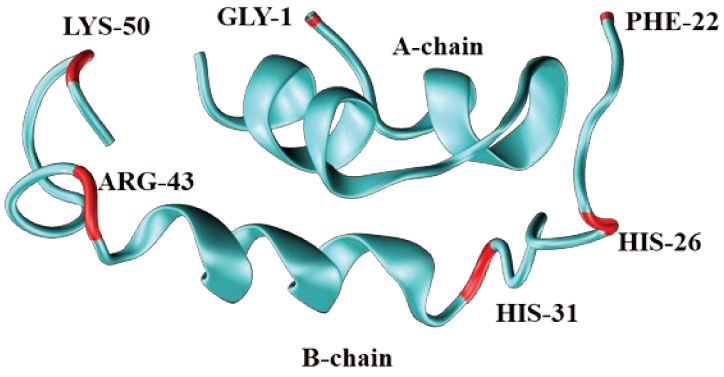

Figure 5-14: insulin structure with the red color representing the location of charged amino acids.

II. MD simulation in vacuum

To derive an insulin structure that closely approximates the experimental one, we conducted a lengthy MD simulation for insulin in the gas phase using the AMBER 18 software package[50]. The FF14SB force field was used for insulin. However, within AMBER, the force field parameters for the C-terminal residues only exist in the zwitterionic form. As a result, we parameterized the force field for the protonated carboxylate group in C-terminal Asparagine (ASN-21) and Threonine (THR-51). Before initiating the MD simulation in vacuum, we minimized the equilibrium structure of the solvated insulin (with a charge of +6) in vacuum through 10,000 steps with restraints on

the heavy atoms, followed by 10,000 steps without restraints. The system was then heated from 100 K to 300 K with a 5 kcal/mol restraint applied for 20 ps, and the process continued for an additional 20 ps without restraints. Subsequently, we launched an NVT ensemble MD simulation for 9.7 µs with a time step of 2 fs. The Langevin thermostat was employed with a collision frequency of 0.5 ps^{-1}, along with an infinite cutoff for the non-bonding terms. The simulation and the system's temperature remained stable throughout the process, as depicted in Figure *5-15*.

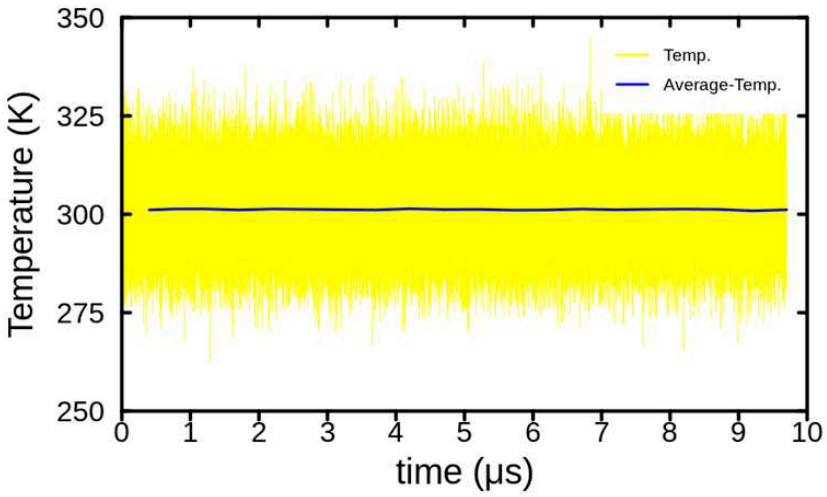

Figure 5-15: Temperature and running average temperature over 0.4 µs during the NVT ensemble MD simulation of insulin with a charge of +6 in vacuum.

The RMSD calculation is illustrated in Figure *5-16*, obtained using the RMSD-based algorithm implemented in cpptraj[51], a utility of Amber Tools. The first frame of the simulation was used as a reference conformation for alignment. The RMSD plot reveals three regions over time. Until about 1.5 µs, $insulin^{+6}$ retains its general form in water, particularly the C and N-termini of the B-chain. Starting at 1.5 µs, significant changes occur, with the C-terminus of the B-chain wrapping around the A-chain, and the N-terminus forming new hydrogen bonds with the B-chain. After 1.5 µs, the new general form is maintained, with only the order and length of the hydrogen bonds between amino acids changing, which do not result in substantial global structural modifications.

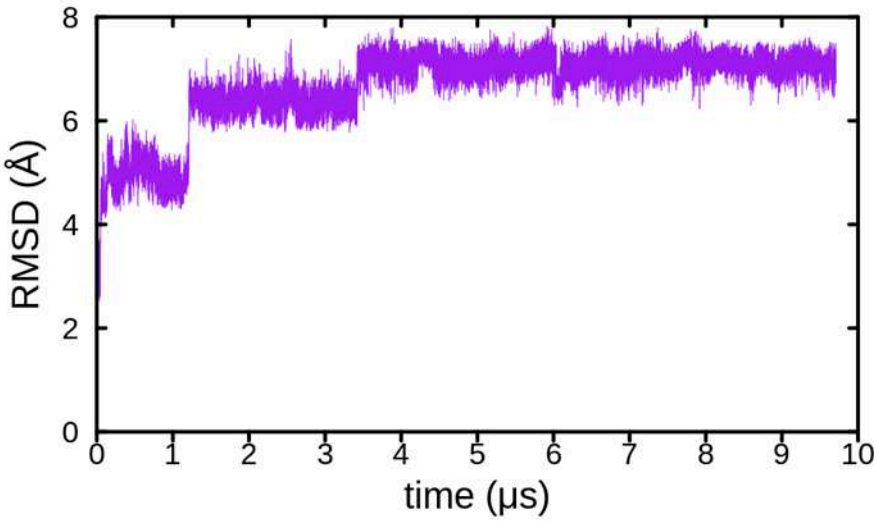

Figure 5-16: RMSD calculation of insulin^{+6} in the NVT ensemble MD simulation in vacuum at 300 K.

III. Clustering

To obtain the most prevalent structures in the MD simulation, we used the DBScan (density-based spatial clustering of applications with noise) clustering algorithm[52], implemented in the cpptraj utility of Amber Tools. When using DBScan, it is crucial to specify the minimum number of points required to form a cluster (minpoints) and the distance cutoff for cluster formation (epsilon). These values can be estimated using the K-distance plot in cpptraj tools. In our system, the minpoints value and cutoff distance were set to 22 Å and 1.35 Å, respectively. The results of the clustering are presented in Table *5-7*.

Table 5-7: Result of the clustering calculation. The Cluster, Frames and Frac are the cluster number, the cluster size and the occurrence fraction respectively. AvgDist and Stdev are the average distance (taken as the RMSD) and standard deviation between two structures in this cluster.

Cluster	Frames	Frac	AvgDist	Stdev
C0	62371	0.642	1.690	0.283
C1	21811	0.225	1.682	0.396
C2	10454	0.108	2.105	0.496
C3	666	0.007	1.392	0.308

Table 5-7 collects information on the four most popular clusters. The most populated one, C0, includes 62,371 frames, representing 64.2% of the total processed frames. Following C0, cluster C1 represents 22.5% and cluster C2 corresponds to 10.8%. To illustrate the differences among these clusters, we aligned the three most populated structures using the Multiseq program in VMD. The results, shown in Figure 5-17, use a color scale based on RMSD values when C0 is used as the reference.

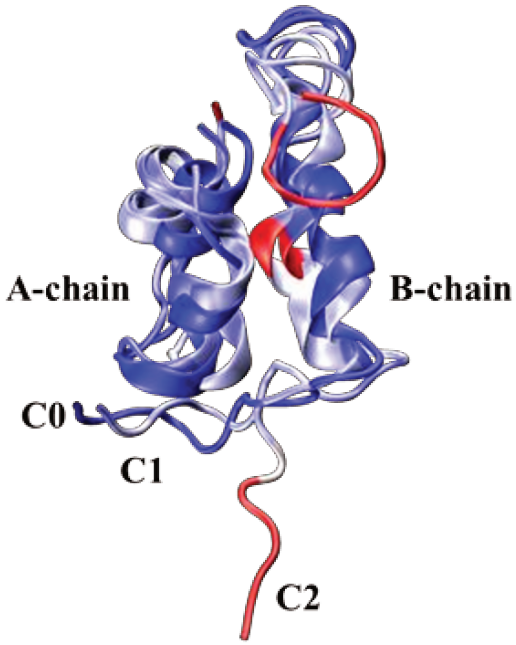

Figure 5-17: the alignment result of the three most populated structures in MD simulation. the color scale with RMSD value.

The alignment demonstrates that the differences among the three clusters are not substantial in the A-chain. Indeed, the A-chain is rigidly bound together by an intrachain sulfide bridge and a hydrogen bond between the C-terminal and N-terminal residues. Most of the differences are thus found in the B-chain. The discrepancy between clusters C0 and C1 is relatively minor, but a significant difference is observed for C2 compared to C0, particularly at the N- and C-termini. This is confirmed by the RMSD calculations per residue shown in Figure 5-18, where calculations were conducted for C1, C2, and insulin in water (the final structure of solvated insulin), using C0 as a reference. As observed, C2 maintains its structure in water, especially at the N- and C-termini of the B-chain. This observation is further supported by the potential energy (PE) curve (Figure 5-19), where the PE steadily decreases until about 2 μs, after which it stabilizes. Therefore, we decided

to select C0 as the reference for RT-TDDFT simulations because C1 (around 1 µs) and C2 (around 2 µs) appear to correspond to the transition structures.

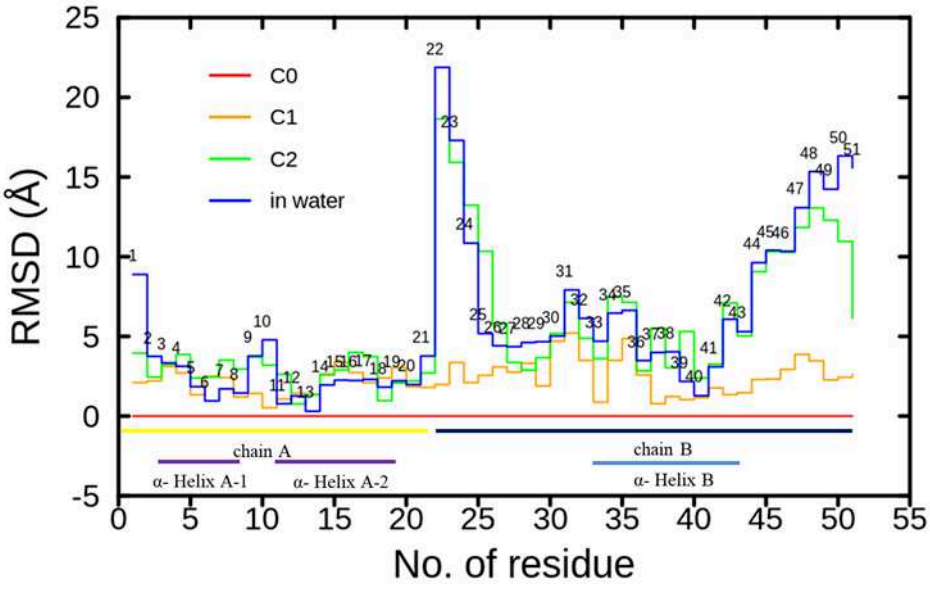

Figure 5-18: RMSD calculations for heavy atoms (excluding hydrogen) per residue for clusters C1, C2, and insulin in water, in reference to C0.

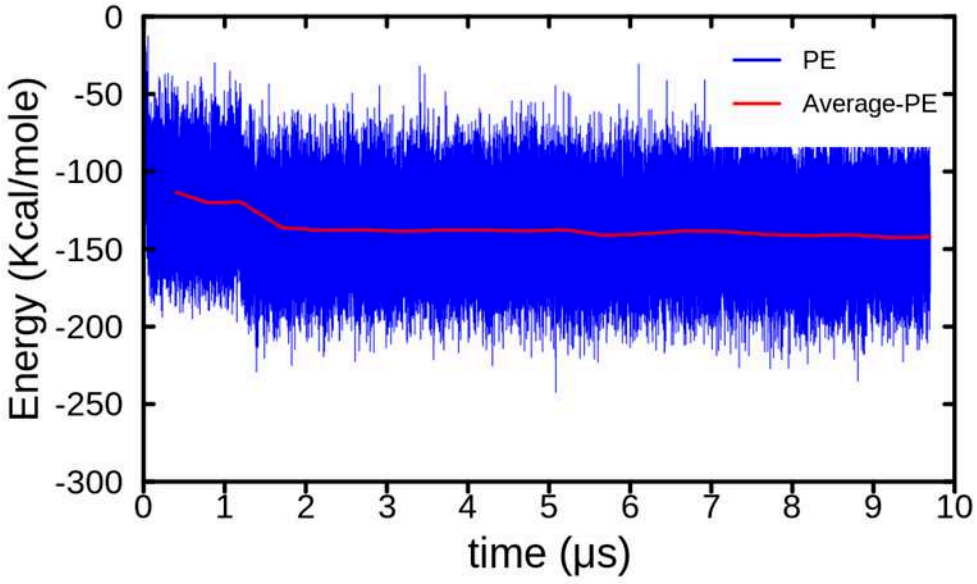

Figure 5-19: Potential energy profile and it average over 0.4 µs of an NVT ensemble MD simulation for insulin^{+6} in vacuum over 9.7 µs.

IV. Optimization of C0 in vacuum by quantum mechanics calculations

The molecular weight of insulin^{+6} is 5,813.676 g/mol and it comprises 794 atoms. The total number of electrons in the system amounts to 3098. To relax the chemical bonds of the reference

structure of insulin derived from the MD simulation, we performed several cycles of geometrical optimization using DFT in deMon2k (version 6.1.6)[33]. The optimization was carried out in Cartesian coordinates using the Auxiliary Density Functional Theory (ADFT) level. We employed the PBE[34] correlation and exchange functional with the DZVP (double zeta valence polarization) basis set and GEN-A2 auxiliary sets[35]. Moreover, the D3 correction to include the dispersion forces was added[53]. The tolerance of the maximum root mean square (RMS) force was set to 10^{-4} Ha/bohr.

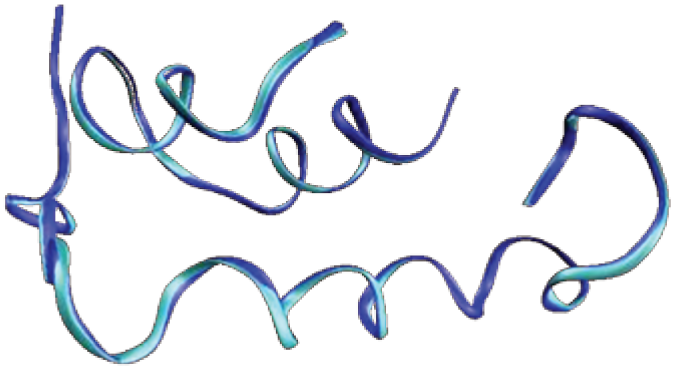

Figure 5-20: NewRibbons representation of MD reference structure of insulin +6 in cyan color and after 480 DFT optimization cycles in blue.

While the conformation of the reference insulin has not undergone crucial modifications, it was sufficient to relax chemical bonds to their equilibrium values. The differences in the secondary structure of the reference and optimized insulin^{+6}, as illustrated in Figure *5-20*, are negligible.

5.2.2 Substance PH^+

Substance P[15] is a neuropeptide comprised of 11 amino acids, constructed from 8 distinct types: Arginine (Arg), Proline (Pro), Lysine (Lys), Glutamine (Gln), Phenylalanine (Phe), Glycine (Gly), Leucine (Leu), and Methionine (Met). The sequence of these amino acids ranges from the N-terminal Arg to the C-terminal Met, as follows: Arg1-Pro2-Lys3-Pro4-Gln5-Gln6-Phe7-Phe8-Gly9-Leu10-Met11. Substance P has three groups with high proton affinity: the side chain of Arg1, the Arg-NH_2 terminal, and Lys. When Substance P undergoes protonation to produce Substance PH^+

in the gas phase (SP), it is anticipated that the added proton will reside on the Arg side chain due to its significant proton affinity. In addition, during the DFT geometry optimization, no proton transfer or negative modes were observed from the Arg side chain.

In a few nanoseconds of MD simulations, we observed significant changes in the hydrogen bond network and conformations of SP, in contrast to insulin^{+6}. To explore the conformations and hydrogen bonds with the Arg side-chain, we carried out extended replica-exchange molecular dynamics simulations (REMD)[54]. REMD is an advanced MD simulation technique aimed at enhancing the sampling of a system's configurational space. This technique involves running multiple parallel simulations (replicas) of the same system at varying temperatures or other thermodynamic parameters. The primary benefit of replica-exchange simulations is their enhanced sampling capability, offering a more comprehensive exploration of a molecular system's phase space. This is especially crucial in scenarios where the system may remain trapped in local energy minima for extended durations in traditional MD simulations. Four 100 ns REMD simulations were carried out with the CHARMM package by some of my colleagues from ICP. 16 windows spanning regularly spanning the temperature from 300 to 956 K were defined. Temperature exchanges were attempted every ps. The REMD simulations highlighted that SP adopts diverse conformations, each characterized by distinct hydrogen bonding patterns, as outlined in Table 5-8.

Table 5-8: Percentage of hydrogen bonds between residues and Arg1. A hydrogen bond is defined when the hydrogen atom and the electronegative atom are separated by less than 2 Å. Residues are identified by name and residue number, along with the specific atom forming the hydrogen bond with Arg1 and its classification as either side-chain (SC) or backbone (BB).

Residues	% HB
Gln6-O-SC	26
Leu10-O-BB	10
Met11-O-BB	9
Pro2-O-BB	7
Lys3-N-SC	4
Lys3-O-BB	1

As indicated in the table, the side chain of Arg most frequently forms a hydrogen bond with the side-chain oxygen (O) atom of Gln6, with a probability of 26%. Following Gln6, Leu10, Met11, Pro2, and Lys3 are candidates for forming hydrogen bonds with Arg, having probabilities of 10%, 9%, 7%, and 5%, respectively.

The diversity of conformations and hydrogen bond networks in SP complicates the selection of a single conformation for XUV simulations. As a result, we selected six different conformations of SP, each with distinct conformations and hydrogen bond networks, as depicted in Figure 5-21. The first four structures, referred to as SP_a, SP_b, SP_c, and SP_d, were extracted at the end of the 100 ns of simulation. The final two conformations, labeled as SP_e and SP_f, were derived from the most probable structures in clustering during REMD simulations.

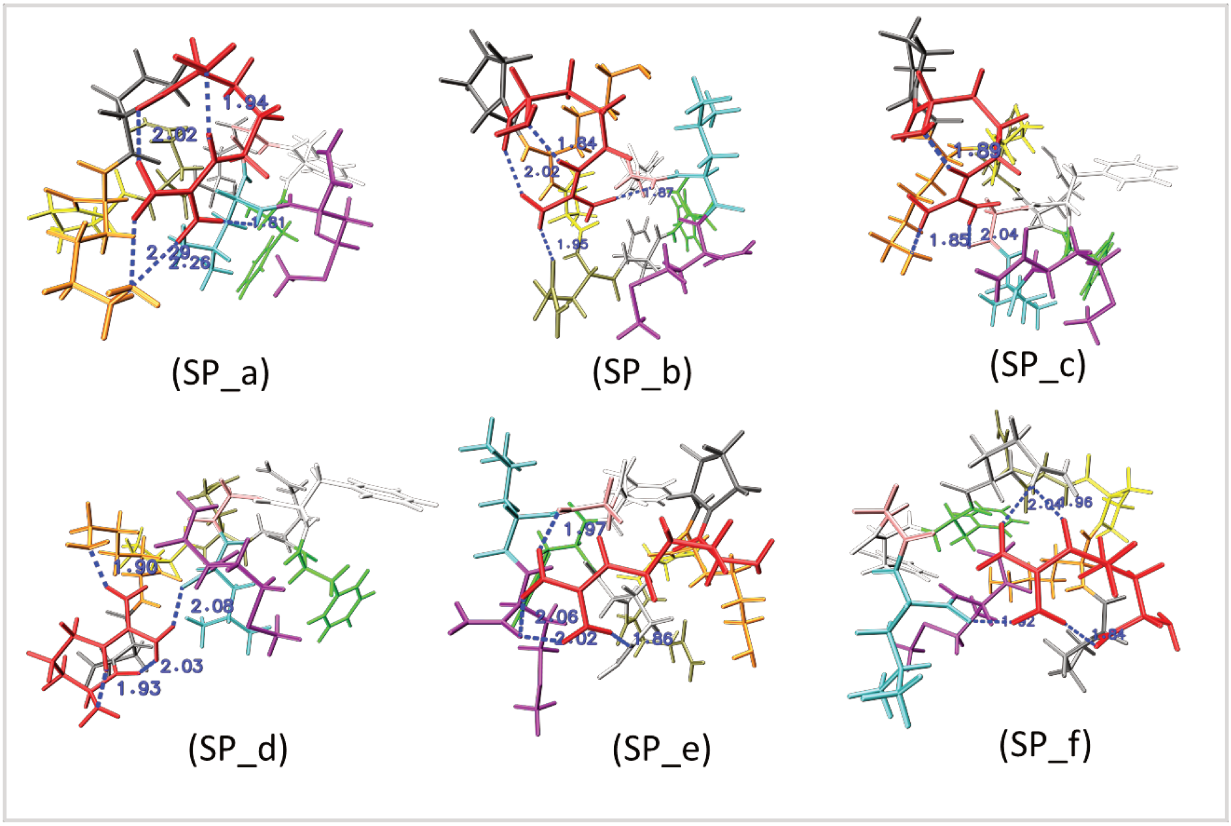

Figure 5-21: Visualization of different conformations of Substance PH⁺ (SP) selected for XUV simulations. The hydrogen bond network with Arg1 is represented by dashed blue lines. Residues are color-coded as follows: Arg1 (red), Pro2 (gray), Lys3 (orange), Pro4 (yellow), Gln5 (tan), Gln6 (silver), Phe7 (green), Phe8 (white), Gly9 (pink), Leu10 (cyan), and Met11 (purple).

As done for insulin, we optimized the geometry of SP conformations at the DFT level to relax the chemical bonds and hydrogen bonds derived from the MD simulation, using the same

methodology as for insulin. The tolerance for the maximum root mean square (RMS) force was set to 10^{-4} Ha/bohr. The total energy of the optimized conformations, relative to SP_a, is presented in Table 5-9.

Table 5-9: relative DFT energies of optimized Substance PH^+ conformations with respect to SP_a.

Structure	ΔE (eV)
SP_a	0.00
SP_b	-0.31
SP_c	0.21
SP_d	0.04
SP_e	-3.18
SP_f	-3.53

As seen in the table, the relative DFT energies of the four conformations derived from traditional MD simulations are comparable. However, in the case of SP_e and SP_f compared to SP_a, there is a significant difference, indicating that they are more stable. This can be attributed to the use of short MD simulations for SP_a to SP_d, whereas in the case of REMD, SP is well-sampled.

In general, after DFT optimization, the conformations retain their overarching structure and hydrogen bonds. No major changes in conformation or the order of hydrogen bonds are observed. The optimized conformations of SP are illustrated in Figure 5-21. Hydrogen bonds involving the Arg side-chain are highlighted with dashed blue lines. The specific hydrogen bond interactions for each conformation are summarized as follows:

- SP_a: the Arg side-chain forms hydrogen bonds with Lys3-N-SC, Leu10-O-BB, Arg1-O-BB, and Arg-N-BB,
- SP_b: the hydrogen bond interactions are with Gln5-O-SC, Gly9-O-BB, Arg1-O-BB, and Arg1-N-BB,
- SP_c: Arg forms hydrogen bonds with Lys3-N-SC, Gly-O-BB, and Arg1-N-BB,
- SP_d: Arg interacts via hydrogen bonds with Lys3-N-SC, Leu10-O-BB, Arg1-O-BB, and Arg1-N-BB,

- SP_e: Arg forms hydrogen bonds with Gln6-O-SC, Gly9-O-BB, and Met-O-BB,
- SP_f: Arg establishes hydrogen bonds with Gln6-O-SC, Leu10-O-BB, and Arg1-O-BB.

In summary, through MD simulations and REMD, it is evident that SP in the gas phase adopts a variety of conformations, each with its own unique hydrogen bonding network. This has led us to select multiple structures for XUV simulations.

5.2.3 Partial conclusion

In this section, we detailed the conformational structures of insulin^{+6} and substance PH^{+} in gas phase. While we utilized long MD simulations for insulin^{+6}, our approach for substance PH^{+} was based on REMD.

Our findings indicate that insulin^{+6} maintains a general form in the gas phase even after 2 µs. Despite the presence of six charges on insulin, its conformational structure remains relatively stable post-2 µs, especially in the A-chain. This consistency enabled us to confidently select a single reference conformation for XUV simulations. Conversely, substance PH^{+} in the gas phase exhibits a variety of conformations and hydrogen bond networks, complicating the selection of a single reference structure. The smallest structural fluctuations in insulin^{+6} are likely due to the three disulfide bonds that fortify the protein's scaffold.

5.3 Interaction of XUV-pulse

In this section, we detail results from the RT-TD-ADFT simulations for insulin^{+6} and substance PH^{+} under the influence of an XUV pulse. We begin with the methodological details of the simulations and then proceed to examine the effects of the electric field strength and pulse orientation. The section culminates with main results on the primary ionization events observed in the simulations.

5.3.1 Numerical detail

Describing large systems, such as insulin^{+6} or SP entirely within RT-TD-ADFT poses a significant challenge in terms of computational cost. We initially tested several parameters in RT-TD-DFT to increase the speed of the simulation without affecting the stability of the electron dynamics and the results.

We have determined that our simulations are stable and require less computational cost when using the second-order Magnus propagator[36] coupled with a predictor/corrector algorithm[37]. We employed a time step of 1 as and a Taylor expansion[38] to evaluate the exponential of the Kohn-Sham matrix with 45 terms. The PBE exchange-correlation functional was used. As discussed in Section 5.1, due to the robustness of the ECP basis set in describing the ionization process induced by XUV-pulse, we incorporated it in our simulations. This basis set substantially reduced the computation time. For density fitting, the auxiliary basis set differs for insulin^{+6} and SP due to the size of insulin^{+6}. For SP we used GEN-A2* on all atoms, whereas for insulin^{+6}, we used GEN-A2 for H and C atoms and GEN-A2$^{\delta}$* for O, N, and S. The difference between GEN-A2 and GEN-A2* is that the latter contains *f* and *g* Hermite Gaussian functions in addition to *s*, *p*, and *d* functions. In the modified GEN-A2$^{\delta}$*, we removed only *g* functions. We tested this auxiliary basis set, as well as GEN-A2, with peptides used in Section 5.1. We determined that compared to GEN-A2*, the effect of the auxiliary basis set was negligible.

To introduce the XUV-pulse, we utilized a squared cosine-shaped pulse in RT-TD-ADFT, maintaining the same parameters as those described in the "basis set tests" section. The only

exception was the maximum electric field strength. The total duration of the pulse was set at 30 fs, with the pulse's centroid energy at 30 eV. In this setup, the maximum electric field was adjusted to remove a single electron. For SP, this condition was met as we tested the effect of the electric field. However, for insulin^{+6}, testing this was challenging; a single simulation with insulin^{+6} requires 50 days, even with optimal parameters and CPUs. For SP, we set the maximum electric field strength at 0.00375 Ha/e.bohr (4.93 x 10^{+11} W/cm^2), while for insulin^{+6}, it was set at 0.0025 Ha/e.bohr (2.19 x 10^{+11} W/cm^2). For insulin^{+6} and each conformation of SP, we conducted three simulations corresponding to the xyz-direction of the pulse.

In all simulations, we employed the complex absorbing potential (CAP) in energy space to remove electrons in the continuum. The parameters of CAP were set at 0.15 Ha^{-1} energy scale for γ_0 and a 0.05 Ha^{-1} damping strength for ξ. The vacuum energy cut-off was approximated using the electron affinity for each system.

5.3.2 Effect of electric field strength and pulse direction

In the experiments, they ensured that each molecule absorbs no more than one photon. This ensures a low probability of multi-photon absorption, maintaining linearity in absorption. Additionally, the molecule is sampled by the XUV-pulse from all directions.

Kraus *et al*.[7] demonstrated that the orientation of the pulse significantly affects ionization and charge migration in smaller molecules, such as iodoacetylene. However, the impact on larger molecular systems, as in our case, is not well-understood.

Therefore, in this sub-section, we investigate the influence of electric field strength and pulse orientation on the interaction with the XUV-pulse. To elucidate these effects, we particularly focus on the results of the SP_a conformation.

a. Effect of electric field strength

To investigate the impact of electric field strength (F_0), we conducted five RT-TD-ADFT simulations for SP_a, introducing the XUV-pulse in the x-direction with F_0 values of 0.001, 0.002, 0.003, 0.004, and 0.005 (Ha/e.bohr).

We computed the variations in energy and charge relative to the ground state over time for each F_0, as illustrated in Figure 5-22.

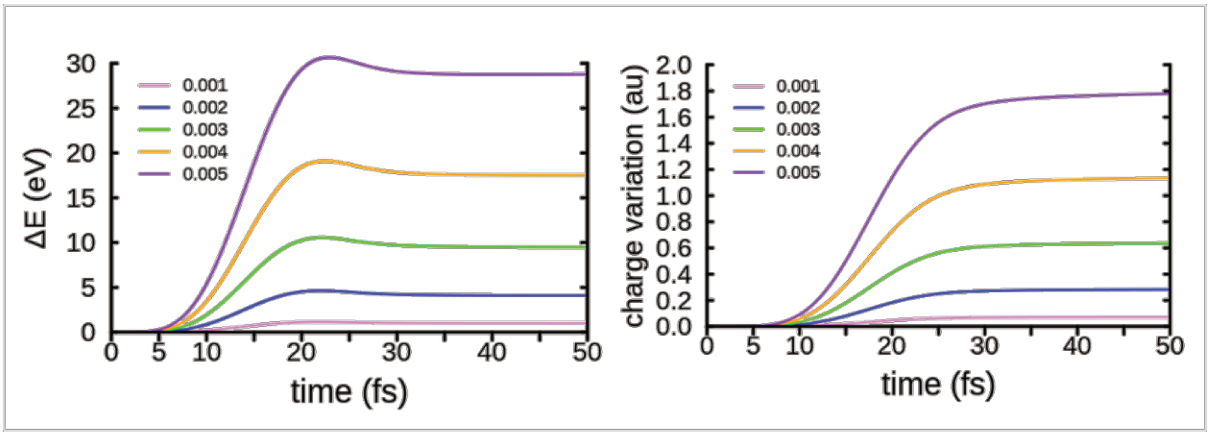

Figure 5-22: Variation in total energy (left) and charge relative to the ground state as a function of time for SP_a, depending on the electric field strength (Ha/e.bohr) on the x-direction.

As observed in the figure, the rate of energy variation increases with the electric field strength during the pulse. Despite the dome-shape observed around 20 fs, as explained in Section 5.1, this is related to the pulse energy which is included in the total energy. This dome-shape is not evident in the case of charge variation. Importantly, after the pulse, both the energy and charge variations remain stationary, confirming the stability of the simulation. Additionally, as F_0 increases, there is a noticeable increase in energy variation. The energy variations are 1, 4.1, 9.5, 17.5, and 28.8 (eV) at 50 fs, successively with increasing electric field strength. Correspondingly, in terms of energy variation magnitude, the charge variation or ionization state also increases with F_0, with values of 0.07, 0.28, 0.64, 1.13, and 1.78 electrons, respectively.

Furthermore, to examine the linear response to F_0, we illustrated the relationship between energy and charge variation at 50 fs as a function of the square of the electric field strength in Figure 5-23.

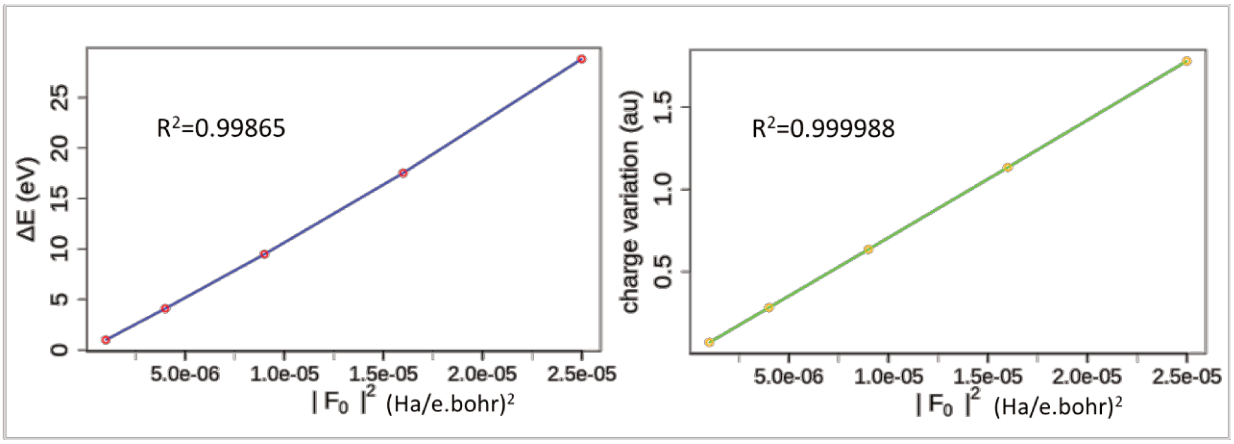

Figure 5-23: Variation in total energy (left) and charge (right) at 50 fs as a function of the square of maximum electric field strength in the x-direction for SP_a.

As observed in the figure, the response of both energy and charge variation is linear within our range of F_0, even when more than one electron is ionized. For energy variation, a slight deviation from a straight line is observed at higher F_0 values. However, in the case of charge variation, the relationship remains perfectly linear.

It is essential to highlight that, as mentioned in Chapter One, the ionization of a single electron in small molecules like N_2 is particularly challenging. When increasing the CAP strength, the electron dynamics becomes unstable, and when the electric field strength is augmented, the response becomes non-linear, leading to the depopulation of electrons in the continuum state. However, as demonstrated here, for larger systems, we were able to ionize more than one electron without any effects on the stability of the simulation.

Traditionally, in XUV-pulse simulations, a single hole is generated to examine the pulse's interaction. Based on our findings related to energy variation and total ionization of the molecule, we can confirm it is not truly required, given that the response is linear. Thus, anyone calculating the energy variation and ionization state with one value of F_0 can use it to predict others. For more, to explore the influence of electric field strength on the nature of ionization within the molecules, we computed the percentage of normalized hole numbers on the residues based on F_0 values, as presented in Table 5-10.

Table 5-10: Percentage of normalized number of holes in residues of SP_a as a function of electric field strength (Ha/e.bohr) on the x-direction.

Residues	$\%Nh$				
	0.001	0.002	0.003	0.004	0.005
Arg1	13.0	13.0	13.1	12.8	13.0
Pro2	7.3	7.2	7.4	7.3	6.9
Lys3	11.3	11.1	11.3	10.8	10.6
Pro4	8.9	8.9	8.9	8.9	9.4
Gln5	8.2	8.4	8.3	8.1	8.4
Gln6	8.8	8.9	8.9	9.2	8.9
Phe7	10.5	10.6	10.6	10.5	10.7
Phe8	10.1	10.4	10.3	10.7	11.4
Gly9	3.2	3.0	3.0	3.0	3.0
Leu10	7.0	6.9	6.8	7.0	6.7
Met11	11.7	11.6	11.4	11.7	11.0

Importantly, as observed in the table, the nature of ionization within the molecule is independent of the electric field strength. Minor differences are observed, for example, in the cases of Pro2, Lys3, and Phe8 at F_0 = 0.005 (Ha/e.bohr). However, when compared to the accuracy of the percentage predictions for decimal numbers, these differences are negligible. This result confirms that the electric field strength does not affect the nature of ionization.

Overall, our findings indicate that in the linear regime, the traditional practice of including one hole in the simulation is not essential. Relying on a strong electric field to remove one electron has its drawbacks. For instance, those using the CAP in simulations for small molecular systems may encounter issues. Furthermore, due to the adiabatic approximation in RT-TD-DFT, when molecular systems are strongly perturbed from the ground state, tracking the exchange and correlation of electrons over time is expected to become challenging, potentially affecting the dynamics[55]. For larger molecular systems, at least in our study, non-adiabatic effects seem to respond similarly across the range of electric field strengths examined. This similarity is evident in the ionization process, which displays a consistent response to F_0.

b. Effect of pulse direction

To investigate the effect of pulse orientation, we conducted three simulations for SP_a, directing the pulse along the x, y, and z axes. The electric field strength was set at 0.00375 (Ha/e.bohr), which corresponds to the removal of approximately one electron from the molecule. The results, which show the percentage of normalized number of holes in residues as a function of pulse direction, are presented in Table 5-11. While we carried out simulations for all SP conformations and insulin^{+6} in three directions, the effects were consistent for all SP conformations and insulin^{+6}. Thus, we have primarily focused on the results from SP_a.

Table 5-11: Percentage of normalized holes in SP_a residues based on XUV-pulse direction, with an electric field strength of 0.00375 Ha/e.bohr.

Residues	$\%Nh$			
	x	**y**	**z**	**<xyz>**
Arg1	13.0	12.0	13.3	12.7
Pro2	7.2	6.6	7.1	7.0
Lys3	11.1	11.2	10.2	10.9
Pro4	9.3	9.0	8.7	9.0
Gln5	8.5	8.2	8.6	8.4
Gln6	8.7	9.0	9.3	9.0
Phe7	10.5	11.0	10.3	10.6
Phe8	10.5	11.9	12.0	11.4
Gly9	3.1	3.3	3.0	3.2
Leu10	6.8	7.5	6.6	7.0
Met11	11.4	10.9	10.9	11.1

As observed in the table, while the effect of pulse orientation is not significant, it is also non-negligible. The difference is more pronounced than when the electric field strength is modified. Depending on the residue, the general differences in $\%Nh$ hover around 1%. The most significant difference can be observed for Phe8 between the x and z directions, with an absolute percentage difference amounting to 12.5%. This could be attributed to the polarity of the molecular orbitals (MOs). Even though the system is extensive, the polarity of the MOs is not uniformly distributed

across the x, y and z directions. A similar difference is also observed for insulin^{+6}. Thus, for the rest of our results, we take into account the average across the x, y and z directions.

5.3.3 Primary ionization events

As mentioned in the introduction, the impact of the size or secondary structure of a protein on ionization and electronic relaxation in the presence of an XUV-pulse remains to be clarified. Both insulin and substance P serve as ideal models to explore this topic since they possess primary, secondary, and, tertiary structures.

Given the extended duration of the pulse in simulation, nuclear motion might influence the dynamics. Therefore, to ensure the robustness of our findings, we concentrate solely on the primary ionization events.

This sub-section delves into the analysis of total energy variation, overall ionization, hole distribution across residues, charge migration, and the correlation of ionization with factors like ionization potential, valence electrons, and the environment of the residues. To streamline our presentation, for substance P, we primarily discuss results pertaining to SP_e, as it doesn't influence our overarching conclusions.

a. Energy variation and total ionization state

To provide an overview of the energy variation and the overall ionization state at the end of the simulations (50 fs), we computed the average energy variation and the total number of holes across the xyz-directions of the pulse for insulin^{+6} and all conformations of SP. These results are summarized in Table *5-12*.

Table 5-12: Average energy variation and number of ionizations across xyz-directions for insulin^{+6} and substance PH^{+}.

system	$\overline{\Delta E}$(eV)	No. of ionization
Insulin^{+6}	38.0	1.83
SP_a	15.4	1.01
SP_b	15.6	1.02
SP_c	15.3	1.03
SP_d	15.2	1.02
SP_e	15.3	1.00
SP_f	15.9	1.06

As seen in the table, the ionization state of SP aligns perfectly with the removal of one electron from the molecule, and this is consistently observed across all conformations. However, for insulin^{+6}, The molecule loses 1.83 electrons. Due to computational constraints, we couldn't adjust the electric field strength to remove just one electron. Notably, as demonstrated earlier, the molecular response remains linear within this range, so it should not influence our conclusions. This is particularly true for insulin^{+6}, which has about 4.26 times more valence electrons than SP. Specifically, insulin^{+6} has 2240 valence electrons, while SP has 526.

The energy remaining after the pulse is a crucial parameter that can be utilized, for instance, in molecular dynamics (MD) simulations to comprehend certain aspects of ultra-fast structural relaxation or energy dissipation. In our simulation, energy variation after pulse represents the internal energy retained in the system due to the excitation of electrons and the molecule's new ionic form. For SP, energy increase reached 15 eV, mirroring its ionization state. however, despite insulin^{+6} being 4.26 times larger, once the energy variation is normalized by its ionization state, the outcome is 20.7 eV. This is only 5 eV greater than SP, whereas a projection based on the number of valence electrons would place it at 63.9 eV. Comparing these findings, for instance,

with Met-enkephalin (P1) evaluated in Section 5.1, the energy variation normalized for a single hole in P1 is 12.4 eV, given its 216 valence electrons. This underscores that energy variation isn't strictly proportional to the number of valence electrons.

This phenomenon is anticipated due to the fact that, in all cases, a single electron is removed, theoretically leading to consistent retained energy. However, the slight relative increase observed with larger systems is attributed to variations in electronic structure, ionization characteristics type of atoms to each system, and an elevated likelihood of excitation as the system size expands. However, what is crucial is that the remaining energy does not substantially increase with the size of the system, which could be helpful in understanding the processes of structural relaxation and fragmentation within larger systems.

b. Ionization of residues and charge migration

To demonstrate the hole distribution across residues and charge migration in insulin^{+6} and SP_e, we computed the normalized hole distribution over residues at two distinct times: 20 fs and 50 fs. This is depicted in Figure 5-24, with the distribution presented in accordance with the sequence number of residues in insulin^{+6} and SP.

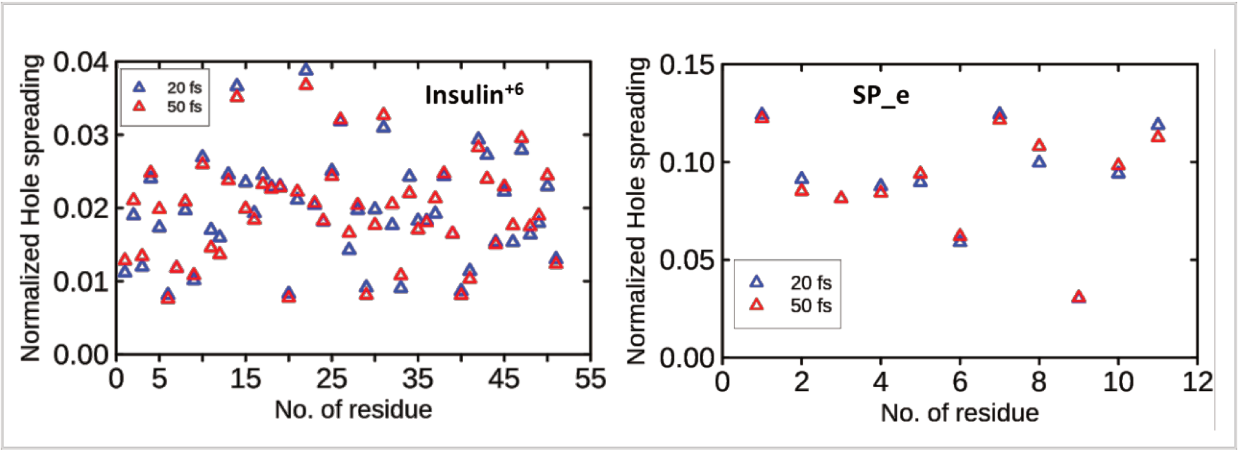

Figure 5-24: Normalized holes spreading over residues at 20 fs and 50 fs for insulin^{+6} (left) and SP_e (right)

Upon first observation of the figure, it can be seen that for both systems, charge is strongly delocalized over all residues. However, some residues are more impacted than others. Insulin^{+6} showing a larger delocalization of holes due to its greater number of residues. Several reasons can account for this extensive charge delocalization during XUV-pulse interaction. The energy uncertainty of the pulse, given its duration, exceeds the energy difference between MOs, leading

to the ionization of multiple electrons across various MOs. Moreover, even for a single distinct energy, the ionization of several electrons from HOMO and deeper valence states is possible, especially in larger systems. Additionally, the XUV-pulse predominantly targets deeper valence electrons, especially within the energy range of our pulse. Deep valence states are known for their strong superposition due to high correlation, which can further spread the charge in the molecular system.

The analysis of residues charge at different times corresponds to the movement of charge among residues due to charge migration, ICD, or auto-ionization, as explained in Chapter One. Because we cannot separate these processes in our analysis, we call all processes "charge migrations". As seen in the figure, charge migration occurs in nearly all residues. However, the amplitude of charge migration, compared to the initial ionization by the pulse, is significantly low. There could be two reasons for this.

First, the nature of the interaction of the XUV-pulse which is different from, for example, UV rays or X-rays. UV rays or X-rays selectively ionize electrons in the HOMO or core electrons of a specific atom in the molecule. Consequently, the probability of post-ionization such as charge migration compared to the probability of initial ionization in the molecule is expected to be significant. On the contrary, an XUV-pulse depopulates nearly all valence electrons, leading to many channels of charge migration being activated simultaneously. Perhaps an atom at the same time receives and transfers electrons. Moreover, because the probability of initial ionization distributes over all atoms in the molecule, it is expected that the likelihood to observe charge migration would be low. This low probability of charge migration and complex charge migration in one fragment have been observed in the case of phenylalanine by F. Calegari *et al.*[8,11,20]. If this reason holds true, it establishes a fundamental difference between XUV-pulse interactions and those of UV-rays and X-rays in terms of charge migration.

Second, this might be an artifact due to the adiabatic approximation in RT-TD-DFT. As discussed in Chapter One, the adiabatic approximations could impact electron dynamics, especially in cases where the system is strongly perturbed. This effect has been observed for the Hubbard model[56] and glycine[55]. However, for larger molecular systems, the impact is not yet clear. It is challenging

to verify this because a reliable code that takes this effect into account doesn't exist to date, especially for large systems.

Even though the amplitude of charge migration is small, we attempted to identify correlations of charge migration between residues. For this purpose, we employed correlation matrix. As illustrated in Chapter Four, the correlation matrix can effectively identify correlations of charge migration among residues. In Figure 5-25, we calculated the correlation matrix for SP_e and SP_a based on charge variations on residues within the 40-50 fs range. We chose this time frame due to the potentially complex charge migration processes in the system, which might lead to rapid changes in correlations between residues.

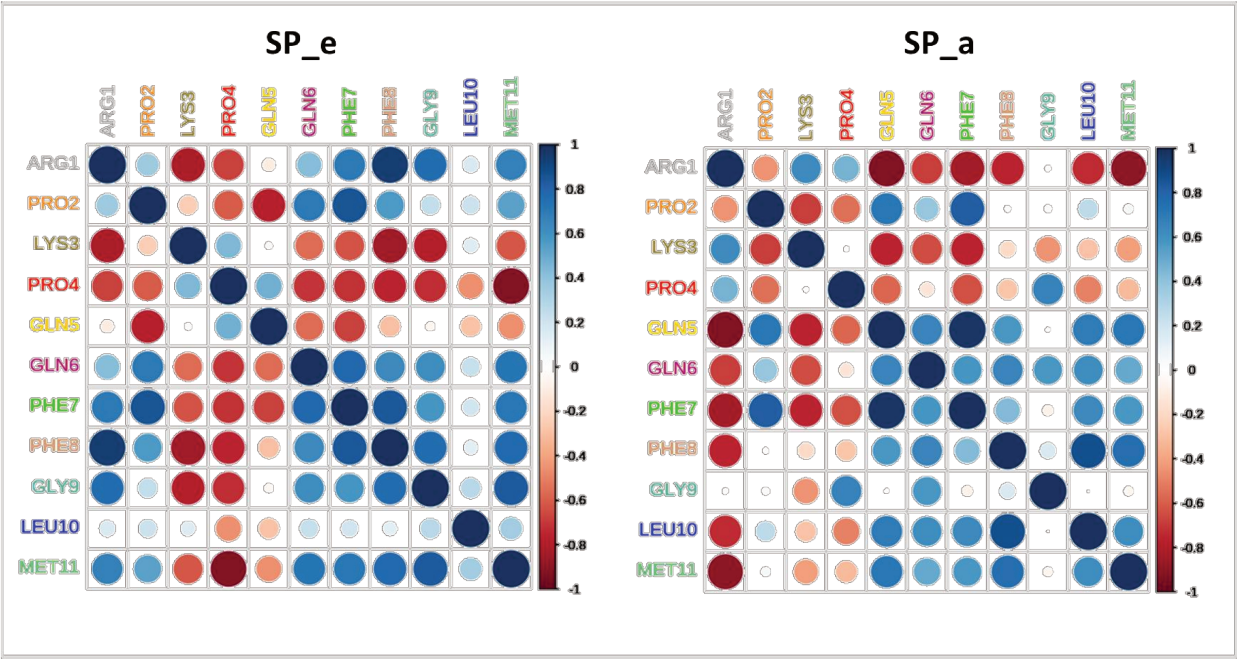

Figure 5-25: Correlation matrix of charge variation between residues in the 40-50 fs interval for SP_e (left side) and SP_a (right side).

As observed in the figure, we can identify several strong anti-correlations between residues, which correspond to charge migration between the two residues. The charge migration correlations between residues differ in SP_e and SP_a, which is expected since the electronic structure and the environment of the residues are quite different in the two systems. In SP_e, Arg1 exhibits a strong anti-correlation with Lys3, and Pro4 has an anti-correlation with Met11. For SP_a, Arg1 shows strong anti-correlation with several residues: Gln5, Phe7, and Met11. These

results confirm that even if the amplitude of charge migration is small, we can still identify charge migration between residues.

c. Correlation between ionization of residues and ionization potential

As previously observed, ionization is spread across all residues, though some residues experience more ionization than others. This disparity might be associated with the ionization potential (IP) of the residues. To investigate this, we calculated the IP for all residues in insulin^{+6} in two ways: First, we calculated the IP for each residue in a vacuum. Second, we determined the IP for each residue in insulin^{+6} using the QM/MM technique, where the environment is considered at the MM level. For this test, we employed the B3LYP[57] exchange-correlation functional which is general hybrid functional. The IP calculation is based on the ionization of electrons in the highest HOMO.

The results, shown in Figure *5-26*, are presented in correlation with the normalized number of holes for each residue.

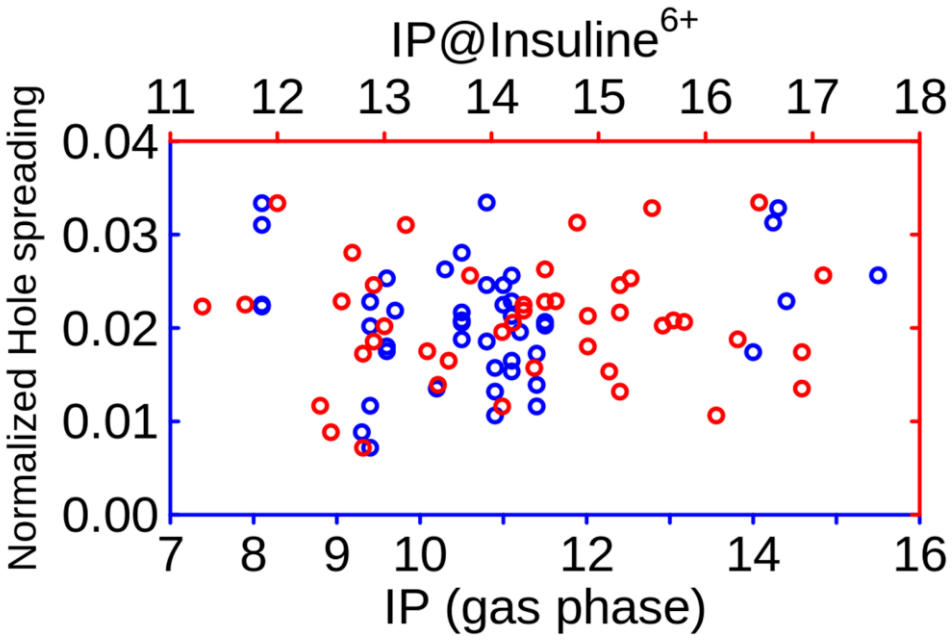

Figure 5-26: Correlation of the normalized number of holes on residues in insulin with their ionization potential in a vacuum (blue circles) and within the insulin+6 environment (red circles).

As evident from the figure, no correlation is observed between the charge on the residue and IP, even when considering the insulin^{+6} environment. This is expected because the XUV-pulse ionizes not only the HOMO but also the deeper valence electrons.

d. Correlation between ionization of residues and valence electrons

For further insight into the correlation of residue ionization, in Figure *5-27*, we plotted the normalized number of holes on the residues with respect to their number of valence electrons for insulin^{+6} and SP_e.

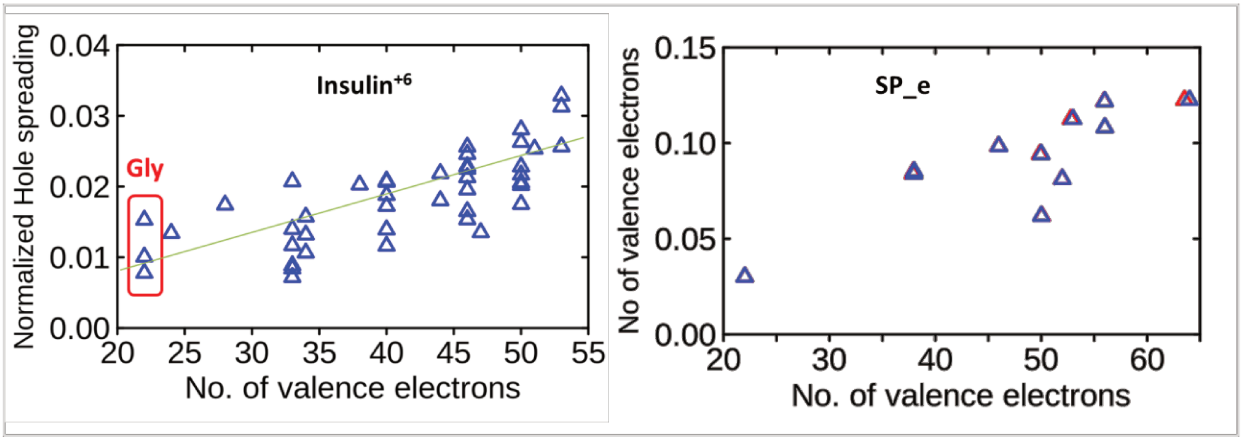

Figure 5-27: Correlation between the normalized number of holes on residues and their number of valence electrons for insulin^{+6} (left) and SP_e (right). The red triangles in the SP_e figure correspond to the number of valence electrons calculated within the SP_e molecule.

As illustrated in the figure, there is a discernible trend between the number of valence electrons and the charge on the residue. This is anticipated because the XUV-pulse interacts with all valence electrons; hence, as the number of valence electrons increases, the likelihood of the XUV-pulse interaction also rises. However, in both systems, while the probability of ionization does increase with the number of valence electrons, it remains variable and depends on the localization of the residue in the structure. For example, insulin^{+6} contains three Gly residues, as visualized in Figure *5-28*, and each Gly displays a distinct ionization state. As evident in Figure *5-28*, these three Gly residues are situated in different environments, which could influence their ionization.

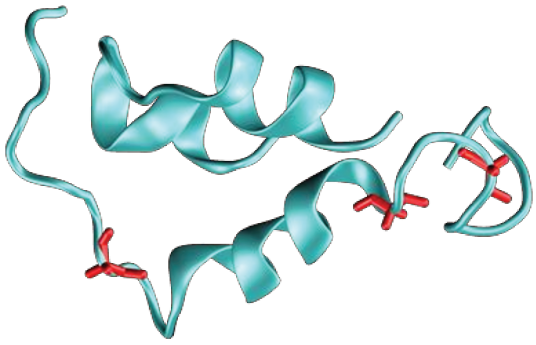

Figure 5-28: Visualization of the insulin structure with glycine residues highlighted in red.

e. Effect of residues environment

To illustrate the effect of the residue environment, we visualized the normalized number of holes on the residues using a structural representation, scaling the color of each residue according to the number of holes for both insulin^{+6} and SP_e. Given that the number of valence electrons in residues and the number of residues in insulin^{+6} and SP_e are different, the direct value of the normalized number of holes on the residue is not directly comparable when assessing the effect of the environment. For instance, larger residues might inherently exhibit high values, which would not necessarily correspond to the environmental effect. Therefore, to make the values of normalized number of holes on the residues comparable across residues, we first normalized these values by the number of valence electrons. Then, to ensure comparability between insulin^{+6} and SP_e, we scaled the obtained values between 0-1. To distinguish this value from the normalized number of holes, we refer to it as NNh. The results are illustrated in Figure 5-29.

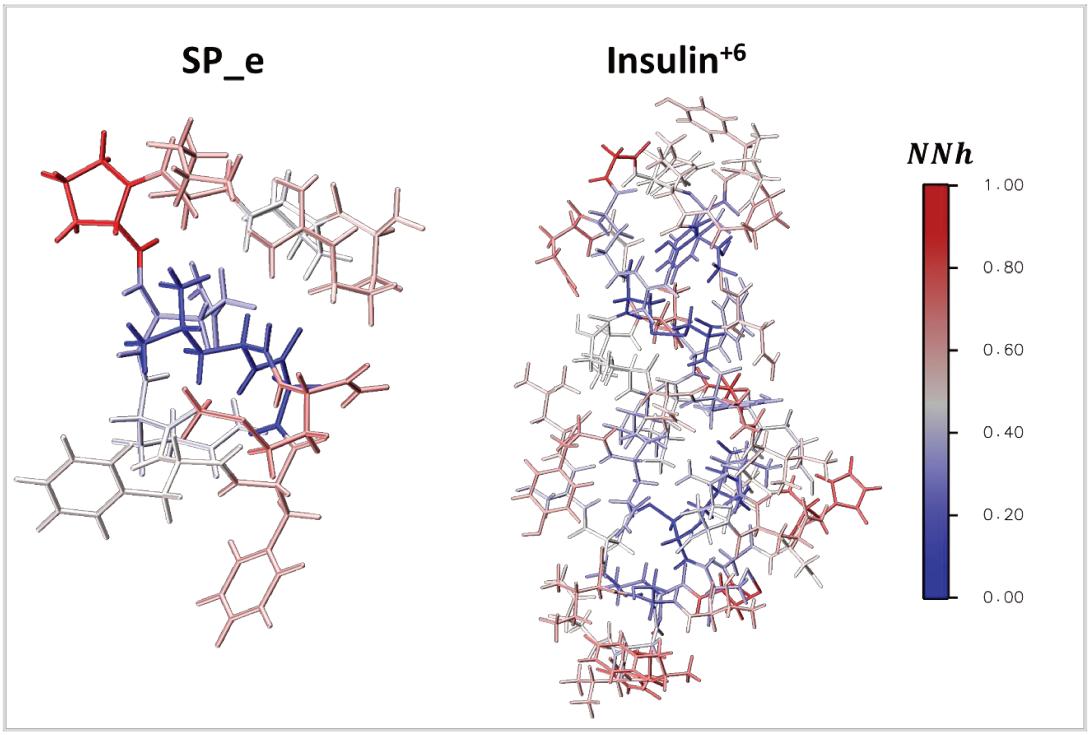

Figure 5-29: Representation of hole spreading over residues for SP_e (left) and insulin^{+6} (right). The color scale of each residue corresponds to the NNh value. For details on NNh, refer to the accompanying text.

Interestingly, the figure reveals that superficial residues are more susceptible to ionization compared to those buried within the molecule. The red color, which indicates a higher probability of ionization, predominantly covers the surface of the molecules, while the blue color, signifying a lower probability of ionization, is distinctly located inside the molecules. The probability of ionization for residues scales with their location from the interior to the surface of the molecule. Additionally, as observed in the figure. the separation zone in ionization probability based on residue location is more pronounced in insulin^{+6} compared to SP_e.

This remarkable result is not an artifact, for instance, from the basis set. Although we demonstrated in Section One the capability of our basis set to describe ionization processes, if there were a problem, the buried residues should be the most susceptible to ionization. This is because they are surrounded by more atoms compared to superficial residues. As a result, the better description of the continuum state would increase, and the probability of ionization would also increase. However, this is not the case.

One possible explanation for this observation is that superficial residues in the gas phase lack interactions compared to buried residues. Consequently, the MOs within superficial residues

undergo a slight destabilization, resulting in elevated energy levels and expanded spatial distribution. One hypothesis posits that the probability of electron depopulation from these high-level MOs increases with the particular energy of the pulse. Another hypothesis suggests that the expanded spatial profile of superficial residues contributes to an increased ionization cross-section. To validate first hypothesis, we introduced a new descriptor in our code to calculate the proximity of MOs to the molecule's surface, which we termed the 'surface proximity index (SPI)' This is achieved by integrating the electron density of each MO outside of the van der Waals radius of the molecular surface. A higher value of SPI for MOs indicates that the MO is closer to the surface.

In Figure *5-30*, we collected the intrinsic properties and probability of ionization at the MO level alongside the value of SPI for insulin^{+6}. The diagram illustrates the energy of each MO, and its color is scaled according to the SPI value, which signifies the proximity of the MO to the surface of insulin^{+6}. The blue and dark blue curves represent the normalized number of holes on the MOs in intervals of 0.05 eV and 0.35 eV energy, respectively. The green line depicts the density of states (DOS) within a 0.35 eV interval, with the value of DOS normalized by the total number of states.

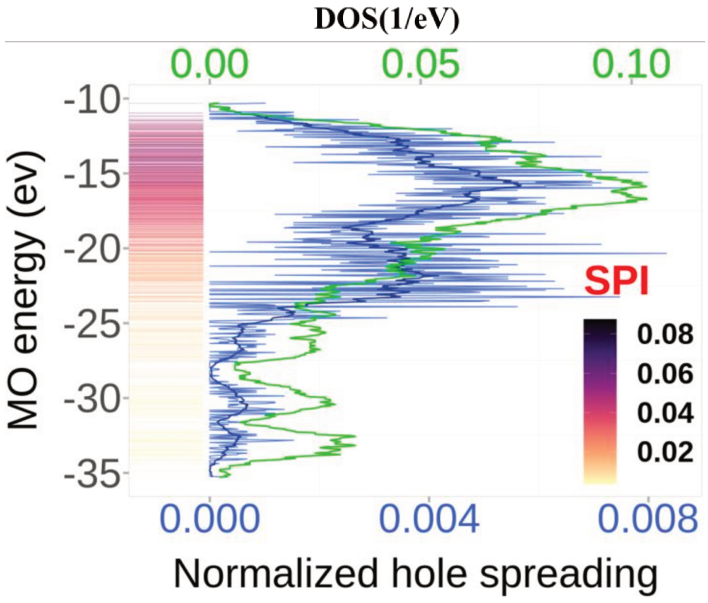

Figure 5-30: Relative probability of hole formation on insulin^{+6} as a function of the ground state Kohn-Sham molecular orbital energies. MO levels are colored in the energy diagram according to the proximity of each molecular orbital to the surface as quantified by the surface proximity index (SPI). Hole distribution and density of Kohn-Sham states are displayed in blue and green respectively.

Interestingly, as observed in the energy diagram, as the energy of the MOs increases, their proximity to the surface also increases. The MOs on the surface of the molecule are destabilized due to a lack of interactions compared to the ones buried in the molecule, which is why they have higher energy. The DOS (green line) is not a smooth curve; it is discretized into five major peaks around -33, -30, -26, -22, and -16 (eV). However, the majority of states are located in the energy range of -25 to -12 eV. As shown by the blue line, for the pulse energy used in this test, the holes spread over all MOs roughly following the same profile as the DOS, with the exception of the peak at -22 eV, which is shifted and more pronounced. Since we normalized the number of holes and the DOS, the difference in amplitude of peaks between both curves indicates the susceptibility of MOs in that energy range to ionization. Comparing the amplitude differences between the DOS and normalized number of holes, it is evident that the difference is more pronounced below -25 eV compared to higher energies. This suggests that our pulse is less susceptible to ionize the MOs below -25 eV. Particularly, the difference between both curves is very small, or coincident, in the range of -24 eV to -20 eV, indicating that the pulse's centroid lies within this region. However, the depopulation of MOs at higher energies remains significant, corresponding to the MOs' greater proximity to the surface. This, in turn, corresponds to the MOs of surface residues in insulin^{+6}. This explains why, with the pulse that we used, the surface residues are more ionized compared to residues inside insulin^{+6}. Furthermore, an additional observation can be made in the curve depicting the normalized number of holes. A significant reduction becomes noticeable at -18 eV. This phenomenon implies that the ionization of molecular orbitals (MOs) does not strictly align with the energy of the pulse. This behavior is expected in a heterogeneous molecular system, as each atom possesses a distinct ionization cross section. For instance, in the case of nitrogen, the measurement of partial ionization cross sections for HOMO and deeper valence levels fluctuates as a function of photon energy. Specifically, after a decrease in the proportion of ionization of the HOMO with increasing photon energies, a subsequent rise occurs above 30 eV[58].

This finding is crucial for understanding the interaction of XUV-pulses, especially with large molecular systems. It suggests that the XUV-pulse exhibits some selectivity towards surface residues or atoms. This selectivity may change by adjusting the energy of the pulse. For instance, a more energetic pulse that depopulates deeper valence electrons might ionize more buried residues, while a low-energy pulse that primarily depopulates the HOMO might ionize more

surface residues. Additionally, altering the protein's surface could also change the probability of surface residue ionization. For instance, by solvating the protein in water, the energy of the MOs on the molecule's surface could be modified, potentially impacting this feature. The second hypothesis, which suggests an increase in the ionization cross-section of surface residues due to the expansion of their MOs' spatial contribution, is also feasible. In this scenario, the susceptibility of surface residues to ionization is not contingent on the energy of the pulse; rather, it relies on the alteration of the protein's surface potential. However, a more in-depth analysis and further testing are required to draw a conclusive decision.

f. Ionization of backbone and side-chain

As explained in the introduction, the protein's backbone is responsible for its primary and secondary structures. It connects the sequence of amino acids, and by forming hydrogen bonds between the amino acid backbones, it gives rise to structures like alpha-helices and beta-sheets. On the other hand, the side chains play a pivotal role in determining the tertiary structure of a protein. Through interactions like electrostatic forces, hydrogen bonds, and disulfide bridges between different amino acid side chains, they shape the three-dimensional structure of the protein. Consequently, radiation damage to the backbone can result in breaking bonds between amino acids, leading to protein fragmentation or alterations in its secondary structure. Damage to the side chains can lead to the loss of these chains and disruption of the bonds between them, compromising the 3D structure of the protein. Thus, damage to the backbone and side-chain can distinctly alter the protein's structure.

To explore the impact of the XUV-pulse on the ionization of the backbone and side-chain, we calculated the percentage of holes on the backbone and side-chain across all peptides and insulin^{+6} discussed in this chapter, as depicted in Figure *5-31*.

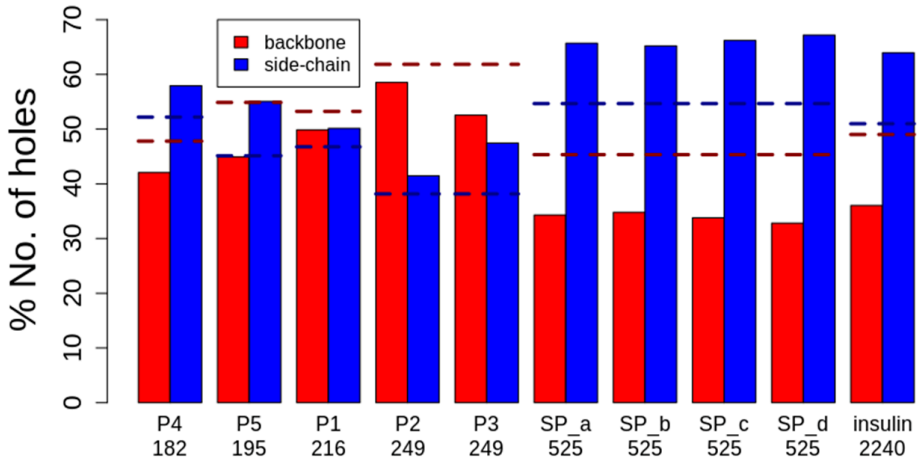

Figure 5-31: Percentage of holes on the backbone (red bar) and side-chain (blue bar) for peptides and insulin^{+6}, arranged based on their number of valence electrons. The dashed red and blue lines indicate the percentage of valence electrons in the backbone and side-chain, respectively.

In all the examined systems, the side-chains consistently show a higher probability of ionization compared to the backbone if considering the number of valence electrons. Interestingly, the percentage of holes does not directly correlate with either the number of valence electrons or the system's size. For example, while P2 and P3 contribute equally in terms of valence electrons, their respective backbones and side-chains respond differently to the XUV-pulse. Moreover, when considering the percentage contribution of valence electrons, the backbone and side-chain in both P4 and insulin^{+6} display similar ionization patterns, even though insulin^{+6} is approximately 12 times larger. The ionization responses of the backbone and side-chain to the XUV-pulse seem to hinge more on their positioning within the system's 3D structure or the specific types of side-chains present, as evidenced by the consistent response across all SP conformations. This distinction in ionization susceptibilities between side-chains and backbones can provide valuable insights into understanding the behavior of proteins when exposed to XUV radiation.

5.3.4 Partial conclusion

In this section, we investigated the interaction between XUV-pulses and the ultra-fast responses induced by these pulses in insulin^{+6} and SP using RT-TD-DFT simulations. Our analysis revealed several key insights, particularly relevant to understanding how XUV-pulses interact with large biomolecules:

- The internal energy remaining after the XUV-pulse interaction does not strictly increase with the size of the system, as one might expect due to the increase in the number of electrons in the system. This observation could provide insights into fragmentation and ultra-fast structural relaxation in large molecular systems.
- We established that ionization nature does not rely on the electric field strength in the linear response regime. However, it exhibits a slight dependence on pulse orientation.
- Our analyses reveal that XUV-pulse-induced ionization of residues is strongly delocalized across all residues. Moreover, the post-ionization charge migration between residues appears negligible when compared to the initial ionization of the residues themselves.
- We discovered that the ionization probability of residues does not hinge on their ionization potential. Instead, it generally correlates with the number of valence electrons in residues, albeit influenced by the residue's environment.
- A notable finding was the higher propensity for ionization in surface residues compared to those buried within the molecule. This observation can be attributed to the energy levels and spatial contribution of molecular orbitals in the surface residues.
- Lastly, our research indicates that, when exposed to XUV-pulses, the side chains of peptides or proteins have a higher likelihood of ionization than their backbones. This revelation can shed light on specific damage mechanisms in proteins triggered by XUV-pulses.

5.4 Conclusion and perspectives

In this chapter, we delved into the ultra-fast electronic responses of larger proteins, such as insulin^{+6} and the peptide substance PH^{+}, to XUV-pulse interactions using RT-TD-ADFT. Our journey commenced with an evaluation of the local basis set's applicability in RT-TD-ADFT simulations for capturing XUV interactions. Subsequently, we analyzed the conformational structure of insulin^{+6} and substance PH^{+} in the gas phase employing MD simulations. Our final exploration centered on the interaction of XUV-pulse and the primary ionization events, based on the results from our RT-TD-ADFT simulations.

During our assessment of the basis set, our conclusions indicated that, for smaller molecular systems, incorporating AOC and/or the use of ghost atoms substantially improves the continuum state description. For larger systems, a standard basis set effectively and qualitatively captures the XUV-pulse-induced ionization processes. Notably, for large systems, we showed that ECP predict XUV-pulse interactions both qualitatively and quantitatively in comparison with ref-AOC.

In our study of the structure of insulin^{+6} and substance PH^{+} in the gas phase, we saw that insulin^{+6} keeps its shape after 2 μs and its structure stays mostly the same. This made it easy for us to pick one structure for XUV simulations. On the other hand, substance PH^{+} had many different shapes and hydrogen bond patterns, making it hard to choose just one structure.

From our RT-TD-ADFT simulations, we observed that the residual energy after the interaction with the XUV-pulse doesn't necessarily increase with the size of the molecule. The character of ionization remains largely unaffected by the electric field strength, though there is a minor dependency on pulse orientation. The XUV-pulse induces ionization that is evenly distributed across all residues. Our findings suggest that the likelihood of a residue's ionization doesn't solely hinge on its ionization potential; instead, it is influenced by the number of valence electrons and the residue's positioning within the structure. A notable observation was that surface residues have a higher propensity for ionization compared to those situated deeper within the molecule. Finally, our results indicate that the side chains of peptides or proteins are more susceptible to ionization by XUV-pulses than the backbone. This offers insights into potential mechanisms of protein damage due to XUV-pulses.

This research opens numerous avenues for future exploration, and I would like to emphasize two pressing areas of focus.

Firstly, within the realm of ultra-fast simulations of XUV-pulses, there's a prevailing hesitance in the community towards using local basis sets. This stems from concerns about inadequate representation of the continuum state when the pulse is directly introduced into simulations, potentially compromising the ionization processes. We showed that, particularly for larger

systems, even standard basis sets can be effectively used in XUV-pulse simulations. To fortify the validity of our findings, we aim to enhance our tests in several ways. We plan to improve the quality of our reference basis set by refining parameters in AOC fitting. Furthermore, we'll extend our simulations in the xyz-directions for all basis sets, increase the variety of basis sets examined to identify optimal choices, and test smaller molecular systems. Additionally, by varying the pulse energy, we aim to expand the scope of our tests across a broader spectrum of XUV.

The second avenue for future research pertains to the susceptibility of surface residues to ionization by XUV-pulse in comparison to those residues buried within molecules. We determined that this characteristic hinges on the energy of the molecular orbitals in the surface residues, a phenomenon reminiscent of behaviors in nanoparticles. The propensity for surface residues to ionize could be manipulated by adjusting the energy of the XUV-pulse and by modifying the interactions at the protein surface.

Consequently, we aim to test our systems by varying the pulse energy for finding the effect of XUV energy. To delve into the effects of surface interactions within proteins, we propose the use of QM/MM techniques. By immersing our systems in a solvated environment, we can discern the impacts of these interactions. Addressing these questions holds paramount importance within the ultra-fast reaction community. This knowledge will significantly deepen our comprehension of how XUV-pulses interact, especially with large molecular systems. Additionally, it will enhance the accuracy of interpretations made from ultra-fast events in biomolecules observed in gas phase experiments when considering their native biological contexts.

REFERENCES:

(1) *Frontiers in Protein Structure, Function, and Dynamics*, Corrected publication.; Singh, D. B., Tripathi, T., Eds.; Springer: Singapore, 2020.

(2) Reisz, J. A.; Bansal, N.; Qian, J.; Zhao, W.; Furdui, C. M. Effects of Ionizing Radiation on Biological Molecules—Mechanisms of Damage and Emerging Methods of Detection. *Antioxidants & Redox Signaling* **2014**, *21* (2), 260–292. https://doi.org/10.1089/ars.2013.5489.

(3) Minkoff, B. B.; Bruckbauer, S. T.; Sabat, G.; Cox, M. M.; Sussman, M. R. Covalent Modification of Amino Acids and Peptides Induced by Ionizing Radiation from an Electron Beam Linear Accelerator Used in Radiotherapy. *Radiation Research* **2019**, *191* (5), 447. https://doi.org/10.1667/RR15288.1.

(4) Platzman, RL. The Physical and Chemical Basis of Mechanisms in Radiation Biology. *Radiation Biology and Medicine. Selected Reviews in the Life Sciences* **1958**, 15–72.

(5) Hunt, C. R.; Ramnarain, D.; Horikoshi, N.; Iyengar, P.; Pandita, R. K.; Shay, J. W.; Pandita, T. K. Histone Modifications and DNA Double-Strand Break Repair after Exposure to Ionizing Radiations. *Radiation Research* **2013**, *179* (4), 383–392. https://doi.org/10.1667/RR3308.2.

(6) Kumar, R.; Horikoshi, N.; Singh, M.; Gupta, A.; Misra, H. S.; Albuquerque, K.; Hunt, C. R.; Pandita, T. K. Chromatin Modifications and the DNA Damage Response to Ionizing Radiation. *Front. Oncol.* **2013**, *2*. https://doi.org/10.3389/fonc.2012.00214.

(7) Kraus, P. M.; Mignolet, B.; Baykusheva, D.; Rupenyan, A.; Horný, L.; Penka, E. F.; Grassi, G.; Tolstikhin, O. I.; Schneider, J.; Jensen, F.; Madsen, L. B.; Bandrauk, A. D.; Remacle, F.; Wörner, H. J. Measurement and Laser Control of Attosecond Charge Migration in Ionized Iodoacetylene. *Science* **2015**, *350* (6262), 790–795. https://doi.org/10.1126/science.aab2160.

(8) Calegari, F.; Trabattoni, A.; Palacios, A.; Ayuso, D.; Castrovilli, M. C.; Greenwood, J. B.; Decleva, P.; Martín, F.; Nisoli, M. Charge Migration Induced by Attosecond Pulses in Bio-Relevant Molecules. *J. Phys. B: At. Mol. Opt. Phys.* **2016**, *49* (14), 142001. https://doi.org/10.1088/0953-4075/49/14/142001.

(9) Palacios, A.; Martín, F. The Quantum Chemistry of Attosecond Molecular Science. *WIREs Comput Mol Sci* **2020**, *10* (1). https://doi.org/10.1002/wcms.1430.

(10) Castrovilli, M. C.; Trabattoni, A.; Bolognesi, P.; O'Keeffe, P.; Avaldi, L.; Nisoli, M.; Calegari, F.; Cireasa, R. Ultrafast Hydrogen Migration in Photoionized Glycine. *J. Phys. Chem. Lett.* **2018**, *9* (20), 6012–6016. https://doi.org/10.1021/acs.jpclett.8b02089.

(11) Calegari, F.; Ayuso, D.; Trabattoni, A.; Belshaw, L.; De Camillis, S.; Anumula, S.; Frassetto, F.; Poletto, L.; Palacios, A.; Decleva, P.; Greenwood, J. B.; Martín, F.; Nisoli, M. Ultrafast Electron

Dynamics in Phenylalanine Initiated by Attosecond Pulses. *Science* **2014**, *346* (6207), 336–339. https://doi.org/10.1126/science.1254061.

(12) De Meyts, P.; Whittaker, J. Structural Biology of Insulin and IGF1 Receptors: Implications for Drug Design. *Nat Rev Drug Discov* **2002**, *1* (10), 769–783. https://doi.org/10.1038/nrd917.

(13) Schulingkamp, R. J.; Pagano, T. C.; Hung, D.; Raffa, R. B. Insulin Receptors and Insulin Action in the Brain: Review and Clinical Implications. *Neuroscience & Biobehavioral Reviews* **2000**, *24* (8), 855–872. https://doi.org/10.1016/S0149-7634(00)00040-3.

(14) Hua, Q.-X.; Gozani, S. N.; Chance, R. E.; Hoffmann, J. A.; Frank, B. H.; Weiss, M. A. Structure of a Protein in a Kinetic Trap. *Nat Struct Mol Biol* **1995**, *2* (2), 129–138. https://doi.org/10.1038/nsb0295-129.

(15) Tregear, G. W.; Niall, H. D.; Potts, J. T.; Leeman, S. E.; Chang, M. M. Synthesis of Substance P. *Nature New Biology* **1971**, *232* (29), 87–89. https://doi.org/10.1038/newbio232087a0.

(16) Harrison, S. Substance P. *The International Journal of Biochemistry & Cell Biology* **2001**, *33* (6), 555–576. https://doi.org/10.1016/S1357-2725(01)00031-0.

(17) Datar, P.; Srivastava, S.; Coutinho, E.; Govil, G. Substance P: Structure, Function, and Therapeutics. *CTMC* **2004**, *4* (1), 75–103. https://doi.org/10.2174/1568026043451636.

(18) O'Connor, T. M.; O'Connell, J.; O'Brien, D. I.; Goode, T.; Bredin, C. P.; Shanahan, F. The Role of Substance P in Inflammatory Disease. *J. Cell. Physiol.* **2004**, *201* (2), 167–180. https://doi.org/10.1002/jcp.20061.

(19) Hervé, M.; Boyer, A.; Brédy, R.; Allouche, A.-R.; Compagnon, I.; Lépine, F. On-the-Fly Investigation of XUV Excited Large Molecular Ions Using a High Harmonic Generation Light Source. *Sci Rep* **2022**, *12* (1), 13191. https://doi.org/10.1038/s41598-022-17416-4.

(20) Calegari, F.; Ayuso, D.; Trabattoni, A.; Belshaw, L.; De Camillis, S.; Frassetto, F.; Poletto, L.; Palacios, A.; Decleva, P.; Greenwood, J. B.; Martin, F.; Nisoli, M. Ultrafast Charge Dynamics in an Amino Acid Induced by Attosecond Pulses. *IEEE J. Select. Topics Quantum Electron.* **2015**, *21* (5), 1–12. https://doi.org/10.1109/JSTQE.2015.2419218.

(21) Labeye, M.; Zapata, F.; Coccia, E.; Véniard, V.; Toulouse, J.; Caillat, J.; Taïeb, R.; Luppi, E. Optimal Basis Set for Electron Dynamics in Strong Laser Fields: The Case of Molecular Ion H_2^+. *J. Chem. Theory Comput.* **2018**, *14* (11), 5846–5858. https://doi.org/10.1021/acs.jctc.8b00656.

(22) Coccia, E.; Mussard, B.; Labeye, M.; Caillat, J.; Taïeb, R.; Toulouse, J.; Luppi, E. Gaussian Continuum Basis Functions for Calculating High-Harmonic Generation Spectra. *Int. J. Quantum Chem.* **2016**, *116* (14), 1120–1131. https://doi.org/10.1002/qua.25146.

(23) Krause, P.; Sonk, J. A.; Schlegel, H. B. Strong Field Ionization Rates Simulated with Time-Dependent Configuration Interaction and an Absorbing Potential. *The Journal of Chemical Physics* **2014**, *140* (17), 174113. https://doi.org/10.1063/1.4874156.

(24) White, A. F.; Heide, C. J.; Saalfrank, P.; Head-Gordon, M.; Luppi, E. Computation of High-Harmonic Generation Spectra of the Hydrogen Molecule Using Time-Dependent Configuration-Interaction. *Molecular Physics* **2016**, *114* (7–8), 947–956. https://doi.org/10.1080/00268976.2015.1119900.

(25) Nestmann, B. M.; Peyerimhoff, S. D. Optimized Gaussian Basis Sets for Representation of Continuum Wavefunctions. *J. Phys. B: At. Mol. Opt. Phys.* **1990**, *23* (22), L773–L777. https://doi.org/10.1088/0953-4075/23/22/007.

(26) Faure, A.; Gorfinkiel, J. D.; Morgan, L. A.; Tennyson, J. GTOBAS: Fitting Continuum Functions with Gaussian-Type Orbitals. *Computer Physics Communications* **2002**, *144* (2), 224–241. https://doi.org/10.1016/S0010-4655(02)00141-8.

(27) Tennyson, J.; Brown, D. B.; Munro, J. J.; Rozum, I.; Varambhia, H. N.; Vinci, N. Quantemol-N: An Expert System for Performing Electron Molecule Collision Calculations Using the R-Matrix Method. *J. Phys.: Conf. Ser.* **2007**, *86*, 012001. https://doi.org/10.1088/1742-6596/86/1/012001.

(28) Laporta, V.; Celiberto, R.; Tennyson, J. Dissociative Electron Attachment and Electron-Impact Resonant Dissociation of Vibrationally Excited O 2 Molecules. *Phys. Rev. A* **2015**, *91* (1), 012701. https://doi.org/10.1103/PhysRevA.91.012701.

(29) Foglia, N. O.; Morzan, U. N.; Estrin, D. A.; Scherlis, D. A.; Gonzalez Lebrero, M. C. Role of Core Electrons in Quantum Dynamics Using TDDFT. *J. Chem. Theory Comput.* **2017**, *13* (1), 77–85. https://doi.org/10.1021/acs.jctc.6b00771.

(30) Burkatzki, M.; Filippi, C.; Dolg, M. Energy-Consistent Pseudopotentials for Quantum Monte Carlo Calculations. *The Journal of Chemical Physics* **2007**, *126* (23), 234105. https://doi.org/10.1063/1.2741534.

(31) Burkatzki, M.; Filippi, C.; Dolg, M. Energy-Consistent Small-Core Pseudopotentials for 3d-Transition Metals Adapted to Quantum Monte Carlo Calculations. *The Journal of Chemical Physics* **2008**, *129* (16), 164115. https://doi.org/10.1063/1.2987872.

(32) Calaminici, P.; Janetzko, F.; Köster, A. M.; Mejia-Olvera, R.; Zuniga-Gutierrez, B. Density Functional Theory Optimized Basis Sets for Gradient Corrected Functionals: 3d Transition Metal Systems. *The Journal of Chemical Physics* **2007**, *126* (4), 044108. https://doi.org/10.1063/1.2431643.

(33) Köster, A.; Geudtner, G.; Calaminici, P.; Casida, M.; Dominguez, V.; Flores-Moreno, R.; Gamboa, G.; Goursot, A.; Heine, T.; Ipatov, A.; others. DeMon2k, Version 6, The DeMon

Developers, Cinvestav, Mexico City (2018). *Google Scholar There is no corresponding record for this reference*.

(34) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865–3868. https://doi.org/10.1103/PhysRevLett.77.3865.

(35) Calaminici, P.; Flores–Moreno, R.; Köster, A. M. A Density Functional Study of Structures and Vibrations of Ta3O and Ta3O−. *Computing Letters* **2005**, *1* (4), 164–171. https://doi.org/10.1163/157404005776611420.

(36) Gómez Pueyo, A.; Marques, M. A. L.; Rubio, A.; Castro, A. Propagators for the Time-Dependent Kohn–Sham Equations: Multistep, Runge–Kutta, Exponential Runge–Kutta, and Commutator Free Magnus Methods. *J. Chem. Theory Comput.* **2018**, *14* (6), 3040–3052. https://doi.org/10.1021/acs.jctc.8b00197.

(37) Cheng, C.-L.; Evans, J. S.; Van Voorhis, T. Simulating Molecular Conductance Using Real-Time Density Functional Theory. *Phys. Rev. B* **2006**, *74* (15), 155112. https://doi.org/10.1103/PhysRevB.74.155112.

(38) Castro, A.; Marques, M. A. L.; Rubio, A. Propagators for the Time-Dependent Kohn–Sham Equations. *The Journal of Chemical Physics* **2004**, *121* (8), 3425–3433. https://doi.org/10.1063/1.1774980.

(39) Samson, J. A. R.; Angel, G. C. Single- and Double-Photoionization Cross Sections of Atomic Nitrogen from Threshold to 31 Å. *Phys. Rev. A* **1990**, *42* (3), 1307–1312. https://doi.org/10.1103/PhysRevA.42.1307.

(40) Samson, J. A. R.; Lyn, L.; Haddad, G. N.; Angel, G. C. RECENT PROGRESS ON THE MEASUREMENT OF ABSOLUTE ATOMIC PHOTOIONIZATION CROSS SECTIONS. *J. Phys. IV France* **1991**, *01* (C1), C1-99-C1-107. https://doi.org/10.1051/jp4:1991113.

(41) Woon, D. E.; Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. V. Core-Valence Basis Sets for Boron through Neon. *The Journal of Chemical Physics* **1995**, *103* (11), 4572–4585. https://doi.org/10.1063/1.470645.

(42) Becke, A. D. A Multicenter Numerical Integration Scheme for Polyatomic Molecules. *The Journal of Chemical Physics* **1988**, *88* (4), 2547–2553. https://doi.org/10.1063/1.454033.

(43) Hirshfeld, F. L. Bonded-Atom Fragments for Describing Molecular Charge Densities. *Theoret. Chim. Acta* **1977**, *44* (2), 129–138. https://doi.org/10.1007/BF00549096.

(44) De La Lande, A.; Clavaguéra, C.; Köster, A. On the Accuracy of Population Analyses Based on Fitted Densities#. *J Mol Model* **2017**, *23* (4), 99. https://doi.org/10.1007/s00894-017-3264-5.

(45) Igel-Mann, G.; Stoll, H.; Preuss, H. Pseudopotentials for Main Group Elements (IIIa through VIIa). *Molecular Physics* **1988**, *65* (6), 1321–1328. https://doi.org/10.1080/00268978800101811.

(46) Cole, R. B. Electrospray and MALDI Mass Spectrometry: Fundamentals, Instrumentation, Practicalities, and Biological Applications. **2011**.

(47) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *The Journal of Chemical Physics* **1983**, *79* (2), 926–935. https://doi.org/10.1063/1.445869.

(48) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99SB. *J. Chem. Theory Comput.* **2015**, *11* (8), 3696–3713. https://doi.org/10.1021/acs.jctc.5b00255.

(49) Phillips, J. C.; Hardy, D. J.; Maia, J. D. C.; Stone, J. E.; Ribeiro, J. V.; Bernardi, R. C.; Buch, R.; Fiorin, G.; Hénin, J.; Jiang, W.; McGreevy, R.; Melo, M. C. R.; Radak, B. K.; Skeel, R. D.; Singharoy, A.; Wang, Y.; Roux, B.; Aksimentiev, A.; Luthey-Schulten, Z.; Kalé, L. V.; Schulten, K.; Chipot, C.; Tajkhorshid, E. Scalable Molecular Dynamics on CPU and GPU Architectures with NAMD. *The Journal of Chemical Physics* **2020**, *153* (4), 044130. https://doi.org/10.1063/5.0014475.

(50) Case, D.; Ben-Shalom, I.; Brozell, S.; Cerutti, D.; Cheatham III, T.; Cruzeiro, V.; Darden, T.; Duke, R.; Ghoreishi, D.; Gilson, M.; others. AMBER 2018; 2018. *University of California, San Francisco* **2018**.

(51) Roe, D. R.; Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9* (7), 3084–3095. https://doi.org/10.1021/ct400341p.

(52) Ester, M.; Kriegel, H.-P.; Sander, J.; Xu, X.; others. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In *kdd*; 1996; Vol. 96, pp 226–231.

(53) Grimme, S. Density Functional Theory with London Dispersion Corrections. *WIREs Comput Mol Sci* **2011**, *1* (2), 211–228. https://doi.org/10.1002/wcms.30.

(54) Nymeyer, H. How Efficient Is Replica Exchange Molecular Dynamics? An Analytic Approach. *J. Chem. Theory Comput.* **2008**, *4* (4), 626–636. https://doi.org/10.1021/ct7003337.

(55) Delgado, J.; Lara-Astiaso, M.; González-Vázquez, J.; Decleva, P.; Palacios, A.; Martín, F. Molecular Fragmentation as a Way to Reveal Early Electron Dynamics Induced by Attosecond Pulses. *Faraday Discuss.* **2021**, *228*, 349–377. https://doi.org/10.1039/D0FD00121J.

(56) Fuks, J. I.; Maitra, N. T. Challenging Adiabatic Time-Dependent Density Functional Theory with a Hubbard Dimer: The Case of Time-Resolved Long-Range Charge Transfer. *Phys. Chem. Chem. Phys.* **2014**, *16* (28), 14504–14513. https://doi.org/10.1039/C4CP00118D.

(57) Becke, A. D. Becke's Three Parameter Hybrid Method Using the LYP Correlation Functional. *J. Chem. Phys* **1993**, *98* (492), 5648–5652.

(58) Plummer, E. W.; Gustafsson, T.; Gudat, W.; Eastman, D. E. Partial Photoionization Cross Sections of N 2 and CO Using Synchrotron Radiation. *Phys. Rev. A* **1977**, *15* (6), 2339–2355. https://doi.org/10.1103/PhysRevA.15.2339.

GENERAL CONCLUSION

In this thesis, we have investigated the ultra-fast (atto- femto-second) responses of biomolecules subjected to fast charged particles and femtosecond XUV-pulses. We have considered energy deposition and electronic relaxation under various irradiation conditions.

In Chapter One, we introduced the general background and the theoretical concepts used in this thesis, focusing on radiation damage caused by photons and charged particles, especially during the physical stage. This introductory chapter also delved into the computational RT-TD-ADFT framework, a central methodology used in this work.

In Chapter Two, we focused on the study of charged particle interactions, emphasizing the necessity of including relativity effect arising from the high speed of the charged particles in the Kohn-Sham potential. The findings revealed that energy transfer relies essentially on the transverse interaction, further influenced by relativistic effects, which expand it in terms of time and space. The insights underscored the importance of including a relativistic external potential in RT-TD-ADFT calculations above 1 MeV. Furthermore, we explored the effects of relativity on the spatial distribution of energy deposition in agreement with proton energies. The results showed that relativity significantly impacts the distance interaction of protons in both perpendicular and parallel directions of proton propagation.

In Chapter Three, we focused on the effect of hydrogen bonds on energy deposition in the guanine-cytosine nucleobase paired with two water molecules, the whole system being irradiated by protons. A detailed investigation of energy deposition over a broad energy range has led to the understanding that the effects of hydrogen bonds are more pronounced at lower energies, and depend significantly on the trajectories and proton impact on the hydrogen bonds. An unprecedented observation was the fluctuation of the hydrogen bond effect on energy deposition at 2 MeV.

In Chapter Four, we investigated the physical stage of α-irradiation of a protein/DNA complex. While assessing both polarizable and non-polarizable force fields and altering the size of the QM region, we deduced that electrostatic induction and the QM region's size had a minimal effect on energy deposition but had an influence on charge migrations, particularly, the QM region's size. We concluded that the best strategy for dealing with charge migration was to increase the size of the QM region. We also delved into charge migration, revealing complex ultra-fast charge migration and electron-hole recombinations in less than 10 fs, especially concerning the carboxylate group of the aspartate residue. Finally, we calculated correlation matrices, showcasing its prowess as an analytical tool in discerning charge migration correlations among the residues.

In Chapter Five, we examined the ultra-fast electronic responses of larger biomolecules, specifically insulin^{+6} and the peptide substance PH^{+}, in interactions with XUV-pulses. Initially, we assessed the suitability of using a local basis set in RT-TD-ADFT for these kinds of interactions. Our investigation then focused on the primary ionization events triggered by the XUV-pulse, based on the insights we gained from our RT-TD-ADFT simulations.

Our work in Chapter Five has yielded several key insights into the ultra-fast electronic responses of larger biomolecules to XUV-pulses. During our assessment of basis sets, we concluded that for larger systems, a standard basis set provides an effective and qualitative account of the ionization processes induced by the XUV-pulse. Furthermore, effective core potentials showed remarkable capability in providing both qualitative and quantitative descriptions of XUV-pulse interactions when compared to our reference basis set. From the RT-TD-ADFT simulations of insulin^{+6} and substance PH^{+}, we revealed that ionization, induced by the XUV-pulse, is delocalized across all residues, with minimal charge migration occurring between residues. A particularly noteworthy observation was that surface residues are more prone to ionization than those located inside the molecule. Lastly, our study indicates that the side chains of peptides and proteins are more susceptible to ionization than the backbones when exposed to XUV-pulses.

This thesis covers very different molecular systems of biological interest from nucleobases to peptides, proteins, and DNA/protein complexes, the work encompasses a variety of systems and types of ionization radiation, such as charged particles and XUV-pulses. Our focus extends to ultra-fast events, from energy deposition to charge migration, thereby posing numerous questions and directions for future research.

We found that incorporating the relativity of charged particles is essential in RT-TD-ADFT to accurately calculate energy deposition. This conclusion is drawn from the analysis of a single trajectory. To enhance the statistical significance of our findings, further investigations could involve computing energy deposition across various systems, conducting additional electron dynamics simulations, or extending the tests using the AOC basis set.

We uncovered a previously unnoticed effect of hydrogen bonds on energy deposition. Subsequent research endeavors should seek to strengthen these findings through expanded trajectory sampling or by simplifying the system under investigation.

We demonstrate the impact of polarizable force fields and the size of the quantum mechanical (QM) region on charge migration. To enhance the robustness of our results, further investigations could involve the incorporation of a complex absorbing potential (CAP) into the simulations.

The insights gained in Chapter Five suggest that even standard basis sets or in particularly ECP can be effective in XUV-pulse simulations for larger systems. To reinforce these findings, future work could involve refining the parameters in AOC fitting and extending simulations across various xyz-directions for all basis sets. Furthermore, it is crucial to examine the propensity for surface residues to ionize, as it could be modulated by varying the energy of the XUV-pulse and by altering interactions at the protein surface. Our proposed use of QM/MM techniques aims to

shed light on the impacts of surface interactions within proteins, particularly in solvated environments.

One of the key challenges in the application of RT-TD-DFT to be useful in radiotherapy lies in its limitations regarding trajectory sampling, energy range, and the lack of transportability of trajectories. These issues have significant implications for broader applications such as track structure simulations or range penetration in radiotherapy. However, potential solutions are beginning to emerge, including the incorporation of relativistic effects and the utilization of machine learning for predicting trajectories.

This research could be expanded by encompassing a wider variety of biological systems and exploring different timeframes of irradiation responses. For instance, it would be of high interest to study irradiation of lipids, crucial components of membranes, or enzymes containing active sites sensitive to irradiation. A goal is to integrate nuclear dynamics into simulations, thereby extending the research beyond the physical stage and into the realm of physical-chemical interactions.

We are planning to further expand our research domain, encompassing a wider variety of biological systems and exploring different timeframes of irradiation responses. For instance, we aim to study lipids, crucial components of membranes, and enzymes containing active sites sensitive to irradiation. Our goal is to integrate nuclear dynamics into our simulations, thereby extending our research beyond the physical stage and into the realm of physical-chemical interactions.

These future pursuits have the potential to significantly enhance our understanding of ultra-fast phenomena induced by ionization radiation in biomolecules, improve the practical applicability of RT-TD-DFT, and deepen our comprehension of how XUV-pulses interact with larger molecular systems within their native biological contexts.

université PARIS-SACLAY
ÉCOLE DOCTORALE
Sciences Chimiques: Molécules, Matériaux, Instrumentation et Biosystèmes (2MIB)

Titre : Simulations numériques des réponses ultra-rapides de molécules biologiques soumises à des ions légers rapides et à des impulsions UV extreme

Mots clés : RT-TD-DFT, biomolécules, rayonnement ionisant, étape physique, dépôt d'énergie, et migration de charge

Résumé : Dans cette thèse, nous étudions l'étape physique de l'irradiation ionisante de systèmes biomoléculaires, un processus se déroulant sur une échelle de temps allant des attosecondes à quelques femtosecondes. L'étape physique recouvre le dépôt d'énergie et les migration de charges, jetant les bases de tous les événements physico-chimiques subséquents. Nous utilisons la Théorie de la Fonctionnelle de la Densité Auxiliaire Dépendent du Temps en Temps Réel (RT-TD-ADFT) pour étudier le dépôt d'énergie et les réponses ultra-rapides des électrons des biomolécules exposés à des ions et à des impulsions XUV.

Le Chapitre 1 présente une vue succincte des mécanismes d'endommagement des biomolécules. Nous abordons les principaux modèles théoriques de dépôt d'énergie et les approches de simulations numériques premiers principes dédiés à leur étude. Ce Chapitre décrit les méthodes théoriques utilisées dans la suite de travail.

Nous explorons dans les Chapitres 2 et 3 l'impact des effets relativistes des particules chargées et des liaisons hydrogène sur le dépôt d'énergie pour un système modèle guanine-cytosine en lien hydrogène avec deux molécules d'eau. Nous montrons la nécessité d'incorporer les effets relativistes pour des protons de plus d'un mégaélectronvolt d'énergie cinétique. Nous montrons que l'effet relativiste exacerbe le dépôt d'énergie, via une modification du mécanisme d'interaction ion-électron. Nous constatons que le dépôt d'énergie diminue en présence de liaisons hydrogène. Cette influence diminue avec l'augmentation de l'énergie cinétique du proton jusqu'à 2 MeV, puis elle recommence à augmenter progressivement. Ces tendances n'avaient pas été observées précédemment dans la littérature TD-DFT.

Dans le chapitre 3 nous étudions l'étape physique de l'irradiation α d'un complexe protéine/ADN. Nous adoptons une approche hybride QM/MM laquelle couple la RT-TD-ADFT et la mécanique moléculaire (MM) en champ de forces non- polarisable et polarisable. Nous montrons que l'induction électrostatique dans la partie MM a un effet négligeable sur le dépôt d'énergie mais est de nature à influencer les migrations de charges entre fragments moléculaires. Nous identifions des migrations ultra-rapides (fs) de trous dans le système protéine/ADN. Finalement, dans le but d'analyser les migrations de charges se déroulant au sein d'ensembles comportant plusieurs centaines d'atomes traités en DFT, nous proposons des outils fondés sur les matrices de corrélation. Il s'agit d'un outil puissant qui permet extraire des informations pertinentes des simulations RT-TD-ADFT sur des systèmes de grandes tailles.

Le chapitre 5 porte sur les réponses ultra-rapides de l'insuline et d'un peptide à une impulsion XUV dans le cadre d'une collaboration expérimentale. Nous identifions d'abord que la simulation RT-TD-ADFT de la photo-ionisation de petites molécules comme N_2 est difficile avec des fonctions de base gaussiennes standard. Cependant, elle peut être améliorée en les enrichissant d'orbitales atomiques optimisées pour le continuum. Toutefois, ces difficultés sont largement atténuées lors de la simulation de molécules plus grandes. Nous avons effectué des calculs pour l'éjection d'électrons et la création de trous sur une période de 50 fs pour les peptides. Les résultats montrent que la distribution de trous couvre toute la structure moléculaire, bien qu'il y ait des disparités marquées entre les différents acides aminés. La probabilité d'ionisation augmente en corrélation grossière avec le nombre d'électrons de valence. Au sein d'un acide aminé spécifique, l'environnement joue un rôle clé dans la détermination de sa probabilité d'ionisation. Une observation remarquable de notre étude est la susceptibilité accrue à l'ionisation des acides aminés présents sur la surface par rapport à ceux profondément intégrés dans la structure.

Le manuscrit se conclue par une synthèse générale des travaux et la proposition de quelques voies de recherche futures possibles.

université PARIS-SACLAY

ÉCOLE DOCTORALE
Sciences Chimiques: Molécules, Matériaux, Instrumentation et Biosystèmes (2MIB)

Title : Numerical simulation of ultrafast response of biological molecules subjected to light fast ions and XUV pulses

Keywords : RT-TD-DFT, biomolecules, ionization radiation, physical stage, energy deposition, and charge migration

Abstract : In this thesis, we study the physical stage of ionizing irradiation of biomolecular systems, a process taking place on a time scale ranging from attoseconds to a few femtoseconds. The physical stage covers energy deposition and charge migration, laying the foundations for all subsequent physico-chemical events. We use Real-Time Time-Dependent Auxiliary Density Functional Theory (RT-TD-ADFT) to study energy deposition and ultrafast electron responses of biomolecules exposed to ions and XUV pulses.

Chapter 1 presents a brief overview of biomolecule damage mechanisms. We discuss the main theoretical models of energy deposition and first-principles numerical simulation approaches dedicated to their study. This chapter describes the theoretical methods used in the rest of the work.

In Chapters 2 and 3, we explore the impact of relativistic effects of charged particles and hydrogen bonds on energy deposition for a model guanine-cytosine system in hydrogen bond with two water molecules. We show the need to incorporate relativistic effects for protons of more than one mega-electronvolt kinetic energy. We show that the relativistic effect exacerbates energy deposition, via a modification of the ion-electron interaction mechanism. We find that energy deposition decreases in the presence of hydrogen bonds. This influence diminishes with increasing proton kinetic energy up to 2 MeV, before gradually increasing again. These fluctuations had not previously been observed in the TD-DFT literature.

In chapter 3 we study the physical stage of a protein/DNA complex irradiated by an α-particle. We adopt a hybrid QM/MM approach that couples RT-TD-ADFT and molecular mechanics (MM) in non-polarizable and polarizable force fields. We show that electrostatic induction in the MM part has a negligible effect on energy deposition, but is likely to influence charge migrations between molecular fragments. We identify ultra-rapid (fs) hole migrations in the protein/DNA system, occurring at the fs scale. Finally, in order to analyze charge migrations occurring within systems comprising several hundred quantum atoms, we propose tools based on correlation matrices. This is a powerful tool for extracting relevant information from RT-TD-ADFT simulations on large-scale systems.

The final chapter focuses on the ultrafast response of insulin and a peptide to an XUV pulse in an experimental collaboration. We first identify that RT-TD-ADFT simulation of the photo-ionization of small molecules like N2 is difficult with standard Gaussian basis functions. However, it can be significantly improved by enriching them with continuum-optimized atomic orbitals. However, these difficulties are largely mitigated when simulating larger molecules. We have carried out calculations for electron ejection and hole creation over a period of 50 fs for peptides. The results show that the hole distribution covers the entire molecular structure, although there are marked disparities between the different amino acids. The probability of ionization increases in rough correlation with the number of valence electrons. Within a specific amino acid, the environment plays a key role in determining its ionization probability. A remarkable observation of our study is the increased susceptibility to ionization of amino acids present on the surface compared to those deeply embedded in the structure.

The manuscript concludes with a general summary of the work and a proposal for possible future research avenues.