Università degli Studi di Napoli *Federico II*

DOTTORATO DI RICERCA IN FISICA FONDAMENTALE ED APPLICATA

Ciclo XXV

Coordinatore: prof. Raffaele Velotta

Model system approach to the study of UV induced DNA-proteins crosslink reactions

Settori Scientifici Disciplinari FIS/03 – FIS/07

Dottorando
Marco Micciarelli

Tutori
Prof. Carlo Altucci
Dr. Ivano Tavernelli

Anni 2010/2014

Contents

Introduction

The goal of this work is the atomistic description of the photoreaction leading, in methanol solution, from 5-benzyluracil (5BU) to a cyclized compound (5,6BU) via computer simulation. This photoreaction is interesting because it will be shown that it can be used as effective model to study crosslink reactions of a DNA base with a close lying protein induced by UV light in a biological context. DNA-protein crosslink consists in the establishment of covalent bonds between the nucleic acid and a proteins upon contact. An efficient methodology to induce this kind of photochemical perturbation is of interest to one of the fundamental problems of modern molecular biology: understanding the mechanism behind the effects of the environment on gene expression regulation. It is known that a major role in this phenomenology is played by transient interactions between DNA and a given set of proteins, influencing the DNA transcriptional activity. The adoption of UV light irradiation in living cells can play a major role in this context, stabilizing these interactions and facilitating their study.

Recently, the rapid advances in time-resolved spectroscopic techniques and computer simulations techniques have provided fundamental insights about the complex photophysical behavior of DNA. In particular, on the computational side, Molecular Dynamics (MD) simulations are a powerful tool allowing to follow the temporal evolution of molecular systems via the numerical integration of the evolution equations. The analysis of the atomistic MD trajectories allows the derivation of macroscopic properties using statistical mechanics. Time Dependent Density Functional Theory (TDDFT) gives a model for the interactions among the nuclei that is accurate, numerically efficient and allows the description of chemical reactivity.

A microscopic description of the molecular mechanisms underlying UV induced crosslink, however, is a formidable task that is far from completion. The model system approach adopted in this work aims at improving the comprehension of such a fundamental process via a sequence of steps leading to a rational description of its spectroscopic and dynamical characteristics. In particular, TDDFT based computational techniques will be applied to characterize the excited states electronic structure properties of the 5BU and its photoproduct. Adiabatic and nonadiabatic Trajectory Surface Hopping (TSH) MD methods will be then applied to investigate the photoreaction mechanism in solution at a molecular level. The computational results will be discussed together with experimental data obtained via spectroscopic measurements allowing a complete characterization of the photoreaction model of crosslinking.

The work is structured as follows. In the first two chapters the theoretical and compu-

tational methods that will be applied in last three Chapters will be reviewed. In particular in the first chapter the density functional paradigm for the solution of the static and time dependent Schrödinger equations of a many electron system will be discussed. It will be also shown how the application of the linear response theory for quantum system to the TDDFT allows the characterization and treatment of the excited eigenstates of a many body electronic Hamiltonian. In the second Chapter the background concepts and derivations for the description of molecules (chromophores) in solution in their ground and excited states will be provided. In doing this, particular emphasis will be given to the adiabatic and nonadiabatic (Trajectory Surface Hopping) MD methods used in this work. In Chapter 3 the basic biological phenomenology regarding the structure, functions and photostability properties of DNA, will be shorty reviewed, highlighting the importance of DNA proteins interactions. In the same chapter, a study of the UV induction of DNA-protein crosslinking in living cells is phenomenologically approached on the basis of cell irradiation experiments. The experimental findings discussed will motivate and contextualize the studies reported in subsequent chapters. In particular, in Chapter 4 a detailed characterization of the excited states of 5-benzyluracil in the gas phase and in methanol solution as obtained applying TDDFT based calculations is reported. In Chapter 5 the photophysical and photochemical properties of 5BU and 5,6BU are investigated and characterized combining the spectroscopic measurement and the TDDFT calculations results. Finally, in Chapter 6, the mechanism of the photoreaction transforming 5BU to 5,6BU after UV irradiation will be discussed applying MD methods.

Chapter 1

Density Functional Theory approach to the electronic problem

In this Chapter, the Density Functional Theory approach will be discussed to solve the static and, with some approximations, the dynamic Schrödinger equation problem for a system of N interacting electrons in presence of an external field. The concepts and methods discussed are preliminary to the modeling of the static and dynamic properties of a general molecular system discussed in next chapter.

1.1 The many-body electronic problem

The heart of every *ab initio* method is the electronic structure calculation, i. e. the solution of the static Schrödinger equation for a system of N nonrelativistic electrons,

$$\hat{H}\psi_k(\mathbf{x_1},.....,\mathbf{x_N}) = E_k\psi_k(\mathbf{x_1},.....,\mathbf{x_N}) \tag{1.1}$$

Where $\psi_k(\mathbf{x_1},.....,\mathbf{x_N})$ is the antisymmetric N-electrons wave function k^{th} eigenstate of the Hamiltonian $\hat{H}$ with eigenvalue E_k. The shorthand notation $\mathbf{x_j} = (\mathbf{r_j}, \sigma_j)$ for the spatial and spin electronic degrees of freedom is used here and in the following. The Hamiltonian operator is given by

$$\hat{H} = \hat{T} + \hat{W} + \hat{V}_0 \tag{1.2}$$

where

$$\hat{T} = -\sum_j \frac{1}{2}\nabla_j^2 \tag{1.3}$$

is the kinetic-energy operator in atomic units,

$$\hat{V}_0 = \sum_j v_0(\mathbf{r}_j) \tag{1.4}$$

is the external potential operator, that can for example be, in the case of a molecular system calculation, the Coulomb interaction between electrons and nuclei treated as fixed point charges (in this case the dependence on the nuclear positions is just parametric) or a coupling potential coming from some external field. Finally

$$\hat{W} = \sum_{i<j} w(|\mathbf{r}_i - \mathbf{r}_j|) \tag{1.5}$$

is the general form of the electron-electron interaction operator. Commonly this is the Coulomb interaction, for the which $w = w_0 \equiv \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}$ but also the case in which $w = 0$ will be of crucial importance. This case will be referred in the following as the non interacting electrons problem.

1.2 The Hohenberg-Kohn theorem

Once the ground state eigenfunction of the many-body Hamiltonian (1.2) is determined in some way, the associated value of any physical observable $\hat{O}$ is obtained from the expectation values $O = \langle \psi_0 | \hat{O} | \psi_0 \rangle$. In other words the wave function of a system contains all the possible information about its quantum state. This concept was concisely stated by Dirac in 1929 [3]:

> "The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completed now, and the difficulty lies only in the fact that application of these laws leads to equations that are much too complicated to be solvable".

Over the years a huge variety of methods has been developed to calculate approximate solutions of the Eq. (1.1) based on the variational principle and different approximations made on the form of the wave function. The density functional theory (DFT) is the result of a completely different approach to the solution of the many body electronic problem. Its basilar point is that the many-body wave function contains definitely too much information, or at least much more than what really needed to calculate any physical observable. The physical object used to describe the electronic quantum ground state instead of the wave function (a function of 3N variables) is the single-particle probability density of the electronic ground state (a function of just 3 variables)

$$n_0(\mathbf{r}) = N \sum_{\sigma_1} \int d\mathbf{x}_2.... \int d\mathbf{x}_N \left| \psi_0(\mathbf{r}, \sigma_1, \mathbf{x}_2....\mathbf{x}_N) \right|^2 \tag{1.6}$$

where the compact notation for the electrons spatial and spin degrees of freedom $\mathbf{x} \equiv (\mathbf{r}, \sigma)$ has been used and for the integrals the equivalence $\int d\mathbf{x} \equiv \sum_{\sigma} \int d\mathbf{r}$ is intended. It is useful at this point to define the N-electron density operator:

$$\hat{n}(\mathbf{r}) = \sum_{i=1}^{N} \delta(\mathbf{r} - \hat{\mathbf{r}}_i) \tag{1.7}$$

using this operator, the single-particle probability density of the electronic ground state is rewritten as

$$n_0(\mathbf{r}) = \langle \psi_0 | \hat{n} | \psi_0 \rangle \tag{1.8}$$

and the external potential operator defined in Eq. (1.4) can be rewritten as

$$\hat{V}_0 = \int d\mathbf{r} \; v_0(\mathbf{r}) \; \hat{n}(\mathbf{r}) \tag{1.9}$$

This equation establishes a general formal map between single-particle external potential operators and external potentials $v_0(\mathbf{r})$ (simple functions of the spatial degrees of freedom) which will be largely used in the following.

In 1964 Hohenberg and Kohn [4], derived the fundamental existence proof of DFT that is now known as Hohenberg-Kohn theorem. The content of this theorem can be described
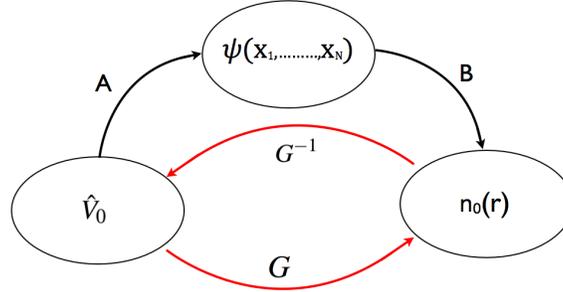
Figure 1.1: A graphical representation of the Hohenberg-Kohn theorem: the map A, morally realized by the Schrödinger equation solution, connects the set of the external potential to the one of the ground state wave function associated to a quantum mechanical system of N electrons. The map B, which is realized by Eq. (1.8), connects the set of ground state wave functions to the one of the electronic densities. The content of the theorem is the existence of the map G^{-1}.

by means of the scheme reported in Figure 1.1: to every given external potential $v_0(\mathbf{r})$ is associated an electron density $n_0(\mathbf{r})$ via the map $G = B \circ A$. The density in fact comes from the ground state wave function applying Eq. (1.8) (the map B in fig. (1.1)). The wave function obeys the Schrödinger equation and is hence associated to the system Hamiltonian which in turn is uniquely determined once the external potential operator $\hat{V}_0$ is given (the map A in the figure); this implies that the ground state density can be written as a functional of the external potential, i.e. $n_0 = n_0[v_0](\mathbf{r})$. Hohenberg and Kohn have shown (by means of a surprisingly very simple proof) that also the converse is true, i.e. $v_0 = v_0[n_0](\mathbf{r})$. Using the language of Figure 1.1 the theorem could be stated as follows: The map G is invertible.

The consequence of this theorem is astonishing powerful: every ground state quantum mechanical observable is completely determined by the ground state density (i.e. is a density functional) following the scheme:

$$n_0 \qquad (G^{-1}) \to \quad v_0[n_0] \ \ (A) \ \to \qquad \psi_0[n_0]$$

$$O_0 = \langle \psi_0[n_0] | \hat{O} | \psi_0[n_0] \rangle = O_0[n_0] \tag{1.10}$$

Among all functional the energy plays a major role. The energy functional is formally expressed as:

$$E_{v_0}[n] = \langle \psi_0[n] | \ \hat{T} + \hat{V}_0 + \hat{W} | \psi_0[n] \rangle \tag{1.11}$$

There's however a subtle and not yet completely solved issue that has to be considered regarding this functional: by construction the energy functional $E_{v_0}[n]$ is only well defined for those functions $n(\mathbf{r})$ to the which is associated, via the G^{-1} map defined in the figure (1.1), an external potential. Such functions are called *v-representable*. The possibility that the total energy functional can be evaluated also on non *v-representable* functions would affect the existence of the functional and of its derivatives (essential in the following of the chapter). This issue is known as the *v-representability* problem and it has been so far only partially solved in 1985 by Chayes *et al.* [5]: they have been able to show that at least every density function, i.e. a positively defined function satisfying the normalization condition (1.14), which is defined on a lattice is associated with the ground state of some external potential. To this day, however, a mathematically rigorous description of the set of *v-representabile* densities does not exist. It is just assumed that this set is sufficiently dense to be identifiable, for all numerical calculations purposes, with the set of all "reasonable" and physically possible densities[1].

The original Hohenberg and Kohn work of 1964 contains two important corollaries related to the energy density functional: as a consequence of the Rayleigh-Ritz principle, Eq. (1.11) has in fact the variational properties:

$$
\begin{aligned}
E_{v_0}[n] > E_0 \qquad & for \;\; n(\mathbf{r}) \neq n_0(\mathbf{r}) \\
E_{v_0}[n] = E_0 \qquad & for \;\; n(\mathbf{r}) = n_0(\mathbf{r})
\end{aligned}
\tag{1.12}
$$

These lead to the Euler equation:

$$
\frac{\delta}{\delta n(\mathbf{r})} \left[E_{v_0}[n] - \mu \int n(\mathbf{r}')d^3\mathbf{r}' \right] = 0
\tag{1.13}
$$

where μ is the Lagrange multiplier ensuring the correct normalization constraint:

$$
\int d^3\mathbf{r} \; n(\mathbf{r}) = N
\tag{1.14}
$$

The energy functional can be rewritten as

$$
E_{v_0}[n] = F[n] + \int d^3\mathbf{r} \; n(\mathbf{r})v_0(\mathbf{r})
\tag{1.15}
$$

where,

$$
F[n] = \langle \psi_0[n]| \; \hat{T} + \hat{W} |\psi_0[n]\rangle = T[n] + W[n]
\tag{1.16}
$$

does not depend anymore on the external potential, and is a common term to any electronic system. The Equation (1.13) now becomes

$$
\frac{\delta F[n]}{\delta n(\mathbf{r})} + v_0(\mathbf{r}) = \mu
\tag{1.17}
$$

---

[1]For a review of this and other subtle issues like the problem of the differentiability of the energy functional see ref [10].

This is the equation that replaces the Schrödinger equation and represents a complete paradigmatic change on the electronic many-body problem in which the wave function is now replaced by the electronic density. However such a dramatic simplification of the problem of course comes at a price: the form of the universal functional $F[n]$ is unknown and has to be approximated.

1.3 The Kohn-Sham equations

The most difficult task, for any direct attempt to find approximate forms of the universal energy functional $F[n]$, is to take into account that the ground state density has to correspond to an antisymmetric many body wave function. In all the modern wave function based methods for the solution of the many-body electronic problem, the many-body wave function is approximated with one (as in the case of the Hartree-Fock approximation) or more Slater determinants of one particle orbitals, so that the correct symmetry behavior for exchange of particle is ensured. Finding such a direct parametrization for the electronic density is still an open problem.

In 1965 however Kohn and Sham found an elegant way to carry in the DFT the power coming from the use of Slater determinants in a sort of indirect way. This idea has made the DFT, still at the present stage, one of the most widely used method for the calculation of properties of complex electronic systems. The first step is to consider the non interacting electrons Hamiltonian

$$\hat{H}_s = \hat{T} + \hat{V}_s \tag{1.18}$$

in which the electron electron interaction operator $\hat{W}$ is set equal to zero and the operator $\hat{V}_s$ is an external potential (the label s is inserted to avoid any confusion with the external potential operator $\hat{V}_0(\mathbf{r})$ which is associated with the interacting Hamiltonian). In this very simple case the many-body problem is solved imposing that the wave function is a Slater determinant of single particle orbitals $\phi_i(\mathbf{r})$ i.e.

$$\psi_s(\mathbf{r}_1, ..., \mathbf{r}_N) = \frac{1}{\sqrt{N!}} \, det \begin{pmatrix} \phi_1(\mathbf{r}_1) & \phi_2(\mathbf{r}_1) & ... & \phi_N(\mathbf{r}_1) \\ \phi_1(\mathbf{r}_2) & \phi_2(\mathbf{r}_2) & ... & \phi_N(\mathbf{r}_2) \\ \vdots & \vdots & & \vdots \\ \phi_1(\mathbf{r}_N) & \phi_2(\mathbf{r}_N) & ... & \phi_N(\mathbf{r}_N) \end{pmatrix} \tag{1.19}$$

Where the orbitals are solutions of the one-particle Schrödinger equations[2]:

$$\left(-\frac{\nabla^2}{2} + v_s(\mathbf{r}) \right) \phi_i(\mathbf{r}) = \epsilon_i \, \phi_i(\mathbf{r}) \qquad i = 1, ..., N \tag{1.20}$$

So, not only the ground state wave function and energy, but the whole spectrum of this non interacting Hamiltonian is easily obtained considering the set of Slater determinants that can be constructed with the whole set of single particle orbitals. The orbitals are ordered

---

[2]For closed shell N-electrons systems the non interacting problem can be effectively reduced to a problem of N/2 particles with no spin degrees of freedom, each one sitting in a single particle orbital. Being this the case for the problems treated in this thesis the spin dependence can be neglected from now on.

according to the ascending order of the single particle energies ϵ_i, and are commonly divided in two classes: the first N are called occupied orbitals, while the others are called unoccupied. The non interacting ground state wave function is simply given by the Slater determinant constructed with the occupied orbitals. The non interacting excited states wave functions can be divided into classes of single excitations, double excitations and so on: the single excitations are the determinants obtained from the ground state with the substitution of an occupied orbital with an unoccupied one, the double excitations are the ones obtained with the substitution of two occupied orbitals with two unoccupied ones and so on. The total energy relative to each wave function is simply the sum of the single particle eigenvalues relative to the orbitals present in the determinant and the electronic density is trivially:

$$n_s(\mathbf{r}) = \sum_{i=1}^{N} |\phi_i(\mathbf{r})|^2 \tag{1.21}$$

The total energy functional (which has a minimum for $n = n_{s0}$, the density relative to the ground state wavefunction) is now

$$E_{v_s}[n_s] = T_s[n_s] + \int d^3\mathbf{r} \; v_s(\mathbf{r})n_s(\mathbf{r}) \tag{1.22}$$

where

$$T_s[n_s] \equiv \langle \psi_s[n_s]| \hat{T} |\psi_s[n_s]\rangle = -\sum_i \langle \phi_i[n_s]| \frac{\nabla^2}{2} |\phi_i[n_s]\rangle \tag{1.23}$$

The Hohenberg-Kohn theorem applies also in this case, and hence also the single particle orbitals can can be considered as functionals of the density as indicated above. The key insight from Kohn and Sham has been to use the Hohenberg-Kohn theorem for the non interacting Hamiltonian to define an external potential $v_s[n_s]$ such that its associated electronic density $n_s(\mathbf{r})$ is the same of the one associated with the full many-body Hamiltonian (i.e. $n_s(\mathbf{r}) = n(\mathbf{r}) \; \forall \; \mathbf{r}$). The logical scheme to be followed to see how this idea is possible is depicted in Figure (1.2): If the electronic density $n(\mathbf{r})$ associated via the application $G = B \circ A$ to the many-body Hamiltonian $\hat{H}$ and the external potential $V(\mathbf{r})$ is *v-representable* for the non interacting Hamiltonian H_s then, the application G_s^{-1} ensures the existence of such potential. So, what is missed to this scheme to represent a complete existence theorem is "just" the *v-representability* problem. Again, while the existence of the Kohn-Sham potential is ensured for every density defined on a lattice (by the proof given by Chayes *et al.* [5]), the condition of the non interacting *v-representability* of the density $n(\mathbf{r})$ has to be in general assumed.

The potential v_s can be conveniently written as

$$v_s[n](\mathbf{r}) = v_0(\mathbf{r}) + v_H[n](\mathbf{r}) + v_{xc}[n](\mathbf{r}) \tag{1.24}$$

where the term

$$v_H[n](\mathbf{r}) = \int d^3\mathbf{r} \; \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \tag{1.25}$$

is the so called Hartree potential and takes into account the electrostatic interaction between the electrons in a mean field way. The term $v_{xc}[n](\mathbf{r})$ is unknown and has to be
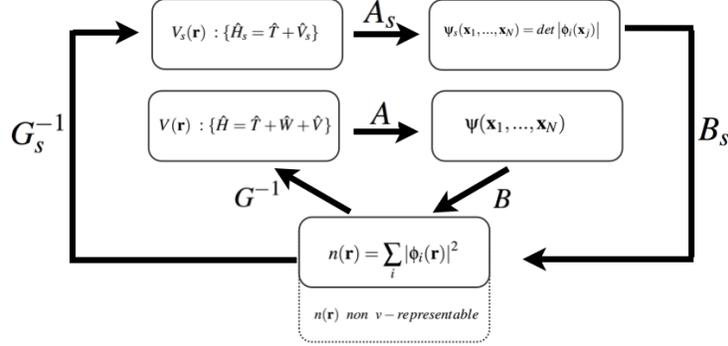
Figure 1.2: Graphical representation of the Kohn-Sham scheme: the Hohenberg-Kohn theorem has to be applied to both the interacting and non interacting case and hence the scheme contains the same sets of the one in fig.(1) but for both the interacting and non interacting quantities. The basic assumption is that the many body ground state electronic density (v-representable by definition for the interacting Hamiltonian) is in the set of v-representable densities for the non interacting Hamiltonian. This subset of non v-representable functions is represented in the figure with the dashed line bordered subset of densities. If the assumption of v-representability for the density is assumed, applying the map G_s^{-1} the existence of the external potential operator $\hat{V}_s$ in the non interacting Hamiltonian reproducing that density is proved.

approximated. It is called exchange correlation (xc) potential because it takes into account the Hartree-Fock exchange and correlation energies. These definitions directly lead to the famous Kohn-Sham equations

$$
\left(-\frac{\nabla^2}{2} + v_0(\mathbf{r}) + v_H[n](\mathbf{r}) + v_{xc}[n](\mathbf{r})\right)\phi_i^{(KS)}(\mathbf{r}) = \epsilon_i \; \phi_i^{(KS)}(\mathbf{r})
$$

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

(1.26)

which can be numerically solved with a self consistent procedure.

The variational principle gives a fundamental additional property of the exchange correlation functional that is essential to find suitable approximations of it. To see this, the total energy functional for the interacting Hamiltonian defined in Eq. (1.15) and (1.16) has to be written in the following way:

$$
E_{v_0}[n] = T_s[n] + \int d^3\mathbf{r} \; v_s[n](\mathbf{r}) \; \hat{n}(\mathbf{r}) =
$$

$$
= T_s[n] + E_H[n] + E_{xc}[n] + \int n(\mathbf{r})v_0(\mathbf{r})d^3\mathbf{r}
$$

(1.27)

where

$$E_H[n] = \frac{1}{2} \int d^3\mathbf{r} \int d^3\mathbf{r}' \ \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \tag{1.28}$$

is the Hartree energy term defined such that

$$\frac{\delta E_H[n]}{\delta n(\mathbf{r})} = v_H(\mathbf{r}) \tag{1.29}$$

The missing term, $E_{xc}[n]$[3], called exchange correlation energy, can be formally derived from the form of the total energy functional in Eq.(1.15) and (1.16)

$$E_{xc}[n] = T[n] - T_s[n] + W[n] - E_H[n] = F[n] - T_s[n] - E_H[n] \tag{1.30}$$

Applying now the variational principle to the energy functional

$$\left. \frac{\delta E_{v_0}[n]}{\delta n(\mathbf{r})} \right|_{n_0} = 0 = \left. \frac{\delta T_s[n]}{\delta n(\mathbf{r})} \right|_{n_0} + v_0(\mathbf{r}) + v_H[n_0] + \left. \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})} \right|_{n_0}$$

$$\left. \frac{\delta E_{v_0}[n]}{\delta n(\mathbf{r})} \right|_{n_0} = 0 = \left. \frac{\delta T_s[n]}{\delta n(\mathbf{r})} \right|_{n_0} + v_s[n](\mathbf{r}) \tag{1.31}$$

and plugging the second relation in the first one yields:

$$v_s[n](\mathbf{r}) = v_0(\mathbf{r}) + v_H[n](\mathbf{r}) + \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})} \tag{1.32}$$

comparing now this relation with the definition of v_s in Eq. (1.24) the useful form for the exchange correlation potential claimed above is found:

$$v_{xc}[n] = \frac{\delta E_{xc}[n]}{\delta n(\mathbf{r})} \tag{1.33}$$

The importance of this relation is crucial: on one side it allows direct numerical implementation of the theory in a computer giving a way to approximate the exchange correlation potential (see next section) and on the other side it makes clear again that the condition for the Kohn-Sham potential to exist is equivalent to the existence of the functional derivative of the Hohenberg-Kohn energy functional for both the interacting and non interacting Hamiltonians.

1.4 Approximating the xc energy functionals: LDA and GGA

At the present stage, in the derivation of the density functional theory, the only ingredient that still is missing to attempt any numerical solution of the many body electronic problem within the Kohn-Sham scheme is a suitable approximation to the unknown exchange

---

[3]The subscript xc will be justified in the following by the fact that the functional derivative of this energy with respect to the density is the exchange correlation potential

correlation potential or equivalently, thanks to Eq. (1.33), the exchange correlation energy. The oldest approximation of $E_{xc}[n]$, is the local density approximation (LDA), proposed by Kohn and Sham in 1965 [16], that is also the starting point of almost all of its modern approximations. The LDA comes from the study of a very important model system (the equivalent for the DFT of the hydrogen atom in atomic physics): the homogeneous electron liquid. This consists of an infinitely extended system, with an infinite number of electrons forming a constant electron density n, that is hence the unique degree of freedom of the system. The negative electronic charge is neutralized by a rigid uniform positive background. For this system the xc energy density $e_{xc}^h(n)$, the density being a number, is just a function of n and the exchange correlation energy density, even if not exactly known in analytic form, can be numerically computed. In fact, very accurate numerical results exist thanks to quantum Monte Carlo calculations [19], [20] for the computation of $e_{xc}^h(n)$ for given values of n and high precision analytical parametrizations fitting the numerical results have been developed [21], [22]. For an exhaustive review on this system and its properties see, for example, the first chapter of ref.[2].

In the LDA the xc energy of a given inhomogeneous system, like an atom, a molecule or a crystal, is expressed as the integral of the homogeneous electron liquid xc energy density evaluated, point by point, at the local value of the system density. In formulae:

$$E_{xc}[n] \simeq E_{xc}^{LDA}[n] \equiv \int d^3\mathbf{r} \; e_{xc}^h[\bar{n}]\big|_{\bar{n}=n(\mathbf{r})} \tag{1.34}$$

where $\bar{n}$ is the uniform density of the electron liquid. The associated exchange correlation potential is hence:

$$v_{xc}^{LDA}[n](\mathbf{r}) = \frac{d}{d\bar{n}} \, e_{xc}^h[\bar{n}]\big|_{\bar{n}=n(\mathbf{r})} \tag{1.35}$$

The LDA has represented an unexpected good level of approximation for the prediction of many structural and energetic properties of a wide range of materials, giving results almost always better than the Hartree-Fock approximation. The fundamental reason of this great success lies in the fact that it is based on a possible and well defined many-body reference model system and therefore satisfies a certain number of exact conditions [23]. Unfortunately however, the LDA, while capturing the right physical trends, is not accurate enough for many chemical applications, nor to predict the correct ground state of many materials. Moreover it predicts wrong dissociation limits of molecules and fractionally charged fragments after dissociation.

Improvements of the LDA can be thought to come from a gradient expansion approximation (known as GEA) of $E_{xc}[n]$ in terms of the density. In fact, if the case of an electron liquid with slowly varying density is considered, i.e. it is assumed that

$$n(\mathbf{r}) = \bar{n} + \delta n(\mathbf{r}) \tag{1.36}$$

where $\bar{n}$ is uniform and $\delta n(\mathbf{r})$ is a density modulation that is both weak and slowly varying in space, then a gradient correction of the xc energy of the kind

$$E_{xc}^{(1)}[n, \nabla n] = \int d^3\mathbf{r} \; B[n(\mathbf{r})](\mathbf{r}) \, |\nabla n(\mathbf{r})|^2 \tag{1.37}$$

can be derived. This idea has a long history, going back to the seminal work by Hohenberg and Kohn in 1964 [4] and over the years much work has been done for the derivation of an expansion series of the xc energy up to several orders. This strategy however has not led to a real improvement of the LDA. The principal reason for this lies in the fact that, even if the convergence of the expansion series is assumed for the case in which all the orders in the derivatives of the density are considered, such convergence can be weak and many terms have to be considered to achieve a real improvement of the LDA. Moreover if one keeps just the first few orders in such expansion, many of the exact conditions fulfilled by the LDA are lost.

A significative improvement over LDA has been achieved with the so called "Generalized gradient expansion" (GGA) class of functionals which still have the structure

$$E_{xc}^{GGA}[n] = \int d^3\mathbf{r}\ e_{xc}^{GGA}\ [n(\mathbf{r}), \nabla n(\mathbf{r})] \tag{1.38}$$

but for their construction the key idea is that an explicit functional expression for e_{xc}^{GGA} is given in such a way to satisfy as many exact properties as possible. In a way this procedure can be thought as coming from a truncation of a gradient expansion series which is a smart reordering of the weakly converging gradient expansion series discussed above. The price to pay for this idea is that there's not a unique prescription to get these functional forms. Sometimes some empirical parameters are used in such derivations, sometimes this is done without any empirical input. Today a whole zoo of hundreds of xc GGA functionals exists making even difficult the attempt of choosing one for a calculation, and sometimes the preference for one of them is just a matter of taste. One of the major criticisms to the (Kohn-Sham) DFT has been in fact that, even if it is in principle an exact theory, the approximations used can not be systematically improved as it happens, for example, in more traditional quantum chemical methods. Nevertheless this fact is also the main advantage of DFT: it allows clever shortcuts able to deliver unexpected degrees of accuracy which are computationally much more expensive in wave function based methods.

Many recent review exists in the literature in which the performances of the GGA functionals are presented and recommended for specific cases. See for example refs [25], [26], [27], [28].

1.5 Beyond the GGA: the Jacob's ladder of DFT

Despite the great success in calculations of electronic structures achieved by the LDA and a handful of the most popular GGA xc functional, much work has recently been done to go beyond the GGA and construct functionals able to reach the so called "chemical accuracy" (i.e. predictions accurate within an error below 1 kcal/mol). Perdew and Schmidt [24] proposed in 2001 a picture in which the different kind of functionals are arranged in a ladder of increasingly better approximations (and higher computational cost) that they called "Jacob's ladder of DFT": the ground level is the Hartree theory, the first two rungs

are represented by LDA and GGA and the final is the one corresponding to the exact xc potential (the heaven of chemical accuracy).

On the third rung of the ladder is located the meta-GGA class of functionals. They are constructed following the same strategy of the GGA class of functionals, so again an explicit functional expression for e_{xc}^{MGGA} is constructed with the aim to satisfy as many exact properties as possible but, to achieve more flexibility, also second order parameters are included: the laplacian of the density and non interacting kinetic energy density $\tau = \frac{1}{2}\sum_{i=1}^{N} |\nabla \phi_i|^2$ [4]

$$E_{xc}^{MGGA}[n] = \int d^3\mathbf{r} \; e_{xc}^{MGGA} \; [n(\mathbf{r}), \nabla n(\mathbf{r}), \nabla^2 n(\mathbf{r}), \tau] \tag{1.39}$$

It is important to stress that the kinetic energy density τ is only an implicit functional of the density and just an explicit functional of the Khon-Sham orbitals. The functional derivative of such xc energy with respect to the density is hence not trivial to do. So, even if the procedure of truncation of the weakly converging gradient expansion series of the xc energy in terms of the density is arbitrary and "justify" any term that is improving the performances of the approximated xc functional, this class of functionals is going outside the pure Khon-Sham theory limits and the total energy has to be considered stationary with respect to the orbitals variations.

The next level of approximation (the fourth rung) is achieved incorporating in the xc functional form also the Hartree-Fock exchange energy density, computed with the Kohn-Sham orbitals:
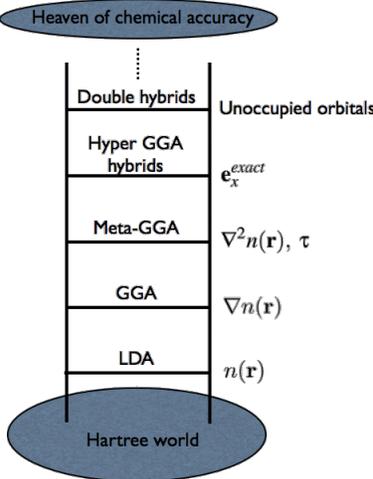

Figure 1.3: The ladder of approximation for the different sets of xc potential, known as "Jacob's ladder of DFT", extending from the Hartree world to the heaven of chemical accuracy.

---

[4]Integrating by parts, it is not difficult to see that $T_s = -\frac{1}{2}\sum_i \langle \psi_{s0}|\nabla_i^2|\psi_{s0}\rangle = \int d^3\mathbf{r} \; \tau(\mathbf{r})$

$$e_x^{hf}(\mathbf{r}) = -\frac{1}{2}\sum_{i,j=1}^{N}\int d^3\mathbf{r}'\ \frac{\phi_i^*(\mathbf{r}')\phi_j(\mathbf{r}')\phi_i(\mathbf{r})\phi_j^*(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|} \tag{1.40}$$

The class of approximated xc functionals which, in addition to the meta-GGA parameters, also depend on e_x^{hf} are called hyper-GGA. Like the meta-GGA these are only implicit functionals of the density and the issues in the calculation of the functional derivatives are the same as above. A widespread use in practical calculations has been made of a particularly successful subclass of hyper-GGA functionals, the so called hybrid functionals. They are just constructed mixing fractions of Hartree-Fock exchange energy with a standard LDA or GGA functional in the following way:

$$E_{xc}^{hybrid} = aE_x^{hf} + (1-a)E_x^{GGA} + E_c^{GGA} \tag{1.41}$$

where the parameter a is determined in a semiempirical way and typically has a value around 0.25.

A refinement of this procedure, comes from the class of functionals called range-separated hybrids: the basic idea is to separate the Coulomb interaction into a short-range (SR) and a long-range (LR) part. Then, standard DFT functionals (coming from the first three rungs of the ladder) are utilized to model the SR part while the Hartree-Fock exchange is utilized for the LR. The main strength of this procedure is that it allows to produce effective potentials with the correct long range behavior. This implies significant improvements in the description of long-chain molecules, bond dissociations and formations. Range separated hybrids also improve the excited states calculation results with TDDFT (see section 1.12) in the case of charge transfer transitions.

Range separated hybrids represent a solution for the description of the asymptotic range behavior of the xc potential preserving the possibilities offered by the error compensation of LDA based functionals. So, what remains then at this point (of the Jacob's ladder) is to increase the accuracy to have a better description of the short distances correlation energy. On the fifth rung of the ladder are located functionals which incorporate the usage of unoccupied orbitals which allows the mix of the canonical Kohn-Sham DFT with other quantum chemical methods. The description of these methods however exceeds the purposes of this chapter.

1.6 Time dependent formalism

The DFT, treated so far, constitutes a tool to find approximate numerical solutions to the many body electronic problem in the ground state. The time dependent version of the DFT (in the following TDDFT) extends its basic ideas to the time dependent regime and allows the treatment of excited states. Of course, by virtue of the Hohenberg-Kohn theorem, the excitation energies, and the whole set of excited wave functions of the many body Hamiltonian are functional of the ground state density[5]. The TDDFT will however be essential to work out a self consistent relation to address the many-body Hamiltonian

---

[5]Being the external potential a functional of the density, so is the Hamiltonian, and hence its whole spectrum

spectrum calculation in terms of just the ground state (time independent) density, i.e. the same physical quantity treated so far.

The many-body problem to be addressed now consists in the solution of the time dependent Schrödinger equation for a system of N nonrelativistic electrons:

$$i\frac{\partial}{\partial t}\psi(\mathbf{x_1}, ..., \mathbf{x_N}, t) = \hat{H}(t)\psi(\mathbf{x_1}, ..., \mathbf{x_N}, t) \tag{1.42}$$

The time dependent many-body Hamiltonian is

$$\hat{H} = \hat{T} + \hat{W} + \hat{V}(\mathbf{r}, t) \tag{1.43}$$

where the kinetic and electron-electron interaction operator are the same as in Eq. (1.3) and (1.5) while the external potential constitutes now the time dependent part of the Hamiltonian[6]:

$$\hat{V}(t) = \sum_{i=1}^{N} v(\mathbf{r}_i, t) \tag{1.44}$$

Finally, by means of the N-electron density operator defined in Eq. (1.7) (whose usage will be again useful in section 1.9), the time dependent single-particle density (the key quantity of TDDFT) can be written as

$$n(\mathbf{r}, t) = \langle \psi(t) | \hat{n}(\mathbf{r}) | \psi(t) \rangle \tag{1.45}$$

1.7   Basic Theorems

To construct a time dependent density functional theory the first essential step is represented by the extension of the Hohenberg-Kohn theorem to the time dependent case. Such extension is a very simple task for the maps A and B of figure (1.1) (in which the time dependence in the arguments of wave function and density has to be intended and $\hat{V}_0$ has to be substituted with $\hat{V}(t)$): the time dependent Schrödinger equation formally defines the first while the second is defined by Eq. (1.45). The non trivial part is to show the invertibility of the map $G = B \circ A$. This result has first been established by Runge and Gross in 1984 [6], and is nowadays known as the Runge-Gross theorem:

> Let $n(\mathbf{r}, t)$ and $n'(\mathbf{r}, t)$ be the electronic densities associated with the temporal evolution of a many-body electron system from a common initial state $\psi(t = t_0)$ and under the influence of the external potentials $v(\mathbf{r}, t)$ and $v'(\mathbf{r}, t)$, respectively. Assuming
>
> 1. $v'(\mathbf{r}, t) \neq v(\mathbf{r}, t) + C(t)$
> 2. $v(\mathbf{r}, t)$ and $v'(\mathbf{r}, t)$ to be Taylor expandable in t around t_0

---

[6]The dynamic problem relative to a time independent Hamiltonian is contained in this formalism as the special case in which the external potential is not time dependent

Then $n(\mathbf{r}, t)$ and $n'(\mathbf{r}, t)$ will start to be different infinitesimally after t_0. There is therefore a one to one map between densities and external potentials for any fixed initial many-body quantum state.

The first condition in the theorem takes into account the case in which the two external potential differ by just a time dependent scalar field $C(t)$: in this case in fact the two associated wave functions will be just scaled by a time dependent phase factor and will give exactly the same densities. The second condition instead is simply coming from the fact that in the theorem proof the Taylor expansion of the potentials has to be considered and there's not a direct physical motivation for it.

This is the fundamental existence theorem of TDDFT implying that all physical time dependent observables are functionals of the time dependent density, as it is immediately seen following a logical scheme identical to the one depicted in Eq. (1.10) for the time independent case.

To be practically useful, however, this scheme still needs the equivalent of the Kohn-Sham formalism for the time independent case, which allows the implementation in a computer code of the density functional paradigm. The question is if it is possible to reproduce the same electronic density $n(\mathbf{r}, t)$ of the many-body Hamiltonian (1.43) with different particle-particle interaction potentials $w'(|\mathbf{r} - \mathbf{r}'|)$ (possibly zero). The positive answer to this question is the content of the van Leeuwen theorem (1999) [7]:

Let $n(\mathbf{r}, t)$ be the time dependent density of a quantum system associated to a many-body system with particle-particle interaction potential $w(|\mathbf{r} - \mathbf{r}'|)$, external potential $v(\mathbf{r}, t)$ and an initial state ψ_0. There exists a unique quantum system with interaction potential $w'(|\mathbf{r} - \mathbf{r}'|)$ and external potential $v'(\mathbf{r}, t)$ (defined up to a time dependent scalar field $C(t)$) reproducing exactly the same time dependent density $n(\mathbf{r}, t)$ provided that

1. $v(\mathbf{r}, t)$ and $v'(\mathbf{r}, t)$ are Taylor expandable in t around t_0

2. $n(\mathbf{r}, t)$ is analytic in time at t_0

3. The initial state ψ_0' is chosen so as to reproduce the given density and its time derivative at the time t_0

This is a very powerful theorem. It in fact contains the Runge-Gross theorem for the special case in which $w = w'$ and $\psi_0' = \psi_0$. In this case in fact the theorem implies the uniqueness of the potential $v'(\mathbf{r}, t) (= v(\mathbf{r}, t))$ reproducing the given density.

The other particular case of strong interest is the one in which w' is chosen to be zero i.e. the second system is chosen to be the non interacting one. The theorem guarantees in this case the existence and uniqueness (up to the usual time dependent scalar field) of an external potential $v_s(\mathbf{r}, t)$ that, under the given conditions, reproduces the time dependent density of the interacting electron problem providing the formal justification for the Kohn-Sham approach also in the time dependent case. The condition 3. of the theorem is however not trivial to realize: the problem of finding a non interacting wave function reproducing the time dependent density and its time derivative at a given initial time is equivalent to the static problem discussed in the first five sections. If the density to reproduce however is

not the one of the ground state wave function, the advantages coming from the variational principle are lost and all the methods derived for its calculation become inservible.

Moreover, also the necessity of the condition 2. of the theorem is under debate since it has been recently discovered [14] the existence of physically reasonable densities coming from Taylor expandable potential that are not analytic in time. At present in fact the attempt to find alternative proofs of this theorem not relying on the Taylor expandability of potential and density is an active area of research[14]

1.8 Time-dependent Kohn-Sham equations

The general functional form of the effective external potential of a non interacting system able to reproduce the interacting time dependent density of a quantum system predicted by the van Leeuwen theorem is

$$v_s = v_s[\psi_0, \psi_{s0}, n](\mathbf{r}, t) \tag{1.46}$$

i.e. it has a functional dependence on the density but also on the initial many body wave function and on the initial non interacting wave function (that, according to the requirements of the van Leeuwen theorem, has to generate the initial correct density and density time derivative). The situation becomes immediately much simpler for the class of potentials of the kind

$$v(\mathbf{r}, t) = v_0(\mathbf{r}) + v_1(\mathbf{r}, t) \; \theta(t - t_0) \tag{1.47}$$

if the system is also assumed to be in its ground state at the initial time t_0 (and for all the previous times). In this case in fact the Hohenberg-Kohn theorem can be invoked for the initial state and hence both the wave function ψ_0 and ψ_{s0} are functional of the ground state density $n_0(\mathbf{r})$. The non interacting effective potential becomes hence just a density functional $v_s = v_s[n_0](\mathbf{r}, t)$. So, the initial non interacting wave function can be built as a Slater determinant made of the Kohn-Sham orbitals, solution of the self consistent equations (1.26) which also define the initial ground state density. Immediately after the instant t_0 the time dependent potential $v_1(\mathbf{r}, t)$ start to act and the density start to evolve under its influence. The time dependent density can be written as

$$n(\mathbf{r}, t) = \sum_{i=1}^{N} |\phi_i(\mathbf{r}, t)|^2 \tag{1.48}$$

where the single particle orbitals are the solutions of the (one particle) effective Schrödinger time dependent equation:

$$\left\{ -\frac{\nabla^2}{2} + v_s[n](\mathbf{r}, t) \right\} \phi_i(\mathbf{r}, t) = i \frac{\partial}{\partial t} \phi_i(\mathbf{r}, t) \qquad i = 1, ..., N \tag{1.49}$$

with the initial condition:

$$\phi_i(\mathbf{r}, t_0) = \phi_i^{(KS)}(\mathbf{r}) \tag{1.50}$$

As it can be noted, the unoccupied Kohn-Sham orbitals (as their time evolution) have absolutely no meaning. They'll come useful and will acquire meaning only within the linear response theory applied to the TDDFT discussed later. Finally, as for the time independent case, the effective external potential can conveniently be written as

$$v_s[n](\mathbf{r}, t) = v_0(\mathbf{r}) + v_1(\mathbf{r}, t) \; \theta(t - t_0) + \int d^3\mathbf{r}' \frac{n(\mathbf{r}', t)}{|\mathbf{r} - \mathbf{r}'|} + v_{xc}[n](\mathbf{r}, t) \qquad (1.51)$$

where the unknown time dependent exchange correlation potential $v_{xc}[n](\mathbf{r}, t)$ is formally defined by this expression.

The Kohn-Sham scheme for this particular time dependent case is just still missing the final ingredient to be complete: a way to give at least some approximate form for the time dependent xc potential. In the static DFT the key to find approximated forms for the xc potential comes from the variational principle but a similar procedure can not be applied in the time dependent regime. The vast knowledge about static xc potential can be however used also to find approximated time dependent xc potential if the so called adiabatic approximation is used. It practically consists in fact in approximating the time dependent exchange correlation potential with the static one, of which many approximated forms exist and can be directly plugged it in the time dependent Kohn-Sham equation. In formulae it is:

$$v_{xc}(\mathbf{r}, t) \simeq v_{xc}^A(\mathbf{r}, t) = v_{xc}^0[n_0](\mathbf{r})|_{n_0(\mathbf{r}) \to n(\mathbf{r}, t)} \simeq v_{xc}^{0, approx}[n_0](\mathbf{r})|_{n_0(\mathbf{r}) \to n(\mathbf{r}, t)} \qquad (1.52)$$

This approximation is not bad in many cases: it becomes in fact exact in the adiabatic limit i.e. when the electronic system can be considered in each instant in the ground state of the Hamiltonian at that time. In this case the effective Kohn-Sham potential at time t has only a functional dependence on the density at the same time t i.e. there's no memory effect. This happens in the limit in which the time dependent perturbation acting on it is infinitely slow.

The adiabatic approximation is however the combination of two approximation as emphasized in the Eq. (1.52): the first comes from the use of a functional coming from static DFT neglecting any memory effect and the second from the fact that the static xc functional is itself approximated. It is hence very difficult to do any prediction about how much accurate this procedure can be a priori. However, test calculations made on exactly solvable benchmark systems [13] indicate that in the cases in which the external perturbation is sufficiently weak and slowly varying in time the adiabatic and exact correlation potentials are very close to each other for all times and hence that the memory effect in the correlation potential are negligible in these processes.

1.9   Linear response theory

In this section the quantum linear response theory will be reviewed to be then applied to the framework of the TDDFT. As discussed at the beginning of section 1.6, the extension of the DFT formalism to the time dependent regime has been made to address the full many-body Hamiltonian spectrum calculation. In particular, the application of the linear

response theory to the TDDFT will be key to address this problem that will be faced in section 10.

1.9.1 The quantum mechanical linear response function

Let $\hat{H}_0$ be a time independent Hamiltonian of a quantum system. At time t_0 a time dependent external field coupled to some observable $\hat{B}$ is applied so that the associated time dependent Hamiltonian is

$$\hat{H}(t) = \hat{H}_0(\mathbf{r}) + \hat{H}_1(t) \tag{1.53}$$

where

$$\hat{H}_1(t) = \hat{B}F(t) \tag{1.54}$$

The response of a quantum mechanical observable $\hat{A}$ is defined as the difference between the time dependent expectation value of that observable and its static initial value. Assuming that initially the system is in its ground state, we have

$$A(t) - A_0 = \langle \psi(t)|\hat{A}|\psi(t)\rangle - \langle \psi_0|\hat{A}|\psi_0\rangle \tag{1.55}$$

where ψ_0 is the ground state many-body wave function associated with a static Hamiltonian H_0 and $\psi(t)$ is the time propagated wave function of the system under the effect of the time dependent external potential $\hat{H}_1(t)$.

As it will be shown in the immediate following, the response can be expanded in powers of the field $F(t)$:

$$A(t) - A_0 = A_1 + A_2 + A_3 + ... \tag{1.56}$$

where the term A_1 is the linear response of the observable $\hat{A}$ and is (by definition) linear in the field $F(t)$, the term A_2 is the quadratic (or second order) response and so on. One way to see how this works, is to write the evolution of the system under the time dependent Schrödinger equation in terms of the time evolution operator:

$$|\psi(t)\rangle = \hat{U}(t, t_0)|\psi_0\rangle \tag{1.57}$$

The form of the evolution operator can be explicitly given only in the trivial case where the Hamiltonian is time independent

$$\hat{U}_0 \equiv e^{-i\hat{H}_0(t-t_0)} \tag{1.58}$$

whereas for the case of the Hamiltonian in Eq. (1.53) it is convenient to define the operator:

$$\hat{U}_1 \equiv e^{i\hat{H}_0(t-t_0)}\hat{U} = \hat{U}_0^\dagger \, \hat{U} \tag{1.59}$$

which obeys to the equation:

$$i\partial_t\hat{U}_1 = i\partial_t \left\{ e^{i\hat{H}_0(t-t_0)}\hat{U} \right\} = -\hat{H}_0 \, e^{i\hat{H}_0(t-t_0)}\hat{U} + e^{i\hat{H}_0(t-t_0)} \left(\hat{H}_0 + \hat{H}_1 \right) \hat{U} =$$
$$= \hat{H}_{1H_0}(t-t_0) \, \hat{U}_1 \tag{1.60}$$

where in the last equality the following compact notation has been used:

$$\hat{O}_{H_0}(t - t_0) \equiv e^{i\hat{H}_0(t-t_0)} \hat{O} e^{-i\hat{H}_0(t-t_0)} = \hat{U}_0^\dagger \hat{O} \hat{U}_0 \qquad (1.61)$$

Eq. (1.60) can be solved iteratively: in fact, integrating both sides over time yields

$$\hat{U}_1(t) = \hat{U}_1(t_0) - i \int_{t_0}^t d\tau \ \hat{H}_{1H_0}(\tau - t_0) \ \hat{U}_1(\tau - t_0) =$$

$$= \hat{U}_1(t_0) - i \int_{t_0}^t d\tau \ \hat{H}_{1H_0}(\tau - t_0) \left[\hat{U}_1(t_0) - i \int_{t_0}^t d\tau' \ \hat{H}_{1H_0}(\tau' - t_0) \ \hat{U}_1(\tau' - t_0) \right] = \dots$$

$$(1.62)$$

This procedure generates an expansion series for $\hat{U}_1(t)$ in powers of the operator $\hat{H}_{1H_0}(t-t_0)$ (and hence of $F(t)$). The zero order solution is $\hat{U}_1(t_0) = \mathbb{1}$, the unit operator. The first order solution is obtained truncating the expansion series in Eq. (1.62) after the first integral:

$$\hat{U}_1(t, t_0) \approx \mathbb{1} - i \int_{t_0}^t dt' \ e^{i\hat{H}_0(t'-t_0)} \ \hat{H}_1(t') \ e^{-i\hat{H}_0(t'-t_0)} \qquad (1.63)$$

The second order solution is obtained truncating the expansion series in Eq. (1.62) after the second integral and so on. From Eq. (1.63) the first order approximation for $\hat{U}(t)$ is derived substituting back the form of U_1 in Eq. (1.59):

$$\hat{U}(t, t_0) \approx \hat{U}_0 \left\{ \mathbb{1} - i\hat{I} \right\} \qquad (1.64)$$

where

$$\hat{I} = \int_{t_0}^t dt' \ F(t')\hat{B}_{H_0}(t' - t_0) \qquad (1.65)$$

In the last equality the explicit form of Eq. (1.54) for H_1 has been substituted to show the linear dependence on the field. Using the form of Eq. (1.64) for the evolution operator and using the fact that the operator $\hat{I}$ is hermitian, we get:

$$A(t) = \langle\psi_0|(\mathbb{1} + i\hat{I})\hat{A}_{H_0}(t - t_0)(\mathbb{1} - i\hat{I})|\psi_0\rangle =$$
$$= A_0 - i\langle\psi_0| \ [\hat{A}_{H_0}(t - t_0), \hat{I}] \ |\psi_0\rangle + o(F(t))^2 \qquad (1.66)$$

So that the linear term of the response expansion can be easily derived:

$$A_1(t) = -i \int_{t_0}^t dt' \ F(t') \ \langle\psi_0|[\hat{A}_{H_0}(t), \hat{B}_{H_0}(t')]|\psi_0\rangle \qquad (1.67)$$

The higher order terms can be obtained considering the neglected quadratic term in Eq. (1.66) and the higher order terms of the expansion series of the evolution operator. Finally, since the expectation values are calculated on the wave function ψ_0 which is an eigenstate

of the unperturbed Hamiltonian $\hat{H}_0$, the commutator $[\hat{A}_{H_0}(t), \hat{B}_{H_0}(t')]$ can be replaced with $[\hat{A}_{H_0}(t-t'), \hat{B}]$. The retarded quantum mechanical linear response function can now be defined as:

$$\chi_{AB}(t-t') = -i\theta(t-t') \langle\psi_0|[\hat{A}_{H_0}(t-t'), \hat{B}]|\psi_0\rangle \tag{1.68}$$

The word "retarded" refers to the fact that the response at a time t is due to a perturbation at an earlier time $t' < t$ and the Heaviside step function ensures that this causality requirement is fulfilled. A very important feature of this function is that it depends only on the properties of the system in its ground state. The linear response can be written in terms of the response function as

$$A_1(t) = \int_{t_0}^{+\infty} dt' \; \chi_{AB}(t-t') \; F(t') \tag{1.69}$$

Being the field $F(t)$ considered to be zero for all times earlier than t_0, the lower extremum in integral in Eq. (1.69) could also be set to infinity so that $A_1(t)$ would take the convenient form of a convolution. For the following however a crucial case will be the one in which the external perturbation field is harmonic ($F(t) = \exp(i\omega t)$) and the limit $t_0 \to -\infty$ is considered. In this case however the condition $\lim_{t \to t_0} F(t) = 0$ is not satisfied. This issue is solved considering the following limit:

$$A_1(t) = \lim_{\eta \to 0^+} \int_{t_0}^{+\infty} dt' \; \chi_{AB}(t-t') \; e^{-\eta(t-t')} \; F(t') \tag{1.70}$$

Now the field $F(t)$ is adiabatically switched on and in the limit $t_0 \to -\infty$ (that can be taken with the constraint $\eta > \frac{1}{|t_0|}$) the effective field $e^{-\eta(t-t')} \; F(t')$ goes to zero. Fourier transforming this relation and using the convolution theorem yields:

$$A_1(\omega) = \lim_{\eta \to 0^+} \chi'_{AB}(\omega)F(\omega) \tag{1.71}$$

where $\chi'_{AB} = \chi_{AB} \; e^{-\eta(t-t')}$ and, for sake of simplicity, it has been used a shorthand notation in which just the ω presence in the arguments of functions, instead of the time, implies that they are their Fourier transformed. $\chi_{AB}(\omega)$ is the frequency dependent response function, defined as

$$\chi_{AB}(\omega) = -i \lim_{\eta \to 0^+} \int_{-\infty}^{+\infty} d\tau \; \theta(\tau) \; \langle\psi_0|[\hat{A}_{H_0}(\tau), \hat{B}]|\psi_0\rangle \; e^{(\eta+i\omega)\tau} \tag{1.72}$$

This function contains much information about a quantum many body system and in the following will be used to derive the excited states energies and other quantities in the framework of the TDDFT. To do this however Eq. (1.72), has to be recast, by means of some algebra, in a way in which the dependence of $\chi_{AB}(\omega)$ on the excited states of the time independent Hamiltonian $\hat{H}_0$ is explicated.

To achieve this form, the first step is to consider the complete set of eigenstates $\{\psi_n\}$, $n = 0, 1, 2, ...$ of the Hamiltonian $\hat{H}_0$, and their associated excitation energies $\Omega_n = E_n - E_0$. Inserting into Eq. (1.72) the completeness relation $\mathbb{1} = \sum_n |\psi_n\rangle\langle\psi_n|$ yields

Figure 1.4: The two paths on the complex plane that have to be used to compute the integral on the right hand side of Eq.(1.74): they are represented by the real axis and respectively the upper and lower semi-circle of radius r.

$$\chi_{AB}(\omega) = -i \lim_{\eta \to 0^+} \sum_n \int_{-\infty}^{+\infty} d\tau \ \theta(\tau) \quad e^{(\eta+i\omega)\tau} \left\{ \langle \psi_0 | \hat{A} | \psi_n \rangle \langle \psi_n | \hat{B} | \psi_0 \rangle \ e^{-i\Omega_n\tau} + \right.$$
$$\left. - \langle \psi_0 | \hat{B} | \psi_n \rangle \langle \psi_n | \hat{A} | \psi_0 \rangle \ e^{i\Omega_n\tau} \right\} \tag{1.73}$$

At this point, the following integral relation can be conveniently used:

$$e^{\eta\tau}\theta(\tau) = \frac{i}{2\pi} \int_{-\infty}^{+\infty} d\omega' \ \frac{e^{-i\omega'\tau}}{\omega' - i\eta} \tag{1.74}$$

This relation can be proved by direct contour integration of the complex extension of the function $\frac{e^{-i\omega'\tau}}{\omega'-i\eta}$ on the two paths described in Figure 1.4 (real axis plus continuous or dashed lines) in the limit in which $r \to \infty$. In the case $\tau > 0$ the upper semicircle has to be considered: the function goes exponentially to zero on it and hence the integral value on the real axis is proportional to the residue in the point $i\eta$, which is $e^{\eta\tau}$. In the case $\tau < 0$ the lower semicircle has to be instead considered and the value of the integral becomes zero since the function is analytical in the whole interior domain of the path.

Substituting Eq.(1.74) in Eq.(1.73) yields

$$\chi_{AB}(\omega) = \sum_n \int_{-\infty}^{+\infty} \frac{d\omega'}{2\pi(\omega' + i\eta)} \int_{-\infty}^{+\infty} d\tau \ \left\{ e^{i(\omega-\Omega_n-\omega')\tau} A_{0n}B_{n0} - e^{i(\omega+\Omega_n-\omega')\tau} B_{0n}A_{n0} \right\} =$$
$$= \sum_n \int_{-\infty}^{+\infty} \frac{d\omega'}{2\pi(\omega' + i\eta)} \ \left\{ 2\pi\delta(\omega - \Omega_n - \omega')A_{0n}B_{n0} - 2\pi\delta(\omega + \Omega_n - \omega')B_{0n}A_{n0} \right\} = \qquad .$$
$$= \sum_n \left\{ \frac{A_{0n}B_{n0}}{(\omega - \Omega_n + i\eta)} - \frac{B_{0n}A_{n0}}{(\omega + \Omega_n + i\eta)} \right\} \qquad .$$
$$\tag{1.75}$$

where the limit for $\eta \to 0^+$ is understood and the notation $O_{0n} = \langle \psi_0 | \hat{O} | \psi_n \rangle$ for the operators matrix elements has been used. This form of the frequency dependent response function is known as spectral decomposition or Lehmann representation.

1.9.2 The density density response function

The results of the last section can now be applied to study, to first order, the time dependence induced in the density of a many body system of electrons by an external time dependent perturbation $\hat{H}_1(t) = \hat{V}_1(t) = \sum_{i=1}^N v(\mathbf{r}_i, t)$ which switches on at the instant t_0. To do this,

the formalism defined in last section must be adapted to this particular case. So, first of all, it has to be noted that the density response function can be defined as the response of the N-electron density operator defined in Eq.(7)

$$n(\mathbf{r}, t) - n_0(\mathbf{r}) = \langle \psi(t)|\hat{n}(\mathbf{r})|\psi(t)\rangle - \langle \psi_0|\hat{n}(\mathbf{r})|\psi_0\rangle \tag{1.76}$$

and that, by means of the same operator, the external time dependent perturbation can be conveniently written as

$$\hat{V}_1(t) = \int d\mathbf{r} \ \hat{n}(\mathbf{r})v_1(\mathbf{r}, t) \tag{1.77}$$

Now the connection with the formalism and the results derived in last section is almost immediate to do: the difference between the right hand side of Eq. (1.77) and the time dependent perturbation considered in Eq. (1.54) is just the fact that in this last case the integral over $\mathbf{r}$ has to be considered. So, let's consider for a moment the following external potential acting in only a given space point $\mathbf{r}'$:

$$\hat{V}'_1(\mathbf{r}', t) = \hat{n}(\mathbf{r}')v_1'(\mathbf{r}', t) \tag{1.78}$$

such that

$$\hat{V}_1(t) = \int d\mathbf{r}' \ \hat{V}'_1(\mathbf{r}', t) \tag{1.79}$$

Applying the same formalism and derivation of last sections, the density linear response would be in this case

$$n_1(\mathbf{r}, t) = \int_{-\infty}^{+\infty} dt' \ \chi_{\hat{n}\hat{n}}(\mathbf{r}, \mathbf{r}', t - t') \ v_1'(\mathbf{r}', t') \tag{1.80}$$

where

$$\chi_{\hat{n}\hat{n}}(\mathbf{r}, \mathbf{r}', t - t') = -i\theta(t - t') \ \langle \psi_0|[\hat{n}_{H_0}(\mathbf{r}, t - t'), \hat{n}(\mathbf{r}')]|\psi_0\rangle \tag{1.81}$$

Using the linearity of the linear response equation, the integral over $\mathbf{r}'$ can be considered now to go back from the external potential $\hat{V}'(t)$ to the external potential $\hat{V}(t)$, getting:

$$n_1(\mathbf{r}, t) = \int_{-\infty}^{+\infty} dt' \int d\mathbf{r}' \ \chi_{\hat{n}\hat{n}}(\mathbf{r}, \mathbf{r}', t - t') \ v_1(\mathbf{r}', t') \tag{1.82}$$

where the response function χ_{nn} is the same of Eq. (1.81). Finally, repeating the same algebra of last section, the frequency dependent density linear response and the density density linear response function are found to be:

$$n_1(\mathbf{r}, \omega) = \int d\mathbf{r}' \ \chi_{\hat{n}\hat{n}}(\mathbf{r}, \mathbf{r}', \omega) \ v_1(\mathbf{r}', \omega) \tag{1.83}$$

$$\chi_{\hat{n}\hat{n}}(\mathbf{r}, \mathbf{r}', \omega) = \lim_{\eta \to 0^+} \sum_{n=1}^{\infty} \left[\frac{\langle \psi_0|\hat{n}(\mathbf{r})|\psi_n\rangle \ \langle \psi_n|\hat{n}(\mathbf{r}')|\psi_0\rangle}{\omega - \Omega_n - i\eta} - \frac{\langle \psi_0|\hat{n}(\mathbf{r}')|\psi_n\rangle \ \langle \psi_n|\hat{n}(\mathbf{r})|\psi_0\rangle}{\omega + \Omega_n - i\eta} \right] \tag{1.84}$$

In the following, only the density density linear response function will be used, while the response functions relative to other observable will not be treated. The subscript $\hat{n}\hat{n}$ will be then omitted in future for sake of simplicity.

1.10 Linear response in TDDFT

In this section the results coming from the linear response theory will be applied to the Kohn-Sham non interacting system. This task can not be trivially achieved because the Kohn-Sham time dependent potential $v_s[n](\mathbf{r}, t)$ of Eq. (1.51), being a functional of the "whole" time dependent density, has a dependence on the external time dependent potential $v_1(\mathbf{r}, t)$ at all the orders. There's a trick however that can be used: the formal expansion of the density response derived in last section

$$n(\mathbf{r}, t) = n_0(\mathbf{r}) + n_1(\mathbf{r}, t) + n_2(\mathbf{r}, t) + ... \tag{1.85}$$

can be used to get the corresponding functional expansion of the time dependent Kohn-Sham potential:

$$v_s[n](\mathbf{r}, t) = v_{s0}(\mathbf{r}) + v_{s1}(\mathbf{r}, t) + ... \tag{1.86}$$

where

$$v_{s0}(\mathbf{r}) = v_{s0}[n_0](\mathbf{r}) = v_s[n = n_0](\mathbf{r})$$

$$v_{s1}(\mathbf{r}, t) = v_{s1}[n_0, n_1](\mathbf{r}, t) = \int_{-\infty}^{+\infty} dt' \int d\mathbf{r}' \left. \frac{\delta v_s[n](\mathbf{r}, t)}{\delta n(\mathbf{r}', t')} \right|_{n=n_0} n_1(\mathbf{r}', t') \tag{1.87}$$

and so on. The non interacting time dependent Hamiltonian

$$\hat{H}_s[n](t) = \hat{T} + \int d^3\mathbf{r} \, \hat{n}(\mathbf{r}) v_s[n](\mathbf{r}, t) \tag{1.88}$$

can hence be written in the form

$$\hat{H}_s[n](t) = \hat{H}_{s0}[n_0] + \hat{H}_{s1}[n_1](t) + o(v_1(\mathbf{r}, t))^2 \tag{1.89}$$

where

$$H_{s0}[n_0] = \hat{T} + \int d^3\mathbf{r} \, \hat{n}(\mathbf{r}) v_s[n_0](\mathbf{r}) \tag{1.90}$$

is the Kohn-Sham non interacting static Hamiltonian of the unperturbed system and

$$H_{s1}[n_1] = \int d^3\mathbf{r} \, \hat{n}(\mathbf{r}) v_{s1}[n_0, n_1](\mathbf{r}, t) \tag{1.91}$$

is the linearized time dependent perturbation.

What remains to do at this point is to apply the usual Kohn-Sham scheme: being by definition the time dependent electron density generated by the non interacting Hamiltonian of Eq. (1.88) the same as the one generated by the interacting Hamiltonian $\hat{H} = \hat{H}_0 + \hat{V}_1(t)$, the application of the linear response theory to the linearized time dependent Hamiltonian in Eq. (1.89) will return the correct interacting linear response density $n_1(\mathbf{r}, t)$.

Writing the time dependent Kohn-Sham potential as in Eq. (1.51) and grouping the terms of the same order in the external potential $v_1(\mathbf{r}, t)$ (keeping in mind that $n_0(\mathbf{r})$ is of order zero, $n_1(\mathbf{r}, t)$ is of order one and so on) yields:

$$v_{s0}(\mathbf{r}) = v_0(\mathbf{r}) + \int d^3\mathbf{r}' \frac{n_0(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + v_{xc}[n_0](\mathbf{r}, t)$$

$$v_{s1}(\mathbf{r}, t) = v_1(\mathbf{r}, t)\,\theta(t - t_0) + \int d^3\mathbf{r}' \frac{n_1(\mathbf{r}', t)}{|\mathbf{r} - \mathbf{r}'|} + \int_{-\infty}^{+\infty} dt' \int d\mathbf{r}' \left. \frac{\delta v_{xc}[n(\mathbf{r}, t)]}{\delta n(\mathbf{r}', t')} \right|_{n=n_0(\mathbf{r})} n_1(\mathbf{r}', t')$$

$$(1.92)$$

So, v_{s0} is nothing else than the unperturbed Kohn-Sham potential, as it had to be, while in the time dependent potential v_{s1} the dependence on the external potential and on the Hartree potential has been explicitated. Finally, the quantity

$$f_{xc}(\mathbf{r}, \mathbf{r}', t, t') = \left. \frac{\delta v_{xc}[n](\mathbf{r}, t)}{\delta n(\mathbf{r}', t')} \right|_{n_0(r)} \qquad (1.93)$$

which is the so called time dependent xc kernel, is the new unknown part that has to be approximated. However it is important to stress that it is a functional of just the ground state density. This kernel is the key quantity in the linear response TDDFT.

What remains now is just to apply the linear response theory to this linearized non interacting system. It yields:

$$n_1(\mathbf{r}, t) = \int_{-\infty}^{+\infty} dt' \int d^3\mathbf{r}' \; \chi_s(\mathbf{r}, \mathbf{r}', t - t') v_{s1}[n](\mathbf{r}', t') \qquad (1.94)$$

where now,

$$\chi_s(\mathbf{r}, \mathbf{r}', t - t') = -i\theta(t - t') \; \langle \psi_{s0}[n_0] | [\hat{n}_{H_{s0}}(\mathbf{r}, t - t'), \hat{n}_0(\mathbf{r}')] | \psi_{s0}[n_0] \rangle \qquad (1.95)$$

Finally, substituting the form of v_{s1} of Eq. (1.92) in Eq. (1.94) yields:

$$n_1(\mathbf{r}, t) = \int_{-\infty}^{+\infty} dt' \int d^3\mathbf{r}' \; \chi_s(\mathbf{r}, \mathbf{r}', t - t') v_1(\mathbf{r}', t') +$$

$$+ \int_{-\infty}^{+\infty} dt' \int d^3\mathbf{r}' \int_{-\infty}^{+\infty} d\tau \int d^3\mathbf{r}'' \; \chi_s(\mathbf{r}, \mathbf{r}', t - t') f_{Hxc}(\mathbf{r}', \mathbf{r}'', t', \tau) \, n_1(\mathbf{r}'', \tau)$$

$$(1.96)$$

Where the Hartree-exchange-correlation kernel $f_{Hxc}(\mathbf{r}, \mathbf{r}', t, t') = f_{xc}(\mathbf{r}, \mathbf{r}', t, t') + \frac{\delta(t-t')}{|\mathbf{r}-\mathbf{r}'|}$ has been introduced. Starting from this equation, it can be derived that

$$f_{Hxc}(\mathbf{r}, \mathbf{r}', t, t') = f_{Hxc}(\mathbf{r}, \mathbf{r}', t - t') \qquad (1.97)$$

i.e. the Hartree-exchange-correlation kernel depends only on the time differences $(t - t')$. To see this, the following compact notation of integral operators can be used:

$$\hat{O}|h\rangle \equiv \int_{-\infty}^{-\infty} dt' \int d\mathbf{r}' \, O(\mathbf{r}, \mathbf{r}', t - t') h(\mathbf{r}', t') \tag{1.98}$$

so that Eq.(1.96) is written as:

$$|n_1\rangle = \hat{\chi}_s|v_1\rangle + \hat{\chi}_s \hat{f}_{Hxc}|n_1\rangle \tag{1.99}$$

Now, using that $|n_1\rangle = \hat{\chi}|v_1\rangle$ and the fact that this equation holds for every perturbing potential $|v_1\rangle$ yields the following operatorial equality:

$$\hat{\chi} = \hat{\chi}_s + \hat{\chi}_s \hat{f}_{Hxc}\hat{\chi} \tag{1.100}$$

The inverse response operators, formally defined with the relations $\hat{\chi}_s^{-1}\hat{\chi}_s = \mathbb{1}$ and $\hat{\chi}^{-1}\hat{\chi} = \mathbb{1}$ have to be now introduced. Multiplying the last equation by $\hat{\chi}_s^{-1}$ on the left and by $\hat{\chi}^{-1}$ on the right yields:

$$\hat{f}_{Hxc} = \hat{\chi}_s^{-1} - \hat{\chi}^{-1} \tag{1.101}$$

Given the fact that both the kernels χ and χ_s depend on $(t - t')$ and hence so do χ^{-1} and χ_s^{-1} Eq. (1.97) is formally proved. Eq. (1.96) can hence be Fourier transformed using the convolution theorem[7] to give

$$n_1(\mathbf{r}, \omega) = \int d^3\mathbf{r}' \, \chi_s(\mathbf{r}, \mathbf{r}', \omega) \left[v_1(\mathbf{r}', \omega) + \int d^3\mathbf{r}'' \, f_{Hxc}(\mathbf{r}', \mathbf{r}'', \omega) n_1(\mathbf{r}'', \omega) \right] \tag{1.102}$$

This equation contains all the results collected so far in the review of the time dependent formalism: it consists in the application of the linear response theory to the TDDFT in the Khon Sham formalism. It will be now shown that the goal achieved is that now a direct form for the frequency dependent response function can be explicitly given in terms of Khon Sham orbitals.

According to Eq. (1.84), what we must compute is the general matrix element $\langle \psi_{s0}|\hat{n}(\mathbf{r})|\psi_{s,n}\rangle$ for the nth non interacting excited state $\psi_{s,n}$. As discussed in section 1.3, the whole spectrum of the non interacting Hamiltonian is easily obtained considering the set of Slater determinants that can be constructed with the whole set of Khon Sham orbitals and the excited states are divided in classes of excitations: the single excitations are the Slater determinants of the kind $|\psi_{s,h\rightarrow k}\rangle$ in which the hth occupied orbital is substituted with the kth unoccupied one; the double excitations, that are the excited states of the kind $|\psi_{s,h\rightarrow k \atop l\rightarrow m}\rangle$ in which the hth and l^{th} occupied orbital are substituted with the kth and m^{th} unoccupied one and so on.

The density operator is a one electron operator, i.e. is in the class of operators of the kind

$$\hat{O}_1 \equiv \sum_{i=1}^{N} \hat{o}(i) \tag{1.103}$$

---

[7]The convolution in this case is actually double but, as demonstrated in ref. [29], the convolution theorem applies also in this case.

as seen from its definition in Eq. (1.7), and so the general rules for the calculation of one electron operators matrix elements on Slater determinants [30] can be used for the calculation of its matrix elements on the excited determinants considered above. According to these rules in fact, for every given reference Slater determinant $|\Phi_0\rangle = |1, 2, \ldots, N\rangle$, three relevant cases have to be considered to compute the general one electron operator matrix element: i) the matrix element $\langle\Phi_0|\hat{O}_1|\Phi_0\rangle$, ii) the matrix elements connecting the reference determinant with a single excitation $\langle\Phi_0|\hat{O}_1|\psi_{s,h\to k}\rangle$ and iii) the matrix element connecting the reference determinant with a double or higher order excitation. The results are:

$$\langle\Phi_0|\hat{O}_1|\Phi_0\rangle = \sum_{i=1}^{N} \langle i|\hat{o}|i\rangle$$

$$\langle\Phi_0|\hat{O}_1|\psi_{s,h\to k}\rangle = \langle h|\hat{o}|k\rangle \qquad (1.104)$$

$$\langle\Phi_0|\hat{O}_1|D, T, ..\rangle = 0$$

where with the notation $|D, T, ..\rangle$ the general excitation of order greater then one (double, triple, etc.) has been indicated. So, coming back to the density operator matrix elements, in the case the ground state is coupled with a single excitation, according to the rules above we have:

$$\langle\psi_{s0}|\hat{n}(\mathbf{r})|\psi_{s,h\to k}\rangle = \int d\mathbf{r}' \ \langle\phi_h|\mathbf{r}'\rangle\delta(\mathbf{r} - \mathbf{r}')\langle\mathbf{r}'|\phi_k\rangle = \phi_h^{(KS)}(\mathbf{r})\phi_k^{*(KS)}(\mathbf{r}) \qquad (1.105)$$

All the other matrix element of the density operator are coupling the ground state with higher order excitations and hence, according to the rules reported in Eq.(1.104), they are all zero. Collecting these results, and using the fact that the excitation energy relative to single excitations $\psi_{s,h\to k}\rangle$ are simply $\omega_{hk} = \epsilon_k - \epsilon_h$, the frequency dependent non interacting response function can be finally written in terms of Khon Sham orbitals and energies starting from Eq. (1.84) as follows:

$$\chi_s(\mathbf{r}, \mathbf{r}', \omega) = \lim_{\eta\to 0^+} \sum_{h=1}^{N} \sum_{k=N+1}^{\infty} \left[\frac{\langle\psi_{s0}|\hat{n}(\mathbf{r})|\psi_{s,h\to k}\rangle \langle\psi_{s,h\to k}|\hat{n}(\mathbf{r}')|\psi_{s0}\rangle}{\omega - \omega_{hk} - i\eta} - \right.$$

$$\left. - \frac{\langle\psi_{s0}|\hat{n}(\mathbf{r}')|\psi_{s,h\to k}\rangle \langle\psi_{s,h\to k}|\hat{n}(\mathbf{r})|\psi_{s0}\rangle}{\omega + \omega_{hk} - i\eta} \right] =$$

$$= \lim_{\eta\to 0^+} \sum_{h=1}^{N} \sum_{k=N+1}^{\infty} \left[\frac{\phi_h^{(KS)}(\mathbf{r})\phi_k^{*(KS)}(\mathbf{r})\phi_h^{*(KS)}(\mathbf{r}')\phi_k^{(KS)}(\mathbf{r}')}{\omega - \omega_{hk} - i\eta} - \right.$$

$$\left. - \frac{\phi_h^{(KS)}(\mathbf{r}')\phi_k^{*(KS)}(\mathbf{r}')\phi_h^{*(KS)}(\mathbf{r})\phi_k^{(KS)}(\mathbf{r})}{\omega + \omega_{hk} - i\eta} \right] \qquad (1.106)$$

This form of the non interacting frequency dependent response function can be contracted in a way that will be useful for the algebra in the following section if the occupation number f_k are introduced:

$$f_k = \begin{cases} 0, & \text{if } k > N \\ 1, & \text{if } k \leqslant N \end{cases} \tag{1.107}$$

and also the fact that $\omega_{hk} = -\omega_{kh}$ is noted. It reads:

$$\chi_s(\mathbf{r}, \mathbf{r}', \omega) = \lim_{\eta \to 0^+} \left\{ \sum_{h,k=1}^{\infty} (f_k - f_h) \frac{\phi_h^{(KS)}(\mathbf{r}) \phi_k^{*(KS)}(\mathbf{r}) \phi_h^{*(KS)}(\mathbf{r}') \phi_k^{(KS)}(\mathbf{r}')}{\omega - \omega_{hk} - i\eta} \right\} \tag{1.108}$$

Here both the indexes h and k run over all the Khon Sham orbitals and the fact that only the correct terms with the correct signs are taken into account is guaranteed by the $f_k - f_h$ term.

The last missing ingredient to start to really operate on the Eq. (1.102) is to give a working form (even if approximated) for the xc kernel defined in Eq. (1.93). In section 1.8 the adiabatic approximation for the time dependent xc potential has been defined (Eq. (1.52)) and discussed: as it as been already pointed there all the memory effect are neglected in this approximation and the approximated xc potential at time t has a functional dependence only on the density at the same time t. So the xc kernel in this approximation becomes:

$$f_{xc}(\mathbf{r}, \mathbf{r}', t - t') \simeq f_{xc}^A(\mathbf{r}, \mathbf{r}', t - t') = \left. \frac{\delta v_{xc}^A[n_0; n_0 = n(\mathbf{r}, t)]}{\delta n(\mathbf{r}', t')} \right|_{n_0} = \delta(t - t') \left. \frac{\delta v_{xc}^A[n_0](\mathbf{r})}{\delta n(\mathbf{r}')} \right|_{n_0} \tag{1.109}$$

while the frequency dependent xc kernel becomes:

$$f_{xc}(\mathbf{r}, \mathbf{r}'\omega) \simeq f_{xc}^A(\mathbf{r}, \mathbf{r}') = \left. \frac{\delta v_{xc}^A[n_0](\mathbf{r})}{\delta n(\mathbf{r}')} \right|_{n_0} = \frac{\delta^2 E_{xc}[n_0](\mathbf{r})}{\delta n_0(\mathbf{r}) \delta n_0(\mathbf{r}')} \tag{1.110}$$

so that the frequency dependence in the kernel drops out. Within the adiabatic approximation hence, starting from any approximated static exchange correlation potential, the corresponding exchange correlation kernel can be straightforwardly derived by means of just a functional derivative. This is the approximation that will be used for all the calculations presented in this thesis. Now the only unknown in Eq. (1.102) is the linear density response and we can look at this expression as an equation to determine this quantity.

Summarizing, equation (1.102) formally gives (apart from the approximations made for the xc kernel) exactly the same frequency dependent density linear response of the many body interacting Hamiltonian[8] with the enormous advantage that now the response function $\chi(\mathbf{r}, \mathbf{r}', \omega)$ is replaced by the non interacting one, the $\chi_s(\mathbf{r}, \mathbf{r}', \omega)$ which is easily calculated by means of its sum over states representation, while the calculation of the interacting one is an overwhelming task. The price to pay consists in the fact that the xc kernel is a new quantity to deal with and that, in addition to the static xc potential, has to be approximated.

---

[8]Strictly speaking what is guarantied here is that the total time dependent non interacting density equals the interacting one (that is the goal of Khon Sham scheme), and that both of them can be expanded in series of power in the external field (as it has been shown in this section). The fact that the two series are equal term by term is however given in virtue of the polynomial equivalence principle.

1.10.1 Excitation energies from a matrix equation

In this section a self consistent matrix equation will be derived for the calculation of the excited states energies of a many body system using just functionals of the ground state density. By means of this equation the full solution of the linear response problem will be addressed and solved in next sections.

The first step is to rewrite Eq.(1.102) in the form

$$\int d^3\mathbf{r}'' \ \Xi(\mathbf{r},\mathbf{r}'',\omega) \ n_1(\mathbf{r}'',\omega) = \int d^3\mathbf{r}'\chi_s(\mathbf{r},\mathbf{r}',\omega)v_1(\mathbf{r}',\omega) \tag{1.111}$$

with

$$\Xi(\mathbf{r},\mathbf{r}'',\omega) = \delta(\mathbf{r}-\mathbf{r}'') - \int d^3\mathbf{y} \ \chi_s(\mathbf{r},\mathbf{y})f_{Hxc}(\mathbf{y},\mathbf{r}'') \tag{1.112}$$

The following two important facts can now be used: in the limit $\omega \to \Omega_n$ (the true excitation energies of the many body Hamiltonian) i) the frequency dependent response function χ diverges with first order poles (see. Eq. (1.84)) and hence so does $n_1(\mathbf{r},\Omega_n)$ and ii) the non interacting response function (that has poles for the Kohn-Sham excitation energies $\omega_{ij} = \epsilon_i - \epsilon_j$) remains instead finite in this limit. This facts are sufficient to deduce that the integral operator acting on $n_1(\mathbf{r},\omega)$ on the left hand side of Eq. (1.111) can not be invertible in this limit. In fact this operator has to cancel out the first order pole of $n_1(\mathbf{r},\Omega_n)$ to give a finite right hand side of the Eq. (1.111) and hence its inverse operator has to diverge.

The operator non invertibility statement is in all equivalent to the condition of existence of a zero eigenvalue for the operator $\hat{\Xi}$. So for $\omega = \Omega_n$ the following relation must hold:

$$\int d^3\mathbf{r}'' \ \Xi(\mathbf{r},\mathbf{r}'',\Omega_n) \ \tilde{n}_{1n}(\mathbf{r}'',\Omega_n) = 0 \tag{1.113}$$

Moreover it can be now shown that Eq. (1.113) actually holds only for these excitation energies. This fact is easily proved by reductio ad absurdum considering possible the case that the for some energy $\omega* \neq \Omega_n$ the operator $\hat{\Xi}$ is not invertible; in this case the right hand side of Eq. (1.111) would become zero, in fact Eq. (1.84) implies that $n_1(\mathbf{r},\omega)$ does diverge only for the excitation energies Ω_n. The right hand side of Eq. (1.111) instead, containing the term $v_1(\mathbf{r}',\omega*)$ which is arbitrary, can always be chosen to be non zero and hence Eq. (1.111) would not be fulfilled in this case. Eq. (1.113) can now be read as an implicit equation for Ω_n via $\tilde{n}_1$

Substituting back the form (1.112) for the kernel Ξ in Eq. (1.113) yields:

$$\tilde{n}_{1n}(\mathbf{r},\Omega_n) = \int d^3\mathbf{x} \ \chi_s(\mathbf{r},\mathbf{x},\Omega_n) \int d^3\mathbf{x}' \ f_{Hxc}(\mathbf{x},\mathbf{x}',\Omega_n) \ \tilde{n}_1(\mathbf{x}',\Omega_n) \tag{1.114}$$

This is the key equation from the which the excitation energies Ω_n will be derived. It is important to stress here that the density response $\tilde{n}_{1n}(\mathbf{r},\Omega_n)$ in this equation is not fulfilling Eq. (1.102) and is remaining finite also for the exact excitation energies. The solutions $\tilde{n}_{1n}(\mathbf{r},\Omega_n)$ can be physically interpreted as stationary oscillation eigenmode of the system

(see ref [1], at page 145 for a discussion on this fact) and will be referred eigenoscillations in the following.

Eq. (1.114) will be now transformed into a matrix equation in the Khon Sham orbitals Hilbert Space basis. To do this, the first step is to compact a bit more the compacted form of Eq. (1.106) for the non interacting response function written in term of the Kohn-Sham orbitals:

$$\chi_s(\mathbf{r}, \mathbf{r}', \Omega_n) = \sum_{j,k=1}^{\infty} \alpha_{kj} \frac{\Phi_{jk}^*(\mathbf{r})\Phi_{jk}(\mathbf{r}')}{\Omega_n - \omega_{jk}} \tag{1.115}$$

Where the following shorthand notations have been used to achieve this expression

$$\alpha_{kj} \equiv f_k - f_j \tag{1.116}$$
$$\Phi_{jk}(\mathbf{r}) \equiv \phi_j^{*(KS)}(\mathbf{r}) \quad \phi_k^{(KS)}(\mathbf{r})$$

and also the limits $\omega \to \Omega_n$ and $\eta \to 0^+$ have been here safely taken because $\omega_{jk} \neq \Omega_n$[9]. The second step is to define the following auxiliary object:

$$g(\mathbf{x}, \Omega_n) \equiv \int d^3\mathbf{r} \ f_{Hxc}(\mathbf{x}, \mathbf{r}, \Omega_n)\tilde{n}_1(\mathbf{r}, \Omega_n) \tag{1.117}$$

With this definition, the multiplication of Eq.(1.114) on both sides for $f_{Hxc}(\mathbf{x}, \mathbf{r}, \Omega_n)$, the integration over $\mathbf{r}$ and the substitution of the form of Eq.(1.115) for χ_s gives:

$$g(\mathbf{x}, \Omega_n) = \sum_{jk} \frac{\alpha_{jk}}{\Omega_n - \omega_{jk}} \int d^3\mathbf{r} \ f_{Hxc}(\mathbf{x}, \mathbf{r}, \Omega_n)\Phi_{jk}^*(\mathbf{r}) \int d^3\mathbf{r}' \ \Phi_{jk}(\mathbf{r}')g(\mathbf{r}', \Omega_n) \tag{1.118}$$

Finally, defining

$$\beta_{hl}(\Omega_n) \equiv \frac{1}{\Omega_n - \omega_{hl}} \int d^3\mathbf{x} \ g(\mathbf{x}, \Omega_n)\Phi_{hl}(\mathbf{x}) \tag{1.119}$$

multiplying both sides for $\Phi_{hl}(\mathbf{x})$ and integrating over $\mathbf{x}$ yields:

$$\sum_{jk} \alpha_{jk}K_{hljk}(\Omega_n) \ \beta_{jk}(\Omega_n) = (\Omega_n - \omega_{hl}) \ \beta_{hl}(\Omega_n) \tag{1.120}$$

where

$$K_{hljk}(\Omega_n) \equiv \int d^3\mathbf{r} \int d^3\mathbf{r}' \ \Phi_{hl}(\mathbf{r})f_{Hxc}(\mathbf{r}, \mathbf{r}', \Omega_n)\Phi_{jk}^*(\mathbf{r}') \tag{1.121}$$

Eq. (1.120) is the wanted matrix equation. This system of equation is however non linear since in general the xc kernel does depend on Ω_n, and hence so do the matrix elements K_{jkhl}.

---

[9]The case $\Omega_n = \omega_{ij}$ for some excited state ψ_n and some pair

The solution should than be found by means of some iterative procedure. In the adiabatic approximation however the xc kernel takes the form of Eq. (1.110) which is ω independent, and so Eq. (1.120) becomes just a linear matrix equation (of infinite dimension).

Finally, once the equation is solved for each excitation energy Ω_n, from the associated coefficients $(\beta_{hl})_n$ the density eigenoscillation $\tilde{n}_1(\mathbf{r}, \Omega_n)$ can be derived: in fact, substituting back in Eq.(1.119) the term $g(\mathbf{x}, \Omega_n)$ in Eq.(1.117) yields:

$$\beta_{hl}(\Omega_n) = \frac{1}{\Omega_n - \omega_{hl}} \int d^3\mathbf{x} \; \Phi_{hl}(\mathbf{x}) \int d^3\mathbf{r} \; f_{Hxc}(\mathbf{x}, \mathbf{r}, \Omega_n) \tilde{n}_1(\mathbf{r}, \Omega_n) \tag{1.122}$$

From this expression, multiplying both sides for $\alpha_{jk}\Phi_{jk}^*$, summing over jk, using the form of Eq. (1.115) for the χ_s and comparing the resulting expression with Eq.(1.114) yields:

$$\tilde{n}_1(\mathbf{r}, \Omega_n) = \sum_{jk} \alpha_{jk} \Phi_{jk}^*(\mathbf{r}) \beta_{jk}(\Omega_n) \tag{1.123}$$

1.10.2   The Casida equations

The matrix equations in the basis of the Khon Sham orbitals derived in the last section (i.e. Eq. (1.120)) allow the calculation of the excitation energies of a many body Hamiltonian by means of just ground state density functionals. They are in fact the fundamental result of the application of the quantum linear response theory to the TDDFT. In this section these equations will be rewritten in a more convenient and compact form using the symmetries and the properties of the algebraic elements that haven't been used so far. Finally they will be also cast as a computationally efficient eigenvalue problem by means of some linear algebra manipulation. The results and the formalism derived will be used in the next section to finally address the full solution of the linear response problem.

The first symmetry that can be used is that, if the Kohn-Sham orbitals are chosen real (as it can always be done for molecular systems) then $K_{hljk} = K_{jkhl}$ (see Eq. (1.121)). Moreover the presence in the equations of the coefficients α_{jk} defined in Eq. (1.117) implies that on the right hand side of Eq. (1.120) only the coefficients β_{jk} connecting an occupied orbital to an unoccupied one are non zero. To efficiently include this property in the equations, it is convenient to adopt a notation in which occupied an unoccupied orbitals are called with different indexes; in the following, the occupied orbitals will then be labeled by the indexes i, i', etc. while the unoccupied ones by a, a' etc. So, considering that $\alpha_{ia} = 1$, $\alpha_{ia} = -1$, $\alpha_{ii} = \alpha_{aa} = 0$. Eq. (1.120) can be rewritten as:

$$\sum_{i'=1}^{N}\sum_{a'=N+1}^{\infty}\left\{K_{iai'a'}(\Omega_n)\,\beta_{i'a'}(\Omega_n)-K_{iai'a'}(\Omega_n)\,\beta_{a'i'}(\Omega_n)\right\}=(\Omega_n-\omega_{ia})\,\,\beta_{ia}(\Omega_n)$$

$$\sum_{i'=1}^{N}\sum_{a'=N+1}^{\infty}\left\{K_{iai'a'}(\Omega_n)\,\beta_{i'a'}(\Omega_n)-K_{iai'a'}(\Omega_n)\,\beta_{a'i'}(\Omega_n)\right\}=(\Omega_n-\omega_{ai})\,\,\beta_{ai}(\Omega_n)$$

$$\sum_{i'=1}^{N}\sum_{a'=N+1}^{\infty}\left\{K_{ii''i'a'}(\Omega_n)\,\beta_{i'a'}(\Omega_n)-K_{ii''i'a'}(\Omega_n)\,\beta_{a'i'}(\Omega_n)\right\}=(\Omega_n-\omega_{i''i})\,\,\beta_{i''i}(\Omega_n)$$

$$\sum_{i'=1}^{N}\sum_{a'=N+1}^{\infty}\left\{K_{a''ai'a'}(\Omega_n)\,\beta_{i'a'}(\Omega_n)-K_{a''ai'a'}(\Omega_n)\,\beta_{a'i'}(\Omega_n)\right\}=(\Omega_n-\omega_{aa''})\,\,\beta_{aa''}(\Omega_n)$$

$$(1.124)$$

where on the right hand sides of these equations only the non zero terms (proportional to α_{ia} and α_{ia}) have been left and the right hand sides are specialized for the four subset of coefficients resulting from the division of the indexes in occupied and unoccupied (ia, ai, $i''i$, aa''). It has to be noted now that, according to Eq. (1.123) only the coefficients $\beta_{ai}(\Omega_n)$ and $\beta_{ia}(\Omega_n)$ of the first two Equations (1.124) are necessary for the construction of the density eigenoscillations $\tilde{n}_1(\mathbf{r},\Omega_n)$. As a result, the coefficients $\beta_{aa''}(\Omega_n)$ and $\beta_{i''i}(\Omega_n)$ do not have any physical meaning (they could take any value, without affecting the $\tilde{n}_1(\mathbf{r},\Omega_n)$)[10] and hence the two last block of equations in (1.124) can be neglected.

Finally, from an algebraic point of view, aiming to arrive to get a more compact form of these equations, it is convenient to consider the coefficients β_{ia} as vector component in a vectorial space, the excitations space. This vectorial space is constructed as N replicas (the index i that is running over the occupied orbitals) of the space spanned by the unoccupied Khon Sham orbitals index a. A given vector $\mathbf{v}$ in this space is hence constructed as $\mathbf{v} = \sum_{ia} v_{ia}\delta_{\mathbf{ia}}$ where the vectors $\delta_{\mathbf{ia}}$ are the normal basis vector in this whose components are $(\delta_{\mathbf{ia}})_{i'a'} = \delta_{ii'}\delta_{aa'}$. In this way, the object with four indexes (e.g. the term $K_{iai'a'}$) can be considered as linear operators in this space (usually referred as super-operators). It is convenient at this point to make the following definitions:

$$A_{iai'a'}(\Omega_n) = \delta_{ji'}\delta_{aa'}\omega_{a'i'} + K_{iai'a'}(\Omega_n)$$
$$B_{iai'a'}(\Omega_n) = K_{iai'a'}(\Omega_n)$$
$$X_{ia}(\Omega_n) = -\beta_{ia}(\Omega_n)$$
$$Y_{ia}(\Omega_n) = \beta_{ai}(\Omega_n)$$

$$(1.125)$$

So that Eq. (1.124) can be rewritten in the following compact form:

$$\mathbb{A}X + \mathbb{B}Y = -\Omega_n X$$
$$\mathbb{B}X + \mathbb{A}Y = \Omega_n Y$$

$$(1.126)$$

---

[10]Their presence in Eq. (1.124) is caused by the fact that Eq. (1.118) has been multiplied by the general term Φ_{hk} to get Eq. (1.120), while it would have been sufficient to multiply Eq. (1.118) for just the terms that were necessary to get the solution $\tilde{n}_1(\mathbf{r},\Omega_n)$ (i.e. for Φ_{ia} and Φ_{ai}). As a result, even if a compact notation has been kept in the equations, arriving to Eq. (1.120) with a less cumbersome algebra, these extra (and unnecessary) terms have been carried trough the algebraic steps.

where the notations for the super-operators algebra $\mathbb{A}\mathbf{X} = \sum_{i'a'} A_{iai'a'} X_{i'a'}$ etc. have been used. Eq.(1.126) can be finally written in the following compact block matrix form:

$$\begin{pmatrix} \mathbb{A} & \mathbb{B} \\ \mathbb{B} & \mathbb{A} \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} = \Omega_n \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} \tag{1.127}$$

Written in this form these equations are known as Casida equations [15]. Finally also the density eigenoscillations in Eq. (1.123) can be rewritten (considering that now the Khon-Sham orbitals are now real) in terms of these new variables:

$$\tilde{n}_1(\mathbf{r}, \Omega_n) = \sum_{ia} \{X_{ia}(\Omega_n) + Y_{ia}(\Omega_n)\} \Phi_{ia}(\mathbf{r}) \tag{1.128}$$

It is important to stress that, as discussed already in the last section, the Casida equations are in general a set of equations formally exact: however they give the true excitation energies of the many-body Hamiltonian only if the ground state density of the system is exactly calculated and the exact form of the exchange-correlation kernel defined in Eq. (1.93) is given. In this exact case however they would be non linear since in general the xc kernel depend on Ω_n, and hence so do the super-operators $\mathbb{A}$ and $\mathbb{B}$. It is just in the adiabatic approximation that the xc kernel, taking the form of Eq. (1.110), is ω independent and the Casida equations become a linear matrix equation.

The Casida equations can be recast, by means of some elementary matrix algebra, in a computationally more efficient form: in fact, adding member by member the Casida equations yields

$$\begin{cases} (\mathbb{A} + \mathbb{B})(X + Y) = -\Omega_n(X - Y) \\ (\mathbb{A} - \mathbb{B})(X - Y) = -\Omega_n(X + Y) \end{cases} \tag{1.129}$$

and plugging the $(X + Y)$ term of the second equation in the first one:

$$(\mathbb{A} + \mathbb{B})(\mathbb{A} - \mathbb{B})(X - Y) = \Omega_n^2(X - Y) \tag{1.130}$$

Now, using the fact that the matrix $\mathbb{A} - \mathbb{B}$ is diagonal, the following matrix can be defined:

$$\mathbb{C} = (\mathbb{A} - \mathbb{B})^{\frac{1}{2}}(\mathbb{A} + \mathbb{B})(\mathbb{A} - \mathbb{B})^{\frac{1}{2}}, \tag{1.131}$$

With this definition, multiplying Eq. (1.130) for $(\mathbb{A} - \mathbb{B})^{\frac{1}{2}}$ on the left yields

$$\mathbb{C}\,\mathbf{e}_n = \Omega_n^2\,\mathbf{e}_n \tag{1.132}$$

where

$$\mathbf{e}_n = \mathbf{e}(\Omega_n) = (\mathbb{A} - \mathbb{B})^{\frac{1}{2}}(X(\Omega_n) - Y(\Omega_n)) \tag{1.133}$$

and the subscript n is introduced to indicate that they refer to the solution Ω_n in a more immediate way. Going back to the component form of vectors and super-operators in the excitations space, from the definitions in Eq. (1.125):

$$\sum_{i,i',a,a'} \left[\delta_{ii'}\delta_{aa'}\omega_{a'i'}^2 + 2\sqrt{\omega_{ai}\omega_{a'i'}}K_{ia,i'a'} \right] (\mathbf{e}_n)_{i'a'} = \Omega_n^2\,(\mathbf{e}_n)_{ia} \tag{1.134}$$

This is the more convenient form of the Casida equation claimed above. Being in fact in the form of an eigenvalue equation it can be solved by means of standard numerical methods. Moreover Eq. (1.132) implies that the set of vectors $\{\mathbf{e}_n\}_{n=1,..,\infty}$, known as Casida vectors, is complete in the excitations space, and hence they fulfill the completeness relation

$$\sum_{n=1}^{\infty} \mathbf{e}_n \mathbf{e}^{\dagger}{}_n = \mathbb{1} \tag{1.135}$$

that will turn to be very useful in the following.

1.10.3 The general linear density response problem solution

In this section, the calculation of the density linear response problem will be finally addressed and solved in terms of the Casida vectors defined in last section. To do this it will be necessary to use just some linear algebra. So, Eq. (1.102) has to be considered now. In this case $v_1(\omega) \neq 0$ but starting from equation Eq. (1.102), the same procedure used in section 1.10.1 to derive the matrix equations (1.120) from Eq.(1.114) can be followed. As a result, the Casida equations in their block form (i.e. Eq.(1.127)) become now:

$$\begin{pmatrix} \mathbb{A} & \mathbb{B} \\ \mathbb{B} & \mathbb{A} \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} - \omega \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} v_1 \\ v_1 \end{pmatrix} \tag{1.136}$$

where now the value of ω is arbitrary (but different from Ω_n for all n) and the terms $(v_1)_{ia} = \int d^3\mathbf{r} \ \Phi_{ia}(\mathbf{r}) v_1(r,\omega) = (v_1)_{ai}$ are coming from the fact that step by step the term with the external potential has to be carried. These equations can be now manipulated as it has been made to derive the form in Eq. (1.132) of the Casida equations to get

$$\mathbb{C} \ \mathbf{e} = \omega^2 \ \mathbf{e} - 2\omega(\mathbb{A} - \mathbb{B})^{\frac{1}{2}} \mathbf{v}_1 \tag{1.137}$$

Now, the following algebraic steps have be followed:
i) rewrite Eq. (1.137) in the following way:

$$(\mathbb{C} \ - \omega^2 \mathbb{1}) \ \mathbf{e} = -2\omega(\mathbb{A} - \mathbb{B})^{\frac{1}{2}} \mathbf{v}_1 \tag{1.138}$$

ii) multiply both sides by $(\mathbb{C} \ - \omega^2 \mathbb{1})^{-1}$ on the left and use the definition (1.133) for the Casida vector $\mathbf{e}$

$$(\mathbb{A} - \mathbb{B})^{\frac{1}{2}}(X - Y) = \mathbf{e} = -2\omega \ (\mathbb{C} \ - \omega^2 \mathbb{1})^{-1}(\mathbb{A} - \mathbb{B})^{\frac{1}{2}} \mathbf{v}_1 \tag{1.139}$$

iii) multiply both sides by $(\mathbb{A} - \mathbb{B})^{\frac{1}{2}}$

$$(\mathbb{A} - \mathbb{B})(X - Y) = -2\omega(\mathbb{A} - \mathbb{B})^{\frac{1}{2}}(\mathbb{C} \ - \omega^2 \mathbb{1})^{-1}(\mathbb{A} - \mathbb{B})^{\frac{1}{2}} \mathbf{v}_1 \tag{1.140}$$

iv) use the following relation:

$$(\mathbb{A} - \mathbb{B})(X - Y) = -\omega(X + Y) \tag{1.141}$$

(which comes from Eq. (1.136) adding member by member the two equations) for the left hand side of Eq. (1.140) and the following result is found:

$$(X + Y) = 2 \, (\mathbb{A} - \mathbb{B})^{\frac{1}{2}} \left(\mathbb{C} - \omega^2 \mathbb{1}\right)^{-1} (\mathbb{A} - \mathbb{B})^{\frac{1}{2}} \, \mathbf{v}_1 \tag{1.142}$$

Now, using the completeness of the Casida vectors, the following relation in the excitations space can be proved:

$$\left(\mathbb{C} - \omega^2 \mathbb{1}\right)^{-1} = \sum_n \frac{\mathbf{e}_n \mathbf{e}^\dagger_n}{\Omega_n^2 - \omega^2} \tag{1.143}$$

in fact, multiplying this equation by $\left(\mathbb{C} - \omega^2 \mathbb{1}\right)$ on the right yields:

$$\mathbb{1} = \sum_n \frac{\mathbf{e}_n \mathbf{e}_n^\dagger}{\Omega_n^2 - \omega^2} (\mathbb{C} - \omega^2 \mathbb{1}) = \tag{1.144}$$

$$= \sum_n \left\{ \frac{\mathbf{e}_n (\mathbf{e}_n^\dagger \mathbb{C})}{\Omega_n^2 - \omega^2} - \omega^2 \frac{\mathbf{e}_n \mathbf{e}_n^\dagger}{\Omega_n^2 - \omega^2} \right\} =$$

$$= \sum_n (\Omega_n^2 - \omega^2) \frac{\mathbf{e}_n \mathbf{e}_n^\dagger}{\Omega_n^2 - \omega^2} = \sum_n \mathbf{e}_n \mathbf{e}_n^\dagger$$

Using this relation, Eq. (1.142) can be set in the more convenient form

$$(X + Y) = 2 \sum_n (\mathbb{A} - \mathbb{B})^{\frac{1}{2}} \frac{\mathbf{e}_n \mathbf{e}^\dagger_n}{\Omega_n^2 - \omega^2} (\mathbb{A} - \mathbb{B})^{\frac{1}{2}} \, v_1 \tag{1.145}$$

in which the vector $X + Y$ is expressed just in terms of the Casida vectors (which are assumed to be computed by means of Eq.(1.132)) and the external field components on the Khon Sham orbitals v_{1ia}. The solution of the problem is finally found just noting that, Eq. (1.128) reduces in this case to

$$n_1(\mathbf{r}, \omega) = \sum_{ia} (X + Y)_{ia}(\omega) \, \Phi_{ia}(\mathbf{r}) \tag{1.146}$$

So, the solution of the density density linear response equation is completely determined now for every external potential $v_1(\mathbf{r}, t)$ once the Casida equations are solved.

1.11 Oscillator strength and auxiliary excited states wave functions

Being now the density linear response problem morally completely solved for every external potential, the results obtained can be used to compute the response properties of an electrons system in cases of interest. In particular, this section will be dedicated to the calculation of photoresponse properties, one of the central tasks of linear response theory, but the results obtained will lead to a general equation to compute in a formally exact way the matrix element of one electron operators coupling the ground state with the general n^{th} excited state of the many body unperturbed system.

To start with, the first-order induced electronic dipole polarization tensor $\underline{\alpha}$ will be analyzed. It is defined by the relation:

$$\mathbf{p}_1(t) = \int dt' \underline{\alpha}(t-t')\mathbf{E}(t') \tag{1.147}$$

where $\mathbf{p}(t) \equiv \int \mathbf{r}\, n(\mathbf{r},t)\, d^3\mathbf{r}$ and $\mathbf{E}$ is a time dependent electric field. Fourier transforming this relation one gets:

$$\mathbf{p}_1(\omega) = \int dt\, e^{i\omega t}\, \mathbf{p}_1(t) = \int d^3\mathbf{r}\, \mathbf{r}\, n_1(\mathbf{r},\omega) = \underline{\alpha}(\omega)\mathbf{E}(\omega) \tag{1.148}$$

where again, for the sake of simplicity, the Fourier transformed functions are just distinguished by their ω argument. Assuming, with no loss of generality, that the electric field is a monochromatic wave linearly polarized long the $\mathbf{z}$ axis, i.e. $\mathbf{E}(t) = \hat{\mathbf{z}}\,\mathcal{E}\sin\omega t$, this case reduces to the one discussed for the application of the linear response theory in which the system is described with a time independent Hamiltonian and an external time dependent perturbing potential

$$v_1(\mathbf{r},t) = \mathcal{E}z\sin\omega t \tag{1.149}$$

Being moreover the field polarized on the z axis Eq. (1.148) can be set in the scalar form:

$$p_{1z}(\omega) = \alpha_{zz}(\omega)\mathcal{E} \tag{1.150}$$

and hence from the second right hand side of Eq. (1.148)

$$\alpha_{zz}(\omega) = \frac{1}{\mathcal{E}}\int d^3\mathbf{r}\, zn_1(\mathbf{r},\omega) = \tag{1.151}$$

$$= \int d^3\mathbf{r}\int d^3\mathbf{r}'\, z\,\chi(\mathbf{r},\mathbf{r}',\omega)z' \tag{1.152}$$

It can be used now that the one-electron operator $\hat{z} = \sum_{i=1}^{N}\hat{z}_i$, written in terms of the usual density operator defined in Eq. (1.7) is

$$\hat{z} = \int dz\, z\, \hat{n}(\mathbf{r}) \tag{1.153}$$

and its matrix elements can be written as:

$$\langle\psi|\hat{z}|\phi\rangle = \int d\mathbf{r}\, z\, \langle\psi|\hat{n}(\mathbf{r})|\phi\rangle \tag{1.154}$$

The Lehmann representation of the frequency dependent response function of Eq. (1.84) can be used now in Eq. (151), together with Eq. (1.154), to give:

$$\alpha_{zz}(\omega) = \sum_n |\langle\psi_0|\hat{z}|\psi_n\rangle|^2 \left(\frac{1}{\omega-\Omega_n} - \frac{1}{\omega+\Omega_n}\right) = \tag{1.155}$$

$$= \sum_n \frac{2\,\Omega_n\,|\langle\psi_0|\hat{z}|\psi_n\rangle|^2}{\omega^2 - \Omega_n^2}$$

where, it has been used the fact that:

$$\int d\mathbf{r} \int d\mathbf{r}' \; zz' \; \langle\psi_0|\hat{n}(\mathbf{r})|\psi_n\rangle \; \langle\psi_n|\hat{n}(\mathbf{r}')|\psi_0\rangle = \tag{1.156}$$

$$= \int d\mathbf{r} \int d\mathbf{r}' \; zz' \; \langle\psi_0|\hat{n}(\mathbf{r}')|\psi_n\rangle \; \langle\psi_n|\hat{n}(\mathbf{r})|\psi_0\rangle =$$

$$= |\langle\psi_0|\hat{z}|\psi_n\rangle|^2$$

The average polarizability, defined as

$$\bar{\alpha}(\omega) \equiv \frac{1}{3} \left(\alpha_{xx}(\omega) + \alpha_{yy}(\omega) + \alpha_{zz}(\omega) \right) \tag{1.157}$$

can be now written using Eq. (155) as

$$\bar{\alpha}(\omega) = \sum_n \frac{f_n}{\omega^2 - \Omega_n^2} \tag{1.158}$$

where the quantities f_n are dimensionless quantity expressing the strength of a given electronic transition called oscillator strengths[11]. Their value is:

$$f_n \equiv \frac{2}{3} \; \Omega_n \sum_{i=x,y,z} |\langle\psi_0|\hat{r}_i|\psi_n\rangle|^2 \tag{1.159}$$

The results of last section, can be now applied to rewrite Eq. (1.151) using the form of Eq. (1.146) for $n_1(\mathbf{r},\omega)$ and the formalism of the excitations space:

$$\alpha_{zz}(\omega) = \frac{1}{\mathcal{E}} \int d^3\mathbf{r} \; z \; n_1(\mathbf{r},\omega) = \frac{1}{\mathcal{E}} \sum_{ia} (X+Y)_{ia} \int d^3\mathbf{r} \; z \; \Phi_{ia}(\mathbf{r}) = \frac{1}{\mathcal{E}} \; z^\dagger (X+Y) \tag{1.160}$$

and hence, using Eq.(1.145) for $(X+Y)$ and the fact that, in this case, $v_1(\mathbf{r},\omega) = z \, \mathcal{E}$,

$$\alpha_{zz}(\omega) = 2 \sum_n z^\dagger (\mathbb{A} - \mathbb{B})^{\frac{1}{2}} \; \frac{\mathbf{e}_n \mathbf{e}^\dagger_n}{\Omega_n^2 - \omega^2} (\mathbb{A} - \mathbb{B})^{\frac{1}{2}} \; z = \tag{1.161}$$

$$= 2 \sum_n \frac{|z^\dagger (\mathbb{A} - \mathbb{B})^{\frac{1}{2}} \mathbf{e}_n|^2}{\Omega_n^2 - \omega^2}$$

So, from a direct comparaison of Eq.(1.162) and Eq.(1.158) it is immediately seen that the oscillator strengths have been found again and their value is:

$$f_n = \frac{2}{3} \sum_{i=x,y,z} \left| \mathbf{r}_i^\dagger (\mathbb{A} - \mathbb{B})^{\frac{1}{2}} \mathbf{e}_n \right|^2 \tag{1.162}$$

---

[11]To be consistent with the existing literature on these topics, the symbol adopted for the oscillator strengths, is the same as the one adopted for the occupation numbers defined in Eq. (1.107) and this could generate some confusion; nevertheless the two quantities can be distinguished from the subscript (n in the case of the oscillator strengths and $\{ijhk...\}$ in the case of occupation numbers).

This is a very useful relation that allows direct computation and comparison with experimental data coming from photo-absorption measurements. But its implications are even stronger: equating the two expression obtained for the oscillator strengths in Eq. (1.162) and Eq. (1.159) one get:

$$\frac{2}{3} \, \Omega_n \sum_{i=x,y,z} |\langle\psi_0|\hat{r}_i|\psi_n\rangle|^2 = \frac{2}{3} \sum_{i=x,y,z} \left|\mathbf{r}_i^\dagger(\mathbb{A}-\mathbb{B})^{\frac{1}{2}}\mathbf{e}_n\right|^2 \qquad (1.163)$$

which implies:

$$z^\dagger(\mathbb{A}-\mathbb{B})^{\frac{1}{2}}\mathbf{e}_n = \sqrt{\Omega_n} \, \langle\psi_0|\hat{z}|\psi_n\rangle \qquad (1.164)$$

and going back to the component form:

$$\sum_{ia} z_{ia} \left[(\mathbb{A}-\mathbb{B})^{\frac{1}{2}}\mathbf{e}_n\right]_{ia} = \sqrt{\Omega_n} \, \langle\psi_0|\hat{z}|\psi_n\rangle \qquad (1.165)$$

It has to be noted now that for the derivation of this equation, and in particular for the derivation of the two expression for the oscillator strengths in Eq. (1.162) and Eq. (1.159), the density linear response to the external time dependent potential $v_1 = \mathcal{E}z\sin\omega t$ has been considered. The fact that for this particular choice, v_1 is coupled with the position operator is completely arbitrary and the same procedure could have been applied for every other potential of the kind $\tilde{v}_1 = o(\mathbf{r})\sin\omega t$ associated to the time dependent external filed operator $\tilde{V}_1(t) = \hat{O}\sin\omega t$ where, as usual it has been used the connection $\hat{O} = \int d\mathbf{r} \, o(\mathbf{r})\hat{n}(\mathbf{r})$; in this case the following relation would have been found:

$$\sum_{ia} o_{ia} \left[(\mathbb{A}-\mathbb{B})^{\frac{1}{2}}\mathbf{e}_n\right]_{ia} = \sqrt{\Omega_n} \, \langle\psi_0|\hat{O}|\psi_n\rangle \qquad (1.166)$$

This is the equation claimed at the beginning of this section. It allows the direct computation of many body systems excited states properties without any knowledge about the form of the many body excited states wave functions: in fact thanks to this result, not only the excitation energies but now also every one-electron operator matrix elements involving the ground state and the general n^{th} excited state wave functions can be practically computed. All what is necessary is to solve the Casida equations. Moreover, from a practical point of view this equation gives also a direct way to compare wave function based theories and TDDFT results in the calculation of many body excited states properties.

Once derived the relations (1.166) in 1995, Casida [15] defined the following auxiliary excited states wave functions can be now defined as

$$\psi_{sn}(\mathbf{x}_1,...,\mathbf{x}_N) = \sum_{ia} c_{ia}^{(n)} \psi_{s(i\to a)}(\mathbf{x}_1,...,\mathbf{x}_N) \qquad (1.167)$$

where[12]

---

[12]The nomenclature for the single excitation non interacting excited states wave functions has been given in section 1.10.

$$c_{ia}^{(n)} = \frac{\left[(\mathbb{A} - \mathbb{B})^{\frac{1}{2}} \mathbf{e}_n\right]_{ia}}{\sqrt{\Omega_n}} \tag{1.168}$$

These wave functions, even if representing a very poor approximation for the interacting excited states wave functions, thanks to Eq. (1.166) have the property that their expectation values of single particle observables are equal to the one computed on the real many body wave functions. To see this it is sufficient to substitute back the definitions (1.167) and (1.168) in the matrix element $\langle \psi_{s0} | \hat{\mathcal{O}} | \psi_{sn} \rangle$, use the fact that, according to the rules reported in Eq. (1.104) for the matrix element of one electron operators between Slater determinants, $\langle \psi_{s0} | \hat{O} | \psi_{s(i \to a)} \rangle = o_{ia} = \langle \phi_i | \hat{O} | \phi_a \rangle$ and compare the result with Eq. (1.166):

$$\langle \psi_{s0} | \hat{\mathcal{O}} | \psi_{sn} \rangle = \frac{1}{\sqrt{\Omega_n}} \sum_{ia} o_{ia} \left[(\mathbb{A} - \mathbb{B})^{\frac{1}{2}} \mathbf{e}_n\right]_{ia} = \langle \psi_0 | \hat{O} | \psi_n \rangle \tag{1.169}$$

It is important to stress that relations (1.169) are formally exact and that the use the auxiliary wave functions is just a way to use Eq. (1.166). As a natural extension of Eq. (1.169) one could think to compute also matrix elements of the kind $\langle \psi_m | \hat{O} | \psi_n \rangle$ by means of auxiliary wave functions assuming approximately true that

$$\langle \psi_m | \hat{O} | \psi_n \rangle \simeq \langle \psi_{sm} | \hat{O} | \psi_{sn} \rangle \tag{1.170}$$

Recently Tavernelli [17] has formally showed, by means of second order perturbation theory, that relation (1.170) is effectively a good approximation of the relative excited states matrix element computed with the exact many bodies wave functions.

1.12 Performances of the adiabatic approximation for the Casida equations

The quality of a concrete numerical calculation made using the Casida equation with the adiabatic approximation for the xc kernel is very difficult to asses a priori. In fact, before arriving to the adiabatic approximation for the xc kernel, three other underlying approximation have to be taken into account: i) the ground state density calculation is approximate because so is the static potential $v_{xc}[n_0]$ ii) the Kohn-Sham orbitals can be expanded in just a finite subset (even if, in principle, ad libitum large) of a complete set of functions and iii) also the Casida equations are exact only if the full set of infinite occupied and unoccupied Kohn-Sham orbitals is used.

Generally speaking however good level of accuracy from the Casida formalism can be expected for the cases in which $\Omega_n \simeq \omega_{ai}$ for some given orbitals pair ϕ_a and ϕ_i: in these cases in fact, the role of the coupling matrix elements $K_{ii'aa'}$ is just to give a small correction to the "bare" non interacting excitation energies ω_{ai} (see Eq. (1.120)).

Several benchmark studies on small and medium size molecules have shown that TDDFT calculation give correct singlet valence excitation energies predictions within an accuracy of 0.2-0.5 eV. Unfortunately however, it can happen that the accuracy of a calculation depends strongly on the type of approximate xc potential (from which the xc kernel is derived) and

again, as in the static DFT case, the choice of the "best" functional could be a daunting task. Much help in this comes from the fact that it has been observed that the Jacobs ladder subdivision of the different class of xc potential discussed in section 1.5 is effective (even if just on average) and gives good indications for the prediction of the accuracy for a functional of a given kind [18].

The emerging picture is hence the following: the TDDFT approach for the calculation of excitation energies and excited states properties of molecules can represent an important tool, but can not be used as a black box. Several possible difficulties can come from many sides and the risk to get a failure is always present. However, thanks to the wealth of data collected over the years with tests on different xc functionals, it is nowadays known where standard xc functional are expected to work and at what level of accuracy; on the other hand, the major sources of failures are also known and there are intense efforts under way to find cures for these problems.

Chapter 2

Molecular Dynamics and zero temperature simulation methods

In this chapter the background concepts and derivations for the description of molecules in solution will be provided giving the conceptual root underlying the calculations and simulations that will be discussed in the following chapters.

Since the early development of quantum mechanics in the 1920s it was appreciated that the Schrödinger equation, together with the probabilistic interpretation of its solutions provided a powerful tool for tackling an enormous variety of problems in the chemical and physical realms. Stepping from the description of atoms to the one of molecules, the overwhelming complexity associated to the solution of a many-body problem involving the coordinates of all the electrons and all nuclei of the considered molecule has however to be faced. Its solution can be approached using the fact that nuclei and electrons have very different masses and their motion can be adiabatically separated. This concept was first introduced by Born and Oppenheimer in 1927. The adiabatic Born-Oppenheimer (BO) ansatz in fact is an excellent approximation to describe the behavior of molecules in the vicinity of their ground state equilibrium configuration. It represents the conceptual basis for the computational description of molecules and consists in the foundational approach underlying the application of Molecular Dynamics (MD) for the study of large molecular systems in solution. However the Born-Oppenheimer approximation is not uniformly applicable and its breakdown has to be taken into account when the nonadiabatic couplings influence the dynamics of a molecular system. Their effects have to be taken into account to explain many phenomena including molecular radiationless electronic decays and photochemical reactions, central topics of the present work.

The Chapter is structured as follows: after the introduction of the Hamiltonian operator for a molecule, the Born-Huang ansatz will be introduced to discuss the BO approximation and the concept of Potential Energy Surface (PES) highlighting the effects implied by the presence of the neglected nonadiabatic terms. At this point, assuming the knowledge of all

relevant PESs of a given molecule acquired, the methods for the calculations of equilibrium molecular geometries and characterization of reaction paths will be discussed. Then the different Molecular Dynamics (MD) methods used in this work will be reviewed. In particular, apart from conventional classical and *ab initio* BO MD schemes, also nonadiabatic MD will be discussed using the Trajectory Surface Hopping (TSH) method, an *ab initio* MD based approach useful to model the dynamics of a molecule in the electronic excited states.

2.1 Molecular Schrödinger Equation

In the following, the attention will be focused on situations in which atoms can be considered as made of charged pointlike quantum particles, nuclei and electrons, and are sufficiently close to form stable molecules. Let us consider hence the general molecule composed of N_N atoms having atomic numbers $\{Z_1 \cdots Z_{N_N}\}$; the Cartesian coordinates and momenta of the N_e electrons will be denoted as $(\mathbf{r}_j, \mathbf{p}_j)$ and of the N_N nuclei as $(\mathbf{R}_n, \mathbf{P}_n)$. The Hamiltonian operator of such a molecule has the form:

$$\hat{H}_{mol} = \hat{T}_N + \hat{T}_e + \hat{W}_{ee} + \hat{V}_{eN} + \hat{V}_{NN} \tag{2.1}$$

in which $\hat{T}_e$ is the kinetic energy of the electrons and is given by

$$\hat{T}_e = \sum_{j=1}^{N_e} \frac{\hat{p}_j^2}{2m_e} \tag{2.2}$$

where m_e is the mass of electrons and the usual correspondence principle $\hat{\mathbf{p}} \equiv -i\,\hbar\overrightarrow{\nabla}$ is intended. In the same way, $\hat{T}_N$, the kinetic energy of the nuclei, is given by

$$\hat{T}_N = \sum_{a=1}^{N_N} \frac{\hat{P}_a^2}{2M_a} \tag{2.3}$$

M_a being the mass of the ath nucleus. The other terms in the Hamiltonian take into account for the Coulomb interactions between all the particles: the pair repulsive interaction between electrons

$$W_{ee} = \frac{1}{2} \sum_{i \neq j} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{2.4}$$

between nuclei

$$V_{NN} = \frac{1}{2} \sum_{a \neq b} \frac{Z_a Z_b e^2}{|\mathbf{R}_a - \mathbf{R}_b|} \tag{2.5}$$

and the attractive interaction between electrons and nuclei

$$V_{eN} = -\sum_{j,a} \frac{Z_a e^2}{|\mathbf{r}_j - \mathbf{R}_a|} \tag{2.6}$$

Within this model, all the quantum mechanical stationary properties of a molecule can be extracted from the solution of the time-independent Schrödinger equation

$$\hat{H}_{mol}\boldsymbol{\Psi}_n(r,\sigma,R) = \mathcal{E}_n \ \boldsymbol{\Psi}_n(r,\sigma,R) \tag{2.7}$$

and its time dependent version

$$i\hbar\frac{\partial}{\partial t}\boldsymbol{\Psi}(r,\sigma,R,t) = \hat{H}_{mol}\boldsymbol{\Psi}_0(r,\sigma,R,t) \tag{2.8}$$

where the multi-index notation $r \equiv (\mathbf{r}_1....\mathbf{r}_{N_e})$, $\sigma \equiv (\sigma_1....\sigma_{N_e})$, $R \equiv (\mathbf{R}_1....\mathbf{R}_{N_N})$ for the electrons coordinates, electrons spin and nuclei coordinates has respectively been used. Since the Hamiltonian does not depend on spin, for the total wave function the following ansatz can be used:

$$\boldsymbol{\Psi}(r,\sigma,R) = \Psi(r,R)\zeta(\sigma) \tag{2.9}$$

where $\zeta(\sigma) = \zeta_1(\sigma_1)...\zeta_{N_e}(\sigma_{N_e})$ is the electronic spin function and the individual electron spin states $\zeta_j(\sigma_j)$ describe electrons whose spin is parallel or antiparallel with respect to some direction in the coordinate space.

Equations 2.7 and 2.8 are the foundation of the description of molecular structure and reactivity and will serve as reference equation to be (morally) solved in the following.

2.2 The Born-Huang ansatz and its implications

The large mass difference between electrons and nuclei $\left(\frac{m_e}{M_n} < 10^{-3}\right)$ implies that the electronic degrees of freedom can be considered, with a good level of accuracy, to be able to respond almost instantaneously to the changes in the nuclear configuration. This approximation, called adiabatic approximation, efficiently allows, in many situations, to separate adiabatically the motion of these two molecular subsystems. To see how this works it is convenient to define the so called electronic Hamiltonian:

$$\hat{H}_{el}(R) = \hat{H}_{mol} - \hat{T}_N - \hat{V}_{NN} = \hat{T}_e + \hat{W}_{ee} + \hat{V}_{eN}(R) \tag{2.10}$$

in which the dependence on the nuclear coordinates is just carried parametrically in the external potential operator $\hat{V}_{eN}(R)$. The eigenfunctions corresponding to this operator are called adiabatic electronic states and are also parametrically dependent on the nuclear coordinates. Leaving aside the electron's spin, these states can hence be defined as

$$\hat{H}_{el}(R)\phi_n(r;R) = E_n(R) \ \phi_n(r;R) \tag{2.11}$$

The set of functions $\{\phi_n(r;R)\}_{n\in\mathbb{N}}$ is a complete basis in the electronic Hilbert space for every given molecular nuclear configuration that for every molecular system can be chosen to be real without loss of generality. Hence the molecular wave function can be expanded in this basis set as follows:

$$\Psi(r,R) = \sum_n \chi_n(R)\phi_n(r;R) \tag{2.12}$$

It is worth to stress that in this expansion, known as the Born-Huang ansatz for the molecular wave function, the coefficients χ_n are functions of the nuclear configurations. An equation for their determination can be derived inserting this ansatz in Eq. (2.7):

$$
\begin{aligned}
\hat{H}_{mol}\Psi(r,R) &= (\hat{H}_{el} + \hat{T}_N + \hat{V}_{NN})\Psi(r,R) = \\
&= \sum_n \ [E_n(R) + V_{NN}(R)]\,\chi_n(R)\phi_n(r;R) + \\
&+ \sum_n \hat{T}_N(\chi_n(R)\phi_n(r;R)) = \\
&= \mathcal{E} \sum_n \ \chi_n(R)\phi_n(r;R)
\end{aligned}
\tag{2.13}
$$

Multiplying Eq. (2.13) by $\phi_m(r;R)$, integrating over all electronic coordinates and using the orthogonality of the adiabatic basis yields:

$$
\begin{aligned}
\int dr \ \phi_m(r;R)\hat{H}_{mol}\Psi(r,R) &= [E_m(R) + V_{NN}(R)]\chi_m(R) + \\
&+ \sum_n \int dr \ \phi_m(r;R)\hat{T}_N\phi_n(r;R)\chi_n(R) = \\
&= \mathcal{E} \ \chi_m(R)
\end{aligned}
\tag{2.14}
$$

Since the electronic adiabatic states depend on the nuclear coordinates, the definition in Eq. (2.3) for $\hat{T}_N$, and the correspondence principle have to be used for the evaluation of the second term in the second equality of Eq. (2.13)

$$
\begin{aligned}
\hat{T}_N\phi_n(r;R)\chi_n(R) &= \sum_a \left\{ \frac{1}{2M_a}\left[\hat{P}_a^2\phi_n(r;R)\right]\chi_n(R) + \right. \\
&+ \frac{1}{M_a}\left[\hat{\mathbf{P}}_a\phi_n(r;R)\right]\hat{\mathbf{P}}_a \ \chi_n(R) + \\
&+ \left. \frac{1}{2M_a}\phi_n(r;R)\hat{P}_a^2\chi_n(R) \right\}
\end{aligned}
\tag{2.15}
$$

The last term is simply the kinetic energy operator acting on the function $\chi_a(R)$. The other terms can be combined into the so called *nonadiabaticity* operators:

$$
\begin{aligned}
\hat{\Theta}_{nm} &= \int dr \ \phi_n(r;R)\hat{T}_N\phi_m(r;R) + \\
&+ \sum_a \frac{1}{M_a}\left[\int dr\phi_n(r;R)\mathbf{P}_a\phi_m(r;R)\right]\mathbf{P}_a
\end{aligned}
\tag{2.16}
$$

Using Eq. (2.15) and (2.16), equation (2.13) for the $\{\chi_n\}$ becomes:

$$
(\hat{T}_N + E_n(R) + V_{NN}(R) + \hat{\Theta}_{nn}(R) - \mathcal{E})\chi_n(R) = -\sum_{n\neq m}\hat{\Theta}_{nm}(R)\chi_m(R)
\tag{2.17}
$$

Equation (2.17) can be interpreted as stationary Schrödinger equation for the nuclei, the collection of coefficients $\chi_n(R)$ in the Born Huang ansatz representing nuclear wave functions, one for each adiabatic electronic state. If the following effective potential is introduced

$$U_n(R) = E_n(R) + V_{NN}(R) + \hat{\Theta}_{nn}(R) \tag{2.18}$$

and the *nonadiabaticity* operators are neglected for a moment, Eq. (2.17) takes the standard form of the stationary Schrödinger equation in which the function $U_n(R)$, an hypersurface in the space of nuclear coordinates, plays the role of potential energy surface (PES) on the which the nuclear motion is located. Following this point of view, the presence of the *nonadiabaticity* operators has the effect to couple the equations relative to each nuclear wave function $\chi_n(R)$.

It is worth to stress that all the derivations made so far are exact, i.e. Eq. (2.17) formally correspond to Eq. (2.7) and no approximations have been made. Addressing the direct solution of Eq. (2.17) however, is a task far too complex in almost all practical cases but the adiabatic approximation can be now used to neglect all the terms involving the derivative of adiabatic states with respect to the nuclear coordinates:

$$\frac{\partial}{\partial R}\phi_n(r; R) \simeq 0 \tag{2.19}$$

this is in fact true in all the cases in which the electrons can be assumed to be moving so fast compared to the nuclei to see them almost frozen. In this case all the non adiabatic terms become zero and Eq. (2.17) becomes:

$$\hat{H}_n\chi_{nM}(R) = (\hat{T}_N + U_n^{(ad)}(R))\chi_{nM}(R) = \mathcal{E}_{nM}\chi_{nM}(R) \tag{2.20}$$

where $\hat{H}_n$ defines the *nuclear Hamiltonian* relative to the adiabatic electronic state $\phi_n(r; R)$, the index M labels the possible nuclear (vibrational) related states and

$$U_n^{(ad)}(R) = E_n(R) + V_{NN}(R) \tag{2.21}$$

The dynamics of the nuclear wave functions as regulated by the time dependent Schrödinger equation 2.8 can be written now as:

$$i\hbar\frac{\partial}{\partial t}\chi_n(R, t) = \hat{H}_n\chi_n(R) + \sum_{n \neq m} \hat{\Theta}_{nm}(R)\chi_m(R) \tag{2.22}$$

that in the adiabatic approximation simply becomes:

$$i\hbar\frac{\partial}{\partial t}\chi_n(R, t) = \hat{H}_n\chi_n(R) \tag{2.23}$$

while the total *adiabatic* wave function simply becomes:

$$\Psi_{nM}^{(ad)}(r, R, t) = \chi_{nM}(R, t)\ \phi_n(r; R) \tag{2.24}$$

which is the standard form of the Born Oppenheimer approximation. As it will be clear in the following however, cases in which this approximation is no longer valid are likely to be

faced. In these cases the non adiabatic couplings effects can no longer be neglected and the electronic and nuclear motions are not longer adiabatically separable.

In next sections. assuming the knowledge of at least an electronic adiabatic state (e.g. the ground state) as known quantities, the computational methods used in this work to approximately solve equation 2.8 will be discussed.

2.3 Zero temperature exploration of Potential Energy Surfaces

Much chemical informations about the behavior of a given molecular system can be extracted without really solving eq. 2.8 but just analyzing the characteristics of a given state PES. In this section some of these fundamental chemical and physical quantities will be derived and discussed assuming to have somehow acquired the knowledge of all necessary adiabatic PESs. To this end it is convenient to define the gradient of the PES as

$$\nabla U_n^{(ad)}(R) \equiv \left\{ \frac{\partial}{\partial R_1} U_n^{(ad)}(R), \dots, \frac{\partial}{\partial R_{3N_N}} U_n^{(ad)}(R) \right\} \tag{2.25}$$

The negative of this vector points in the direction of the steepest descent of the PES and corresponds to the force acting on a classical particle in the PES. An other quantity of primary interest is the $3N_N \times 3N_N$ force constant matrix, or Hessian matrix, the elements of the which are defined as

$$K_{ab} \equiv \frac{\partial^2 U_n^{(ad)}(R)}{\partial R_a \partial R_b} \qquad \{a, b = 1, \dots N_N\} \tag{2.26}$$

The points $\bar{R}$ in the nuclear configurational space for the which the gradient of a given PES vanishes

$$\nabla U_a^{(ad)}(R)\Big|_{R=\bar{R}} = 0 \tag{2.27}$$

are called *stationary* points. The problem of the search of stationary points for a given real multivariable function via iterative methods has a long history going back to Newton and Gauss. Many efficient numerical algorithm are available nowadays to locate stationary points of PESs with many degrees of freedom. These algorithms use as input parameters

- An initial guess point in the nuclear configurational space (i.e. a starting nuclear structure)

- The PES value at each iteration step

- The PES gradient and hessian matrix (or some its approximated form, due to the big computational effort required for its calculation on every intermediate step).

Once a stationary point on a PES has been found, the nature of the PES in the vicinity of a stationary point can be investigated with a direct computation of the eigenvalues of the Hessian matrix. For nonplanar molecules, six eigenvalues will equal to zero reflecting
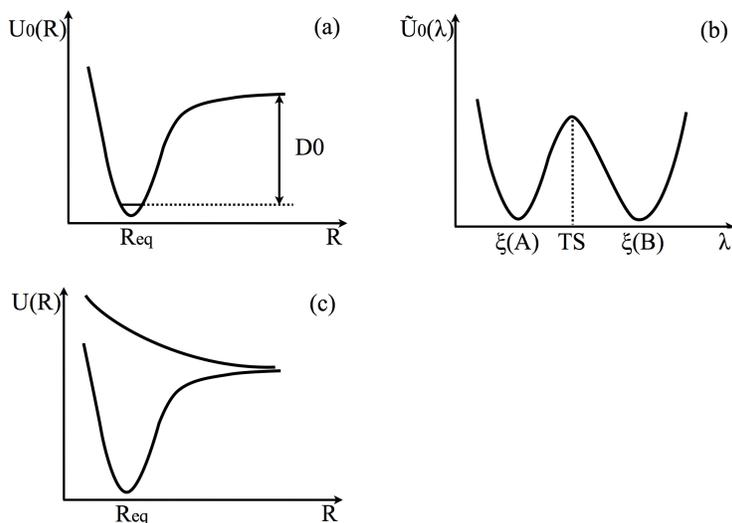
Figure 2.1: Schematic view of typical PESs for some explicative cases are reported in the different panels. In (a) the ground state PES of a diatomic molecule in depicted: R is the bond distance, R_{eq} the equilibrium distance between the two atoms and D_0 the dissociation energy. In (b) is taken in consideration a typical case in which a one dimensional reaction coordinate is taken into account to characterize an isomerization reaction; In (c) a typical case for ground and first excited state PESs of a diatomic molecule is reported. If the molecule is excited on the first electronic adiabatic state, for example, upon interaction with of an external light field, dissociation can occur.

the fact that the three independent global rotations and translations of the center of mass will leave the energy unchanged. In fact there are only $3N_N - 6$ independent coordinates (called internal coordinates) necessary to describe a given nuclear structure. If the remaining eigenvalues of the Hessian matrix are all positive (negative), then a minimum (maximum) of the potential energy surface has been found.

Stable states of molecular systems correspond to global and local minima on their ground state PES. As explicative example a diatomic molecule can be taken in consideration. In Figure 6.10a is reported a typical ground state PES as function of the bond length in a diatomic molecule. The minimum of $U_0(R)$ at R_{eq} gives the equilibrium distance between the two atoms. As a consequence of the quantum mechanical zero-point energy motion, the lowest possible energy of the molecule is above the bottom of the potential minimum of a quantity corresponding to the lowest nuclear bound state energy in this potential. The molecule is said to be stable if the difference between this zero-point energy and the energy of the separate atoms, $U_0(\infty)$, is finite. This energy is called dissociation energy (D_0 in Figure 6.10a).

For polyatomic molecules the presence of multiple minima on the ground state PES is often encountered. These minima in $U_0(R)$ correspond to different isomers of the given

molecule[1]. To change isomeric form, the molecule has to pass a maximum of the potential curve that corresponds to a saddle point of $U_0(R)$ in which the Hessian matrix will have just one negative eigenvalue. These points are called *transition states.* A standard example of such situation is the umbrella vibration of NH_3 but more generally, it happens for every molecule in which intramolecular hydrogen transfer are possible.

Once the two minima of the ground state PES, that can be conveniently identified as A and B, have been found, in order to learn more about the path of an isomerization reaction, the PES values along possible paths connecting the two stable states can be investigated. For polyatomic molecules it is of course not desirable (an not computationally feasible) to compute the PES as a function of all internal coordinates. Quite often however only a few coordinates are important to describe the reaction. Then it becomes possible to take into account only the motion along a set of so called *reaction coordinates* ξ_i, a set of functions of the nuclear coordinated

$$\xi_i = \xi_i(R) \tag{2.28}$$

by means of the which the sates A and B are unambiguously identified by two distinct values of the reaction coordinates ξ_i, namely $\xi_i(A)$ and $\xi_i(B)$. The *minimum energy path* connecting the states A and B, the minimum potential barrier to be crossed and its associated transition state can be then for example identified by means of a set of constrained stationary point search aimed at reconstructing the function

$$\tilde{U}_0^{(ad)}(\lambda_i) \equiv \min_{\xi_i(R)=\lambda_i} U_0^{(ad)}(R) \quad \lambda_i \in \{\xi_i(A), \xi_i(B)\} \tag{2.29}$$

As explicative example in Figure 6.10b is reported the case in which a one dimensional reaction coordinate is taken into account.

Finally, a case typically encountered when excited states PESs are considered can be mentioned. In Figure 6.10c is schematically reported the ground state and a possible excited adiabatic PESs for a diatomic molecule as function of the atoms bond length. In this case the excited state PES has no minimum. This implies that if the nuclei, e.g. after the absorption of a photon (see the following), start to "ride" the excited surface, the molecule will simply dissociate (photodissociation).

2.4 Molecular Dynamics on ground and excited PES

To model the nuclear motion in a way consistent with the general molecular model introduced in section 2.1, Eq. (2.17) should be taken into account. As discussed in section 2.2, if the Born Oppenheimer approximation is adopted, all the non-diagonal terms of the nonadiabaticity operators $\hat{\Theta}_{nm}$ defined in Eq. (2.16) are neglected and Eq. (2.23) instead of Eq. (2.17) can be considered. To evaluate quantitatively the relevance of the neglected terms in the description of nuclear motion the amplitude of their leading terms, involving the first derivative of the adiabatic states with respect to the nuclear coordinates, can be examined:

---

[1]the case of a chemical reaction involving more than one molecule is equivalent to this case if the coordinates of all the molecules involved are taken in consideration
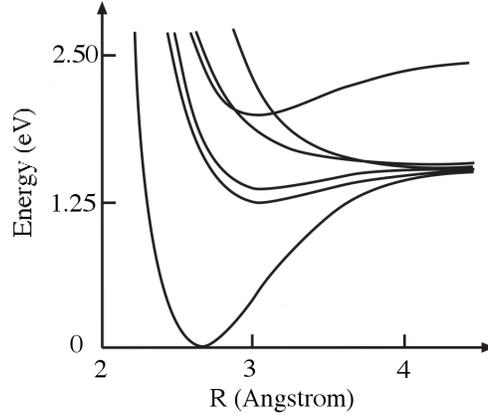
Figure 2.2: Potential Energy Surfaces (PESs) for a diatomic molecule (I_2). R is the distance between the two atoms and the energies are reported in eV. Note that at room temperature (300 K) $k_B T$ is 0.026 eV. Image taken from ref [31]

$$\hat{\Theta}_{nm} \sim \sum_a \frac{1}{M_a} \left[\int dr \ \phi_n(r;R) \mathbf{P}_a \phi_m(r;R) \right] \mathbf{P}_a \sim$$

$$\sim \sum_a \frac{1}{M_a} \int dr \ \phi_n(r;R) \nabla_{R_a} \ \phi_m(r;R) \ \nabla_{R_a} \equiv \sum_a \frac{1}{M_a} \mathbf{\Theta}^a_{nm} \ \cdot \nabla_{R_a} \tag{2.30}$$

where the terms $\mathbf{\Theta}^a_{nm}$, called *nonadiabatic coupling vectors*, are

$$\mathbf{\Theta}^a_{nm} = \int dr \ \phi_n(r;R) \nabla_{R_a} \ \phi_m(r;R) \equiv \langle \phi_n | \nabla_a | \phi_m \rangle \tag{2.31}$$

In the last equality of Eq. (2.30) the Dirac shorthand notation has been used together with the notation $\nabla_{R_a} \equiv \nabla_a$. To estimate the amplitudes of the coefficients Θ_{nm}, considering that by hypothesis $n \neq m$ and hence $< \phi_n | \ \hat{H}_{el} \ | \phi_m > = 0$, the following equality can be used:

$$0 = \nabla_a \langle \phi_n | \hat{H}_{el} | \phi_m \rangle = \langle \phi_n | \nabla_a \hat{H}_{el} | \phi_m \rangle + E_n \langle \phi_n | \nabla_a | \phi_m \rangle + E_m \langle \phi_n | \nabla_a^\dagger | \phi_m \rangle \tag{2.32}$$

so that, considering that $\nabla_a^\dagger = -\nabla_a$ the following alternative form for the nonadiabatic vectors is derived

$$\mathbf{\Theta}^a_{nm} = \frac{1}{E_n - E_m} \langle \phi_n | \nabla_a \hat{H}_{el} | \phi_m \rangle \tag{2.33}$$

Now, as explicative case, the PESs of a diatomic molecule reported in Figure 2.2 (I_2 in this case) as function of the atoms bond distance can be taken in consideration. In

particular, assuming the nuclei in a bound state of the ground state PES $\chi_{0M}(R)$ at a given time, Eq. (2.23) yields $\frac{\partial}{\partial t}\chi_n(R) = 0 \; \forall t > t_0$, i.e. the excited electronic states will never be populated. This is not true however, if the nonadiabatic terms are not zero. The amplitude of their leading terms however in this case can be estimated as follows:

$$
\begin{aligned}
\hat{\Theta}_{n0}\chi_0(R) \sim & \frac{1}{M}\frac{1}{h\Omega_n} \langle\phi_n| \nabla_R\hat{H}_{el} |\phi_0\rangle \cdot \nabla_R\chi_0(R) \lesssim \\
\lesssim & \frac{1}{M}\frac{1}{h\Omega_n} \langle\phi_n| \nabla_R\hat{H}_{el} |\phi_n\rangle \cdot \nabla_R\chi_0(R) = \\
= & \frac{1}{M}\frac{1}{h\Omega_n} \nabla_R \langle\phi_n| \hat{H}_{el} |\phi_n\rangle \cdot \nabla_R\chi_0(R) = \\
= & \frac{1}{M}\frac{1}{h\Omega_n} \nabla_R E_n(R) \cdot \nabla_R\chi_0(R)
\end{aligned}
\tag{2.34}
$$

where the third equality is the Hellmann-Feynman theorem. It can be noted now that the amplitude of the electronic energy gradient $\nabla E_n(R)$ is in the order of $k_B T_0$ (T_0 being the room temperature of 300 K) per tenth of Angstrom (the typical amplitudes of nuclear vibration motions) and is of the same order of the other terms in Eq. (2.23) for $\chi_0(R)$. Instead, as it can be observed from Figure 2.2, the excitation energies for bond distances close to the equilibrium position are in the order of the eV which is $\sim 10^3 \; k_B T_0$. The leading part of the nonadiabatic couplings is hence, in normal situations, at least three order of magnitude smaller than the oder quantities involved in the time evolution equation and the adiabatic approximation in which the nuclei moves on the electronic ground state PES is excellent. This is however not true anymore even for ground state nuclear dynamics if the bond distance lies between 3.5 and 4.0 $\mathring{A}$ (e. g. during a molecular dissociation, at very high temperature or as a consequence of a collision) or if the nuclear dynamics on some excited adiabatic state PES is considered. In these last cases in fact the energy differences between the different PESs can become of the order of $k_B T_0$ and amplitudes exchange between different electronic adiabatic states have to be taken into account. The same arguments can be used in the case of bigger (polyatomic) molecules in which the excitation energies in the vicinity of the ground state equilibrium structure is around 5 eV, but in nuclear configurational space regions sufficiently far from the one of equilibrium and if the dynamics on some electronic excited state is taken into account the Born-Oppenheimer separation of the nuclear and electronic degrees of freedom could easily be not valid anymore.

Regarding the diagonal terms of the nonadiabaticity operators $\hat{\Theta}_{nn}$,

$$
\hat{\Theta}_{nn} = \int dr \; \phi_n(r;R)\hat{T}_N\phi_n(r;R)
\tag{2.35}
$$

(note that the second term of Eq. (2.16) is identically zero when the adiabatic states are chosen real, as it has been chosen here), it has to be mentioned that they are retained in the "pure" Born-Oppenheimer approximation and are called Born-Oppenheimer diagonal corrections [32] [33]. These terms, depending on the nuclear masses, induce an isotopic dependence of the PESs. However they represent a correction of the adiabatic PES which is almost always small enough to be negligible [34] [35] [36]. If also the diagonal corrections are neglected, the approximation for the nuclear motion is called Born-Oppenheimer adiabatic

approximation[2].

Finally, in all cases in which the Born-Oppenheimer is valid and hence Eq. (2.23) can be used to efficiently approximate the nuclear time evolution, the large masses of the nuclei can be again used to further simplify the model. In fact, to all nuclei are associated de Broglie wavelengths of the order of the percent of the bond distance at room temperature[3] and the classical limit of quantum mechanics already provides an adeguate description of the nuclear dynamical processes. In this approximation, the nuclei can be treated as classical charged point like particles whose dynamics is obtained solving the dynamical problem

$$M_a \ddot{R}_a(t) = -\nabla_a U_n(R)$$ (2.36)

in which $\{R_a(t)\}_a$ is the classical trajectory of the nuclei moving under the effect of the nth electronic PES. The average position probability distribution can be approximated with its classical equivalent concept, i.e. the average permanence time in a given configurational space point and the expectation value of a given observable can be evaluated using the ergodic hypothesis:

$$< O(R) > = \lim_{T \to \infty} \frac{1}{T} \int_0^T O(R(t)) dt$$ (2.37)

The numerical integration of the classical equations of motion (2.36) to model the dynamics of molecules is called Molecular Dynamics (MD) and, as pointed out above, can be used to sample the nuclear configurational space of molecular systems and compute the expectation values of dynamical observable. The method, originally born within theoretical physics in the late 1950s [37] [38] has today a vast range of applications such as chemical physics, materials science and the modelling of biomolecules in solution.

Two main different approaches are possible to model the electronic PES and correspond to two main types of MD simulations. In one of them, called *ab initio* MD, the PESs are approximately computed *on the fly* (i.e. for each nuclear configuration encountered in the numerical integration of the Newton's equations 2.36) solving numerically the corresponding stationary electronic Schrödinger Equation 2.11. Even if for molecules of sizes of up to some tens of atoms the electronic energies can be efficiently numerically computed (e.g. by means of the Density Functional Theory discussed in Chapter 1), the application of ab initio MD is ampered by the computational cost of the electronic structure calculations. In the other approach, called *Classical* MD, the electronic PES (usually of the ground state) is empirically approximated. Such empirical potentials used in chemistry are frequently called force fields and are constructed to model the behaviour of a molecule in the proximity of its equilibrium configuration (or, in the case in which more different isomers are possible, of one given isomer equilibrium configuration) in which the ground state PES can be approximated with quadratic functions. The parameters of the force fields are extrapolated using ab initio calculations or experiments. For systems involving more than one molecule (as it is

---

[2]Very often however the name Born-Oppenheimer approximation is also used for what is called here Born-Oppenheimer adiabatic approximation

[3]An exception is the the lightest nucleus, the hydrogen, in which the quantum nature has to be taken into account to accurately describe its dynamics. Effects of tunneling in proton transfer processes, for example, may play a not always negligible role.

practically always the case in classical MD) intramolecular interactions can be modelled taking into account the long range electrostatic and Lennard-Jones potentials[4].

As force filed example, the AMBER [41] force field (used in the present work) can be taken into account. It has the form:

$$
\begin{aligned}
U_0^{ad}(R) \simeq \sum_{Molecules} &\left\{ \sum_{B=1}^{N_B} K_B(r_B - r_{Beq})^2 + \sum_{A=1}^{N_A} K_A(\theta_A - \theta_{Aeq})^2 + \right.\\
&\left. \sum_{D=1}^{N_D} \sum_n K_{Dn}(1 + \cos{(n\omega_D + \gamma)}) \right\} + \\
&+ \sum_{ab} \left(\frac{1}{4\pi\epsilon_0} \frac{q_a q_b}{r_{ab}} \right) + \left(\frac{A_{ab}}{r_{ab}^{12}} - \frac{B_{ab}}{r_{ab}^6} \right)
\end{aligned}
\tag{2.38}
$$

where the terms in curly brackets represent the intramolecular interactions: the index B runs over the chemical bonds in the molecule, r_B is the value of a given bond distance, r_{Beq} is the equilibrium bond distance and K_B is the respective empirical constant. The same model is applied for the angles A of each consecutive set of three bonded atoms θ_A of each molecule. The last intramolecular term involves the dihedral angle ω_D of each set of four consecutive bonded atoms. Its functional form can not be trivially approximated with quadratic functions and hence the dependence of the equilibrium configuration potential on the dihedral angles is approximated considering the first terms of its Fourier series expansion[5]. The last two terms take into account the long range electrostatic Coulomb and Van der Waals intramolecular interactions and the sum runs over all pair of atoms. In particular, to take into account the Coulomb interactions, each atom is considered as charged point like particle whose charge has to be empirically assigned. In this work it has been done using the so called RESP procedure [43]. It consists in a fit of the long range electrostatic potential generated by the molecule:

$$
V(\mathbf{r}) = \frac{e}{4\pi\epsilon_0} \int d\mathbf{r}' \; \frac{1}{|\mathbf{r} - \mathbf{r}'|} \left\{ n(\mathbf{r}', \bar{R}) + \sum_a z_a \delta(\mathbf{r} - \bar{\mathbf{R}}_\mathbf{a}) \right\}
\tag{2.39}
$$

with the one corresponding to the point charge in Eq. (2.38). In Eq. (2.39) with $\bar{R} = \{\bar{\mathbf{R}}_a\}_a$ has been indicated the nuclear equilibrium structure used to derive the force field and with $n(\mathbf{r}, \bar{R})$ the electronic density associated to the adiabatic ground state electronic wave function in $\bar{\mathbf{R}}$.

Classical MD has been successfully used in the last decades to describe an overwhelming variety of situation. Of course in all cases in which the molecules can not be assumed to lie in the proximity of the ground state PES or chemical changes are involved (e.g. in all cases in which the chemical bonds structure changes) the application of empirical methods to model

---

[4]The latter takes into account the Van der Waals electrostatic interactions coming from dipole-dipole, dipole-induced dipole interactions and short range repulsive forces.

[5]The number of terms necessary to have a good representation of the potential depends from the case in exam.

the ground state PES is not trivial and, when computationally feasible, the application of ab initio methods is preferred.

The quantum effects deriving from the presence of the nonadiabatic coupling can become relevant in the dynamic of molecules far from their equilibrium structures and, as will be pointed out in the following of this section, when this happens a simple adiabatic MD approach is not anymore adequate for its description. This is particularly true when photochemical processes involving the dynamics of chromophores in the electronic excited states are taken into account. Nevertheless, in most cases the whole reaction path of isomerization reactions happening on electronic excited PESs is almost down the hill and hence even a single (nonadiabatic) MD simulation of few picoseconds can possibly be sufficient to capture the reaction mechanism. In fact, the time scale associated to the occurrence of molecular photoreactions is limited by electronic recombination phenomena (e.g. fluorescence emission or collisions with solvent molecules) that are usually fast (on the nanosecond time scale) and slow isomerization reactions involving the crossing of relatively high potential energy barriers in the excited states can be hence almost always disregarded. This consideration makes appealing the construction of an appropriate nonadiabatic algorithm even if its associated computational cost is higher than that of a normal ground state MD. In this section the most important quantum effect deriving from the breaking of the adiabatic approximation in the nuclear dynamics, namely the nuclear wave functions splittings, will be discussed together with a method, the Trajectory Surface Hopping (TSH), which allows to approximately take it into account using ab initio Born Oppenheimer MD.

To begin with, it is useful to discuss qualitatively this effect starting from a prototype example. In particular the case of a diatomic molecule whose two lowest lying PESs are reported in the inset of Figure 2.3 as function of the bond distance can be conveniently considered. In the same plot the amplitude of the nonadiabatic coupling as defined in Eq. (2.31) is also reported. As it can be noted it becomes significantly different from zero only in the small region in which the two PESs get close.

It can be imagined that, as a consequence of an electronic transition in the Frank Condon region induced by the interaction with some external electromagnetic field, the nuclear wave function starts to move on the excited PES. In Eq. (2.22) this case is realized considering that, for a time t^*, $\chi_0(R, t^*) = 0$ and $\chi_1(R, t^*) \neq 0$. For $t \gtrsim t^*$ the dynamic will be essentially adiabatic until the region of coupling between the ground state and excited PESs is encountered. At that point, the nonadiabatic couplings start to act as a "source of amplitude" for the ground state nuclear wave function χ_0. As a result, when the nuclear wave function is completely localized over the coupling zone, the amplitude of $\chi_0(R, t)$ is not negligible anymore and even if the motion is again adiabatic with good approximation, both the equations for the time evolution of $\chi_0(R, t)$ and $\chi_1(R, t)$ have to be taken into account. As a consequence, the total molecular wave function has to be now approximated as $\psi(R, t) \simeq \chi_0(R, t)\phi_0(R, t) + \chi_1(R, t)\phi_1(R, t)$ and the concept of probability for the molecule to be found on a given electronic state can be considered. In particular, using the orthogonality of the adiabatic electronic states this probability can be written as:

$$P_n^{QM} = | < \chi_n \phi_n | \psi > |^2 = \int d\mathbf{R} \ |\chi_n(R)|^2 \tag{2.40}$$

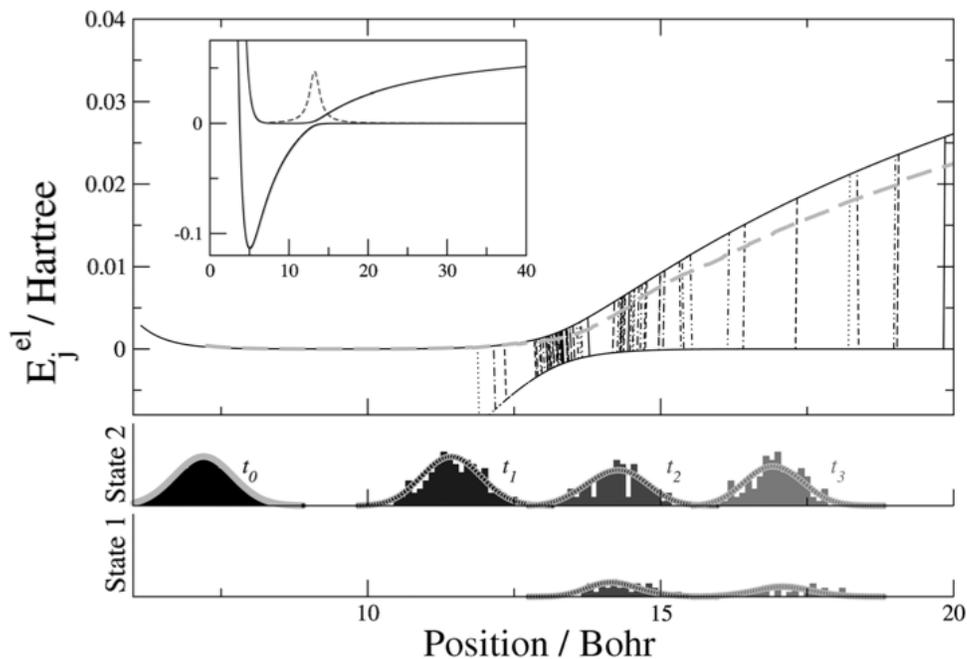This phenomenon is particularly important for the study of chemical photo-reactions. In

Figure 2.3: Nonadiabatic dynamics of NaI. Top) Ground state and excited PESs as a function of the interatomic distance. In the inset are reported the full PESs profiles (continuous lines) together with the nonadiabatic couplings (dashed line). The nonadiabatic dynamics starts in the first excited state of NaI. In TSH, all trajectories lie on an adiabatic PES, and therefore, they are only distinguishable at the instant in time at which a surface hop occurs (vertical lines). Bottom) The two lower panels show the TSH trajectories distribution (histograms) which are compared with nuclear probability density at four different times along the dynamics obtained integrating the full nonadiabatic time dependent Schoredinger equation using the Non Adiabatic Bohmian Dynamics (NABDY) method [57] (State 1 is the ground state and State 2 is the first excited state). The picture has been taken from ref [58].

fact, even if the adiabatic (and hence also classical) approximation is valid for the most part of the dynamics when these effect are neglected the results obtained can be even qualitatively wrong. To see this it is sufficient to take in consideration the case in which the evolution of a molecule on one of two different adiabatic surfaces (coupled as in example above) leads to dissociation, while the other doesn't. In this case, only one of these two cases will be captured by any adiabatic MD model.

The basic idea underlying the TSH method is to take into account the possible wave packet splittings involved in a nonadiabatic dynamic using a swarm of adiabatic trajectories. During these trajectories the nuclei are susceptible to hop from a PES to an other when nonadiabaticity regions are crossed according to a stochastic algorithm. The resulting nuclear motion approximately reproduces the nonadiabatic one provided that:

1. The nonadiabatic coupling regions are sufficiently localized in just some small "hot spots" in the nuclear configurational space and hence the adiabatic MD approximation is almost always valid.

2. The hopping probabilities are constructed by means of some ad hoc procedure in a way that if N adiabatic trajectories pass by a strong coupling zone between two given PESs n and m, the ratio of trajectories $\frac{N_i}{N}$, $i = n, m$ running on one of the two surfaces fulfills the following relation:

$$\lim_{N \to \infty} \frac{N_i}{N} = P_i^{QM} \qquad i = n, m \tag{2.41}$$

where the probability P_i^{QM} have been defined in Eq. (2.40).

3. The total energy is conserved in each trajectory taking also into account the jumps of potential energy occurring in the hops between PESs.

The first condition is hoften fulfilled in molecular systems. The PESs are in fact almost always made of adiabatic states and only when a couple of electronic states becomes almost degenerate in energy the coupling between them becomes strong, like in the exemplar case of fig 2.3.

Regarding the second condition, a large number of different possible phenomenological or "ad hoc" algorithms approximately realizing it has been proposed over the years [54] [55] . In this work, the "fewest switches" algorithm [56], a simple but powerful scheme proposed by Tully in 1990 will be used and briefly reviewed in the following. The basic idea underlying this hopping algorithm is to consider a sort of adiabaticity test of a given nuclear trajectory that can be evaluated on the fly and by means of the which the "hot coupling spots" on the PESs can be found and taken into account. To derive this test the time dependent Hamiltonian operator obtained considering the electronic Hamiltonian along a given nuclear trajectory has to be considered:

$$\hat{H}(r, t) = \hat{H}_{el}(r; R(t)) \tag{2.42}$$

together with its associated time dependent Schrödinger equation

$$ i\hbar \frac{\partial}{\partial t} \psi(r,t) = \hat{H}(r,t)\psi(r,t) \tag{2.43} $$

which defines the time dependent electronic wave function $\psi(r,t)$ that can be expanded in the basis of the adiabatic electronic states along the same trajectory:

$$ \psi(r,t) = \sum_n c_n(t)\phi_n(r; R(t)) \tag{2.44} $$

thanks to this expansion, both the time derivative and Hamiltonian operators effects can be easily evaluated. In fact:

$$ \frac{\partial}{\partial t}\psi(r,t) = \sum_n \left[\dot{c}_n(t)\phi_n(r; R(t)) + c_n(t)\frac{\partial}{\partial t}\phi_n(r; R(t)) \right] $$
$$ \hat{H}(r,t)\psi(r,t) = \sum_n c_n(t)E_n(R(t))\phi_n(r; R(t)) \tag{2.45} $$

Plugging these terms into the Schrödinger equation (2.42) the following equations regulating the time evolution of the coefficients $c_n(t)$ are obtained

$$ \dot{c}_n(t) = \sum_m d_{nm}c_n(t) \tag{2.46} $$

where

$$ d_{nm} = -\frac{i}{\hbar}\delta_{nm}E_m(R(t)) - D_{nm} \tag{2.47} $$

and

$$ D_{nm}(t) = \langle \phi_m | \partial_t \phi_n \rangle = \langle \phi_m | \nabla_a | \phi_n \rangle \, \dot{\mathbf{R}}_a(t) = \mathbf{\Theta}_{nm}^a \dot{\mathbf{R}}_a(t) \tag{2.48} $$

where it has been emphasized that these terms are directly related to the nonadiabatic coupling vectors defined in Eq. (2.31).

It can be assumed now that for a given initial time t_0 the nuclei move on a given surface $\bar{n}$ and that the expansion coefficients c_n are accordingly set as $c_n(t_0) = \delta_{n\bar{n}}$. As a consequence, for $t > t_0$, if the resulting nuclear trajectory is adiabatic, i.e. the nonadiabatic coupling vectors are essentially zero during the whole trajectory, then according to Eq. (2.46) the coefficients $c_n(t)$ will be all zero apart for the case $n = \bar{n}$ in which $c_{\bar{n}}(t) = e^{-\frac{i}{\hbar}E_{\bar{n}}t}$. This implies that $|c_n|^2(t) = \delta_{n\bar{n}} \; \forall t > t_0$. If a nonadiabatic coupling region is instead encountered during the nuclear dynamics, then the square modulus of coefficients relative to other states will start to be significantly different from zero and other electronic states will start to be populated.

In the fewest switches algorithm the hopping probabilities are constructed considering that if N adiabatic trajectories pass by a nonadiabatic coupling region between two given PESs n and m, the ratio of trajectories $\frac{N_i}{N}$, $i = n, m$ running on one of the two surfaces will asymptotically fulfill the relation:

$$\frac{N_i}{N} = |c_i|^2 \tag{2.49}$$

It is important to stress here that, even if it might seem intuitively true, it has not been found so far any rigorous derivation of the fact that $|c_i|^2 \simeq P_i^{QM}$. Nevertheless this condition can be easily realized using as ingredients just quantities derived on the fly during an MD trajectory. To show this, let us consider a given TSH trajectories ensemble formed by N equivalent systems. By hypothesis, the number of trajectories on a given state n at a time t will be $N_n(t) = |c_n|^2(t)N$ and hence the variations of the ensemble occupation distribution in an MD time step τ will be

$$\frac{\Delta N_n(t)}{N} = \Delta |c_n|^2(t) \simeq \tau \frac{d}{dt}|c_n|^2(t) = \sum_m 2\tau Re[c_n^*(t)c_m(t)d_{nm}(t)] \tag{2.50}$$

where in the last equality Eq. (2.46) has been used. On the other side, the ensemble state occupations can be considered to be governed by a master equation of the kind:

$$\frac{\Delta N_n}{N} = \sum_m g_{m\to n}(t)|c_n|^2(t) - g_{n\to m}(t)|c_m|^2(t) \tag{2.51}$$

where the coefficients $g_{m\to n}$ are the wanted hopping probabilities from the adiabatic state m to the adiabatic state n. These can be determined equating Eq. (2.50) and Eq. (2.51). In particular as it can be noted both Eq. (2.50) and Eq. (2.51) involve a sum over the same adiabatic states indexes m and hence the equality will be fulfilled in the particular case in which the single terms of the total sum are equal, i.e.:

$$g_{m\to n}(t)|c_n|^2(t) - g_{n\to m}(t)|c_m|^2(t) = 2\tau Re[c_n^*(t)c_m(t)d_{nm}(t)] \tag{2.52}$$

A solution of these equations is provided by:

$$g_{m\to n} = \frac{2\tau Re[c_n^* c_m d_{nm}]}{|c_n|^2}\theta(Re[c_n^* c_m d_{nm}]) \tag{2.53}$$

In fact, as it can be noted from Eq. (2.47) and (2.48), $d_{nm} = -d_{mn}$. This symmetry would imply that also $g_{nm} = -g_{mn}$ if the θ functions were not considered. Thanks to the presence of these Heaviside functions hence, for each given couple of states n and m at most one of the two terms in Eq. (2.51) is not zero and at this point the correctness of the transition probabilities in Eq. (2.53) is obtained straightforwardly.

Regarding the third point, in the implementation of the TSH used in this work the forces acting on the nuclei in the coupling regions are hypothesized to be small enough (i.e. not impulsive) to consider negligible the variations in the nuclear momenta directions induced on the nuclei when they pass trough them and a simple isotropic velocity rescale procedure is applied in order to compensate the PES jumps consequent to a surface hop.

Finally it is worth mentioning that even if the validity of TSH and, in particular, of the fewest switches algorithm has never been rigorously proved, the Tully's algorithm, as presented here and used in this work, has been much tested on simple model systems and the results obtained have been compared with the ones obtained integrating the full nonadiabatic time dependent Schrödinger equation and they have been found to be numerically

consistent [56]. In figure 2.3 for example, taken from ref [58], the results relative to the modeling of a nonadiabatic dynamics of NaI which starts in the first excited state are reported. In particular in the two lower panels are shown the TSH trajectories distribution (histograms) which are compared with nuclear probability density at four different times along the dynamics obtained integrating the full nonadiabatic time dependent Schoredinger equation using the Non Adiabatic Bohmian Dynamics (NABDY) method [57].

The reason behind the lack of any rigorous derivation of the validity of the TSH lies in the fact that a reasonable classical limit for the nonadiabatic nuclear motion has not been found so far and actually much effort is dedicated to its research. Alternative methods involving the direct integration of the full nonadiabatic quantum temporal evolution of the nuclei on the fly (e.g. the above cited NABDY method) or other methods [ref Federica, ma devo capire quale..] still involving a semiclassical treatment of the nuclei but improving the phenomenological approach underlying any TSH algorithm are currently under investigation. They will probably be applicable to the study of photochemical reactions of medium and large molecules in the near future but not at present.

2.4.1 Basic numerical integration methods

In addiction to the way in which the nature of the nuclei is approximated and the model utilized for the description of the nuclear interactions (i.e. the many body PES), there are at least two other principal aspects determining the quality of an MD simulation: the algorithm used to integrate the equations of motion and the accuracy and efficiency in the energies and forces calculation.

For the numerical integration of the classical nuclear dynamics equations (2.36), symplectic integrators of a given order in the numerical integration time step τ[6] are commonly preferred. This class of integrators in fact has the property to conserve (numerically) a quantity $\tilde{H}_\tau$ which approximate, during the whole dynamic integration, the total energy of the system

$$H = T_N(\dot{R}) + U_0^{ad}(R) \tag{2.54}$$

with an error of order τ elevated to the power of the order of the integrator [44]. This property avoids numerical long term drifts in the total energy (total energy fluctuations within a given numerical noise are however unavoidable). As example, one of the most commonly used integrators is the second order symplectic algorithm called *velocity Verlet* [69]:

$$\mathbf{R}_a(t+\tau) \simeq \mathbf{R}_a(t) + \mathbf{v}_a(t) + \frac{\tau}{2M_a}\mathbf{F}_a(t)$$

$$\mathbf{v}_a(t+\tau) \simeq \mathbf{v}_a(t) + \frac{\tau}{2M_a}\left[\mathbf{F}_a(t) + \mathbf{F}_a(t+\tau)\right] \tag{2.55}$$

$$\mathbf{F}_a(t) = -\left.\nabla_a U_0(R)\right|_{R=R(t)}$$

---

[6]A symplectic integrator of order n consists of a symplectic map $S_{n\tau}(R,P) = (R', P')$ in the nuclear phase space that approximate the exact classical time evolution $(R'(t), P'(t)) = (R(t+\tau), P(t+\tau))$ up to the order τ^n.

where the time step τ has to be small enough to be i) only a small fraction (~ 0.1) of the smallest vibrational period of the molecular system that has to be simulated and ii) avoid numerical drift in the total energy.

In most cases (as in the case of this work) the aim of a given MD simulation is to describe the bulk liquid properties of a given molecule. However, the size of any computationally affordable simulated system is very far from the thermodynamic limit. In these cases this issue can be solved using appropriate boundary conditions for the numerical time evolution equations in a way to mimic the presence of an infinite bulk surrounding the simulated N-particle model system. A common way to realise this condition is the use of periodic boundary conditions: a box containing a volume V in which all particles of the system lie at the initial integration time is treated as the primitive cell of an infinite periodic lattice and hence every particle in the systems is considered to interact with all particle of the unit cell and also with the ones belonging to all its infinite replicas. Of course this procedure implies that the sum over the long range terms of the force field becomes infinite and hence computationally unaffordable. In practice then, the electrostatic long range interactions are explicitly computed only for distances smaller than a given cutoff radius r_c and the long range interactions for distances $r > r_c$ are taken into account in an approximated way. As example of method of approximation for the long range interactions calculation, which has been used in this work and is particularly suited when periodic boundary conditions are used, the so called Ewald summation [46] algorithm can be cited. According to this method, the interactions energies over the cutoff radius are approximately computed with a summation in Fourier space (taking advantage from their intrinsic periodicity that comes from the use of periodic boundary conditions). For a detailed discussion of this algorithm see e.g. the Appendix B of ref. [45].

2.4.2 MD in different statistical ensembles

The numerical integration of Eq. (2.36) allows the calculation of expectation values of dynamical observables via Eq. (2.37) in the micro canonical ensemble in which, under the ergodicity hypothesis, all the microscopic possible isoenergetic states are equiprobable. In this case the expectation value of a given observable formally correspond to the evaluation of the integral

$$< O(R) > = \int dR \ O(R) \ \delta(H(R) - E) \tag{2.56}$$

in the total nuclear configurational space. Here E is the value of the total energy and $H(R)$ is the Hamiltonian function of the system defined in Eq. (2.54). The thermodynamic fixed variables in this case are the total number of nuclei N_N, the volume of the periodically replicated box V (which fixes the total volume of the system in the thermodynamic limit) and the total energy E.

Even if the access to the direct numerical calculation of the integral (2.56) is already a powerful tool, the main disadvantage of the microcanonical ensemble is that it is often very far from the conditions in which experiments are performed. The canonical ensemble instead, in which the fixed thermodynamic variables are N_N, V and the temperature T, already represents a fairly good approximation of the experimental case of a flask of liquid

at opened air. The total energy of the system is however not conserved in this case (because of the interactions with a reservoir in thermal contact with it) and to reproduce the correct energy fluctuations, other dynamical evolution schemes have to be taken into account in a way to realize the following relation:

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T O(R(t)) dt = \int dR \, O(R) \, e^{\frac{H(R)}{k_B T}} \tag{2.57}$$

The correct total energy fluctuations can be realized by the integration, instead of Eq. (2.36), of the following *Nosé-Hoover chain* equations [47]

$$
\begin{aligned}
\dot{\mathbf{R}}_{\mathbf{a}} &= \frac{\mathbf{P}_{\mathbf{a}}}{M_a} \\
\dot{\mathbf{P}}_{\mathbf{a}} &= \mathbf{F}_a - \frac{p_{\eta_1}}{Q_1} \mathbf{P}_{\mathbf{a}} \\
\dot{\eta}_j &= \frac{p_{\eta_j}}{Q_j} & j = 1, \ldots, M \\
\dot{p}_{\eta_1} &= \left[\sum_{a=1}^{N_N} \frac{P_a^2}{M_a} - 3N_N K_B T \right] - \frac{p_{\eta_2}}{Q_2} p_{\eta_1} \\
\dot{p}_{\eta_i} &= \left[\frac{p_{\eta_{i-1}}^2}{Q_{i-1}} - K_B T \right] - \frac{p_{\eta_{i+1}}}{Q_{i+1}} p_{\eta_i} & i = 2, \ldots, M-1 \\
\dot{p}_{\eta_M} &= \left[\frac{p_{\eta_{M-1}}^2}{Q_{M-1}} - K_B T \right]
\end{aligned}
\tag{2.58}
$$

In this temporal evolution scheme, the $3N_N$ dimensional phase space of the particles system is extended adding $2M$ dynamic variables (η_i, p_{η_i}), called thermostats variables, that also have to be evolved in time. These represent a sort of "agent" whose role is to rescale the velocities in a way to let the instantaneous kinetic energy fluctuate around the value corresponding to a desired temperature T. In fact the first two set of equations describe the time evolution of the particle in the system with an additional term in the momentum equation who depends on p_{η_1} and acts as a sort of friction term that can be either positive or negative. The thermostat variables (η_2, p_{η_2}) ensure that the momentum variable p_{η_1} have the correct Maxwell distribution as the physical momenta have to do. The thermostat variables (η_3, p_{η_3}) have the role to thermostat the momentum variable p_{η_2} and so on. The process should be in principle infinitely repeated but the correct Maxwell distribution for the physical momenta is numerically realized already for $M = 2$ or 3. For the parameters $Q_1, \ldots, Q_M$, the (empirical) prescription to have optimal performances is to set them according to the relations $Q_1 = 3N_N K_B T \Delta t^2$ and $Q_j = K_B T \Delta t^2$ for $j > 1$, where Δt has the dimension of a time and has to be of the order of few tens of the integration step.

At this point, the only missing ingredient to set up a simulation in the canonical ensemble reproducing a given desired experimental conditions is to find a way to control the pressure P or, equivalently, the particle density N_N/V. This can be practically realized for example by means of a so called "equilibration" MD simulation in which not only the temperature

of the system is controlled but also the pressure, that can be evaluated monitoring the expectation value of the virial:

$$P = \langle\ \mathbf{R}_a \mathbf{F}_a\ \rangle \tag{2.59}$$

The periodic box dimensions (which control the molecules density) can be varied until the desired (stable) value of the pressure (and hence of the density) is reached. At that point a good representative equilibrium configuration of the system can be considered to be obtained and the proper MD simulation, called of production, can begin.

An other useful scheme of integration of the dynamical equation describing the evolution in a statistical ensemble different from the microcanonical one is the one corresponding to the temporal evolution subject to an olonomic constraints:

$$M_a \ddot{R}_a(t) = -\nabla_a U_n(R) + \lambda \nabla_a \sigma$$
$$\sigma_k(R, t) = 0 \tag{2.60}$$

that can be numerically integrated using the SHAKE algorithm [48]. The constraint can be used to fix the value of a reaction coordinates during an MD simulation and the results obtained to compute the reversible work required for a molecular system to move along a reaction path defined by varying a reaction coordinate $\xi(R)$. In fact, the case in which

$$\sigma(R) = \xi(R) - s \qquad s \in (s_0, s_f) \tag{2.61}$$

can be conveniently considered. In particular, as in the minimum energy reaction path discussion reported in section X, $\xi(R) = s_0$ has to identify the reactant state and $\xi(R) = s_f$ the product state. In this way the probabilistic interpretation of the Landau free energy A of a molecular system in the canonical ensemble can be used:

$$A(s) = -K_B T \ln P_\xi(s) \tag{2.62}$$

where the probability density $P_\xi(s)$ is given by

$$P_\xi(s) = \ <\delta(\xi(R) - s)> \ = \frac{1}{Z} \int d\mathbf{R}\ e^{-\beta U(R)} \delta(\xi(R) - s) \tag{2.63}$$

The free energy variations along the reaction path can be characterised computing

$$\Delta A(s) = \int_{s_0}^{s} ds'\ \frac{dA}{ds'}(s') = -\int_{s_0}^{s} ds'\ \frac{K_B T}{P_\xi(s')} \frac{dP_\xi}{ds'}(s') \tag{2.64}$$

In fact, when Eq. (2.60) are used to fix the reaction coordinate to given values s, the free energy variation can be expressed as [49]

$$\Delta A(s) = \int_{s_0}^{s} ds' \frac{<\ z^{-\frac{1}{2}}(R)[\lambda + K_B T G(R)]\ >_{s'}^{constr}}{<\ z^{-\frac{1}{2}}\ >^{constr}} \tag{2.65}$$

where λ is the olonomic constraint force in Eq. (2.60),

$$z(R) = \sum_a \frac{1}{M_a} (\nabla_a \xi(R))^2 \tag{2.66}$$

and

$$G(R) = \frac{1}{z^2(R)} \sum_{ab} \frac{1}{M_a M_b} \nabla_a \xi(R)(\nabla_a \nabla_b \xi(R))\nabla_b \xi(R) \tag{2.67}$$

It has to be mentioned that in the particular case in which the reaction coordinate is a distance $\xi = |\mathbf{R}_i - \mathbf{R}_j|$ between two atoms i and j, the function $z(R)$ is a constant, $G(R) = 0$ and Eq. (2.65) simply becomes:

$$\Delta A(s) = \int_{s_0}^{s} ds' < \lambda >_{s'}^{constr} \tag{2.68}$$

Constraint MD can hence be applied to study isomerization reactions mechanisms improving the minimum energy path analysis description discussed in section **??**. The statistical ensemble associated with the constrained dynamical integration scheme is in fact called blue moon[7] ensemble because it allows an efficient sampling of configurational space portions that would be very rarely met in the course of a normal MD simulation.

2.5 Coupling with the environment

In order to properly describe the behaviour of a chromophore in solution, the perturbations implied by the presence of the solvent molecules on the PESs of the molecule have to be taken into account. In this work, two main approximated way to take into account the solvation effects have been utilised. In the first model, called Polarisable Continuum Model (PCM), the solvent is treated implicitly as a dielectric continuum and is shortly reviewed in Appendix B while in the second one, called Quantum Mechanics/Molecular Mechanics (QM/MM), the solvent molecules are explicitly taken into account via a mixed ab initio-classical MD approach.

The basic idea underlying the QM/MM approach was laid out in a seminal paper by Levitt and Warshel [50]. It consists in partitioning of the system into two subsystems: the QM one, which consists of the central region which is described with ab initio MD. The other subsystem, referred as MM, contains the surrounding environment (in this case the solvent molecules) and is described with classical MD using an appropriate force field. Many different implementations have been developed realizing the QM/MM scheme over the years. In this section will be discussed the additive scheme used in the present work [51], which consists in considering the whole nuclear system described by the total Hamiltonian:

$$H_{tot} = H_{QM} + H_{MM} + H_{QM-MM} \tag{2.69}$$

where H_{QM} is the Hamiltonian of the QM system in which the necessary PESs are computed on the fly ab initio, H_{MM} is the Hamiltonian of the MM subsystem which is treated with classical MD and H_{QM-MM} is the term describing the potential energy of the interaction between the two subsystems and is modelled as:

---

[7]The form "once in a blue moon" is a common english saying referring to the occurrence of a rare event.

$$H_{QM-MM} = \sum_{a \in MM} \left\{ \frac{eq_a}{4\pi\epsilon_0} \left[\int d\mathbf{r}' \; n(\mathbf{r}, R) v_{SC}(|\mathbf{r} - \mathbf{R}_a|) + \sum_{a' \in QM} \frac{z_{a'}}{|\mathbf{R}_a - \mathbf{R}_{a'}|} \right] + \right.$$
$$\left. + \sum_{a' \in QM} \left(\frac{A_{aa'}}{(\mathbf{R}_a - \mathbf{R}_{a'})^{12}} - \frac{B_{aa'}}{(\mathbf{R}_a - \mathbf{R}_{a'})^6} \right) \right\} \quad (2.70)$$

The first two terms represent the electrostatic interaction of the QM electrons (described by the electronic density, $n(\mathbf{r}, R)$) and the QM nuclear charges z_a with the point charges of the classical MM part q_a. The third term describes the steric non bonded van der Waals (vdW) interactions between QM atoms and MM atoms.

To model the interaction between (quantum) electrons and (classical) point charges in the MM system the Coulomb potential is replaced with a suitable screened potential which has the form

$$v_{SC}(|\mathbf{r} - \mathbf{R}_a|) = \frac{r_{c,a}^4 - (|\mathbf{r} - \mathbf{R}_a|)^4}{r_{c,a}^5 - (|\mathbf{r} - \mathbf{R}_a|)^5} \quad (2.71)$$

where the parameters $\{r_{c,a}\}_{a \in MM}$ are empirical distances called covalent radii and are set for each chemical element using test QM/MM simulations in liquid phase in a way to return correct structural properties of the liquid [52].

The need of the empirical screened potential comes from the large difference in the level of approximation used to describe the two subsystems of the QM/MM approach. In fact the positively charged MM ions (which aren't, by construction, surrounded by any electronic density) can spuriously trap the quantum electrons[8]. The use of the screened Coulomb potential has hence to be intended as a computational expedient adopted to avoid these unphysical effects and that is able to reproduce the correct phenomenological results.

It has to be mentioned that there is an other important computational problem that is soon faced when the interaction potential in Eq. (2.70) is used in a real simulation and is related to the computational cost necessary for the evaluation of its first term. To see this, the complexity of its direct calculation can be roughly evaluated considering the number N_r of space grid points that can be used to describe the electronic density, which is generally in the order of 10^5, and the number N_{MM} of the classical atoms. The number of operations required for its "brute force" calculation would than be of the order of $N_r N_{MM}$ implying an increase in the computational cost of a "gas phase" ab initio simulation of a given molecule of several orders of magnitude. This second issue is solved taking into account the electrostatic interaction between the QM system and the more distant MM atoms (outside a cutoff distance D_c from any QM atom) in an approximated way including just the coupling with the first multiple moments (up to the quadrupole) of the quantum charge distribution with the classical point charges [51].

It is important to stress that, even if the use of the phenomenological screened potential in Eq. (2.70) is required to have a working QM/MM model, within this scheme, the effects

---

[8]This issue is known as "electrons spill out" problem

of the MM atoms on the QM system (i.e. the polarization of the electronic density) is directly taken into account together with possible collision events happening when an atom pair (belonging to whatever subsystem) distance becomes equal to the sum of their vdW radii.

Finally it is important to note that, to take into account the effects coming from the back-reaction (polarization) of the QM system as a consequence of the presence of the MM system, a polarizable force field would be required. These effects are hence neglected when the MM system force field is of the kind of Eq. (2.38) as in the present work.

Chapter 3

DNA-Protein crosslink reactions

In the first three sections of this chapter the basic phenomenology regarding the DNA structural and photophysical properties will be briefly reviewed in order to highlight the biological relevance of its interactions with nearby proteins. This is done in view of the discussion, reported in section 4, of an experimental study of UV driven DNA-protein crosslink reactions (defined and discussed in section 3) in living cells. This experiment represents the first step of the work performed in this thesis. The concepts and open problems discussed in this chapter motivate and contextualize the studies reported in the following chapters.

3.1 DNA-proteins interactions

Nucleic acids are the basic biological macromolecules of heredity. Their structural unit is the nucleotide, which consists of a phosphoric acid, a pentose sugar, and a nucleobase. Depending on different types of the pentose sugar, nucleic acids are classified as the ribonucleic acid (RNA) and deoxyribonucleic acids (DNA). The genetic code is stored in DNA via a symbolic alphabet composed of a set of four icons, corresponding to four different nucleobases (informally, bases): adenine (A) and guanine (G) called pyrimidine, and thymine (T) and cytosine (C), called purines.

In 1953, Watson and Crick discovered the famous double helix structure of the DNA [59] which is sketched in Fig. 3.1. In this structure, a double-strand DNA is built up by two anti-parallel strands running in opposite directions along the same axis. The bases are arranged in stacked configurations and their stacking interactions are at the basis of the stability of this structure. The four DNA bases pair up with each other following the so called base-pair rule: (A) only bonds with (T) and (C) only bonds with (G). The base pair structures are reported in Fig. 3.1. In each base pair, specific hydrogen bonds are formed (dashed lines in Fig. 3.1) which connect the two DNA strands together. The two adjacent base pairs are on average separated by 3.4 Å and a complete turn covers 10 base pairs and is about 35 Ålong, while the radius of the helix is about 10 Å.

The information contained in DNA is expressed transforming the genetic code (ordered sequence of bases in DNA) in specific polymers, the proteins, via a map between nucleobases triples and a set of 21 aminoacids, the building blocks (monomers) of proteins. This happens
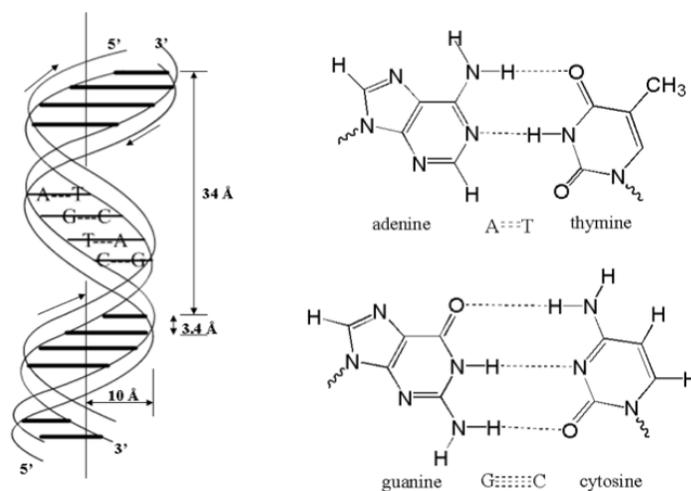
Figure 3.1: The double helix molecular structure of DNA is sketched in left panel. In the right panel the structure of the nucleobases pairs are reported. The hydrogen bond structure they form is also shown using dashed lines.

via an astonishing complex sequence of biochemical processes that starts with the separation of the two DNA strands and the copying of the DNA code in RNA strands and ends with translation, the process in which cellular ribosomes create proteins. The different proteins, characterized by a given (ordered) sequence of aminoacids, perform an endless array of functions within the living organisms, e.g. catalysis of metabolic reactions, replication of DNA and the transport of molecules from one location to another.

The nature of a given cell lineage (e.g. part of a human hand or hair) comes from the collection of proteins it produces more than its DNA content (the same for every cell in an organism). The mechanism at the basis of control of gene expression (different sections of DNA are expressed in different cell lineage) is one of the foundational issues of modern bio-sciences. It is known that the activity of (sequence-specific) DNA sequences mediator (activators or repressors) influence transcriptional activity. In fact, during transcriptional activation, the electrostatic binding of ubiquitous proteins, called transcription factors, to defined DNA sequences triggers a cascade of spatially and temporally coordinated reactions able to inhibit or enhance the transcription of given DNA sequences.

DNA-protein intramolecular interactions play hence a major role in modern molecular biology investigations. They are involved not only in structural organization of the DNA within the cell nucleus, in DNA replication, transcription and repair but also in the regulation of the gene expression.

3.2   Photostability of DNA

The structure of DNA is very well suited for biological information storage: the backbone is resistant to cleavage, and both strands of the double-stranded structure store the same biological information. The stability of the chemical structure of DNA is however crucially important. DNA damage, i.e. an alteration in the structure of DNA such as a break in a strand, a base missing from the backbone, or a chemically changed base, occurs frequently in the natural environment. Even if just a chemical bond in one of three billion nucleotides of human being is modified, the outcome may be catastrophic. DNA damages are however continuously repaired via complex biochemical reactions preventing the loss of information contained in the genetic code.

One of the most important external perturbations possibly inducing DNA damage is UV radiation. The DNA bases are in fact good chromophores and, as shown in Fig. 3.2, have intense absorptions in the UV around 260 nm [60]. Possible photoreactions driven by the UV absorption consist in intra and extra strand covalent new bond formations. For example, adjacent cytosine and thymine bases can bind with each other (creating the so called pyrimidine dimers) or with nearby molecular complexes (reactions called crosslink, see next section). In addition to these direct DNA damage, also indirect DNA damages are possible starting with electron detachment phenomena in the bases and leading to the formation of reactive free radicals. These photoreactions result in the formation of harmful photo-lesions potentially bringing to mutations and carcinogenesis.
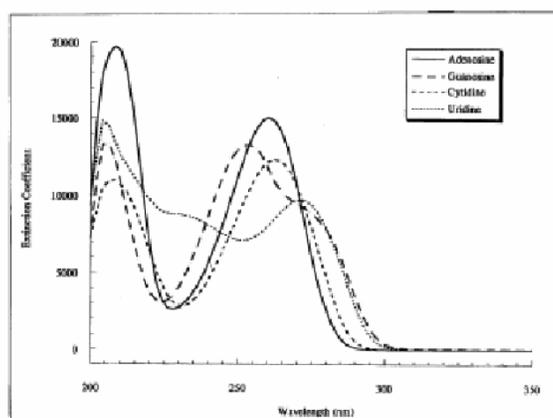

Figure 3.2: Absorption spectra of nucleobases

DNA however has its own resistance to UV-light induced photochemical damage. In fact, it is known from the 70's (by means of steady state fluorescence measurements) that the DNA/RNA nucleic bases fluorescence quantum yields (QY) at room temperature in water solution are of the order of 10^{-4} and the fluorescence lifetimes are on the order of a picosecond or less [62]. This implies the existence of extremely efficient non-adiabatic energy

internal conversion processes bringing the excited nucleobases back on the (chemically unreactive) electronic ground state.

The much longer excited states lifetime of other analogous bases not forming the DNA or RNA [63],[64],[65] suggests that the choice of nature for the actual form of the DNA bases could even be the result of an evolutionary process in the early biotic ages. Among the bases, the uracil (the nucleobase present in RNA instead of thymine) is by far the fastest decayer [66]: in practice, its lifetime determination is limited by the experimental time-resolution, which is about 100 fs, meaning that it could be substantially faster.

Several quantum chemical calculations have shown that the ultrafast DNA bases decay of the excited states is much helped by the existence of barrierless paths in the excited states leading to conical intersections[1] (CI) between the first and second singlet excited state and the ground state allowing ultrafast and very efficient non radiative decay processes. The geometrical changes involved in reaching the CIs are different for different molecules but there's a general trend: the internal conversions through the CI generally involve out of plane motions leading to puckered structures in the nucleobases rings [63], [69].

In recent years much effort has been also dedicated to the study of the photo-deactivation processes of the nucleobases also by means of *ab initio* nonadiabatic excited states MD simulations [67], [70], [71],[72], [73], and a general survey picture of the deactivation paths has been made by Barbatti et al. in ref [74]. In figure 3.3, reported here from this last work, is shown a schematic description of the photo-dynamical simulations results for the thymine and uracil. In this figure, the thick lines represent the eterocyclic ring plane of the molecule and the atoms moved out of plane (see Fig 4.1, in next chapter, for the pyrimidines atoms labeling) are connected with thin lines. After the electronic excitation on the first bright excited state these molecules usually remain trapped in a shallow minimum for few picoseconds (the channel P1 in the figure). Also the direct pathway leading (in an ultrafast way) to the ground state via the CI can be activated (the channel P2 in the figure) but it has never been observed in the excited states simulations of thymine while it corresponds to a consistent percentage of the trajectories in the case of the uracil. The presence in the thymine of an extra methyl group linked to the eterocycle, partially inhibits this pathway. This fact is consistent with the longer lifetime of the thymine experimentally observed.

From the first bright excited state PES reached upon excitation, uracil and thymine can possibly relax also to a lower excited state, via non-adiabatic transitions, reaching the minimum energy region of this state. Once there however, they can deactivate through a CI connecting this state to the ground state (path P1b) or go back to a the higher state crossing a barrier, and consequently deactivate via an other CI with the ground state (the paths P1c and P1d in figure).

3.3   DNA-proteins Crosslink reactions

Over the years several chemical and enzymatic methods have been developed to study DNA-protein interactions. Many times however, these interactions are not strong enough to be easily detectable. Several methods have been used to stabilize them and one of the most

---

[1]In the quantum chemistry language, for conical intersection is intended the crossing of two PESs which gets degenerate in a given set of configurational space points.
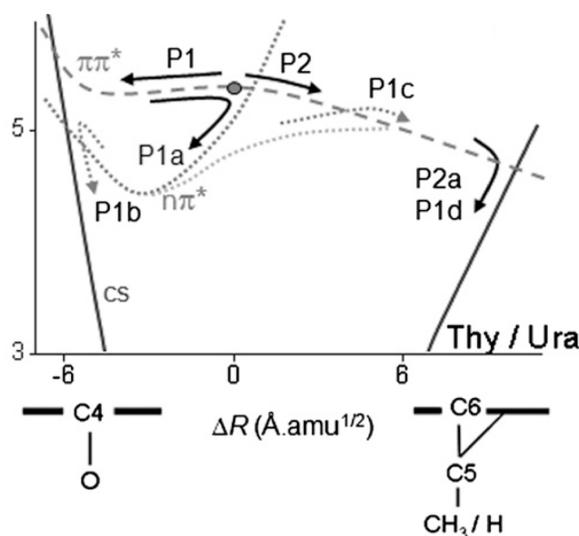
Figure 3.3: Schematic description of the dynamics of the thymine and uracil nucleobases based on dynamics simulations results. The image is taken from ref [74]. The thick lines represent the molecular ring plane and the atoms moved out of plane are connected with thin lines. The label assigned in this figure to the atoms and to the excited states (via the corresponding excited electronic states symmetry, $n\pi^*$ and $\pi\pi^*$) are consistent with the ones adopted in next chapters.

popular consists in the induction of DNA-protein crosslink reactions. A crosslink reaction in fact consists in the formation of a covalent bond that links one polymer chain to another and can stabilize a transient proximity relation between the two reactant macromolecules. These reaction are in fact reversible only in extended incubation at $65\,^{\circ}$C.

The standard crosslink catalysis agent used in bio-molecular experiments is formaldehyde. It can traverse biological membranes, allowing crosslinking to be induced on intact cells. The crosslink reaction induced by formaldehyde is a two steps process whose mechanism is sketched in Fig. 3.4. In the first step (upper panel of Fig. 3.4) formaldehyde binds aminoacids immino and amino chemical groups; in this way it forms an activated compound that is able to connect, in the second reaction step, another amino group forming a stable bridge (lower panel of Fig. 3.4). Formaldehyde hence produces protein-nucleic acid as well as protein-protein crosslink reaction.

In biological experiments the first desired feature of the crosslink induction agent is that it has to be able to link to the DNA as many proteins as possible (high reaction yield). At the same time however it is also fundamental to avoid the observation spurious crosslink events between DNA and physiological unrelated proteins. Formaldehyde is able to connect large protein multi-complexes to DNA but false positive interaction observations are possible.

An other used crosslink reaction induction agent is UV light: as mentioned above in fact, one of the possible UV photo-damages of DNA is the formation of covalent bonds
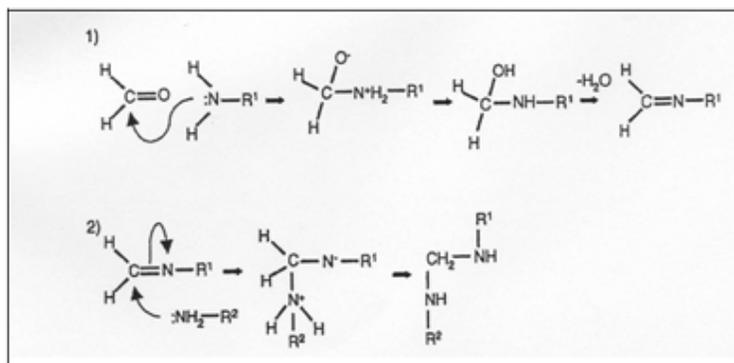
Figure 3.4: The two step mechanism of crosslink reactions induced by formaldehyde. In the first reaction step (upper panel) the formaldehyde binds aminoacids immino and amino chemical groups via a chemical reaction called creation of a Schiff base. In the second reaction step this activated compound binds an other amino group forming a stable bridge between two molecules.

between nucleotides and close lying aminoacids. Photochemical crosslinking provides an attractive alternative to formaldehyde-based protocols, but irradiation with conventional UV sources typically yields inadequate product amounts. Cross-linking with pulsed UV lasers can represent instead a potentially revolutionary technique to increase photochemical yield and avoid the observation of non-existent interactions. A study of the performances and possibilities offered by the use of UV lasers in this context will be discussed in next section.

3.4   UV-laser crosslink induction in living cells

In this section a study of the UV induction of DNA-protein crosslinking directly in living cells is phenomenologically approached on the basis of a cells irradiation experiment. This experimental work, in which I was involved at the beginning of my Ph.D. studies, is the seed of my theoretical and computational thesis work. Its discussion helps in contextualizing and motivating the computational results that will be discussed in next chapters and will be hence briefly reviewed. The experimental set up details (that can be found in ref [75]) will not however be reported because their discussion exceeds the purposes of this chapter.

The scheme of the experiment is summarized in Fig. 3.5 from the initial excitation induced by UV laser pulses to the final measurement of the amount of crosslinked material, in seven steps. A solution of human leukemia cells (U937) was irradiated at 263 nm. Various irradiation times and pulse energies have been tested. Shining laser pulses onto the target can cause two-photon absorption by the DNA bases followed by the formation of covalent bonds with aminoacids lying in the proximity of the excited base (step (1) of Fig. 3.5). Successively, as shown in steps (2)-(7) of Fig. 3.5, samples were treated by a ChIP protocol

[76, 77]. It was, then, possible to isolate and measure, by quantitative polymerase chain reaction (qPCR) (see ref [78] for a review), the amount of a specific DNA region, named TRAIL-gene-promoter, linked to a specific nuclear protein, named H3K4me3 [79]. In step (2) living cells are broken and the DNA is sheared, whereas in step (3) an antibody is added to selectively recognize only those DNA fragments cross-linked with the protein of interest (H3K4me3), thus enabling to fish them on. During an incubation time (Fig. 3.5, step (4)) the immunoprecipitation protocol is carried out, whereas in steps (5) and (6) selected DNA fragments are inversely crosslinked (by incubation of the samples at 65 °C) and washed by removing the protein-antibody complexes. Finally, in step (7) the appropriate DNA fragments are amplified and quantified by qPCR.

In order to analyze the obtained results a simple phenomenological model for the response of the system to the pulsed UV-irradiation was developed. It is based on the following assumptions: i) only DNA-bases of the cell chromatin get excited upon irradiation, ii) the dominant excitation is initiated by two-photon absorption as reported for ps pulses in [80, 81], and iii) damage to the cells caused by irradiation with intense UV laser pulses occurs at relatively high UV light intensities (larger than 2 GW/cm^2) in agreement with a typical threshold-like behavior with the pulse energy [83]. As long as the first hypothesis is concerned, it is well known that aromatic amino-acids (e.g. tryptophan, phenylalanine, tyrosine) which are abundant in proteins, absorbs at 260 nm. Thus, in principle, UV light could also trigger crosslinking by exciting proteins rather than DNA (or both). Nevertheless, the static UV ($\sim$ 260 nm) absorption spectra of bases and aromatic amino-acids show much stronger absorption of bases compared to aminoacids. The influence of solvent and intra-cellular water-like liquid has also been discarded in view of the negligible water non-linear absorption at 260 nm at the used intensities [84], the linear absorption being very small as well in this UV window. The second hypothesis, i.e. the assumption of biphotonic excitation mechanisms underlying the crosslink formation, will be discussed in next section.

By relying on the above hypotheses, an expression of the crosslinking process yield, Y, can be easily worked out as a function of the two macroscopic knobs of our experiment, i.e. the pulse energy, E_{pulse}, and the number of laser shots, N_{shot}, per irradiation. For the sake of simplicity we will refer in the following to E_{pulse} although the relevant physical quantity turns out to be the laser pulse intensity. If $\Pi(E_{pulse})$ is the probability that DNA is excited by UV two photon absorption, and $D(E_{pulse})$ is the probability that a crosslinked specie survives after the irradiation of a single pulse (damage function), the overall yield, Y , after N_{shot} is:

$$Y(E_{pulse}, N_{shot}) \propto \Pi D^{N_{shot}-1} + \Pi(1-\Pi)D^{N_{shot}-2} + \ldots \qquad (3.1)$$

where the sum is over N_{shot} terms each of them containing a polynomial in Π whose coefficients form the Pascal triangle. The sum in Eq. (1) can be easily carried out to give:

$$Y(E_{pulse}, N_{shot}) = k \, \Pi(E_{pulse}) \, D^{N_{shot}-1}(E_{pulse}) \, \frac{1 - \xi^{N_{shot}}(E_{pulse})}{1 - \xi(E_{pulse})} \qquad (3.2)$$

where

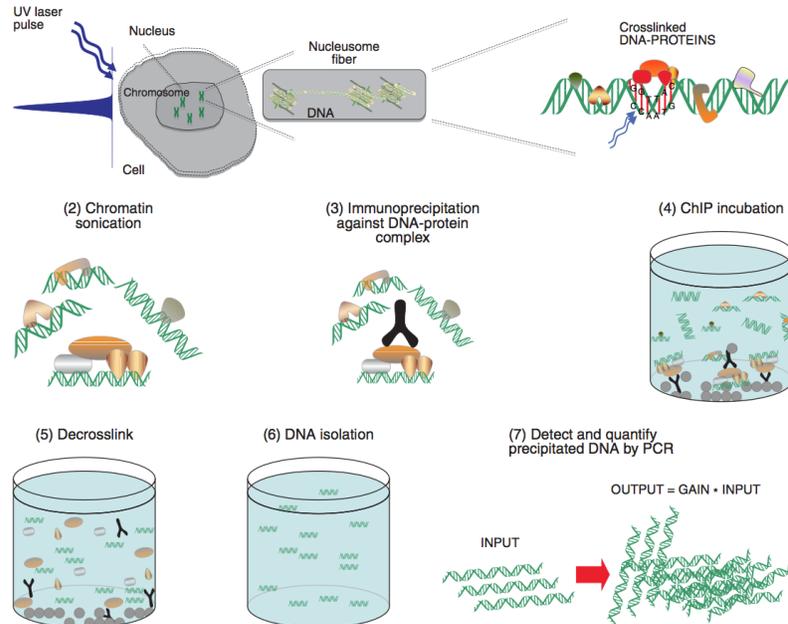$$\xi(E_{pulse}) = \frac{1 - \Pi(E_{pulse})}{D(E_{pulse})} \qquad (3.3)$$

Figure 3.5: Schematic layout of the experiment. In (1) solutions containing living cells are irradiated by fs-UV laser pulses. This causes the formation of a covalent bond between an excited DNA base and a protein nearby. The samples are, then, processed by ChIP protocol [26,27] and qPCR [28] to measure the amount of a specific DNA region, the TRAIL promoter, bound to a specific protein, the H3K4me3 Histone [29]. (2) (6) Scheme of the Chromatin Immunoprecipitation (ChIP) protocol consisting in five serial phases before accessing the final qPCR (7) for detection and quantification. In (2) cells are opened and DNA is sheared, in (3) the antibody of interest, which selectively recognizes only H3K4me3, is added. The last three phases consist in an immunoprecipitation (4) to isolate the protein of interest and, consequently, the crosslinked DNA fragments. A reverse crosslink takes place in (5) and (6) to purify and isolate the DNA fragments from the protein complexes. In (7), by means of qPCR, the appropriate DNA fragments are hugely amplified in number so as to quantify the amount of TRAIL promoter DNA region linked to the H3K4me3 protein.

and k is an instrumental constant.

If σ_1 and σ_2 are the cross sections for the first and the second photon absorption respectively, and p is the probability the excitation leads to a covalent DNA-protein bond, $\Pi(E_{pulse})$ can be written as:

$$\Pi(E_{pulse}) = p \ \sigma_1\sigma_2 \ \frac{E_{pulse}^2}{E_{photon}^2 A^2} \tag{3.4}$$

$E_{photon} = 4.5$ eV being the single photon energy and A the laser spot size onto the sample that, in our experiment, has been measured to be 0.25 cm^2.

The damage function $D(E_{pulse})$ has been chosen so as to reproduce the above mentioned threshold-like behavior of the crosslinked cells damage with the pulse energy [97]:

$$D(E_{pulse}) = \frac{1}{2} \left[1 - Erf\left(\frac{E_{pulse} - E_{th}}{\Delta E}\right) \right] \tag{3.5}$$

In Eq. 3.5, E_{th} stands for the energy threshold and ΔE for the slope of the damage when the energy increases. It is worth stressing that we have not accounted for possible damage to non-crosslinked cells. Such an assumption has been corroborated by checking irradiated non-crosslinked cells by means of the fluorescence activated cell sorting methodology and finding out that cell-damage is negligible in our conditions. Eq. 3.2 can be to best-fit the measurements data which report the crosslink yield versus E_{pulse} for 2 and 5 minutes irradiations, as shown in Fig. 3.6. The best-fit parameter values turned out to be $p \ \sigma_1\sigma_2 = 4.4 \times 10^{-34} cm^4$, $E_{th} = 185 \ \mu J$, and $\Delta E = 30 \ \mu J$.

The estimation of the product $\sigma_1\sigma_2$ depends on the particular macromolecule under consideration. For instance, the absorption cross-section of liquid aqueous solution of thymine at the wavelength of 282 nm for the electronic transition to the first excited state triggered by fs laser pulses is $\sigma_1 \sim 3 - 10 \times 10^{-17} cm^2$ [88], whereas the absorption cross-section of a uridine molecule from the first excited singlet state to higher excited states at 266 nm is $\sigma_2 \sim 10^{-15} cm^2$ [80, 89]. These values would provide $p \sim 0.1$ in the case the two-photon absorption involves two different bases each of them making the first excited state transition, and $p \sim 0.01$ for the case involving one single base making the two photon transition. $E_{th} = 185\mu J$, and $\Delta E = 30\mu J$ are realistic values in the damage dynamics of tissues due to irradiation with intense UV laser pulses [83]. In particular, $E_{th} = 185\mu J$ leads to a threshold intensity of $\sim 3 \ GW/cm^2$ for crosslinked cell damage in good agreement with the value reported for cell damage in Chinese hamster ovary by 345 nm laser fs pulses [90].

The behavior of the crosslinking yield versus the laser pulse energy in Fig. 3.6 demonstrates the nonlinear response of the system to the pulsed laser irradiation. The yield increases quadratically up to a maximum value, after which it quickly drops off due to crosslinked cell damage. The longer irradiation time results in a smaller yield due to damage accumulation. The maximum value falls at about 110 μJ for 2' irradiation, whereas it shifts to lower pulse energies (50 μJ) for 5' irradiation, again due to higher damage. The fair agreement between simulations and measurement is noticeable in view of the high intrinsic fluctuations of the extremely complex system under investigation. The error bars result from the root mean square of the values measured in five experimental runs carried out at different times with different samples of the same cell line. Thus, we can assess that
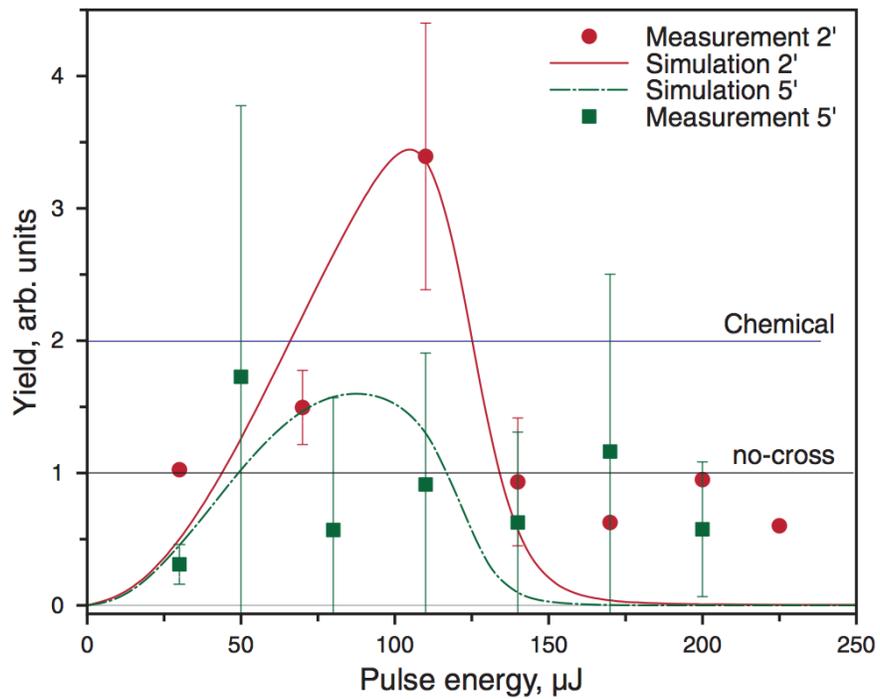
Figure 3.6: Measured (experimental points) and calculated (lines) DNA-protein crosslinking yield in human U937 acute myeloid leukemia cells versus the laser pulse energy for 2' (circle) and 5' (squares) irradiation times. The blue and black horizontal lines represent the yield values obtained with ordinary chemical methods and measured in non-irradiated cells, respectively.

our analysis is reliable even though only few experimental points are above the no-cross level, the latter being the amount of crosslink yield found in non-irradiated cells. This value is about half the yield usually obtained with ordinary chemical methods, which amount to $\sim 1 - 3\%$ [91]. It is remarkable that the maximum yield obtained in the experiment by inducing crosslink with UV fs-pulses is about twice as high as the value achieved with chemical methods. Thus, not only ultrashort UV pulses produce crosslink in time interval much shorter than those required by chemical methods, but even the yield can be definitely larger than that obtained with the conventional methods currently used. This is a striking finding which could be even improved paving the way to detect and study new interactions in molecular biology, which are too weak to be detected [91] with standard techniques.

It worth stressing, however, that the model discussed here is far from being a detailed description of the crosslink mechanism in living cells from a microscopic prospect. In particular, although it has been shown that a fundamental role can be played by the two photon absorption, it cannot be stated whether the absorption is localized in a single DNA base, as suggested in [81, 82, 80], or there is a contribution from a collective excitation involving two bases, rather close to each other in the double strand DNA, as suggested in [92] and theoretically predicted in [93]. Moreover other studies on other DNA-binding transcription factors [86] report that the crosslinking quantum yields obtained using high-peak intensity femtosecond laser are comparable with the ones obtained using UV lamps. This suggests that the photochemical effects of UV radiations on nucleic acid and proteic complexes may depend on the particular systems investigated. Finally, even mechanisms involving two singly-excited bases lying in the lowest triplet state, next to each other, like in the bi-molecular triplet-triplet annihilation [94] are in principle compatible with the experimental results.

The presented model however, reaches three important goals:

1. It reproduces very well the response of a very complex system to laser pulse irradiation in a number of cases.

2. It sheds light on the investigated crosslink mechanism, providing some explanation for the nonlinear behavior, still being of general validity.

3. It relates the measurable DNA-protein crosslink yield of a complex system such as a human cell directly to macroscopic degrees of freedom, such as E_{pulse} and N_{shot}, which are controlled in the experiment.

Finally, after testing the reliability of the model by fitting the measurements, the best possible irradiation conditions can be looked for within the $E_{pulse} - N_{shot}$ plane. The result is shown as a contour plot in Fig. 3.7. The maximum crosslink efficiency is expected at $N_{shot} \simeq 1000$ and $E_{pulse} \simeq 120\mu J$. This corresponds to about half a minute irradiation time and would lead to yield enhancement of a factor of 2.5 over the result achieved via the chemical method. Of course, the shorter the irradiation time the smaller the induced damage in the crosslinked cells.
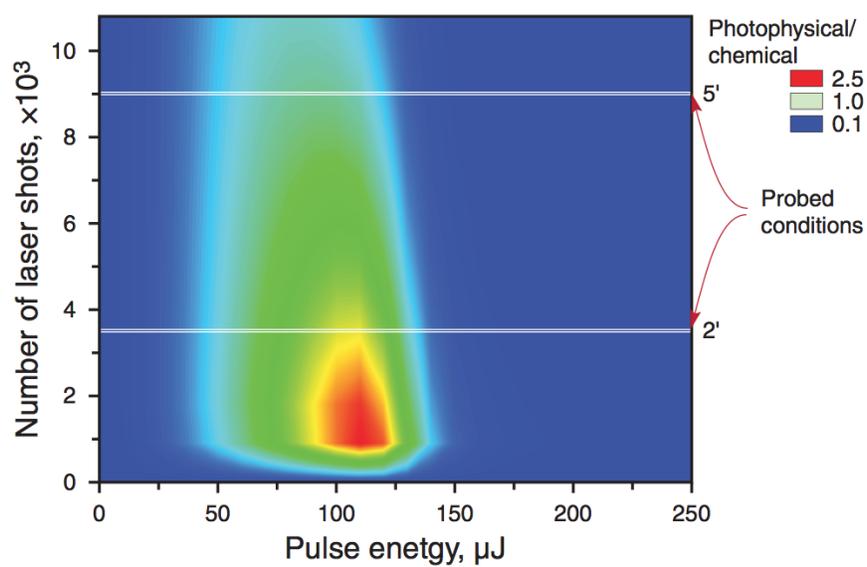
Figure 3.7: Map of the calculated ratio of photophysical/chemical yield for DNA-protein crosslink in living cells, as a function of the pulse energy and the number of laser shots. The white horizontal lines refer to the experienced conditions of the experiment, corresponding to 2' and 5' irradiations, respectively.
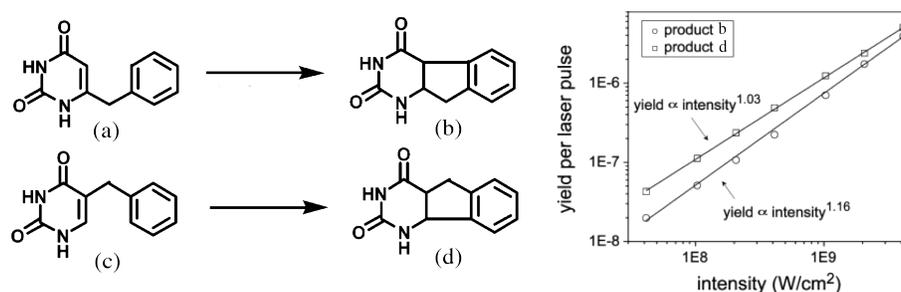
Figure 3.8: Results taken from ref [102]. Photoreaction scheme for simple pyrimidine-aromatic aminoacid crosslink photoreactions are shown in the left panel. The photoreaction leading from compound (a), the 5-bezyluracil, to the cyclized structure in (b), the 5,6ben-zyluracil, will be the detailed studied in next chapters. In the right panel the reaction yields obtained per laser pulse as a function of UV intensity.

3.5   Photochemical crosslink possible mechanisms and conclusions

As mentioned in last section, the mechanism of DNA protein crosslink photochemical formation is essentially unknown. The occurrence of a crosslink upon UV excitation in fact can be correlated with a combination of factors, including the inherent photoreactivity of the monomers that are joined and the correct distance and geometrical arrangement of the molecular moieties at the site of the crosslink. Two photon absorption based mechanisms have been proposed [80, 81] in the 90's and the experimental results obtained are consistent with this hypothesis but other experiments contradict this hypothesis [86] and many issues remain opened. The two photon mechanisms hypothesized are generally based on radical cation formation and rely on the two-photon absorption by the same base. The two photons however can also be absorbed activating different bases of the DNA double strand, separated by two-three bases [92]. Moreover the decay channel of the bases in base-stacking conformations can strongly differ from that of unstacked ones and the phenomenological modeling discussed in last section is not able to discriminate the actual microscopic absorption mechanism.

Experimental chemical studies of these reaction are limited by the inefficiency of the reactions that is also accompanied by protein and nucleic acid damage. The resulting complicated reaction mixtures also make it difficult to identify specific amino acid/base crosslinks and only very few crosslinked structures have been so far characterized [99, 100, 101].

Some simple model systems to explore the crosslinking of aromatic aminoacids (tyrosine, phenylalanine, and tryptophan) to the pyrimidine bases have been proposed in [102]. This model-building strategy involved attaching the pyrimidine base to the aminoacids, using a short linker to mimic the proximity and orientation in DNA-protein complexes. Among them two exemplar photocyclization reactions, have been reported in Fig. 3.8.

As it can be observed in the right panel of Fig. 3.8, measurements of the photochemical reactions yield of these small molecules selected to model the crosslinking reactions exhibit nearly linear dependences on UV intensity. This proves that they are triggered by the absorption of a single photon. This result indicates that biphotonic mechanisms can be activated only in more extended systems and that, in particular, phenomena regarding the two-photon absorption by the same base may play a minor role in the formation of crosslink and lead preferentially to other kinds of DNA photodamage. This hypothesis would explain the increase of damage for the highest tested UV intensities in the phenomenological study reported in last section and the consequent decrease of the crosslinking yields.

The study of model system reactions like in Fig 3.8 however can provide fundamental insights about the still unknown photochemically induced DNA proteins crosslink mechanisms at a molecular level. In particular the study of the photoreaction leading from 5-benzyluracil, structure (a) in Fig. 3.8, to the formation of a cyclized adduct, structure (b) in Fig. 3.8, model of crosslinked species, will be the subject of the following of this thesis. In particular molecular *ab initio* calculations results will be joined to experimental spectroscopic data to fully characterize the properties of reactant and product separately in the next two chapter. The photoreaction mechanism will be instead discussed in the last chapter. On the basis of these results, further indications about the photochemical DNA proteins crosslink mechanism will be given in the conclusions of the last chapter.

Chapter 4

Benchmark calculations for the description of the 5-benzyluracil excited states

5-benzyluracil (5BU, Figure 1-c) is a DNA base derivate obtained from uracil (Figure 1-a) with the chemical substitution of an hydrogen atom with a phenyl group (via a methyl bridge) in position 5 or, equivalently, from thymine (Figure 1-b) adding the phenyl directly to the methyl group. This molecule allows the study of the proximity interactions of a pyrimidine nucleobase with an aromatic chromophore. In a biological context, interpreting the aromatic chromophore as a model for aromatic amino acids (e.g. phenylalanine or tyrosine), 5BU can be used to mimic DNA-proteins interactions. The molecule is photoreactive: after UV irradiation it becomes 5,6BU, a cyclized structure in which a covalent bond between the carbons in position 6 and 10 is formed (Figure 1-d). This reaction can be used as a local model of photo-induced DNA protein crosslinking [102].

In this chapter is reported a detailed characterization of the excited states of 5-benzyluracil in the gas phase and in methanol solution as obtained applying Linear Response Time Dependent Density Functional Theory (LR-TDDFT) based calculations. Electronic structure calculations are particularly challenging for 5BU due to the bi-chromophoric nature of this system. In fact in this kind of molecular systems LR-TDDFT often underestimates excitation energies involving states with strong CT character. In general, testing the accuracy of LR-TDDFT (or DFT) calculations is not trivial, since the results depend on the choice of the exchange correlation (xc) energy functional employed. Most functionals are constructed using physical intuition and/or fitting procedures and it is difficult to systematically assess their relative performances in the absence of reference data. On the other hand, using post Hartree Fock multi-determinant methods to fully characterize the excited states of 5BU is not possible with reasonable computational cost. Indeed, it was recently shown in studies of 5BU in gas phase at the CASSCF-CASPT2 level of theory [105] that the size of the active space affordable allows an adequate description of the first excited state of this molecule in the Frank Condon region, while it fails for the second. These CASPT2 calculations however
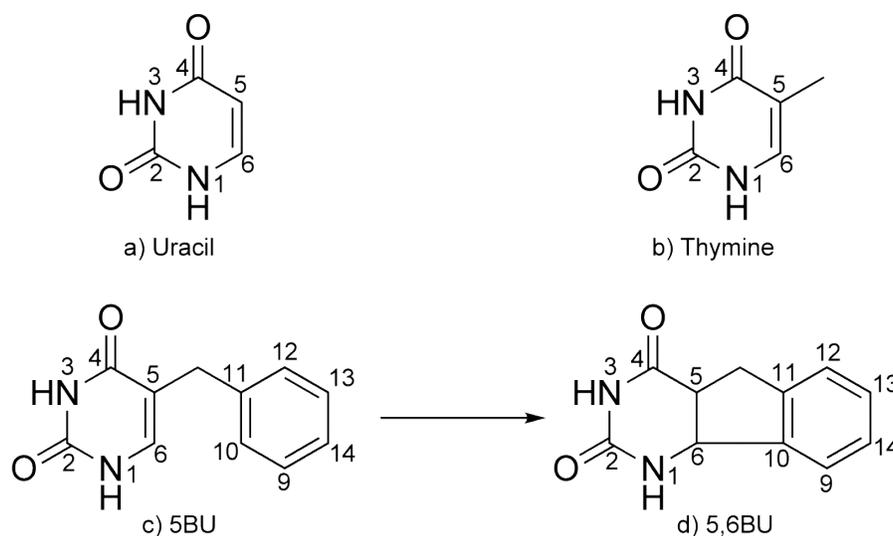
Figure 4.1: Molecular Lewis structure of a) uracil, b) thymine, c) 5-benzyluracil (5BU) and d) its photoproduct (5,6BU). The atoms in the cyclic and eterocyclic rings have been labeled to facilitate discussion in the text (see sec. 3.3).

provide a reference for the first excited state transition energy in the gas phase.

To test the performances of LR-TDDFT, the conceptual ranking of the approximated xc energy functionals in order if increasing complexity proposed in 2001 by Perdew and Schmidt[106] will be followed. This ranking, known as the "Jacob's ladder of DFT/TDDFT", has been discussed in chapter 1. In particular, the performance of popular functionals belonging to different rungs of the ladder will be tested. When comparing their performance the following parameters will be analyzed: i) the absolute and relative energies of the states; ii) the different excited states character's assignment; iii) the accuracy of the CT states energy.

Once the quality of the results obtained with the different functionals is established, the possibility to use the Generalized Gradient Approximation (GGA) for the xc energy functional in the description of the reactive decay path on a CT Potential Energy Surface (PES) will be also investigated. In fact, it will be shown in the following that even if in the Frank Condon region the GGA functionals (in particular the PBE [108]) describe poorly the global manyfold of excited states, the accuracy of this CT PES improves significantly in regions of the molecular configurational space away from the Frank Condon region. This is an interesting result because the low computational cost associated to this approximation in plane waves based *ab initio* codes allows the study of the bonding mechanism with on the fly molecular dynamics techniques. This will be discussed in chapter 6.

4.1 Computational details

All calculations discussed in the following were performed applying DFT and TDDFT methods using the Gaussian09 program package [107]. Benchmarks obtained applying wave function based solutions for the static Schrödinger equation are taken from ref [105]. The LR-TDDFT calculations were performed using six different popular xc functionals belonging to different rungs of the Jacob's ladder that is re-schetched here below for convenience:

$$F_{Hyper}[n, \nabla n, \nabla^2 n, \tau, \{\phi_i\}] \quad || \quad F_{LRC}[n, \nabla n, \quad \nabla^2 n, \tau, \{\phi_i\}]$$
$$F_{Meta}[n, \nabla n, \nabla^2 n, \tau]$$
$$F_{GGA}[n, \nabla n]$$
$$F_{LDA}[n]$$

As discussed in chapter X, at the bottom of the ladder there is the Local Density Approximation (LDA) in which the xc potential depends only on the electronic density n; One rung up is the GGA approximation in which a dependence on the gradient of the density is included. The next rung represents the Meta-GGA functionals for which a dependence on the second order derivatives of the density (the laplacian and the electronic kinetic energy τ) is also considered. Finally at the top of the ladder are the Hyper-GGA functionals in which a direct dependence on the Kohn Sham orbitals $\{\phi_i\}$ is included; at the same level are placed the so called Long Range corrected functionals in which the xc potential is modeled so as to provide the correct asymptotic behavior. In the calculations that will be exposed in the following, in particular, PBE [108] has been chosen in the GGA class of functionals, M06L [109] in the class of Meta-GGA local functionals, B3LYP [110] in the class of Hybrid functionals, the M06 [111] in the Meta-hybrid GGA class of functionals. M062X [111], and CAM-B3LYP [112], with corrected long range behavior for the xc potential, have been also tested to study possible long range effects associated to the CT transitions. In all cases the Dunnings cc-pVDZ [113] correlation consistent basis set was used. For the numerical integrations of the xc energy a grid, with 75 radial shells and 302 angular points per shell has been used. As convergence criterium we used a threshold of 1.00×10^{-06} a.u. in energy and of 1.00×10^{-08} in the Root Mean Square (RMS) of the density matrix.

To compare the different performances, we computed the excited states energy and character using, for all functionals, the same molecular geometry. In particular, it has been chosen the 5BU optimized geometry obtained in ref [105] at the state-averaged CASSCF(20,15) level of theory using the Dunning's double z + polarization [114] basis set with no symmetry restrictions for the molecular geometry. A description of this molecular geometry, in which the most relevant geometrical parameters (bond lengths and bond indices) are discussed, can be found in ref [105]. The use of a single geometry in all comparisons is justified by the observation that geometry optimizations performed with different functionals showed very small variations in the equilibrium structures.

Calculations of 5BU solvated in MeOH were also performed to compare with experimental absorption data and to investigate the decay channel leading to photocyclization. The solvent was represented implicitly and treated as a continuum dielectric medium using the self-consistent reaction-field (SCRF) model in its integral equation formalism (IEFPCM) [115] variant. In this representation, the solvent is characterized solely by its relative static

dielectric permittivity ϵ_r which we set to the value 32.613 [116].

In order to model the absorption in the solvated phase, full geometry optimizations of 5BU in the electronic ground state were performed with different functionals. For each ground state optimized geometry, the electronic excitations energy and character were obtained. To cover the experimental energetic range, the first 10 vertical excitations have been taken into account. Simple Lorentzian functions envelopes have been considered to model the Frank Condon lines broadening.

To further analyze the results obtained for the excited states, the overlap of the orbitals participating to the CT excitations have also been quantified using the Λ diagnostic parameter [117] as implemented in the Gamess program suite [118]. For these calculations the same basis set (cc-pVTZ) was used as well as the same xc functionals adopted in the other calculations performed with Gaussian09. Within Gamess, a Lebedev grid with 96 radial points and 770 angular points was used, for the self consistent field convergence, an energy threshold of 10^{-6} a.u and the MeOH solvation was modeled using a conductor like PCM, with charge scaling set to $\epsilon_r - 1/\epsilon_r$. These changes in the simulation conditions did not cause significative variations in the results. In particular no changes were observed in the energies, transitions assignment and oscillator strengths.

4.2   Results and discussion

4.2.1   Vertical Excitation Energies in gas phase

In ref [105] the first two electronic excited states of 5BU have been characterized in the Frank Condon region using high level post Hartree Fock methods. It has been shown that the first excited state transition has mainly the intrachromophore uracil $n\pi^*$ character while the second excited state transition main character is given by the uracil $\pi\pi^*$ transition. However, as mentioned in the introduction, only the vertical excitation energy to the first excited state is well predicted. In fact, the energy value obtained for the second excited state vertical excitation energy depends strongly on the different approximated forms of the (multideterminamental) wave function considered in the variational calculations. The same effect was already noted in the case of sole uracil [119]. To benchmark the LR-TDDFT results, then, only the lowest $n\pi^*$ transition energy has been considered. In particular the value obtained at the CASPT2(20,14) level has been used as reference (4.94 eV).

The LR-TDDFT vertical excitation energies to the first three excited states obtained with the different tested functionals are reported in Figure 4.2 and their corresponding numerical values (transition energies and oscillator strengths) are reported in Table 1. The different xc functionals are organized following their order in the Jacob's ladder and the CASPT2 reference value is shown with a continuous thin black line to facilitate the comparison with the LR-TDDFT results. First of all, it can be observed that all the predicted energies grow systematically from left to right i.e. as we move up in the Jacob's ladder. This is true in particular for the transition energy to the state with uracil $n\pi^*$ character, reported in the Figure with black continuous lines, whose values can be directly compared with the CASPT2 reference result. As it can be observed, when the GGA approximation is used, the $n\pi^*$ transition is underestimated by almost 0.7 eV with respect to the reference value. This indicates clearly that the most local implementations of the xc functional do

not adequately describe the energy of this transition. Instead, using the Meta-Hybrid-GGA level of approximation (with the M06 functional) the CASPT2 and LR-TDDFT predictions coincide suggesting that for this choice of the xc functional the LR-TDDFT predictions become reliable.

	PBE	M06L	B3LYP
$S_0 \rightarrow S_1$	3.92 (0.02; 0.64)	4.21 (0.02; 0.60)	4.57 (0.04; 0.57)
$S_0 \rightarrow S_2$	4.08 (0.00; 0.42)	4.38 (0.00; 0.59)	4.80 (0.00; 0.48)
$S_0 \rightarrow S_3$	4.15 (0.00; 0.54)	4.59 (0.00; 0.50)	4.87 (0.02; 0.47)
	M06	M062X	CAM-B3LYP
$S_0 \rightarrow S_1$	4.68 (0.04; 0.55)	5.09 (0.01; 0.60)	5.03 (0.01; 0.65)
$S_0 \rightarrow S_2$	4.84 (0.01; 0.57)	5.16 (0.00; 0.48)	5.23 (0.01; 0.43)
$S_0 \rightarrow S_3$	4.95 (0.04; 0.46)	5.30 (0.21; 0.63)	5.27 (0.19; 0.62)

Table 4.1: Gas phase vertical excitation energies of 5BU (in eV) for the first three singlet excited states according to the tested xc functionals. The oscillator strengths and the Λ values corresponding to each transition are reported in the parenthesis (oscillator strength first).

On the other hand, the other two transitions (different from the one discussed so far) involve the π electrons of both chromophores and both have CT character with hole on the benzene and excited electron on the uracil (see Figure 4.3). The energy of these states is usually underestimated by all the local and semi local approximations tested so far and the use of long range corrected functionals (e.g. M062X and CAM-B3LYP) is recommended. In our case, we observe that, with M062X and CAM-B3LYP, the character of these low lying transitions (indicated as the green and blue line in Figure 4.2) changes drastically and they become almost pure intrachromophore transitions typical of the two separated chromophore species (see Figure 4.3). In particular, the character of these transitions is now associated to the benzene dark $\pi\pi^*$ transition (5.10 eV in isolated benzene at the CASPT2 level of theory [105]) and the uracil dark $n\pi^*$ and bright $\pi\pi^*$ transitions (4.54 eV and 5.00 eV in isolated uracil at the CASPT2 level of theory [120]). Their energies are reported in Figure 4.2 with green, black and blue lines, respectively. The energy of the transitions associated to the CT states grows considerably and is not shown in the Figure. Figure 4.2, however, also shows that the transition energy to the uracil $n\pi^*$ state increases beyond the benchmark value when long range corrected functionals are used. Consistently with the argument of Peach et al. [121], this trend may indicate a spurious effect induced by these functionals and requires further investigation. In [117] it was shown that failures of the long range corrections can occur, in particular, when the overlap between the occupied and virtual orbitals participating in the electronic transitions are larger than a given threshold. In the same work, a quantitative measure for this threshold has been introduced using the so called Λ parameter:

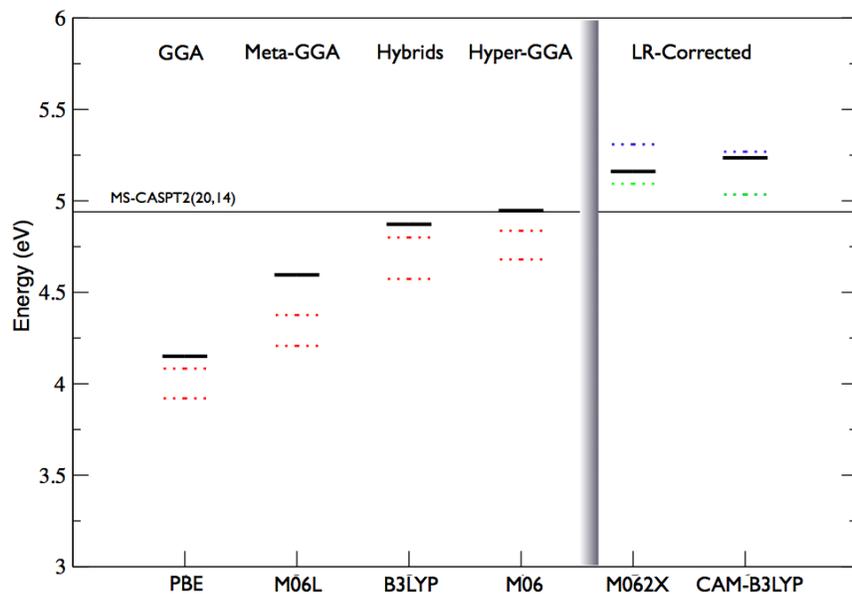$$\Lambda = \frac{\sum_{ia} k_{ia}^2 O_{ia}}{\sum_{ia} k_{ia}^2} \tag{4.1}$$

Figure 4.2: Vertical excitation energies to the first three excited states according to the tested xc functionals. The functionals are arranged following their order on the Jacob's ladder of DFT-TDDFT (as indicated in the upper side of the plot). For all cases, the excitations whose main character is given by the uracil $n\pi^*$ transition have been reported with continuous black lines. The states with intra-chromophore CT character have been reported with dotted red lines. In green and blue are reported the states with intrachromophore benzene and uracil $\pi\pi^*$ character, respectively. The vertical excitation energy of the uracil $n\pi^*$ transition in 5BU computed at the CASPT2 level of theory is indicated with a continuous thin black line, to be used as reference.

where $k_{ia} = X_{ia} + Y_{ia}$ and $O_{ia} = \int d\mathbf{r} \, |\phi_i(\mathbf{r})||\phi_a(\mathbf{r})|$ (see chapter 1 for the definition and derivation of these variables). This parameter takes values $0 \leq \Lambda \leq 1$ and measures the overlap of the orbitals participating in the electronic transition: small values signify long-range excitations and large values signify shod-range excitations. Its calculation has been proposed as a diagnostic tool of the reliability of local and semi-local xc functionals in the evaluation of CT transition energies. In fact, it has been observed that the CT excitation energies are statistically likely to have significant errors if the value of the parameter falls below 0.3 for the GGA (PBE) and below 0.4 for hybrids (B3LYP). When the Λ parameters associated with CT transitions are greater than these values, on the other hand, the CT transitions are frequently well described by standard functionals. The Λ parameters obtained for the transitions of 5BU examined here are reported in Table 4.2.1. As it can be observed, they all lie in the intermediate overlap range between 0.40 and 0.65. While these values are above the critical threshold and seem to indicate that the CT assignment of the states is not overestimated by semi-local xc functionals, they are not large enough to completely resolve the issue. Further analysis will be performed, for the solvated molecule, in the next section.

4.2.2 Solvation effects in the Frank Condon Region

The performance of the different functionals in calculations for solvated 5BU is assessed in this section by direct comparison with experimental absorption spectra. The relevant states in this case are those associated to bright $\pi\pi^*$ transitions. Calculations were performed using: i) the M06 functional which, among the local and semi-local tested functionals, provides the best results in the gas phase ii) the M062X and CAM-B3LYP because the reliability of the CT transition energies could not be assessed with reference calculations in the gas phase; iii) the PBE functional (even though it is known to be less accurate, this functional was also tested in view of future calculations, see next section).

To begin with, ground state geometry optimizations of the 5BU including the SCRF have been performed using these different functionals. The solvated optimized structures show appreciable differences with respect to the gas phase calculation obtained with the same functional as it can be seen by considering panel (a), (b) and (c) of Figure 4.4. Moreover, differently from what observed in gas phase calculations, the equilibrium structures predicted with different functionals also show appreciable differences (see Figure 4.4, panel (d)). The main difference between the molecular structures involves, however, just the relative orientation of the chromophore rings and the different geometries are essentially indistinguishable energetically for all functionals (with the exception of PBE). To quantify this, the following indicator was introduced:

$$\Delta E^\alpha(\{R_\beta\}; \{R_\alpha\}) = E^\alpha(\{R_\beta\}) - E^\alpha(\{R_\alpha\}) \tag{4.2}$$

The indexes α and β run over the tested xc energy functionals and $\{R_\alpha\}$ indicates the minimum energy nuclear structure obtained via geometry optimization using the functional α. As it can be seen in Table 4.2.2, values of the ΔE^α are usually within $k_B T$ at room temperature.

In Figure 4.5, is shown the experimental absorption spectra (continuos black curves), together with the theoretical predictions obtained with different functionals. The numerical
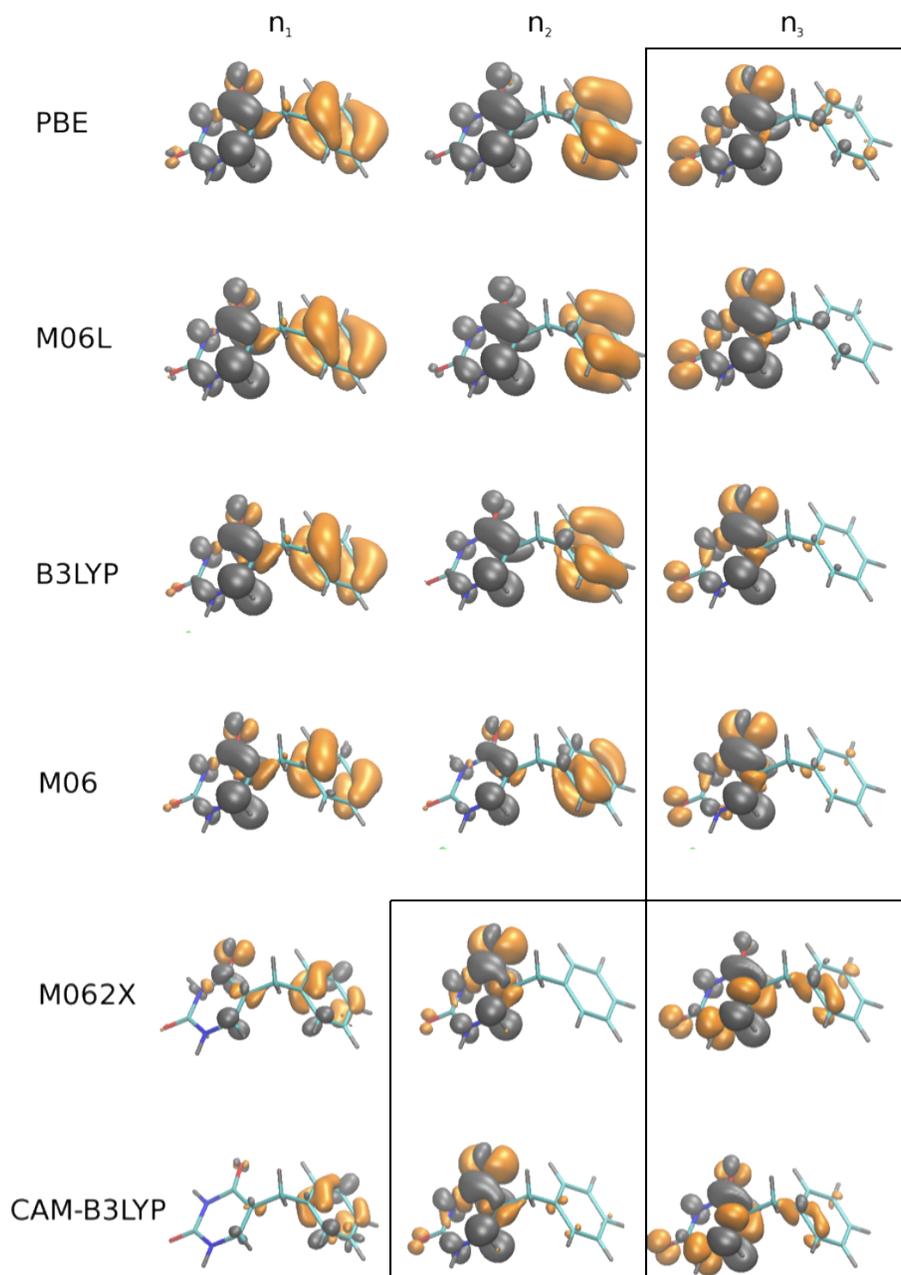
Figure 4.3: Isosurface plots of the first three excited states transition densities obtained, in gas phase, using the xc functionals indicated at the beginning of each line. The positive part is colored in gray while the negative part is in orange. The charge distribution corresponding to the $n\pi^*$ transition (black line in Figure 4.2) is identified by the black frame. Note that, due to the inversion of the order of the transitions discussed in the text, this charge distribution is n_3 for local functionals and n_2 when long range corrections are used. The charge transfer character of the other two transitions is apparent for all local functionals; these transitions become localized on the individual chromophores (see two bottom plots for n_1 and n_3) with long range corrections.

v_{xc} / Rel. Geoms	PBE	M06	M062X	CAM-B3LYP
PBE	0	0.097 (3.8)	0.079 (3.0)	0.093 (3.6)
M06	0.102 (3.9)	0	0.009 (0.3)	0.011 (0.4)
M062X	0.076 (2.9)	0.007 (0.3)	0	0.006 (0.2)
CAM-B3LYP	0.095 (3.7)	0.009 (0.4)	0.010 (0.4)	0

Table 4.2: The computed energy differences $\Delta E^\alpha(\{R_\beta\}; \{R_\alpha\}) = E^\alpha(\{R_\beta\}) - E^\alpha(\{R_\alpha\})$ are reported in eV. The columns in the table correspond to the relaxed geometries (the index β) and the rows to the xc functional used for the calculation (the index α). In parenthesis the energy values in terms of kBT at room temperature (300 K, equivalent to 0.0256 eV) are reported. These values are colored in green if they are smaller than one and red if they are bigger to allow a faster visualization of the results.

values of the first two excited states transitions energies (and oscillator strengths) are reported in Table 4.3. The tested functionals, except for PBE, correctly predict the two band shape of the spectrum (even if they are blue-shifted in all cases). The error in the transition energies can be hence estimated comparing the position of the lower energy peak in the computed spectrum with the experimental one, and results are reported in the last line of Table 4.3. As observed for gas phase calculations, the M06 functional (upper left panel of Figure 4.5) gives the most accurate results with an error of only 0.26 eV. This result confirms that the all transition energies (and in particular the ones with CT character) are well described by this functional. The M062X and CAM-B3LYP results (upper right panel and lower left panel respectively), are almost superimposable and confirm that the long range correction in the xc potential overcorrects the Meta-GGA semi local approximation: excitation energies are overestimated (the error becomes 0.58 eV), consistently with what we observed in the gas phase for the lowest lying dark state excitation energy. The PBE functional (lower right panel) underestimates all the bright transition energies. Furthermore, the red shift of the frequencies is not uniform: higher energy transitions are shifted more than lower energy ones. This results in a merging of the corresponding peaks in the spectrum which loses the experimental bimodal shape. If, however, we focus only on the first two transitions the shape of the lowest energy experimental peak is correctly reproduced even though the peak is redshifted by about 0.57 eV. This indicates that properties that depend only on the first two excited states may be reasonably reproduced also by PBE.

In Figure 4.6 are reported the isosurface plots of the transition densities corresponding to the first two electronic excitations (responsible for the lower energy absorption band formation). The different panels show results for the tested functionals. Differently from what observed in the gas phase results, the interchromophore uracil $n\pi^*$ and $\pi\pi^*$ characters are considerably mixed in both transitions when the M06 functional is used. The origin of this
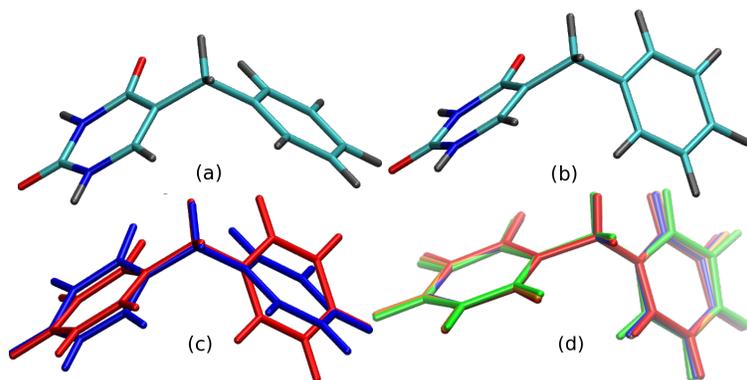
Figure 4.4: Optimized structures of 5BU obtained with the M06 xc energy functional (a) in the gas phase and (b) in the solvated phase. In (c) the two M06 based structures are aligned (the gas phase one is blue and the solvated one in red) to better show the molecular distortion (internal rotation) induced by the SCRF. In (d) are reported the aligned molecular structures obtained with different xc functionals (PBE in green, M06 in red, M062X in yellow and CAM-B3LYP in blue).

	PBE	M06	M062X	CAM-B3LYP	Exp.
$S_0 \to S_1$	4.01 (0.02)	4.88 (0.12)	5.22 (0.01)	5.26 (0.26)	?
$S_0 \to S_2$	4.21 (0.02)	5.01 (0.10)	5.26 (0.27)	5.33 (0.00)	?
Abs. peak	4.11	4.94	5.26	5.26	4.68

Table 4.3: Excitation energies (in eV) for the first two singlet excited states in the solvated phase for the tested xc functionals. Question marks indicate values that cannot be assigned experimentally. Oscillator strengths are reported in parenthesis. In the last line we show the position of the lower energy peak.
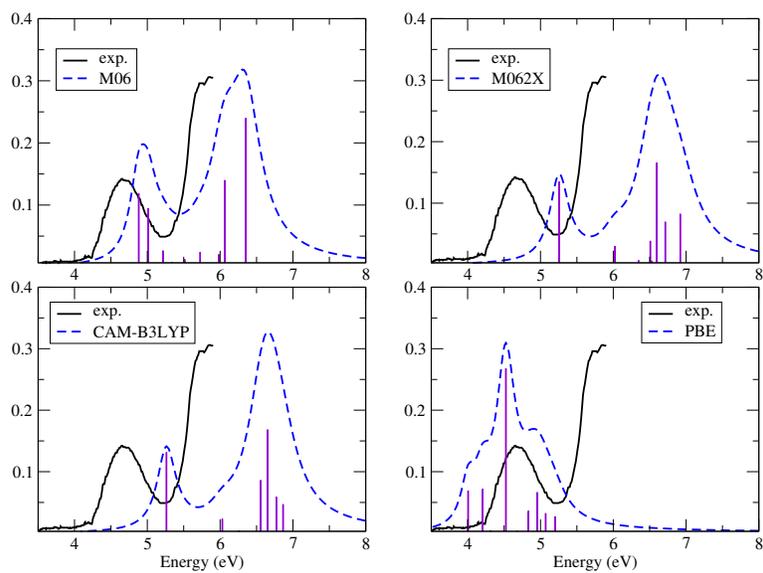
Figure 4.5: The different panels show the solvated phase absorption spectra obtained with the functional indicated in the inset (blue dashed lines). In each panel, the experimental spectrum is also shown as the black solid line. The purple vertical lines (whose height reflects the transitions oscillator strengths) identify the positions of the resonances in the computed spectra.
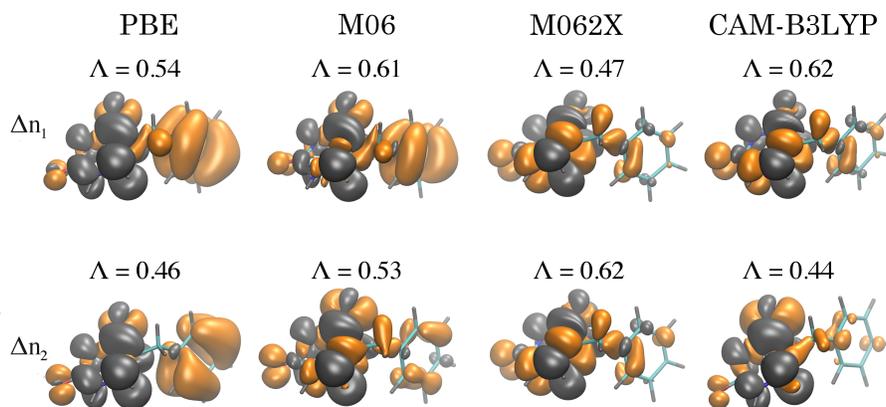
Figure 4.6: Isosurface plots of the transition densities for the first (Δn_1, upper line) and second (Δn_2, bottom line) excitation of 5BU. The different columns refer to the adopted functional, as indicated by the text. In the densities, gray indicates the positive (electron) and orange the negative (hole) densities.

mixing can be attributed to solvent polarization effects on the $\pi\pi^*$ and $n\pi^*$ transitions in uracil already reported in the literature. This solvatochromic effect was in fact observed in studies of isolated uracil in solution performed with both LR-TDDFT [122] and wave function based methods [123]. Moreover, again at variance with the gas phase, the lowest intrachromophore transitions located on the benzene is found at energies higher than 6 eV (results not shown). Finally, Figure 4.6 also shows that, for these transitions, the contribution of the CT component decreases systematically moving up the Jacob's ladder (left to right progression in the Figure). With PBE the CT component is dominant in both excitations. Instead, when the M06 functional is used a strong CT character is observed only in the first excitation (in which the uracil $\pi\pi^*$ transition is also dominant). When the long range corrected functionals are considered, all CT contributions become negligible and the changes in the electronic structure of the molecule are entirely localized on the uracil, with only a slight perturbation induced by the aromatic π electrons of the benzene. The Λ parameters associated to the transitions (reported in Figure 4.6 above each isosurface plot) again reveal that in all cases the orbital overlap lies in the intermediate range in which the long range corrections can worsen the quality of the results.

4.2.3 Reactive state PES

Based on the analysis reported so far, reliable LR-TDDFT results are obtained when the M06 functional is used. In this section, using the M06 energies, a possible decay path, leading to the formation of the C6-C10 covalent bond (see Figure 4.1 panel c for the atoms labeling) in the first excited state of the molecule will be described. These atoms in fact form the extra bond observed in the photoproduct 5,6BU (Figure 4.1, panel d)
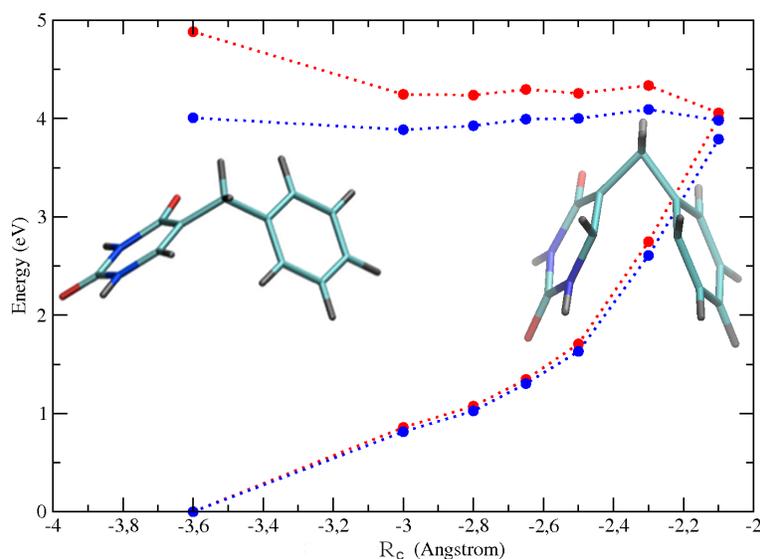
Figure 4.7: Energy scan of the excited zwitterionic PES using the M06 xc functional (in red). The reaction coordinate utilized (R_c) is the negative of the C6 -C10 atoms distance (the minus is used to better read the resulting curve in the left-to-right sense). The results obtained on the same geometries using the PBE xc functional are also reported in blue. A qualitative agreement between the two PESs is observed for values of the C6 -C10 atoms distance smaller than $3\mathring{A}$. In the insets are also reported the molecular structures of solvated 5BU at the ground state equilibrium geometry (left) and when the C6 -C10 atoms distance is $2.1\mathring{A}$ (right).

In the first excited state, due the CT character of the Frank Condon transition, the 5BU assumes a zwitterionic form. This can be seen from the isosurface plot of the electronic transition density of the first excited state obtained using the M06 xc functional in Figure 4.6. As it can be observed the CT character of the excitation induces and electronic charge migration within the molecule from the benzene to the uracil moiety. In order to quantify this charge separation the electron/hole density obtained from the positive/negative part of the transition density has been integrated. A partial charge of -0.42 e on uracil and +0.40 e on benzene was found (the missing 0.02 charge is distributed on the ring substituents). The electrostatic attraction originating from this charge separation, can bring the chromophores closer during this relaxation process towards a stacked configuration.

Following the idea that the excitation of 5BU on this CT state is the responsible for the photo-induced cyclization of 5BU, a scan of the first excited state PES was then performed along the reaction coordinate (R_c) chosen as the negative of the C6-C10 atoms distance. The results obtained are reported in Figure 4.7 (red curve). The point at $R_c = -3.66$ Å is obtained with an unconstrained geometry optimization on the ground state PES (and represent the equilibrium structure in the Frank Condon region). On the other hand, the energy

at $R_c = -2.81$ Å is obtained through an unconstrained geometry optimization on the CT PES and the corresponding structure is a local minimum on this surface (more details on this structure will be given in next chapter). The points at $R_c = -3.0$ Å , $R_c = -2.5$ Å and $R_c = -2.1$ Å have been obtained by means of constrained optimizations (in which all molecular degrees of freedom are allowed to change but the one associated to the distance R_c) while the others have been just extrapolated with single point calculations on the average nuclear coordinate between two optimized geometries. Figure 4.7 shows that there is at least one monotonically decreasing energetic path on this surface that brings the two atoms to a distance of 2.1 Å . The electronic configuration of the two carbon atoms (C6 and C10) changes dramatically along the path. In the reactant state (structure on the left in Figure 4.7) the carbons have a $sp2$ planar configuration which becomes a pyramidal $sp3$ hybridization at the end of the path, allowing for a covalent bond formation among them (structure on the right in Figure 4.7). Moreover, the small energy gap between excited and ground state PES at $R_c \simeq 2.1$ Å (less than 0.1 eV) imply the possibility of nonadiabatic transitions back on the ground state PES. It can be expected that the excess kinetic energy accumulated along the path can drive the two carbon atoms even closer than what is shown in the plot. In this case, if short enough distances are reached during the relaxation motion, the covalent bond formation between the two atoms becomes possible (also in the ground state). Another possible competing reaction path could eventually drive back, after the nonadiabatic transition to the ground state, the system to the starting ground state equilibrium structure (following the ground state PES gradients that are pointing to the starting equilibrium configuration direction). An accurate description of the bond formation mechanism requires the application of ab initio MD methods in which the description of nonadiabatic PES crossings effects are included. This kind of simulations can be efficiently performed using the LR-TDDFT based-Trajectory Surface Hopping approach to nonadiabatic dynamics [125]. The use of hybrid xc functionals is however still beyond reach in this kind of calculations, especially when combined with the use of plane waves basis sets. In view of the nonadiabatic dynamic calculations that will be performed using GGA functionals, the performance of the GGA in the description of this decay path were also tested. In particular, in Figure 4.7 are reported the results obtained computing the PES energy on the same molecular geometries using the PBE functional (blue curve in the figure). As it can be seen, even if the PBE fails in the description of the PES in the FK region, the two PESs become parallel for values of $|R_c|$ shorter than $3\AA$. This observation opens the way to the use of the dynamical ab initio simulations methods as mentioned above to approach the description of the bond formation mechanism. This will be the subject of chapter 6.

4.3   Conclusions

In this chapter it was shown that, when the M06 xc energy functional is used, LR-TDDFT allows a good description of the photophysical and photochemical properties of the 5BU, providing a valid alternative to more sophisticated wave function based calculations, computationally too expensive for this system. In particular, the value obtained for the lowest dark $n\pi^*$ state transition energy in the gas phase is in very good agreement with CASPT2 level calculations. Moreover, we also observe a good agreement with experimental

data in the evaluation of the lowest bright $\pi\pi^*$ states transition energies in the solvated phase (within an error of 0.27 eV).

Tests performed using different xc energy functionals showed that, when the functional is too local, the transition energies are severely underestimated. The reason behind this behavior is the CT character associated to the excited states of the molecule. It was also showed that using long range corrections in the functionals does not improve the results. On the contrary, the transition energies are overestimated and the CT presence in the character of the states is spuriously under-predicted.

The presence of the strong CT character in the first electronic transition as predicted by the M06 functional may play a critical role explaining the photoreactivity of the 5BU: the zwitterionic character of 5BU on the excited state induces a relaxation motion that can lead to the covalent bond formation observed in the photoproduct. The performance of the GGA approximation in computing the PES along this putative reaction path was also investigated in view of the application of TDDFT based nonadiabatic dynamical simulations methods (at present affordable only with this functional class). The use of these methods will allow to study the bonding mechanism at the base of the cyclization photoreaction that, from 5BU, leads to the 5,6BU formation. In particular it was found that, even if the GGA approximation for the xc functional cannot describe the photophysical properties of the molecule in the Frank Condon region, a qualitative agreement with the M06 results is obtained in the description of the CT state PES when regions of the configurational space far from the Frank Condon region, but relevant for the study of the decay path, are considered.

Chapter 5

Photophysical and photochemical properties of 5-benzyluracil and 5,6-benzyluracil

In this chapter, the photophysical and photochemical properties of 5-benzyluracil and 5,6-benzyluracil, the latter produced by photocyclization of the former through irradiation with femtosecond UV laser pulses, are investigated combining results coming from spectroscopic measurement and LR-TDDFT based calculations.

First, the absorption spectra of the two molecules are compared and the principal electronic transitions involved discussed, with particular emphasis on the perturbation induced on the two chromophore species (uracil and benzene) by their proximity. It will be then shown that the steady state fluorescence measurements allow detection and selective identification of 5- and 5,6-benzyluracil. The principal electronic decay paths of the two molecules, obtained through calculations, will explain the features observed in the emission spectra and the photo-reactivity of 5-benzyluracil. The order of magnitude of the lifetime of the excited states will then be derived with steady-state fluorescence anisotropy measurements and rationalized with the help of the computational findings. Finally, the spectroscopic data collected will be used to derive the photocyclization and fluorescence quantum yields.

In obtaining a global picture of the photophysical and photochemical properties of the two molecules, the findings reported will demonstrates that the use of 5-benzyluracil as model system to study the proximity relations of a DNA base with a close lying aromatic aminoacid is valid at a local level. It will be shown, in particular, that the main characteristics of the decay processes from the excited states of the uracil/thymine molecules remains almost unchanged in 5-benzyluracil, the main perturbation arising just from the presence of the close lying aromatic group.

5.1   Experimental materials and computational details

5BU powder was dissolved in methanol and absorbance and emission spectra were acquired using a quartz cuvette with optical length of 1 cm. The absorption spectra were acquired in the range 210-380 nm at several concentrations, both for 5BU and 5,6BU, verifying the linearity of the response. For the fluorescence spectra, the excitation wavelength was 265 nm and the fluorescence was collected in the range 290-380 nm. 5,6BU was obtained by UV photoirradiation of 5BU by a femtosecond laser system (Light Conversion, Lithuania). This device operates at 258 nm, 200 fs pulse duration, 2 kHz repetition rate thereby delivering 0.2 W average power. The sample was constantly stirred with a magnetic bar to have a homogenous solution while irradiating. For the spectral measurements of the 5,6BU, samples obtained after 1 minute exposition to the laser UV laser pulses were used since it was verified that these samples contained only 5,6BU. The characterization of the final product after UV laser pulse irradiation of the 5BU, essential to verify the occurrence of the photocyclization process and the consequent complete transformation of 5BU into 5,6BU was obtained via High-performance liquid chromatography (HPLC) and Nuclear Magnetic Resonance (NMR) spectroscopy of the pure and irradiated 5BU. A more detailed discussion of experimental apparatus and experimental conditions can be found in ref [126].

All calculations were performed applying DFT and TDDFT methods and using the Gaussian 09 package [132]. In chapter 4 the results obtained using the M06 xc functional [133], in the meta-hybrid class of functionals, were systematically tested proving the reliability of the results obtained for the electronic structure prediction of the 5BU. Consistently with what reported in chapter 4, also for the calculations exposed in this chapter the Dunnings cc-pVDZ correlation consistent basis set [134] was used, for numerical integrations a pruned grid has been used, with 75 radial shells and 302 angular points per shell and the MeOH solvation effects are taken into account by means of a self-consistent reaction-field (SCRF) model. Full geometry optimizations of the reactant and product in their singlet ground state were performed to then apply LR- TDDFT and determine the excited state energies and oscillator strengths. To cover the experimental energy range, the first 10 vertical excitations have been computed for both molecules and the absorption spectrum was modeled with Lorentian functions envelopes. Full geometry optimizations of the nuclear positions have been performed also on the excited states using the M06 functional (and, for 5BU, also CAM-B3LYP, see section 5.4 for an explanation of this choice) for the characterization of the electronic decay channels.

5.2   Photophysical properties in the Frank Condon region

In Figure 5.1 the DFT/M06 ground state structures of the 5BU and 5,6BU in MeOH are reported. As discussed in chapter 4, in the case of 5BU the structure shows appreciable differences from the more symmetric geometry obtained in gas phase.This distortion is due to a solvent induced internal rotation around the hinge connecting the two chromophore moieties. For 5,6BU the changes in the molecular geometry predictions upon solvation are instead negligible. The photoreaction product has in fact a much less flexible structure due
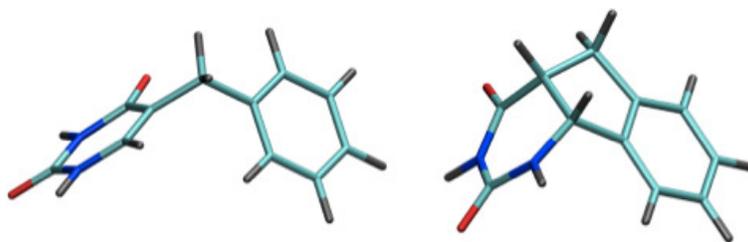
Figure 5.1: Stick representations of the solvated ground state molecular geometries of 5BU (left) and 5,6BU (right). Carbon atoms are colored in cyan, oxygen in red, nitrogen in blue and hydrogen in gray.

to C6-C11 (see Figure 5.2 for atoms label) covalent bond formation.

The UV electronic absorption spectra of 5BU and 5,6BU in MeOH were experimentally recorded and computed with LR-TDDFT to characterize the principal electronic transitions. The experimental and computed absorption profiles of the molecules are reported in Figure 2 (respectively with dashed red line and continuous black line). The data reflect the dramatic electronic structure after cyclization. Although the calculated bands show a systematic blue shift, the overall agreement between theory and experiments is sufficient to allow an unequivocal assignment of the transitions. To analyze the spectrum of the 5BU (left panel) let us recall that this molecule is chemically composed of two chromophores: the benzene (model of aromatic amino acid) and the uracil/thymine moieties. The calculations show that the electronic transitions involve mainly the uracil fragment. In particular, the main contribution to the first two low energy electronic transitions originates from the typical uracil bright $\pi\pi^*$ and dark $n\pi^*$ transitions [136], while no intrachromophore transitions are observed for the benzene moiety. This can be explained as follows: the lowest $\pi\pi^*$ electronic transition of benzene [137], the only in the range of the experimental irradiation energy (see also the following discussion on the decay channels), has an oscillator strength that is much lower than the bright uracil $\pi\pi^*$ transition and its contribution to the absorption spectrum of 5BU is therefore negligible.

The computed uracil transition energies in 5BU are slightly red shifted compared to the ones relative to pure uracil in similar solutions [130] due to the perturbation induced by the chemical bridging to the second chromophore (benzene). In addition to these intrachromophore transitions, the charge transfer (CT) excitations from the benzene to the uracil moiety contribute to the lower energy band centered at 265 nm in the experimental spectrum (around 60% of the TDDFT absorption line at 248 nm). This result show that the reactive zwitterionic CT excited state of 5BU, characterized in chapter 4 can be effectively reached upon UV irradiation of 5BU. Moreover, in the two lowest absorption lines of 5BU the uracil $\pi\pi^*$ and $n\pi^*$ characters are mixed and hence the two transitions have similar oscillator strengths (two separate bright uracil $\pi\pi^*$ and dark $n\pi^*$ transition are observed for sole uracil in gas phase calculations [137]). The origin of the mixing is mainly due to the solvent induced quasi-degeneracy of the $\pi\pi^*$ and dark $n\pi^*$ transitions. This solvatochromic effect was already detected in studies of solvated uracil performed both via LR-TDDFT methods
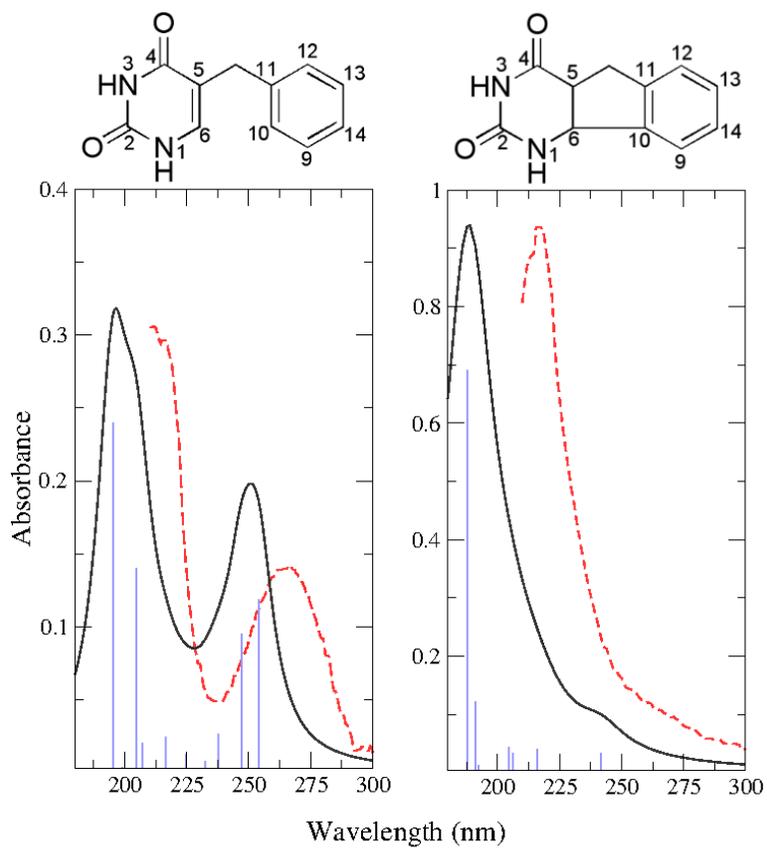
Figure 5.2: Absorption spectra of 5BU (left) and 5,6BU (right). Experimental spectra (red dashed lines) are shown together with LR-TDDFT/M06 convoluted spectra (continuous black lines). The vertical excitations energies (resonances) are represented by blue bars with a height corresponding to the oscillator strength. Above each plot the Lewis structures of the two compounds are reported. The atoms of the two rings have been labeled with numbers.
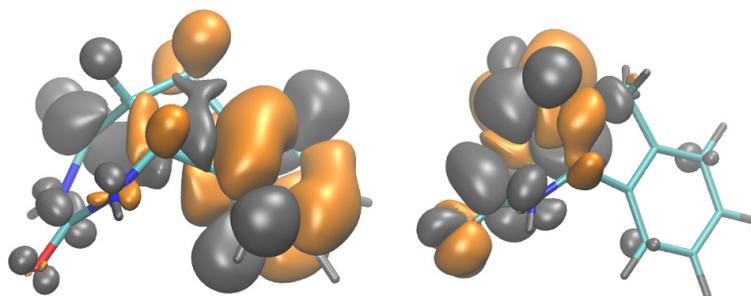
Figure 5.3: Isosurface plots of the transition densities of the first (left) and second (right) excitations of the 5,6BU in the Frank Condon region. Gray indicates positive (electron) and orange negative (hole) contributions.

[130] and wave function based calculations [138].

Concerning 5,6BU (data shown in the right panel of Figure 5.2), an unequivocal assignment of the electronic transitions in the range between 250-300 nm (that includes the laser wavelength, 260 nm) is not possible because of the weak absorption around 240 nm, which just results in a shallow shoulder in the computed spectrum not resolved experimentally (at room temperature). The first excited state transition is the closest and brightest in that spectral region but, together with this transition, higher energy excitations can be activated. The cyclization process induces a change in the π-conjugated system of the uracil moiety (disappearance of the enonic group) with consequent loss of the $\pi\pi^*$transition observed in isolated uracil and 5BU. Characteristics of the 5,6BU are therefore the intrachromophore aromatic $\pi\pi^*$transition in the benzene moiety (at 240 nm with an oscillator strength of roughly one third of the $\pi\pi^*$transition in uracil) and the $n\pi^*$transition of uracil (see Figure 5.3, left and right parts, respectively). As for the $\pi\pi^*$transition in the uracil moiety, the interchromophore CT band observed in 5BU disappears in the photocyclized product. In addition, the second excitation has an almost pure character given by the dark uracil $n\pi^*$transition (that is no longer mixed with the uracil $\pi\pi^*$transition, since the latter is missing in 5,6BU) and, accordingly, its oscillator strength becomes zero.

5.3 Optical characterization of the photocyclization process

We started the optical characterization of 5BU and 5,6BU by measuring, as a function of the irradiation time, their absorption spectra. The obtained results are plotted in Figure 5.4. The change of 5BU into 5,6BU can be considered complete after 60 s irradiation since no more significant changes in the spectra are detectable with longer irradiation times. Figure 5.4 also shows that the 5,6BU absorbance at 220 nm is approximately three times higher than that of 5BU, in agreement with the corresponding oscillator strengths computed with TDDFT. It is worth mentioning that at the laser wavelength (258 nm) both species absorb
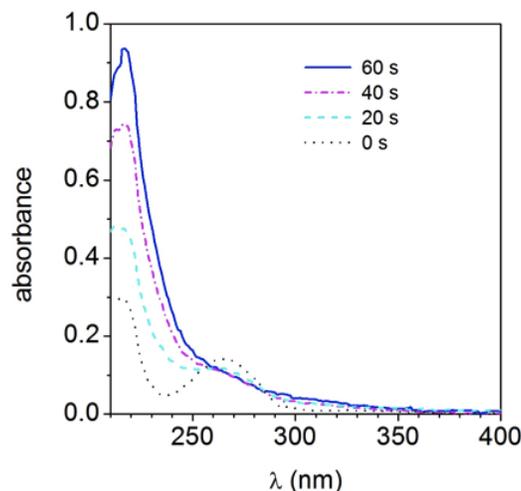
Figure 5.4: Absorbance of irradiated 5BU. The irradiation time is shown in the legend: dotted, 0 s; dashed 20 s; dash-dotted, 40 s and solid line 60 s. 0 s corresponds to 5BU and 60 s corresponds to 5,6BU.

with a comparable efficiency. Within the experimental errors in fact, an isosbestic point is found at 260 nm (very close to the UV laser photons). This indicates also, already at a pure spectroscopic level, that no photoproduct other than 5,6BU is produced by UV irradiation. In fact, the presence of an isosbestic point typically does indicate that only two species that vary in concentration contribute to the absorption (at least in spectral regions around the isosbestic point).

Similarly to what was done for absorption, fluorescence emission spectra were measured as a function of irradiation time, so that the change of 5BU into 5,6BU could be again followed as a function of the irradiation time. Compared to the absorption, the change in the spectral features is much more striking. Figure 5.5 shows that a significant shift of the peak from 310 nm to 360 nm occurs when 5BU is photocyclized into 5,6BU (dotted line t=0 s versus solid line t=60 s, in Figure 5.5).

5.4   Excited states decay channels

In order to rationalize the steady state fluorescence features observed experimentally, the decay channels of reactant and product were computationally studied via minimum energy path analysis (geometry relaxations in the electronic excited state). As discussed in chapter 3, it is known from several quantum chemical calculations [127, 128] that the excited state ultrafast decay of DNA/RNA bases is favored by the existence of barrierless paths that lead to very efficient non radiative decay processes via conical intersections (CI) with the ground state. Moreover, recent nonadiabatic molecular dynamics studies in gas phase [139]
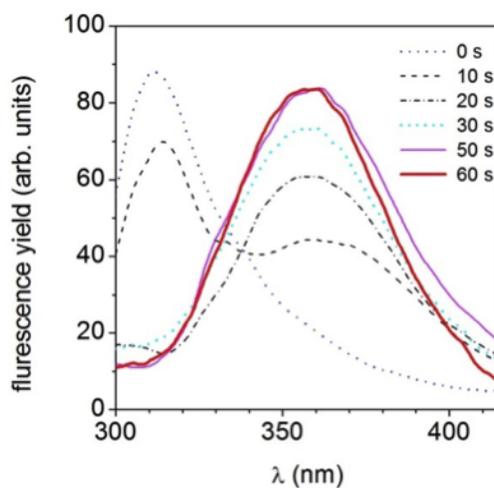
Figure 5.5: Fluorescence of irradiated 5BU ($\lambda_{exc} = 265$ nm). The irradiation time is shown in the legend: dotted, 0 s; dashed 10 s; dash-dotted, 20 s; short-dashed 30s, thin solid line 50 s and thick solid line 60 s. As it will be shown in the following 0 s corresponds to 5BU and 60 s corresponds to 5,6BU.

have shown that, after excitation to the $\pi\pi^*$state, the pyrimidine nucleobases remain mainly trapped into a shallow energy minimum for a time of the order of few picoseconds before finding their way to a CI with the ground state through a non-radiative path. In our calculations on 5BU we found a relaxation path on the $\pi\pi^*$state leading to a stable intermediate state in which a molecular distortion similar to that observed in uracil and thymine occurs (puckering out of plane motions). The molecular structure obtained following this relaxation path is reported in the upper panel of Figure 5.6. It shows that the structural changes typical of the nucleobases relaxations discussed above are observed and that the changes in the bonds and angles structure are confined to the uracil moiety. In Figure 5.6 the electronic transition density for this structure is also reported. It clearly shows the dominance of the uracil $\pi\pi^*$transition in the excited state character. Since experimentally no emission around 330 nm (the computed fluorescence wavelength corresponding to this geometry) is observed, it can be concluded that this metastable intermediate leads to non-radiative decay paths through CI and is non-emitting. This is hence a decay property of nucleobases that is stored in 5BU. It has to be mentioned that for the study of this relaxation pathway, the use in the calculations of the long range corrected CAM-B3LYP xc functional has played a key role. In fact, even if it has been shown in chapter 4 that more reliable results are obtained using the M06 xc functional in the calculations, when the CAM-B3LYP is used the out of plane molecular motions leading to this relaxed structure are predicted completely barrier less; when the M06 functional is used instead a small barrier in the excited PES have to be crossed to arrive to this final relaxed structure which is not reached due to the absence of
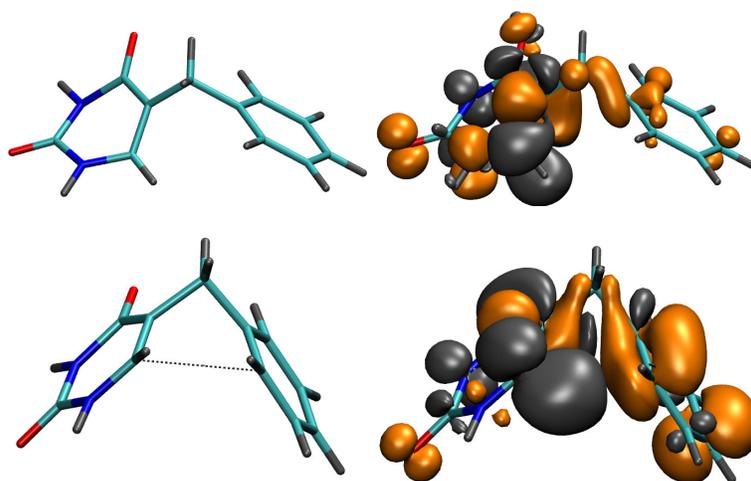
Figure 5.6: Molecular geometries and corresponding transition densities of 5BU relaxed in the two minima on the $\pi\pi^*$state. The minimum found with TDDFT/CAM-B3LYP is reported in the upper panel, while the one found with TDDFT/M06 is shown in the lower panel. The color code for the isosurface plots is the same as Figure 3.

temperature effect in the geometry optimizations steepest descent algorithms.

In 5BU a decay channel not observed in single nucleobases, was found to be activated. This relaxation is driven by the CT $\pi\pi^*$state with hole on the benzene and excited electron on the uracil moiety, which, as discussed in chapter 4, gives a zwitterionic character to the molecule in the excited state. The relaxed molecular structure on this PES is reported in the bottom panel of Figure 5.6, where the corresponding transition density is also reported to highlight its CT character. It can be observed that the attraction originating from this charge separation, brings the chromophores closer during this relaxation process and brings them towards a stacked configuration. The C5-C6-C15 angle (see Figure 5.2 for the atom labels) between the two moieties changes from a value of $113\,^\circ$ in the Frank Condon region to a value of $93\,^\circ$ in the relaxed CT $\pi\pi^*$configuration, while the C6-C10 distance decreases from 3.66 Å to 2.81 Å. As it was shown in chapter 4, the excess kinetic energy acquired during this "down the hill" relaxation motion (~ 1 eV) can lead to a sort of molecular butterfly relaxation motion until the formation of the C6-C10 covalent bond observed in the cyclized compound is going to be formed.

Finally, the relaxed structures of the sole uracil and 5BU on the $n\pi^*$PES are reported in Figure 5.7 together with the isosurface plot of the corresponding transition densities. It can be observed that while the transition in uracil has a pure $n\pi^*$character, in 5BU the $n\pi^*$state strongly mixes with the CT $\pi\pi^*$state. The photoemission energy of this configuration is about 311 nm, in good agreement (considering the global blue shift of the TDDFT energies) with the position of the experimental fluorescence peak. Hence, this is the most likely candidate identified by our calculations for the stable radiative state associate to the
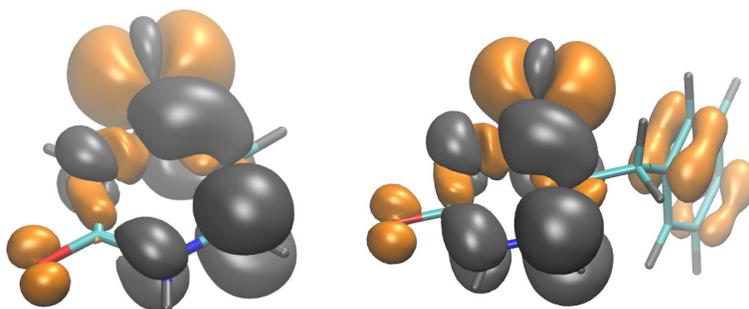
Figure 5.7: Transition densities for the $n\pi^*$ state in the relaxed geometry of uracil alone (left) and the 5BU (right). The color code for the isosurface plots is the same as in Figure 3.

fluorescence measured experimentally. The mixing of the $n\pi^*$ state with the CT $\pi\pi^*$ state is also probably responsible for the efficiency of the emission, which gains oscillator strength from the $\pi\pi^*$ contribution.

In Figure 5.8 a cartoon summarizes the different deactivation channels of 5BU suggested by the calculations performed: the photoreactivity of excited 5BU has be attributed to the presence of a CT state characterized by the formation of a (partial) hole on the benzene moiety and the transfer of a fractional charge to the uracil. In this CT state the two chromophores approach each other because of Coulomb attraction. As shown in chapter 4, following this relaxation path, the molecule can arrive in a configuration in which the energy gap between ground and excited states becomes very small and the new covalent bond observed in the cyclized product 5,6BU can eventually be created once the system relaxes back to the electronic ground state (Figure 5.8b). Other deactivations channels can be activated involving the uracil moiety, but these intrachromophore transitions are not directly involved in the formation of the cyclized product. In particular, the relaxation on the uracil n* state leads to the radiative decay on the ground state (Figure 5.8a) while the relaxation on the uracil * state leads to an ultrafast non radiative decay path similar to those reported in the literature for uracil and thymine [139] (Figure 5.8c).

The relaxation channels described so far do not involve excitations localized solely on the benzene. In fact, while these excitations are clearly possible, they occur at energies higher than those relevant to the experiments performed: the UV laser irradiates at an energy of 4.8 eV while the first bright transition in benzene is at 6.20 eV [140] and multiphotonic absorption processes do not play a role in the formation of the final photoproduct, as it has experimentally shown in ref [102].

Finally, it is also important to stress that the covalent bond formation between the carbons C6 and C10 in 5BU does not coincide with the generation of the final product 5,6BU. The full reaction mechanism requires an additional tautomerization reaction, causing a proton to be transferred from the carbon atom C10 to the carbon atom C5. This part of the reaction, together with the modeling of formation of the covalent bond will be discussed in next chapter.
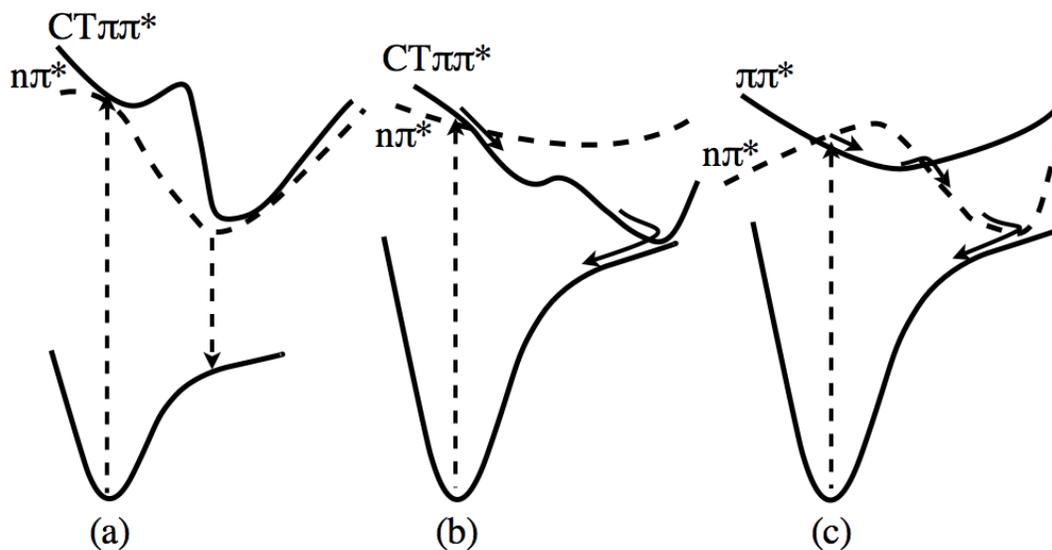
Figure 5.8: Schematic description of the possible decay paths from the electronic excited states of 5BU based on our calculations: a) relaxation path via the intrachromophore uracil $n\pi^*$state (radiative decay channel); b) relaxation path via the inter-chromophore CT $\pi\pi^*$state (non-radiative and reactive decay channel); c) relaxation path via the intra-chromophore uracil $\pi\pi^*$state (non-radiative and non-reactive decay channel).

Ex. state order	Absorption En. (nm)	Fluorescence En. (nm)
1	243 (0.03)	260
2	232 (0.00)	343
3	221 (0.00)	307
4	217 (0.04)	242
5	207 (0.03)	343
6	205 (0.04)	263

Table 5.1: List of absorption and relaxed excitation energies for the first 5 singlet excited states. In parenthesis are reported the oscillator strengths.
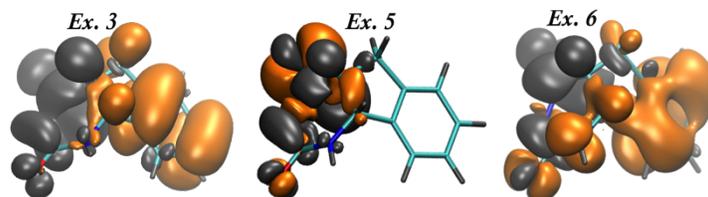
Figure 5.9: 5,6BU electronic transitions relative to the molecular geometries relaxed on the 3rd, 5th and 6th excited states. The central one is relative to the fluorescence peak of the compound and has the same ($n\pi^*$) character of the stable emitting state in 5BU. The color code for the isosurface plots is the same as in figure 3.

The 5,6BU electronic decay channels were also explored to characterize the measured fluorescence. The relaxed photoemission energies are reported in Table 1 in while in Figure 5.9 are reported the main relaxed structures with the corresponding transition densities isosurfaces. Also in this case, a good agreement was found between the position of the experimental fluorescence peak (350 nm) and the photoemission energy from the energy minimum on the n* PES (330 nm). It can be therefore concluded that for both 5BU and 5,6BU the experimentally observed fluorescence in the energy window 300-420 nm corresponds to the emission from two different minima on the same $n\pi^*$ surface (with possible mixing with the CT $\pi\pi^*$ state). However, it has to be be mentioned that in 5,6BU the decay in the $n\pi^*$ state is probably not the only fluorescent channel: in fact, as it can be noted from Figure 5.5, the fluorescence signal remains appreciable (above the base line) at high energies ($\geqslant$ 300 nm) where the emissions from the structures relaxed on the third and sixth excited states are located (see Table 1).

5.5 Excited states decay time

To gain insight into the electronic relaxation time scale of the 5BU and 5,6BU we also measured the fluorescence anisotropy through the polarization coefficient [140]

$$A = (I_{||} - I_{\perp})/(I_{||} + I_{\perp}) \tag{5.1}$$

as a function of the laser irradiation time. In Eq. 5.1 $I_{||}$ and $I_{\perp}$ are the fluorescence yield polarized parallel and orthogonal to the polarization direction of the incident beam, respectively. Results are reported in Figure 5.10 and show that the $A \simeq 0.5$ at $t = 0$, i.e. when only 5BU is in solution, and reduces to approximately 0.17 for long irradiation time when 5BU is completely photocyclized (i.e. for 5,6BU).

It is known that the time scale of the rotational relaxation time of organic dyes having shape and steric hindrance similar to 5BU and 5,6BU in MeOH solution lies in the sub-nanosecond range (70-80 ps) [141, 142]. Therefore, the same time scale can be reasonably assumed for the rotational relaxation time of 5BU and 5,6BU. As a consequence, the high value of anisotropy found for 5BU implies that its fluorescence signal has to decay in a
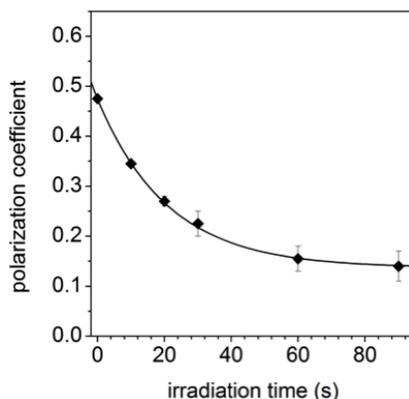
Figure 5.10: Anisotropy of the fluorescence emission as a function of the irradiation time measured at $\lambda = 320$ nm. At t $= 0$ s only 5BU is present which photocyclizes to 5,6BU when irradiated. The continuous line is the best fit with the function (see section 6) that leads to $\tau = 20.9 \pm 0.6$ s. The error bars come from the uncertainty in the measurement of the fluorescence yield propagated in the equation for the polarization coefficient.

time-scale of few picoseconds. This confirms that the proximity of the benzene moiety to the uracil in 5BU does not perturb significantly the electronic structure of the uracil side and the molecule preserves its ultrafast decay properties.

This experimental observation is somewhat surprising since it sets 5BU apart from other compounds obtained via chemical substitution on the uracil carbon C5. In fact, experiments in which the effects of other substituents (e.g. 5- OH-Ura and 5-NH2-Ura) were studied, consistently showed that the excited state lifetime significantly increases upon substitution (to nanosecond or more) [144]

On the other hand, the lower anisotropy in the 5,6BU is a signature of a decay time lying in the range of nanosecond. This can be ascribed to the changes in the electronic structure which, as discussed above, entail the disappearance of the uracil $\pi\pi^*$transition in 5,6BU, which is responsible for the ultrafast decay time of 5BU.

5.6   Photocyclization QY

The features observed in the fluorescence spectra reported in Figure 5.5 allow an indirect measurement of the photocyclization QY (p_{pc} in the following), via the knowledge of the characteristic time for $5BU \rightarrow 5,6BU$ transformation. To show this, the fluorescence yield in the 355-365 nm range is reported as a function of the irradiation time in Figure 5.11. This plot shows that the collected fluorescence signal saturates after approximately 1 minute of irradiation at 0.25 W. This behavior can be explained by considering the occurrence of the reaction $h\nu + 5BU \rightarrow 5,6BU$ to be governed by a first order kinetic together with the conservation of the number of molecules; hence we have
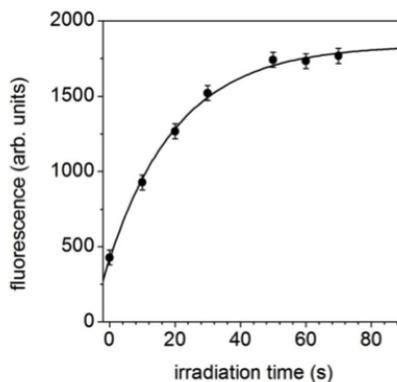
Figure 5.11: Fluorescence yield of the irradiated solution in the range 355-365 nm as a function the irradiation time. The continuous line is the best fit with the function $f(t) = a + be^{-t/\tau}$ that leads to $\tau = 21.3 \pm 1.4$ s.

$$\frac{d}{dt} N_{pc}(t) = \frac{N_{5BU}(t)}{\tau} \tag{5.2}$$

$$N_{5BU}(t) + N_{pc}(t) = N_{5BU}(0) \tag{5.3}$$

where t is the irradiation time and N_{5BU} and N_{pc} are the concentration of 5BU and 5,6BU, respectively. In Eq. 5.2, τ is the characteristic time for the transformation of 5BU into 5,6BU and is given by

$$\tau = \frac{1}{f_L p_{abs} p_{pc}} \tag{5.4}$$

where f_L is the repetition rate of the laser (2 kHz), p_{abs} is the probability that a molecule absorbs one photon in a single laser pulse and p_{pc} is the probability that the excited molecule will photocyclize (i.e. p_{pc} is the photocyclization quantum yield). The assumption $\tau \gg \tau_{eq}$, τ_{eq} being the time required for homogenizing the solution, is necessary for the validity of Eq. 5.2 and is well fulfilled, i.e. it can be reasonably assumed that the gradient in the 5,6BU concentration between irradiated and not irradiated volume is negligible.

Since the fluorescence signal is proportional to the concentration and the latter can be easily derived from Eq. 5.2, the following relation for the overall fluorescence detected in the wavelength range 355-365 nm is derived:

$$FY(t) = k_{inst} \, N_{5BU}(0) \left[FY_{pc} - (FY_{pc} - FY_{5BU}) \, e^{-\frac{t}{\tau}} \right] \tag{5.5}$$

In Eq. 5.5 FY_{pc} and FY_{5BU} are the fluorescence quantum yields of 5,6BU and 5BU, respectively, k_{inst} is an instrumental constant. The best fit of the experimental data in Figure 11 with a function like that in Eq. 5.5 allows to estimate the characteristic photocyclization

time as $\tau = 21.3 \pm 1.4$ s. To check the validity of this estimate we considered also the dependence of the polarization coefficient A on the irradiation time (see Figure 10) that follows a law similar to Eq. 5.5, i.e. we have

$$A(t) = A_{pc} - (A_{pc} - A_{5BU}) \ e^{-\frac{t}{\tau}} \tag{5.6}$$

where A_{5BU} and A_{pc} are the polarization coefficients of the 5BU and 5,6BU, respectively. The best fit of the experimental results in Figure 10 allows to estimate the characteristic photocyclization time as $\tau = 20.9 \pm 0.6$ s in very good agreement with the estimation through fluorescence yield. From the two independent measurements, the combined estimation of this time is $\tau = 21.2 \pm 0.6$ s.

Apart from p_{pc}, the only remaining unknown quantity in Eq. 5.4 is p_{abs}, which can be estimated as follows. Far from the saturation (linear absorption) the probability that a photon is absorbed by a molecule is proportional to the photon number N_f and to the absorption cross section σ_{abs}, i.e.

$$p_{abs} = \frac{N_f \ \sigma_{abs}}{S} = \frac{E_{pulse} \ \sigma_{abs}}{h\nu \ S} \tag{5.7}$$

where S is the effective area illuminated by the laser. In the experimental conditions examined here, this area coincides with the section of the cuvette since the mixing is fast enough to consider the solution homogeneous. Measured the absorptivity ϵ of the 5BU to be 11,900 $Mol^{-1}cm^{-1}$, which in turn provides $\sigma_{abs} = 3.82 \times 10^{-21} cm^2$ and $\epsilon = 4.5 \times 10^{-17} Mol^{-1}cm^{-1}$ and since we had $E_{pulse} = 170\mu J$, $h\nu = 4.8$ eV ($\lambda = 258$ nm) and S $= 1cm^2$, the probability that in our experimental condition a photon is absorbed by a 5BU molecule is $p_abs = 10^-2$. Finally, from Eq. 5.2, by considering $f = 2$ kHz, we obtained $p_{pc} = 2.4 \times 10^-3$ in excellent agreement with Sun et al. [102].

5.7   Fluorescence Quantum Yileds measurements

The fluorescence of 5BU and 5,6BU is provided by [129]:

$$QY_X = QY_R \frac{I_X \dot{O}D_R \dot{n}_X^2}{I_R \dot{O}D_X \dot{n}_R^2} \tag{5.8}$$

where OD stands for "optical density" (absorbance), I for the fluorescence intensity and n for the refraction index. The subscript X and R refer to the molecule whose QY has to be measured and to the reference, respectively. Since 5BU is diluted in methanol and the reference molecule (tryptophan) in water, then $n_X \simeq n_R$, and hence

$$QY_X = QY_{Try} \frac{Slope_X}{Slope_{Try}} \tag{5.9}$$

In Eq. 5.9 Slope refers to the slopes in Figure 12, X stands for 5BU or 5,6BU, respectively. By using $QY_{Try} = 0.13 \pm 0.01$ as reported in ref [143], we obtained $QY_{5BU} = 0.11 \pm 0.02$ and $QY_{5,6BU} = 0.38 \pm 0.02$. These values represent the first measurements of the fluorescence quantum yield for 5BU and 5,6BU. QY for both molecules are four order of magnitude higher than fluorescence QY for thymine and uracil [144].
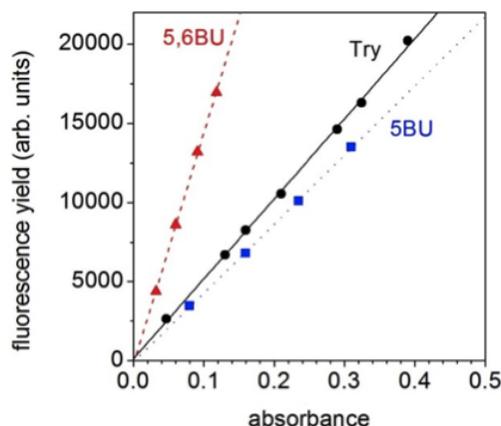
Figure 5.12: Fluorescence vs absorbance for 5BU, 5,6BU and Tryptophan ($\lambda_{exc} = 278$ nm), the latter used as standard.

The relatively high fluorescence QY in 5BU can be explained with the lack of some of the non-radiative decay channels which are active in the DNA and RNA bases. Indeed, for these molecules and in particular for isolated uracil and thymine, non-adiabatic dynamics studies in the gas phase [139] show that after photoexcitation into the bright $\pi\pi^*$state, the molecules can decay via two main non radiative pathways. Following the first pathway the nucleobases relax into the lower lying $n\pi^*$state and from there into the ground state via a CI. The other relaxation pathway leads to the ground state directly from the $\pi\pi^*$state via a CI. The second (faster) relaxation channel has a different relevance for the two bases being more effective for uracil than for thymine (such a difference can also explain the shorter lifetime of the uracil excitation which has been observed experimentally [145]). This is probably due to the need of an out of plane motion of the 5-substituent for its activation [136]. Uracil has a H atom in this position, while thymine has a methyl group. The steric hindrance of the methyl, which limits the out of plane motion, frustrates the role played by the second relaxation channel for thymine compared to its importance in uracil. This argument suggests that for 5BU the second relaxation channel, if not completely suppressed, is even more inhibited than for thymine. In fact, in 5BU, the 5-substituent is an entire phenyl group whose steric hindrance is larger than that of the methyl. This hypothesis is also supported by the calculations performed in ref. [137] on 5BU where no CI has been found directly connecting the $\pi\pi^*$ PES to the ground state. Thus, the suppression of this decay channel favors the radiative decay explaining the high fluorescence QY. Following this argument, the even higher fluorescence QY in 5,6BU as well as its longer fluorescence lifetime result from the lack of the uracil $\pi\pi^*$ transition, and hence also of the ultrafast $\pi\pi^* \rightarrow n\pi^* \rightarrow$ ground state non-radiative decay channel.

5.8 Conclusions

In this chapter the main photophysical and photochemical properties of 5BU and 5,6BU have been reported and discussed by using a combined experimental (spectroscopy) and theoretical (electronic structure calculations) approach. It has been shown that irradiation of 5BU with high intensity UV femtosecond pulses (Figure 1, left panel) in MeOH solution leads to the formation of a cyclized compound: 5,6BU (Figure 1, right panel). This photoreaction mimics the formation of a crosslink between a pyrimidine nucleobase and a close lying aromatic aminoacid.

The spectroscopic data allowed to measure the value of 2.4×10^{-3} for the photocyclization QY in excellent agreement with the previous measurements of Sun et al. [102]. The fluorescence QY and fluorescence anisotropy of 5BU and 5,6BU have been also measured for the first time. Interestingly, the fluorescence QYs of both molecules are found to be very high, i.e. four orders of magnitude larger than the QY of single nucleobases in solution. Moreover, the fluorescence anisotropy measurements revealed that the fluorescence signal of the 5BU decays in the ultrafast time scale of few picoseconds, a property which is lost when the molecule takes the cyclized form for 5,6BU.

The observed ultrafast decay properties of 5BU can be ascribed to the fact that the main excited states properties of the nucleobases remain almost unchanged in 5BU (the presence of a decay path similar to the one responsible for non-radiative decay to the ground state in isolated uracil/thymine was detected). However, the same cannot be said for the different decay channel branching ratios: as proved by the high fluorescence QY, in fact, the radiative decay channel (identified in the relaxation of the molecule on the uracil $n\pi^*$ state) is more favored in 5BU than in the DNA bases alone. This behavior could be explained with the inhibition of the out-of-plane motion of the 5-substituent due to the presence of the phenyl group. This hinders the non-radiative relaxation pathway that in uracil and thymine leads to the ground state directly from the bright $\pi\pi^*$ state via a CI. The higher decay time of 5,6BU has instead to be ascribed to the changes in the electronic structure induced by the loss of the electronic -conjugation associated to the enone chemical group of the DNA bases. As a consequence the bright uracil $\pi\pi^*$ transition is no longer observed in 5,6BU causing the disappearance of all the ultrafast decay channels associated to this transition.

The global picture emerging from the combination of the theoretical and computational results leads to the conclusion that the use of 5BU as model system to study the proximity relations of a DNA base with a close lying aromatic aminoacid is valid at a local level: we observe that the main characteristics of the decay processes from the excited states of the uracil/thymine molecules remain almost unchanged in 5BU and the main perturbation observed is caused by the presence of the close lying aromatic group. In particular, this moiety entails the formation of a CT electronic transition, leading to photoreactivity after the absorption of one photon. Of course, since in this model system only two chromophores are present, alternative mechanisms involving more than two partners in a biological environment cannot be discarded or captured. In particular, in a more extended system, multiphoton absorption processes may play a role. Finally, the high fluorescence QY of both molecules, together with strong differences in their absorption spectra and fluorescence spectra, lifetime and anisotropy, call for spectroscopic time resolved experiments that, together with non-adiabatic excited states molecular dynamics simulations that will

be discussed in next chapter, will allow the study of the photoreaction mechanism.

Chapter 6

5-benzyluracil photocyclization reaction mechanism

The goal of this chapter is to investigate the mechanism of the photoreaction transforming 5-benzyluracil (5BU) to 5,6-benzyluracil (5,6BU) after UV irradiation. In last chapter, it has been shown that this reaction can be effectively used as model for UV induced crosslinking among proximal pyrimidine nucleobasis and aromatic amino acids. The chemical structure of 5BU is shown again in the left panel of 6.1 for convenience. It is composed by two chromophores: the uracil nucleobase, on the left in the figure, and the benzene ring characteristic of aromatic amino acids such as phenylalanine or tryptophan, on the right. The chromophores are linked via a methyl bridge which keeps them at a distance of about 4 Å in the equilibrium structure. At this distance, after light absorption, the electronic densities of the individual chromophores can migrate within the molecule. In particular, upon irradiation at 256 nm (about 4.7 eV for future reference), the molecule transforms in 5,6-benzyluracil (5,6BU). 5,6BU is shown on the right of Figure 6.1. This is a cyclized structure in which a new covalent bond (model of the crosslink) is formed between carbons C6 (in the uracil) and C10 (in the benzene ring). In the reaction, the protonation state of the molecule also changes. In 5BU, in fact, there is an hydrogen bound to C10 while the carbon in position 5 is not protonated. On the other hand, in 5,6BU, C10 does not have an H attached, while C5 becomes protonated.
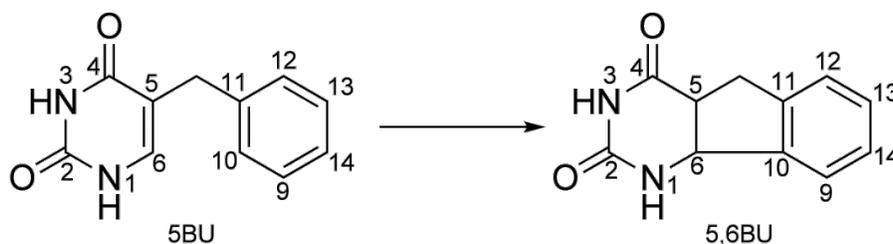
Figure 6.1: Chemical structures of 5BU (on the left) and 5,6BU (on the right). In 5BU the carbon atom in position 10 is protonated, while carbon 5 is not. In 5,6BU the situation is reversed with an H attached to carbon 5 and none bound to carbon 10.

In the first experimental study of 5BU reactivity in methanol solution [102], the main focus was to determine if the reaction mechanism involved one or two photon excitations. Based on measurements of the reaction quantum yield (QY) as a function of the light intensity, it was concluded that single photon absorption was responsible for cyclization. Furthermore, when the experiment was performed in deuterated methanol, it was found that the atom attached to C5 in the product was a deuterium. This finding points to a solvent assisted reaction mechanism in which the H lost by C10 is donated to methanol and the solvent in turn provides the atom that binds to C5.

In chapter 4, a detailed study of the accuracy of results obtained for the relevant excited states of the system with different xc functionals was performed. Comparing gas phase results with the reference calculations in [105] and results in solution with experiments, it has been shown that results obtained using the meta-hybrid GGA class of functionals, in particular M06, are reliable for 5BU. MO6 was adopted in chapter 5 to rationalize the spectroscopic properties of 5BU and 5,6BU. In this chapter, the features of spectroscopic spectra were successfully analyzed on the basis of LR-TDDFT calculations. Three possible relaxation channels of the molecule were then identified: (1) A motion towards an equilibrium structure in the lowest excited potential energy surface (PES) followed by relaxation on the ground state via fluorescence (QY = 0.12); (2) non adiabatic, non-radiative and ultrafast relaxation paths to the ground state similar to those observed in experiments and simulations regarding isolated DNA and RNA bases [145]; (3) formation of the 5,6BU via a putative reaction path characterized by shortening of the C6-C10 distance (QY $\sim 10^{-3}$). (The QYs mentioned here were obtained from the experiments.)

In spite of the investigations mentioned above, there are still several open questions on the microscopic mechanism leading to the formation of 5,6BU via the channel mentioned in point (3). In particular, two main scenarios exist: (a) the reaction (in particular the C6-C10 bond formation) occurs on the excited state and then the product relaxes to the ground state via a radiative transition; (b) during the relaxation dynamics there is a non adiabatic transition to the ground state and the reaction is completed without the emission of light. Furthermore, the time scale of the reaction is uncertain as it is the precise sequence of steps (bond formation and/or breaking, details of the hydrogen detachment and bonding...)

leading to the formation of 5,6BU. Finally, the cause of the low quantum yield of formation of 5,6BU is not known. In the following, the first two questions will be addressed combining the static electronic structure calculations results obtained with surface hopping QM/MM simulations (both based on LR-TDDFT) and *ab initio* Born Oppenheimer dynamics on the ground state to investigate the reaction mechanism along the reaction path suggested by previous calculations. The model adopted for the reaction is discussed in the next section, together with the computational details for the different simulations performed. The reaction mechanism is then described in detail in the Results.

6.1 Model and Methods

Due to the complexity and the size of the system, a rigorous non adiabatic ab-initio MD study of the entire photochemical process that leads to the photociclyzed product cannot be afforded. Thus, several approximations have to be introduced that will now be discussed in detail. First, only the only reaction channel investigated is the one that starting from the Franck Condon region follows the reaction path characterized by the shortening of the distance of the carbon C6 and C10. This choice is motivated by the previous studies with implicit solvent, which showed that the excitation of the system in the first CT state using the M06 xc functional leads to a closure of the structure (approaching of the two aromatic rings) and a shortening of this distance. In the following, for convenience of representation, the negative of this distance, R_C will be used as reaction coordinate. Unfortunately, in chapter 4 it has been showed that the departure from the Franck Condon region along the selected reaction path cannot be accurately described by the exchange-correlation functional typically adopted for the dynamics (PBE) which fails to correctly reproduce the shape of the CT states in this region. However, the quality of the PBE PES cut along the chosen reaction coordinate improves quickly as the system moves away from the initial geometry. This can be also seen in 6.2 which shows the reaction profile and summarizes the different approaches used to reconstruct the relevant electronic PES along the whole reaction path. In the first part of the figure, $R_C < -2.5$ Å, the energy profile for the ground and excited state computed in implicit solvent with M06 (purple curve) is compared with those determined using PBE (in orange). This part of the reaction profile was computed by means of a relaxed energy scan along the chosen reaction coordinate. At each selected value of R_C between -3.6 Å and -2.5 Å (the points in the figure), the corresponding relaxed structures and energies were computed, as obtained via steepest descent minimization of the energy for all molecular degrees of freedom but the constrained distance. Along the scan, the M06 and PBE excited state curves approach and, for $R_C \geq -3$ Å, become parallel. This ensures, in particular, that the gradients (and thus the forces) computed along the two surfaces are compatible. Based on this, the geometry at $R_C = -2.5$ Å(in the following it will be indicated as "state A") was chosen as the starting frame for the non adiabatic dynamics using the trajectory-surface hopping algorithm with PBE based energies. This part of the calculation is performed in explicit solvent (see Computational Details for the set up). For the surface hopping,4 trajectories were initialized starting from the same structure (A) but with different velocities $\{\mathbf{v}_i\}_{i=1}^N$ (N is the number of atoms in the solute, and $\mathbf{v}_i = (v_{ix}, v_{iz}, v_{iz})$). The velocities are obtained in the following way: first the vectors $\{\Delta_i\}_{i=1}^N$ have been computed.
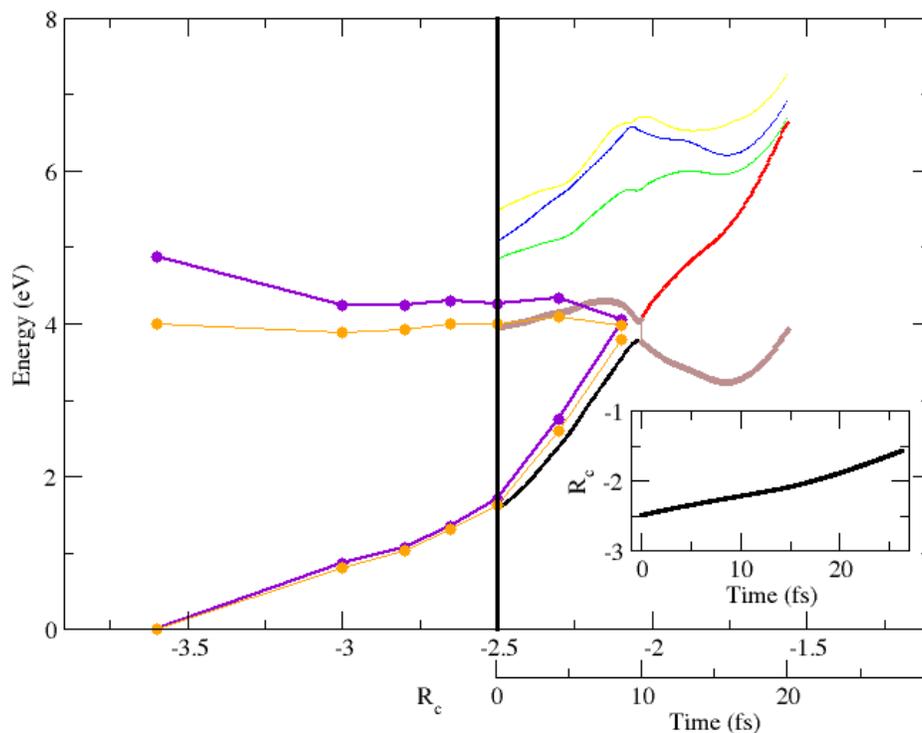
Figure 6.2: Ground and excited state PES along the reaction path. The left part of the figure shows the result of the static steepest descent calculations along the R_C relaxed scan. The purple curve is the MO6 result, in orange the curves obtained with PBE. The right part of the figure shows the first four PES of the system computed at the PBE level along one of the surface hopping trajectories as a function of the same reaction coordinate. This is possible because, as shown in the inset, during the non adiabatic dynamic R_C monotonically grows and is essentially a linear function of time. In the TSH dynamics, the thick brown line identifies the state where the evolution occurs based on the surface hopping algorithm. For $R_C \in [-2.5, -2.1]$ we show simultaneously the PESs obtained with static and dynamic calculations to show the similarity of the curves and the positions where the two avoided crossing points (static and dynamical) are located.
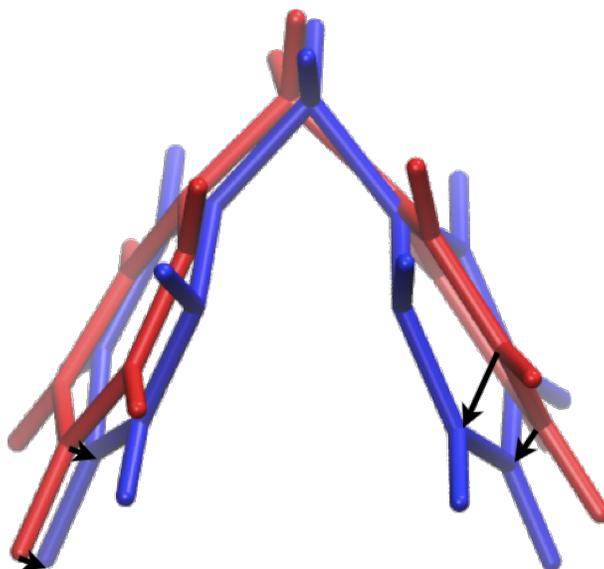
Figure 6.3: Stick representation of the structure A (red) from which the surface hopping dynamics is initialized and of the structure obtained by steepest descent at the avoided crossing (blue). These are the structures used to determine the initial velocities in the excited state non adiabatic dynamics. The black arrows indicate some of the Δ vectors defined in the text to compute such velocities.

These are the displacement vectors between the positions of each atom in A and in the structure at $R_C = -2.1$ Å, i.e. close to the avoided crossing between the first excited and ground state PES. This second structure is shown in blue in 6.3, where is also shown in red the structure A, together with some representative Δ_i vectors indicated as black arrows. To assign the velocities, it was set $\mathbf{v}_i = \alpha \Delta_i$ where α is a constant with dimensions of inverse time. This assignment ensures that all velocities are directed along the putative reactive path. Indicating with K_0 an estimate of the kinetic energy of the system upon excitation and ΔU the difference between the PES value at Frank Condon structure and at A, α is obtained from the conservation of energy:

$$\frac{1}{2} \sum_{i=1}^{N} m_i \alpha^2 |\Delta_i|^2 = K_0 + \Delta U \tag{6.1}$$

(m_i are the masses of the atoms) [1].To get an idea of the effect of temperature on the time scale of the non adiabatic motion, two values of K_0 were selected. The first, $K_0^{(1)} = 0.026$

---

[1]The conservation of energy written as above neglects the rearrangement of the solvent during the motion and variations in the solvent/solute interaction. These approximations are justified by the ultrafast time scale of the solute's motion away from the Frank Condon region and the fact that the solute/solvent interactions are relatively small.

eV corresponds to the average thermal energy for the system (300 K), the second, $K_0^{(2)} = 0$ eV, represents the extreme case of a cold solute. Small perturbations to the set of velocities obtained in A from $K_0^{(1)}$ were also added to obtain two more sets of initial conditions. The systems with geometry A and velocities $\{v_i^k\}_{i=1}^{N_a}$ ($k = 1, \ldots, 4$) were then propagated in time using TSH with initial state amplitudes $C_0 = 0$ and $C_1 = 1$, where the subscripts indicate the ground and first excited state. As described in the second part of 6.2, the dynamics continues on the first excited state for a few femtoseconds and then a transition to the ground state occurs. The prolongation of the non adiabatic dynamics shows a quick departure of the system from the region of strong coupling between the two PESs, which makes the back transition to an excited state very improbable. To investigate the subsequent steps in the reaction, we then continue the the dynamics adiabatically on the ground state. A detailed analysis of the reaction mechanism along the path reconstructed via the different methods illustrated above is given in the Results section.

6.1.1 Computational details

The static calculations (first part of Figure 6.2) were performed via DFT and TDDFT using the Gaussian09 program package [107] with the M06 and PBE exchange correlation functionals as described in chapter 4. In particular it can be conveniently recalled that in these calculations the MeOH solvation effects were treated implicitly and treated as a continuum dielectric medium using the self-consistent reaction-field (SCRF) model.

The dynamical calculations were performed with a QM/MM representation of the system in which the solute (quantum region) is described at the PBE/DFT-LR-TDDFT level and the MeOH solvent molecules are described at a classical MD level. For the QM/MM was used the additive scheme discussed in chapter 2 as implemented in the program package CPMD [39]; The extension of the formalism to describe the coupling of an excited electronic density with a classical environment can be found if ref [40]. The steps followed in order to get the initial conditions for our simulations are as follows:

1. The 5BU was imbedded in a periodically repeated solvation box of side of about 50 Å containing 1567 MeOH molecules. The system was equilibrated with a 10 ns MD simulation in the NPT thermodynamic ensemble at ambient pressure and temperature. During equilibration, the 5BU and MeOH molecules interactions were modeled using the General Amber Force Field. The nuclear charges of 5BU were set using the RESP procedure based on the electrostatic potential derived with a PBE/DFT calculation (in the nuclear equilibrium structure). During the dynamics the 5BU was kept almost fixed in its initial structure using restraints for all the internal degrees of freedom. For the MD integrations we used the Sander program [41]. The solvation box obtained with this procedure is represented in 6.4 in which the solute atoms are represented with sphere (the radius is representative of the Van der Waals radius of the atom) and the MeOH molecules with lines.

2. Translation of the system topological and force field parameters in the Gromos program [42] format and short QM/MM re-equilibration of the system in the NVE ensemble (1 ps).
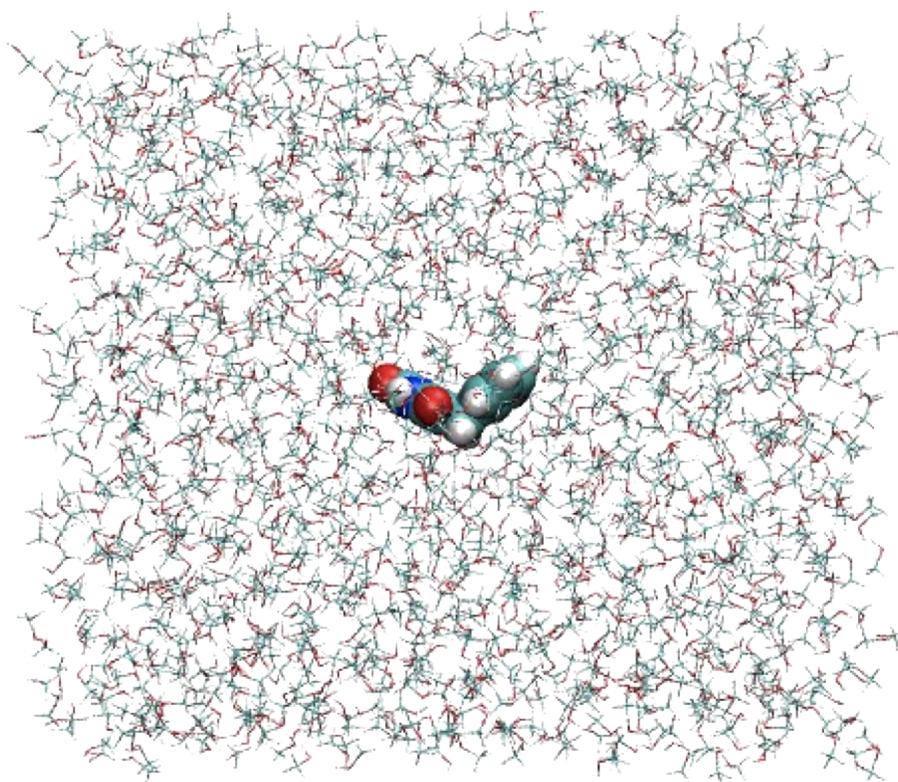
Figure 6.4: Typical snapshot of the QM/MM simulation box after equilibration. The system contains 1567 MeOH, shown here in a ball and stick representation. 5BU is at the center of the box. Its atoms are shown as cyan (C), blue (N), red (O), and white (H) spheres with dimensions representative of their Van der Waals excluded volume.

3. Substitution of the 5BU coordinates with the ones relative to the configuration corresponding to the "state A" and short (1 ps) re-equilibration of just the solvent degrees of freedom (keeping fixed all the positions of the QM system).

The QM/MM simulations on the ground state were performed using Born Oppenheimer molecular dynamics. For these simulations the solute was imbedded in a square quantum box ($L = 12$ Å) in which a plane waves basis set was used for the DFT calculations, with a cutoff of 65.0 Ry. DCACP Kleinman-Bylander pseudopotentials have been used to describe the core electron and take into account the dispersion interactions. As convergence criterium a threshold for the self consistent electronic energy of 10^-6 a.u was used. For the dynamics of the solute in the excited states the Tully's fewest switches algorithm for the TSH [56] discussed in chapter 2 was used. A time step of 5 a.u. ($\simeq 0.12$ fs) was used for the MD integration on the ground state while in the excited states we used a time step of 3 a.u. ($\simeq 0.72$ fs).

6.2   Results

Based on the calculations described in the previous section, the reaction mechanism along the path hypothesized in this work is the following. After irradiation, the molecule is found on the first excited state. In the Frank Condon region, charge density plots show that this state has a pronounced charge transfer character, with the transferral of about 40% of the charge density of one electron from the benzene to the uracil moiety. This charge unbalance causes a Coulomb attraction between the two rings which triggers the motion shortening the C6-C10 distance. As shown by the static calculations (purple curve with dots, Figure 6.2) this motion is energetically favorable on the first excited state, whose PES shows a rapid decrease (of about 1 eV) in going from $R_C = -3.6$ Å (Frank Condon) to $R_C = -3.0$ Å. On the contrary, shortening of the distance is strongly unfavorable on the ground state where it correspond to a sharp increase in the energy (about 3 eV). This is an indication of the key role played by the excited state in the reaction mechanism. Note also that an accurate calculation of the PES is critical to identify the reaction channel: using the PBE exchange correlation functional (orange curve in 6.2) the essentially flat PES would have left the system close to the Frank Condon region, failing to initiate the reaction.

The sharp decrease of the M06 excited PES is also an indication that the motion along the chosen reaction coordinates happens on an ultrafast time scale. This fast time scale is confirmed by the dynamical part of the calculation. The behavior of the system along the 4 SH trajectories is represented in 6.5 where are reported, as a function of time, the first four PESs for the system. The state upon which the motion occurs according to the SH algorithm is indicated with a thick brown line. As expected, all transitions occur in close proximity of the avoided crossing. The actual time of the transition is slightly different for the different trajectories, but always occurs between 10 and 25 fs. The values of the reaction coordinate corresponding to the transitions are very similar (from 6.2 this value is $R_C \approx -2.1$ Å), leading to very similar structures of the molecule when the hop to the ground state occurs. The (typical) conformation of the molecule at this stage will be indicated in the following as A' and is shown in the first snapshot of 6.6.

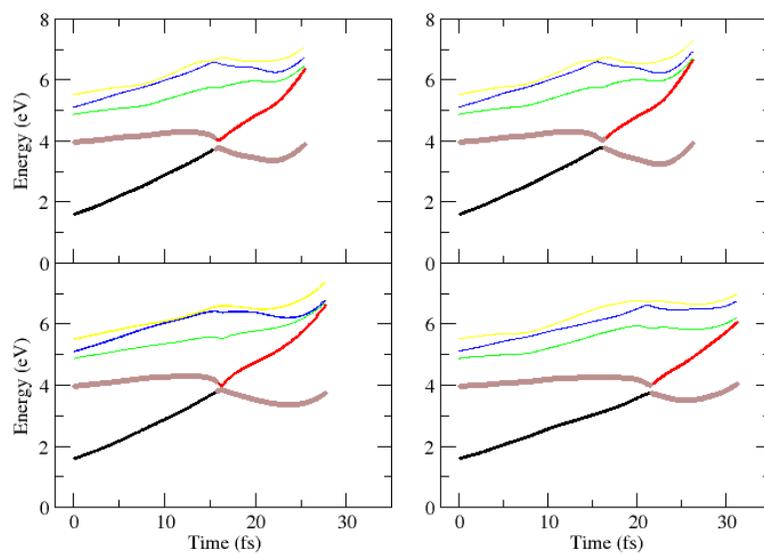This geometry has two interesting features: first it can be observed that the motion

Figure 6.5: PESs of the ground state and of the first 4 excited states along 4 different exemplar TSH trajectories (in black, red, green, blue and yellow). The PES driving the dynamics for each TSH trajectory step is highlighted (the curve as larger width) and colored in brown.
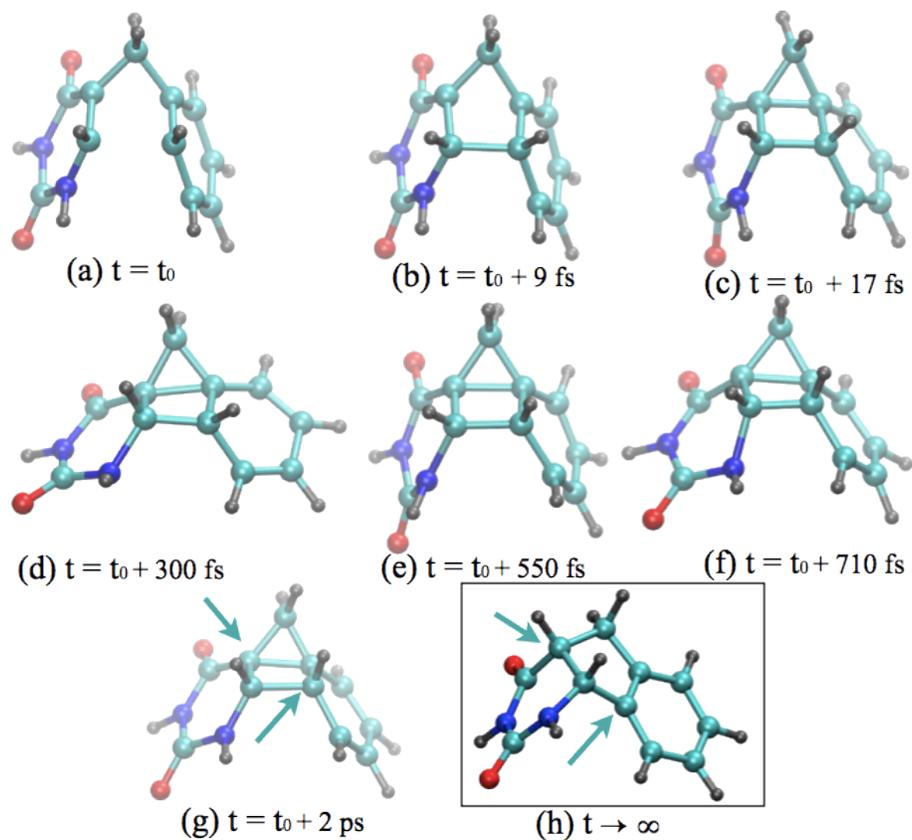
Figure 6.6: Snapshots of the time evolution of the system during the DFT based ground state dynamics. t_0 is the starting time, immediately after the non adiabatic transition described in the text takes place. The images show the different steps leading from the structure of the system upon transition, panel (a) (structure A in the text), to the final stable structure found in our simulations, panel (g) (structure A'' in the text). Snapshots (b) and (c) show the, ultrafast, formation of the two covalent bonds among the C6-C10 and C5-C11 carbons, respectively. Snapshots (d)-(e)-(f) illustrate the butterfly motion occurring during thermalization of the system due to collisions with the solvent (note the longer time scale). Snapshot (h) reports the equilibrium structure of the final product of the photocyclization reaction as obtained from static optimization. The green arrow in this panel and in (g) point to the carbons that must change protonation state to complete the reaction.

accentuates the transition from the planar Frank Condon structure of the rings to the stacked geometry which was already visible in A (see 6.3); second the electronic structure of the carbons in positions 6 and 10 begins to change from the $sp2$ hybridization towards an $sp3$ form which facilitates the formation of the covalent bond between them. This tendency is confirmed by the (static) relaxation of the structure performed with constrained $R_C = 2.1$ Å which shows that in the minimum energy structure, corresponding to the blue molecule in 6.3, the $sp3$ hybridization is completely stabilized leading to a pyramidalization of the bonding geometry of the carbons. Essentially the same pyramidalization occurs along the dynamical trajectories (which, as mentioned in the Methods, after the hop are propagated adiabatically on the ground state). As the potential energy decreases along the trajectory, C6 and C10 come closer, the $sp2 \rightarrow sp3$ transition becomes complete, and the covalent bond forms. The resulting structure is shown in panel (c) of 6.6.

As a consequence of the new bond formation, the C6 C5 bond in the uracil ring, a double bond in the Frank Condon region (see also 6.1), becomes a single bond. The aromatic structure in the benzene ring is also lost with a single bond left between C10 and C11. These changes in the electronic structure of the molecule favor a change in the hybridzation state of C5 and C11 in which these atoms become $sp3$ too, promoting the formation of a covalent bond also between them. This bond formation is indeed observed in the next few femtoseconds of dynamics, and shown in panel (c) of 6.6; its occurrence stabilizes the electronic structure of the molecule which no longer changes significantly along the dynamics. For the next picosecond, in fact, the most relevant feature observed in the dynamics is a butterfly motion in which the rings approach and separate (snapshots (d)-(f) in 6.6). During this motion, collisions with the solvent reduce the kinetic energy of the molecule which reaches, on the picosecond timescale, a rather stable structure shown in panel (g) of the figure. The cooling of the solute after the non adiabatic transition can also be followed via 6.7, where the different phases of the process are reflected in the time history of the solute's instantaneous kinetic energy. At very short times after de-excitation, a rapid drop is observed, corresponding to the formation of the covalent bonds among the carbons ($t \leq 100$ fs). Subsequently, the large amplitude oscillations of the butterfly motion cause several high momentum transfer collisions with the solvent, leading to the steep descent of the instantaneous kinetic energy seeable for $100 < t < 750$ fs. Damping of the large oscillations in the structure is followed by a slower thermalization, eventually taking the solute to thermal equilibrium.

The structure in panel (g) of 6.6 is representative of the final state obtained with our calculations. In the following, this state will be indicated with A''. A'' is identified as the intermediate formed at the end of the ultrafast first step of the full cyclization reaction of 5BU. The final product of the reaction (5,6BU), as obtained in Chapter 5 via static optimization calculations, is shown in the last panel of 6.6. This structure differs from the one we obtained in two crucial features. First the C5 C11 bond is not present, second C10 is no longer protonated while there is an hydrogen attached to C6. Although the calculations performed so far were unable to follow the formation of 5,6BU, some hypothesis on the steps that conclude the reaction can be made on the basis of physical (and/or chemical) arguments. The next event that needs to happen is the detachment of the proton from C10. Since it involves breaking of a chemical bond, this step is expected to be an activated process occurring on time scales much longer then the ps, and requires advanced sampling methods
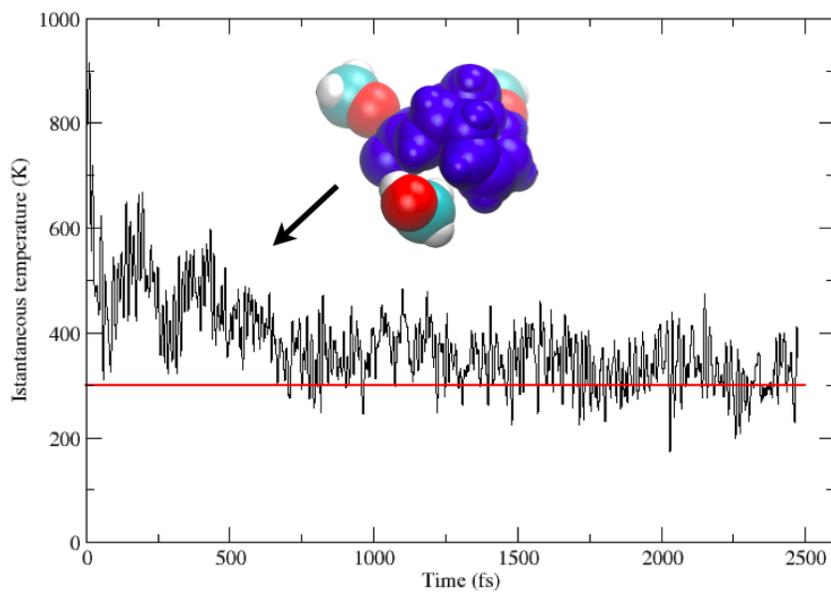
Figure 6.7: Istantaneous temperature of the solute as a function of time. The rapid drop for $t < 50$ fs corresponds to the formation of the bonds among the four carbons, the subsequent slower decay is cooling via collisions with the solvent leading, for $t > 1500$ fs to thermalization of the system. In the inset we show a snapshot of collision among the solute and 3 solvent molecules. The 5BU (in blue) and the MeOH molecules atoms are represented with their Van Der Waals spheres.

to be described. Once the C10-H bond is broken, it becomes quite easy to imagine a sequence of spontaneous changes in electronic structure leading to the formation of 5,6BU. Indeed the deprotonated carbon is now available to go back to the $sp2$ hybridization that stabilizes the aromatic benzene ring. The stabilization of the benzene structure also promotes the breaking of the C11-C5 bond, with C11 going back to an $sp2$ state. The new configuration leaves only C5 in an $sp3$ hybridization state with one lone pair free to form a new bond. This negative charge density acts as an attractor for an H in the solvent which can now be bound to complete the reaction. This sequence of steps does not determine the origin of this last proton which can be either the one released from C10 or be donated by one of the solvent molecules. However, based on the experiment by Sun with deuterated methanol mentioned in the Introduction, the second hypothesis seems to be the correct one.

6.3   Conclusions

In this chapter, calculations combining static geometry optimizations with excited state (surface hopping) non adiabatic dynamics and ground state QM/MM adiabatic propagation were used to investigate a possible reaction mechanism for the first (ultrafast) part of the photcyclization reaction of 5BU. The key steps in the reaction mechanism that have been identified are summarized in 6.8 where are schematized the relevant intermediate structures indicating explicitly the H atoms involved in the C-C bond formation. Upon excitation, the molecule is found on a state with relevant charge transfer state (top left panel of 6.8). The Coulomb attraction between the rings causes deformation of the planar Frank Condon structure towards a stacked configuration originated by a "downhill" motion on the first excited state PES. After a few tens of femtoseconds, this motion takes the system to an avoided crossing where a transition to the ground state occurs. Immediately after this transition, the bond structure of the molecule reorganizes by forming a first intermediate A' (top right panel of 6.8) in which the C6 and C10 carbon have formed a covalent bond and two lone pairs are created at C5 and C11. Further propagation on the ground state shows that a new bond between carbons C5 and C11 is then rapidly formed creating the stable structure indicated with A'' (bottom right structure in 6.8). In this conformation, the hybridization state of the four newly bound carbon atoms is $sp3$ and the aromaticity of the benzene ring is lost. From this state, the final steps that lead from it to the formation of the 5,6BU product have been hypotesized on the basis of the results obtained. The key step in the second part of the reaction is the breaking of the C10-H bond, followed by the solvent mediated capture of an H by C5. Note that this reaction mechanism is fully compatible with single photon absorption mechanism identified by Sun [102]. Verifying the steps in the second part of the reaction requires, however, to modify our simulation set up by including an *ab initio* description of the methanol. Furthermore, since the C10-H bond breaking is most likely an activated event, rare event simulation methods should be applied to estimate the free energy barrier (and/or the timescale) associated to it. These studies will be the object of future investigations.
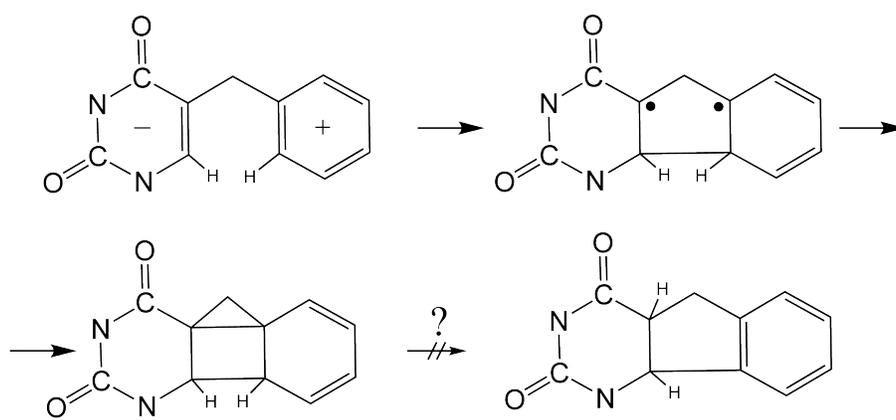
Figure 6.8: Chemical structures relative to the relevant steps in the photo-reaction of 5BU (see discussion in the text for details).

Conclusions

In this thesis work the photoreaction $5BU \rightarrow 5,6BU$ in MeOH was studied using several computational methods. Kohn Sham Density Functional Theory (DFT) and Linear Response-Time Dependent DFT (LR-TDDDFT), reviewed in Chapter 1, were used for the electronic structure calculations. Zero temperature minimization algorithms were used for the characterization of the ground state and excited molecular Potential Energy Surfaces of reactant and product. The non-adiabatic Trajectory Surface Hopping (TSH) Molecular Dynamics (MD) method was used for the description of the most critical step of the reaction mechanism happening in the excited states. In this step the adiabatic Born Oppenheimer (BO) approximation breaks down. *Ab initio* BO MD was used to investigate the relaxation motions subsequent to the non-adiabatic transition of the molecule back on the electronic ground state. The MeOH environment effects were also taken into account in the calculations. In particular a polarizable continuum model was used to model the solvent in zero temperature calculations while the Quantum Mechanics / Molecular Mechanics (QM/MM) method in which the solvent molecules are treated explicitly has been used for the dynamic studies. All these methods have been discussed in Chapter 2.

The global picture emerging from all reported results indicate that the photo-reaction under investigation is an effective local model for the study of the atomistic mechanism at the base of DNA protein crosslink in a biological environment. The phenomenology regarding this problem has been discussed in Chapter 3.

The electronic and optical properties have been benchmarked and discussed in Chapter 4. To validate the TDDFT predictions on the electronic structure, critical for bi-chromophoric molecules as the 5BU, the results obtained were systematically assessed by comparison with high level quantum chemical calculations in the gas phase and with experimental data in solvated phase. In particular the performances of different approximated forms of the exchange-correlation energy functional were tested, by using the conceptual ranking known as the Jacobs ladder of DFT/TDDFT to rationalize the different results obtained. This comparison allowed to identify the best functional (specifically, the M06) for the description of the 5BU.

Once assessed the reliability of optical computational results, in Chapter 5, LR-TDDFT calculations were used to characterize the photophysical and spectroscopic properties of both reactant (5BU) and product (5,6BU). In this Chapter, several spectroscopic experimental results have also been reported and rationalized on the basis of the calculations performed.

Finally the description of a possible reaction mechanism is reported in chapter 6. Results indicate that the photoreaction proceeds via de-excitation into the electronic ground state of

an intermediate compound. The ultrafast time scale associated to the reaction intermediate formation has also been estimated from TSH calculations.

The state of the art computational techniques adopted in this work have been effectively used to approach and (at least partially) solve a concrete problem in photochemistry and biology. Several limitations have been however encountered, and the complete characterization of the photoreaction required the use of various approximations. The two most critical computational points have been: i) xc functionals which are computationally treatable for *ab initio* MD methods not always give reliable results for the electronic structure and ii) too large parts of the molecular configurational space needs to be sampled if a complete characterization of all the possible decay channels of a molecule in solution (and the relative breaching ratios) is wished. The improvement of the existing computational algorithms is currently an active field of research and a major step forward in the simulation possibilities in the near future cannot be excluded.

Appendix A

Spectroscopic observables and lines broadening

In the case in which a sample of molecules is shined with light of frequency resonant with the excitation frequencies

$$\Omega_n(R) = \frac{1}{h}\{U_n(R) - U_0(R)\} = \frac{1}{h}\{E_n(R) - E_0(R)\} \tag{6.2}$$

(notations defined in Chapter 2) in regions of the nuclear configurational space corresponding to some stable nuclear state, electronic transitions from the ground state PES to the excited PES $U_n(R)$ can possibly occur. The large mass difference between electrons and nuclei can be used also in this case to consider the dynamic time scale of such electronic transitions very fast compared to the nuclear motion time scale. It means that the nuclear coordinates and momenta changes involved during the time of an electronic transition can be, with good approximation, disregarded. Referring to the explicative cartoon reported in Figure 6.9 this implies that the electronic transitions are *vertical* and the nuclei are frozen during the excitation process. This scheme for the description of electronic optical transitions is known as *Frank-Condon principle* and the form Frank-Condon region is usually used to refer the excited PES in the vicinity of a stable nuclear configuration.

In order to describe quantitatively these phenomena it can be considered, with no loss of generality, the interaction of a molecule with an electric field of the kind

$$\mathbf{E}(\mathbf{t};\omega) = \mathbf{E}_\omega \; sin \; (\omega t). \tag{6.3}$$

From the electronic point of view, this case has been considered in Chapter 1 as application of the linear response theory in the framework of TDDFT. In particular it has been showed how the induced electronic dipole of an electronic system $\mathbf{p}(t) = \mathbf{p}(\omega) \; sin \; (\omega t)$ at the first-order in the field amplitude can be directly derived using the first-order induced electronic dipole polarization tensor α (both quantities have been defined in Eq. 1.147). It can be recalled here that for their averaged form[2], in the case the electron system is the subsystem of a molecule, i.e. its Hamiltonian operator is the electronic Hamiltonian defined in Eq. (2.10), it holds:

$$p(\omega, R) = \bar{\alpha}(\omega, R)E_\omega \tag{6.4}$$

---

[2]$p = \frac{1}{3}(p_x + p_y + p_z)$, $E = \frac{1}{3}(E_x + E_y + E_z)$, $\bar{\alpha} = \frac{1}{3}(\alpha_x x + \alpha_y y + \alpha_z z)$,
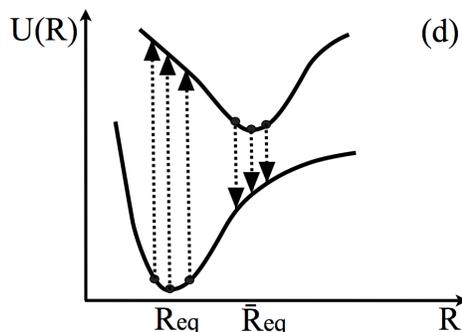
Figure 6.9: Scheme relative to the case in which a biatomic molecule is excited on the first electronic adiabatic state, for example, upon interaction with of an external light field. Different possible vertical transition are shown from the ground state to the first excited (absorption) and vice versa (fluorescence) as possible excitation and de-excitation events corresponding to different nuclear bond distances.

with

$$\bar{\alpha}(\omega, R) = \lim_{\eta \to 0^+} \sum_n \frac{f_n(R)}{(\omega - \Omega_n(R) - i\eta)(\omega + \Omega_n(R) - i\eta)} \tag{6.5}$$

where the quantities f_n, the oscillator strengths, have been defined in Eq. (1.159) and the dependence on the nuclear coordinate R in p and α has been explicitly reported because the adiabatic electronic states and PES depend on the nuclear structure.

Now the macroscopic Maxwell equation for the electric field propagation in a dielectric medium can be considered. In Gaussian units it reads:

$$\left(\frac{\partial^2}{\partial t^2} - c^2 \nabla^2 \right) \mathbf{E}(\mathbf{x}, t) = -4\pi \frac{\partial^2}{\partial t^2} \mathbf{P}[\mathbf{E}](\mathbf{x}, t) \tag{6.6}$$

where $\mathbf{x}$ is a macroscopic coordinate in the dielectric medium and the quantity $\mathbf{P}[\mathbf{E}]$, a functional of the electric filed $\mathbf{E}$, is the average of the induced electronic dipole in a microscopic volume $\Delta V(\mathbf{x})$ around the coordinate $\mathbf{x}$. This volume has to be infinitesimal from a macroscopic point of view but still microscopically sufficiently big to consider the average in the appropriate statistical ensemble. In formulae:

$$\mathbf{P}(\mathbf{x}) \equiv \frac{1}{\Delta V(\mathbf{x})} \sum_{m=1}^{\bar{N}} \mathbf{p}_m = \frac{\bar{N}}{\Delta V(\mathbf{x})} \frac{1}{\bar{N}} \sum_{m=1}^{\bar{N}} \mathbf{p}_m = \rho \langle \mathbf{p} \rangle_{ens} \tag{6.7}$$

in which $\bar{N}$ is the number of molecules contained in the macroscopically infinitesimal volume $\Delta V(\mathbf{x})$ and ρ is the molecules volume density [3]. Passing to averaged quantities

---

[3]The chromophore density ρ is not supposed to depend on the macroscopic position in the medium, i.e
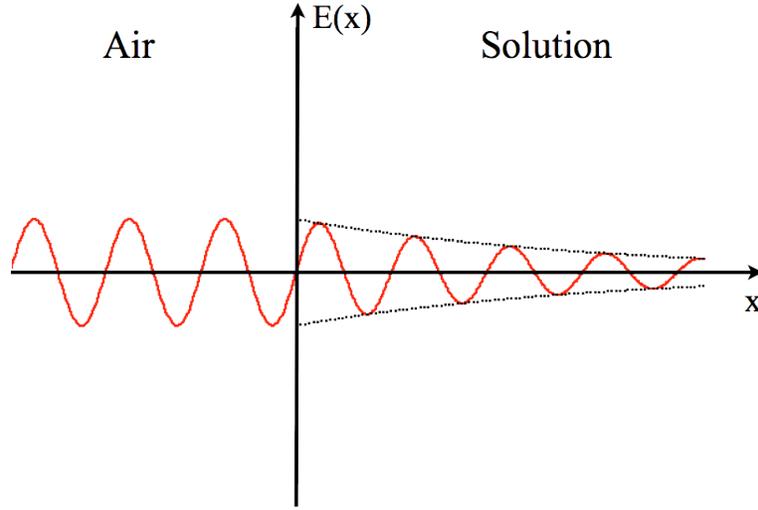
Air   E(x)   Solution

X

Figure 6.10: Schematic description of macroscopic electric field behavior as a function of the macroscopic coordinate $\mathbf{x}$. On the left side of the plot ($\mathbf{x} < 0$) the propagation in air (or vacuum) is sketched, case in which the field amplitude is constant. On the right side instead ($\mathbf{x} > 0$) is the behavior of the field in the solution containing the absorbing chromophores is shown. The field amplitude decades as $\mathbf{E}_{0\omega}e^{-\beta x}$, β being the absorption coefficient.

$$P(\mathbf{x}) = \rho \left\langle p \right\rangle_{ens} = \rho E_\omega(\mathbf{x}) \left\langle \bar{\alpha}(\omega, R) \right\rangle_{ens} = \chi(\omega) E_\omega(\mathbf{x}) \tag{6.8}$$

where $\chi(\omega) = \rho \left\langle \bar{\alpha}(\omega, R) \right\rangle_{ens}$ is the dielectric susceptivity. Using the form of Eq. (6.8), the Maxwell equation (6.6) can be conveniently recast in the form:

$$\left(\frac{\partial^2}{\partial t^2} - v^2 \nabla^2 \right) \mathbf{E}(\mathbf{x}, t) = 0 \tag{6.9}$$

where

$$v^2 = \frac{c^2}{1 + 4\pi\chi} = \left(\frac{c}{n} \right)^2 \tag{6.10}$$

that is the propagation velocity of the field $\mathbf{E}$ in the dielectric medium that now depends on ω and is a complex quantity (because so is α and hence χ) and

$$n(\omega) = 1 + 4\pi\chi(\omega) \tag{6.11}$$

is the refraction index. Imposing the starting assumption

$$\mathbf{E}(\mathbf{x}, t) = \mathbf{E}_\omega(\mathbf{x}) \; sin \; (\omega t) \tag{6.12}$$

___
the medium is considered homogenous

Eq. (6.9) is straightforwardly solved finding

$$\mathbf{E}_\omega(\mathbf{x}) = E_{0\omega} \ e^{i\mathbf{k}\cdot\mathbf{x}} + c.c. \tag{6.13}$$

where $|\mathbf{k}| = \frac{\omega n}{c}$. The light intensity in the medium (proportional to the square of the electric field) will then be (Beer's law)

$$I_\omega \propto |\mathbf{E}_\omega(\mathbf{x})|^2 = I_{0\omega} e^{-\beta x} \tag{6.14}$$

where x is the path length that the electric field wave does in the dielectric medium in the direction of the propagation vector $\mathbf{k}$ and β, the absorption coefficient is:

$$\beta = Im \ [\frac{\omega n}{c}] = \frac{\omega}{c} Im \ [n] = \frac{\omega}{c} Im \left[\sqrt{1 + 4\pi Re[\chi] + i4\pi Im[\chi]} \right] \tag{6.15}$$

This result is sketched in Figure **??**. The last term of Eq. 6.15 can be better evaluated if its first order Taylor expansion is considered, i.e. using the approximated equality $f(x) = \sqrt{A + x} \simeq \sqrt{A} + \frac{1}{2\sqrt{A}} \ x$ where $A = 1 + 4\pi Re[\chi]$ and $x = i \ 4\pi Im[\chi]$. Doing so, Eq. (6.15) becomes:

$$\beta \simeq \frac{4\pi\omega}{2c\sqrt{1 + 4\pi Re[\chi]}} \ Im[\chi] \simeq \frac{2\pi\omega}{cRe[n]} \ Im[\chi] = \frac{2\pi\omega\rho}{cRe[n]} \ \langle Im[\bar{\alpha}(\omega, R)]\rangle_{ens} \tag{6.16}$$

The imaginary part of $\bar{\alpha}(\omega, R)$ can be derived from Eq. (6.5):

$$Im[\bar{\alpha}(\omega, R)] = \sum_n f_n(R) \lim_{\eta \to 0^+} g_{n,\eta}(\omega) \tag{6.17}$$

where

$$g_{n,\eta}(\omega) = \frac{\eta}{[(\omega - \Omega_n)^2 + \eta^2][(\omega + \Omega_n)^2 + \eta^2]} \tag{6.18}$$

the sequence of functions $g_{n,\eta}(\omega)$ has the properties to be a nascent of a combination of delta functions. In fact, noting that:

$$\lim_{\eta \to 0^+; \omega \neq \pm\Omega_n} g_{n,\eta}(\omega) = 0$$

$$\lim_{\eta \to 0^+; \omega = \pm\Omega_n} g_{n,\eta}(\omega) = \infty \tag{6.19}$$

$$\lim_{\eta \to 0^+} \int g_{n,\eta}(\omega) \ d\omega = \frac{\pi}{4\Omega_n^2}$$

the limit in Eq. (6.17) can be evaluated:

$$\lim_{\eta \to 0^+} g_{n,\eta}(\omega) = \frac{4\Omega_n^2}{\pi} \{\delta(\omega - \Omega_n) + \delta(\omega + \Omega_n)\} \tag{6.20}$$

Finally, considering that by definition $\omega > 0$ and $\Omega_n > 0 \; \forall n$, clearly the second delta function in Eq. (6.20) is unimportant. Collecting all these terms in Eq. (6.16), the following final approximated form for the absorption coefficient can be written:

$$\beta \simeq \frac{8\omega\rho}{cRe[n]} \; \sum_n \; \langle \; f_n(R) \; \Omega_n^2(R) \; \delta(\omega - \Omega_n(R)) \; \rangle_{ens} \tag{6.21}$$

The most usual experimental case is to consider some chromophore molecules in some optically inert solvent (i.e. whose electronic absorption resonances $\Omega_n^{(solv)}$ are out of experimental irradiation range) in sufficiently low concentration to consider negligible their interactions. In this case the most appropriate statistical ensemble is the canonical one and Eq. (6.21) becomes

$$\beta \simeq \frac{8\omega\rho}{cRe[n]} \; \frac{1}{Z} \sum_n \; \int dR \; e^{-\frac{U_0^{(ad)}(R)}{kT}} \; |\chi_0(R)|^2 \; f_n(R) \; \Omega_n^2(R) \; \delta(\omega - \Omega_n(R)) \tag{6.22}$$

which result in a collection of sharp peaks (absorption lines) centered at $\Omega_n(\bar{R})$, $\bar{R}$ being the nuclear equilibrium configuration (and hence most probable) corresponding to the minimum of $U_0(R)$. The lines amplitudes are proportional to the chromophores density ρ and the oscillator strengths of the electronic transitions $f_n(\bar{R})$ while the broadening grows with $K_B T$.

The absorption coefficient, as defined in Eq. (6.14), can be directly linked to quantities often measured experimentally: the absorbance for example, or absorption spectrum[4] defined as

$$A(\omega) \equiv -\frac{1}{l} \log_{10} \frac{I(\omega)}{I_0(\omega)} = -\frac{1}{l \ln 10} \ln \frac{I(\omega)}{I_0(\omega)} = \frac{1}{l \ln 10} \beta l = \frac{\beta}{\ln 10} \tag{6.23}$$

where $I(\omega)$ and $I_0(\omega)$ are respectively the incoming and out coming monochromatic light intensity in and from a sample of chromophores and l the length of the path in the direction of the beam propagation that the light travels through the sample (their relation with the absorption coefficient was reported in Eq. (6.14)). This quantity is usually measured in analytical chemistry to characterize some given chromophores (the absorption spectrum shape represents a sort of finger print for a chromophore) or to detect their concentrations. An other quantity often used is the cross section of light absorption by a single particle defined as

$$\sigma \equiv -\frac{1}{\rho l} \ln \frac{I(\omega)}{I_0(\omega)} = \frac{A}{\rho}(\ln 10) = \frac{\beta}{\rho} \tag{6.24}$$

where ρ (as already defined above) is the chromophores number density per unit of volume.

As it goes without saying, a direct calculation of the absorption coefficient via Eq. (6.22) is not feasible analytically. A way to evaluate computationally ensemble averages of physical observable (as in the case of Eq. (6.22)) by means of molecular dynamics methods has

---

[4]The name absorbance is usually preferred when a single frequency is taken into account, while the name absorption spectrum is usually referred to the whole function $A(\omega)$

been discussed in section 2.4. However often the direct calculation of the first bunch of electronic excited states energies (enough to cover the experimental frequency range) and oscillator strengths for a molecule in just its equilibrium configuration already allows an efficient characterization of the absorption spectrum (see results reported in Chapter 4). In fact it can be easily modeled with the envelop of some opportune broadening functions of "bell shape" (e.g. lorentzian or gaussian functions) whose peaks are centered on the resonance frequencies corresponding to the electronic excitation energies $\Omega_a(\bar{R})$ and whose intensities are proportional to the corresponding oscillator strengths. The broadening of the absorption lines can be eventually set in a way to reproduce the spectrum shape obtained experimentally.

Finally, since the life time of the first electronic excited states is in most cases large enough compared to every characteristic time of the vibrational motion, the thermal equilibrium of an excited molecule can be considered to be often established in an excited electronic state before recombination. In these cases, emission from an equilibrium distribution in an excited electronic state PES to the electronic ground state PES can take place spontaneously (fluorescence) or as stimulated process (see Figure 6.9). After recombination it can happen that the molecule fall back on the starting ground state PES minimum. If it is the case, and after de-excitation the same ground state PES minimum is reached (as it would be the case of Figure 6.9) the overall process of electronic excitation and de-excitation does not affect the chemistry of the chromophore. The case, instead, in which the excited molecule falls after recombination in an other ground state PES minimum (corresponding to an other molecular isomeric form) is a chemical photo-reaction.

Appendix B

Model of the solvent as a polarizable continuum

A simple but powerful method to take into account the charge polarization effects on the electronic density of a chromophore (solute) in solution, consists in treating it as a continuum medium characterized by its static permittivity ϵ. Within this model the solute, characterized by the total charge distribution

$$n_T(\mathbf{r}) = n_0(\mathbf{r}) + \sum_\alpha Z_\alpha \delta(\mathbf{r} - \mathbf{R}_\alpha) \tag{6.25}$$

is placed in a vacuum cavity embedded in an infinite continuum medium. The medium is considered polarizable, dielectric, homogeneous, not charged and with static permittivity ϵ. The molecular charge distribution induces in the dielectric a reaction potential which acts back on the charge distribution itself, which in return (self-consistently) changes the total solute charge density $n_T(\mathbf{r})$.

To see how this works in practice, first of all it can be used the fact that the polarized dielectric field can be described by means of a polarization charge ρ_p induced inside the dielectric to the which is associated a polarization vector $\overrightarrow{P}$ such that $\rho_p = \overrightarrow{\nabla} \cdot \overrightarrow{P}$. The first Maxwell equation (in gaussian units) inside the dielectric is

$$\overrightarrow{\nabla} \cdot \overrightarrow{E} = 4\pi \overrightarrow{\nabla} \cdot \overrightarrow{P} \tag{6.26}$$

where the macroscopic density charge is assumed to be zero because the solvent is composed by neutral molecules. As shown in Chapter 1, in the linear response regime the polarization vector is proportional to the electric field, and hence so does the electric displacement field

$$\overrightarrow{D} = \overrightarrow{E} - 4\pi \overrightarrow{P} = \epsilon \overrightarrow{E} \tag{6.27}$$

so that

$$\overrightarrow{P}(\mathbf{r}) = \frac{1 - \epsilon}{4\pi} \overrightarrow{E}(\mathbf{r}) = \frac{\epsilon - 1}{4\pi} \overrightarrow{\nabla} V_0(\mathbf{r}) \tag{6.28}$$

So, to take into account the solvent polarization effect in this model it is sufficient to take into account the creation of an induced surface charge density σ on the cavity boundary surface S around the solute dividing the vacuum space to the continuum. This charge density σ can be derived using the divergence theorem:

$$\int_V \overrightarrow{\nabla} \cdot \overrightarrow{P} = \int_S \overrightarrow{P} \cdot d\overrightarrow{S} \qquad (6.29)$$

where V is the total tridimensional space. The surface density is hence derived with the following relation:

$$\sigma(s) = \left. \overrightarrow{P} \cdot \hat{v}_n \right|_{s_+} = \left. \frac{\epsilon - 1}{4\pi} \frac{\partial V_0}{\partial v_n} \right|_{s_+} \qquad (6.30)$$

where $V_0 = \int \frac{n_T(\mathbf{r'})}{|\mathbf{r} - \mathbf{r'}|} d\mathbf{r'}$ is the electrostatic potential generated by the solute total charge distribution, $\hat{v}_n$ is the unitary vector orthogonal to the cavity surface, s is a coordinate indicating a point on the surface, and with s_+ is meant that for the evaluation of the potential it has to be considered the limit $\mathbf{r} \to \mathbf{r}(s)$ taken from the continuum side of the surface. Finally the following boundary condition

$$\left. \frac{\partial V_0}{\partial v_n} \right|_{s_-} = \epsilon \left. \frac{\partial V_0}{\partial v_n} \right|_{s_+} \qquad (6.31)$$

which expresses the continuity condition for the electric field can be used to get:

$$\sigma(s) = \left. \frac{\epsilon - 1}{4\pi\epsilon} \frac{\partial V_0}{\partial n} \right|_{s_-} \qquad (6.32)$$

so that everything now is expressed in terms of the coordinates in the inner surface volume. On the other side, the surface charge density σ generates an electrostatic potential

$$v_\sigma(\mathbf{r}) = \int_S \frac{\sigma(s)}{|\mathbf{r} - \mathbf{r}(s)|} ds \qquad (6.33)$$

which can be added in the electronic Hamiltonian of the solute generating, being just an external potential term, a corresponding in DFT, simply to an extra term in the Kohn Sham potential:

$$v_{s\sigma}[n_\sigma] = v_{s0}[n_\sigma] + v_\sigma(\mathbf{r}) \qquad (6.34)$$

Equation [6.32], together with the Kohn Sham equations with the potential [6.34] can be solved iteratively: first the DFT problem is solved in gas phase generating $n_T^{(0)}(\mathbf{r})$ from Eq. [6.25]. At this point $\sigma^{(0)}$ can be computed by means of Eq. [6.32]. So, a new Kohn Sham potential is generated by means of Eq. [6.34] and hence a new density $n_T^{(1)}(\mathbf{r})$ is computed. The procedure is repeated until convergency isn't reached.

What remains to complete the model is just a way to determine the cavity surface. There's not an unique prescription to do this and the parameters of the cavity have to be considered as empirical parameters to be added to the model and several different implementations for this cavity surface exist.

Appendix C

Electronic transitions analysis

In a TDDFT calculation in the Casida's formalism discussed in Chapter 1, many insight can come from the study of the electronic transitions character (or symmetry) associated with an excited state. Referring to the section 1.3.8 formalism, by electronic transition character is intended the study of the pairs of occupied and unoccupied molecular orbitals to the which are corresponding Casida's vectors significantly different from zero in a given excited states. These orbitals in fact drive the excited states forces acting on the molecule in the FK region

This proposition is easily proved applying to the Hellman Feynman forces $F_n = \frac{\partial E_n}{\partial R}$ the use of the auxiliary excited states wave functions defined in section 1.3.8:

$$F_n \simeq \langle \psi_{sn} | \frac{\partial H_{el}}{\partial R} | \psi_{sn} \rangle = \sum_{ia} |c_{ia}^n|^2 \langle \psi_{s(i \to a)} | \frac{\partial H_{el}}{\partial R} | \psi_{s(i \to a)} \rangle \simeq \quad (6.35)$$

$$= F_0 - \sum_{ia} |c_{ia}^n|^2 \langle \phi_i^{KS} | \frac{\partial H_{el}}{\partial R} | \phi_i^{KS} \rangle + \langle \phi_a^{KS} | \frac{\partial H_{el}}{\partial R} | \phi_a^{KS} \rangle$$

where F_0 is the Hellman Feynman force relative to the electronic ground state. The result claimed is so achieved noting that the forces are computed with the nuclei on an equilibrium configuration the term F_0 equals zero by definition. It is important to stress that, as stated in section 1.3.8, the use of the excited states functions to compute matrix elements is only exact for the ones connecting excited states to the ground state and they are only approximately true for excited state-excited state matrix elements. Moreover, also the second equality is, in general, only approximately true. This is because in practical numerical implementations of TDDFT the auxiliary wave functions can be built of orbitals expanded in atom centered functions basis sets. In these cases the orbitals depend explicitly on the nuclear position coordinates and the Hellman Feynman theorem is not valid anymore. The forces in fact have not been computed in this way in the calculations discussed in this work and equations 6.36 are just useful to show immediately how the forces approximatively depend on the pairs of molecular orbitals involved in a given excitation.

For the excitation character analysis, the molecular orbitals are conveniently divided in classes: there's for example the class of σ orbitals, which are symmetrical with respect to rotation about the bond axis of the bond where they are mainly localized; usually the

σ orbitals are obtained as superposition of atomic s, $p_{//}$ and $d_{//}$ orbitals (where with $//$ is denoted the direction parallel to some given bond in the molecule). They can have bonding and anti-bonding character if they are respectively increasing or decreasing the bond strength of the bond where they are located; for homodiatomic bonds this property depends respectively on the presence or absence of a nodal plane between these two bonded atoms. An other class is the one of π orbitals, for the which the bond vector of the bond where they are mainly localized is contained in a nodal plane of the orbital; orbitals of this class are usually given by superpositions of $p_\perp$ and $d_\perp$ atomic orbitals (where by $\perp$ is intended a direction perpendicular to some given bond). Finally, there's a third class of molecular orbitals whose occupation by electrons neither increases nor decreases the bond strength between the involved atoms. This orbital class is called non-bonding and is designated by the letter n. The non-bonding orbitals are in a way the equivalent in molecular orbital theory of the lone pairs in Lewis structures and are represented by the orbitals mainly localized in space regions without any bond presence.

Using this division of the molecular orbital in classes, the electronic transitions are commonly divided in the following classes:

$\sigma \to \sigma^*$

$\pi \to \pi^*$

$n \to \sigma^*$

$n \to \pi^*$

aromatic $\pi \to$ aromatic π^*

In this notation the $*$ just indicate that the second orbital is unoccupied. Of course these definitions have just to be intended as qualitative and non quantitative and are useful to give intuitive chemical explanations of geometrical changes induced on a molecule by the light absorption resonant with one of the electronic excited states. The affinity of a given orbital to one of the reported classes can be made by means of an isosurface plot, while the affinity of a transition to one of the given classes can be made analyzing the isosurface plot of the involved orbitals only in the cases in which there's just one dominant orbital pair contribution to the excitation. In all the cases in which there are several coefficients c_{ia} significantly different from zero the assignment cannot be easily made. To avoid these situations in every case the electronic transition discussion in the following will be referred to isosurfaces plots of the following transition densities:

$$\Delta n_n(\mathbf{r}) \equiv \sum_{ia} |c_{ia}^{(n)}|^2 \ \left\{ |\phi_a^{KS}(\mathbf{r})|^2 - |\phi_i^{KS}(\mathbf{r})|^2 \right\} \tag{6.36}$$

so that all the different contributions coming from different orbital transitions are plotted in the same isosurface where are weighted according to their relevance in the excitation; the positive part will be the superposition of all the unoccupied orbitals contributions while the negative part contains the information about the occupied orbitals. Just the information about the phase (i.e. the sign) of every orbital gets sacrificed for the advantage coming from the completeness and immediacy of these plots.

Bibliography

[1] Ullrich, C. A. (2012). *Time-Dependent Density-Functional Theory. Concepts and applications.* Oxford University press Inc., New York.

[2] Giuliani, G.F. and Vignale, G. (2005). *Quantum Theory of the Electron Liquid.* Cambridge University Press, Cambridge.

[3] Dirac, P. A. M. (1929) *Proc. R. Soc. Lond. A, **123**, 714-33*

[4] Hohenberg, P. and Kohn, W. (1064) *Phys. Rev*, **136**, B864-71

[5] Chayes, J. T., Chayes, L., and Ruskay, M.B. (1985) *J. Stat. Phys.*, **38**, 497-518

[6] Runge, E. and Gross, E. K. U. (1984) *Phys. Rev. Lett.*, **52**, 997-1000.

[7] van Leeuwen, R. (1999). *Phys. Rev. Lett.*, **82**, 3863-6.

[8] van Leeuwen, R. (1998). *Phys. Rev. Lett.*, **80**, 1280-3.

[9] van Leeuwen, R. (2001). *Int. J. Mod. Phys. B*, **15**, 1969-2023.

[10] van Leewen, R. (2003). *Adv. Quantum Chem.*, **43**, 25-94.

[11] *Phys. Rev. A.*, **77**, 062511-1-9.

[12] Gross, E. K. U., Ullrich, C. A., and Grossmann, U. J. (1995). In *Density Functional Theory* (ed. E. K. U. Gross and R. M. Dreizler), Volume 337, pp. 149-71. Plenum, New York.

[13] Thiele, M., Gross, E. K. U., and Kümmel, S. (2008) *Phys. Rev. Lett.*, **100**, 153004-1-4.

[14] Maitra, N. T., Todorov, T. N., Woodward, C., and Burke, K. (2010). *Phys. Rev. A.*, **81**, 042525-1-7

[15] Casida, M. E. (1995). In *Recent Advances in Density Functional Methods*, edited by D. P. Chong, World Scientific, Singapore, pp. 155-92

[16] Kohn, W. and Sham, L. J. (1965) *Phys. Rev.*, **140**, A1133-8.

[17] Tavernelli, I. (2010) *J. Chem. Phys.* **133**, 194104

[18] Lars Goerigk and Stefan Grimme (2011) *Phys. Chem. Chem. Phys.*, **13**, 6670-6688 DOI: 10.1039/C0CP02984J

[19] Ceperley, D.M. and Alder, B.J. (1980) *Phys. Rev. Lett.*, **45**, 566-9

[20] Tanatar, B. and Ceperley, D.M. (1989) *Phys. Rev. B.*, **39**, 5005-16

[21] Vosko, S. H. Wilk, L. and Nusair, M (1980) *Can. J. Phys.*, **58**, 1200-11

[22] Perdew, J. P. and Wang, Y (1992) *Phys. Rev. B.*, **45**, 13244-9

[23] Perdew, J. P. and Kurt, S. (2003). In *A primer in density functional theory* (ed. C. Fiolhais, F. Nogueira, and M. A. L. Marques), Volume 620, Lecture Notes in Physics, pp. 1-55. Springer Berlin.

[24] Perdew, J. P. and Schmidt, K. (2001). In *Density Functional Theory and Its Applications to Materials* (ed. V. E. van Doren, C. van Alsenoy and P. Geerlings), Volume 577, API Conference Proceedings, pp.1-20. American Institute of Physics, Melville, New York.

[25] Kurt, S. Perdew, J. P. and Blaha, P. (1999) *Int. J. Quantum. Chem.*, **75**, 889-909

[26] Koch, W. and Holthausen, M. C. (2001) *A Chemist's Guide to Density Functional Theory.* Wiley-VHC, Weinheim.

[27] Sousa, S. F., Fernandes, P. A. and Ramos, M. J. (2007) *J. Phys. Chem.* A ,**11**, 10439-52

[28] Rappaport, D., Crawford, N. R. M., Furche, F. and Burke, K (2009). In *Computational Inorganic and Bioinorganic Chemistry* (Ed. E. I. Solomon, R. B. King, and R. A. Scott), EIC Books, pp. 159-72. Wiley, Chichester.

[29] Morse, P. M. and Feshbach, *H. Methods of Theoretical Physics, Part I.* New York: McGraw-Hill, pp. 464-465, (1953)

[30] Szabo, A., Ostlund N. S., *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory.* Dover Books on Chemistry pp. 69-74 (1996)

[31] May, V. and Kuhn, O. (2011). *Charge and Energy Transfer Dynamics in Molecular Systems.* Wiley-VCH, Berlin.

[32] N. C. Handy, Y. Yamaguchi, H. F. Schaefer III, J. Chem. Phys. 1986, 84, 4481.

[33] W. Cencek, J. Rychlewski, R. Jaquet, W. Kutzelnigg, J. Chem. Phys. 1998, 108, 2831.

[34] P. R. Bunker, J. Mol. Spectrosc. 1968, 28, 422.

[35] J. O. Jensen, D. R. Yarkony, J. Chem. Phys. 1988, 89, 975.

[36] O. L. Polyansky, J. Tennyson, J. Chem. Phys. 1999, 110, 5056.

[37] Alder, B. J.; T. E. Wainwright (1959). "Studies in Molecular Dynamics. I. General Method". J. Chem. Phys. 31 (2): 459.

[38] A. Rahman (1964). "Correlations in the Motion of Atoms in Liquid Argon". Phys Rev 136 (2A): A405A411.

[39] CPMD, http://www.cpmd.org/, Copyright IBM Corp 1990-2008, Copyright MPI fr Festkrperforschung Stuttgart 1997-2001.

[40] I. Tavernelli, B. F.E. Curchod, U. Rothlisberger, (2011). "Nonadiabatic molecular dynamics with solvent effects: A LR-TDDFT QM/MM study of ruthenium (II) tris (bipyridine) in water". Chemical Physics 391, 101.

[41] Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM Jr, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA (1995). A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. J. Am. Chem. Soc. 117.

[42] W. F. van Gunsteren and H. J. C. Berendsen, Groningen Molecular Simulation (GROMOS) Library Manual, BIOMOS b.v., Groningen, 1987

[43] A. Laio, F. L. Gervasio, J. VandeVondele, M. Sulpizi & U. Rothlisberger J. Phys. Chem. B. 2004, 108, 7963-7968.

[44] H. Yoshida (1992) *Symplectic Integrators for Hamiltonian Systems: Basic Theory.* in Chaos, Resonance, and Collective Dynamical Phenomena in the Solar System, Edited by Sylvio Ferraz-Mello, pages 407-411.

[45] Tuckerman, M. (2010). Statistical Mechanics: Theory and Molecular Simulation. Oxford University Press Inc, New York.

[46] Ewald P. (1921) "Die Berechnung optischer und elektrostatischer Gitterpotentiale", Ann. Phys. 369, 253287.

[47] Martyna, Glenn J.; Klein, Michael L.; Tuckerman, Mark, The Journal of Chemical Physics, Volume 97, Issue 4, August 15, 1992, pp.2635-2643.

[48] Ryckaert, J-P; Ciccotti G, Berendsen HJC (1977). "Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes". Journal of Computational Physics 23 (3): 327341.

[49] Sprik, M. and Ciccotti, G. (1998). Free energy from constrained molecular dynamics. J. Chem. Phys., 109, 7737.

[50] Warshel, A; Levitt, M (1976). "Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme". Journal of Molecular Biology 103 (2): 22749.

[51] A. Laio, J. VandeVondele, U. Rothlisberger, J. Phys. Chem. B 2002, 106, 7300.

[52] A. Laio, J. VandeVondele, U. Rothlisberger, J. Chem. Phys. 2002, 116, 6941.

[53] Tavernelli I., Curchod B.F.E., Rothlisberger U. (2011) "Nonadiabatic Molecular Dynamics with Solvent Effects: a LR-TDDFT QM/MM Study of Ruthenium (II) Tris(bipyridine) in Water" Chemical Physics, 391, 101.

[54] K. Takatsuka, T. Yonehara, Phys. Chem. Chem. Phys. 2011, 13, 4987.

[55] T. Yonehara, K. Hanasaki, K. Takatsuka, Chem. Rev. 2012, 112, 499.

[56] J. C. Tully, J. Chem. Phys. 1990, 93, 1061.

[57] Trajectory-Based Nonadiabatic Dynamics with Time-Dependent Density Functional Theory B.F.E. Curchod, U. Rothlisberger and I. Tavernelli ChemPhysChem, 14(7), 1314-1340 (2013)

[58] Curchod B. F. E., Rothlisberger U., Tavernelli I. "Trajectory-Based Nonadiabatic Dynamics with Time-Dependent Density Functional Theory", ChemPhysChem, 14, 1314-1340.

[59] Watson, J.D.; Crick, F.H.C. Nature, 1953, 171, 737-738.

[60] Voet, D.; Gratzer, W. B.; Cox, R. A.; Doty, P. Biopolymers 1963, 1, 193-208.

[61] Daniels M and Hauswirth W 1971 Science 171 675

[62] Daniels M and Hauswirth W 1971 Science 171 675

[63] Gustavsson T.; Banyasz A.; Improta R.; Markovitsi D. 2011 J. Phys.: Conf. Ser. (2011) 261 012009

[64] Canuel C, et al. (2005) J Chem Phys 122:074316.

[65] Blancafort L, Cohen B, Hare PM, Kohler B, Robb MA (2005) J Phys Chem A 109:44314436.

[66] Gustavsson T, Banyasz A, Lazzarotto E, Markovitsi D, Scalmani G, Frisch M J, Barone V and Improta R 2006 J. Am. Chem. Soc. 128 607.

[67] Hudock HR, Levine BG, Thompson AL, Satzger H, Townsend D, Gador N, Ullrich S, Stolow A, Martnez TJ (2007) J. Phys. Chem. A, 111 (34), pp 85008508

[68] Kwok W-M, Ma C and Phillips D L 2008 J. Am. Chem. Soc. 130 5131

[69] Bona?i?-Kouteck? V, Kouteck? J, Michl J (1987) Neutral and charged biradicals, zwitterions, funnels in S1, and proton translocationTheir role in photochemistry, photophysics, and vision. Angew Chem Int Edit 26:170189

[70] Barbatti M, Lischka H (2008) Nonadiabatic deactivation of 9 H-adenine: A comprehensive picture based on mixed quantum-classical dynamics. J Am Chem Soc 130(21): 68316839.

[71] Hudock HR, Martinez TJ (2008) Excited-state dynamics of cytosine reveal multiple intrinsic subpicosecond pathways. ChemPhysChem 9:24862490.

[72] Asturiol D, Lasorne B, Robb MA, Blancafort Ls (2009) Photophysics of the $\pi\pi^*$ and $n\pi^*$ states of thymine: MS-CASPT2 minimum-energy paths and CASSCF on-the-fly dynamics. J Phys Chem A 113:1021110218.

[73] Szymczak JJ, et al. (2009) Photodynamics simulations of thymine: Relaxation into the first excited singlet state. J Phys Chem A 113:1268612693.

[74] Barbatti M.; Aquino A. J. A.; Szymczak J. J.; Nachtigallova D.; Hobza P.; Lischka H.; PNAS 107, 21453 (2010)

[75] C. Altucci, A. Nebbioso, R. Benedetti, R. Esposito, V. Carafa, M. Conte, M. Micciarelli, L. Altucci, and R. Velotta; Laser Physics Letters, 9, 234 (2012)

[76] M. Vermeulen, H.C. Eberl, F. Matarese, H. Marks, S. Denissov, F. Butter, K.K. Lee, J.V. Olsen, A.A. Hyman, H.G. Stunnenberg, and M. Mann, Cell 142, 967 (2010).

[77] S. Denissov, M. van Driel, R. Voit, M. Hekkelman, T. Hulsen, N. Hernandez, I. Grummt, R. Wehrens, and H. Stunnenberg, EMBO J. 26, 944 (2007).

[78] M. Kubista, J.M. Andrade, M. Bengtsson, A. Forootan, J. Jonak, K. Lind, R. Sindelka, R. Sjöback, B. Sjo?green, L. Strömbom, A. Stahlberg, and N. Zoric, Mol. Aspects Med. 27, 95 (2006).

[79] A. Nebbioso, N. Clarke, E. Voltz, E. Germain, C. Am- brosino, P. Bontempo, R. Alvarez, E.M. Schiavone, F. Fer- rara, F. Bresciani, A. Weisz, A.R. de Lera, H. Gronemeyer, and L. Altucci, Nat. Med. 11, 77 (2005).

[80] D.N. Nikogosyan, Int. J. Radiat. Biol. 57, 233 (1990).

[81] C. Russmann, M. Truss, A. Fix, C. Naumer, T. Herrmann, J. Schmitt, J. Stollhof, R. Beigang, and M. Beato, Nucl. Acids Res. 25, 2478 (1997).

[82] Ch. Russmann, M. Beato, J. Stollhof, C. Weiss, and R. Beigang, Nucl. Acids Res. 26, 3967 (1998).

[83] S.H. Chung and E. Mazur, J. Biophoton. 2, 557 (2009).

[84] S. Lejnine, G. Durfee, M. Murnane, H.C. Kapteyn, V.L. Makarov, and J.P. Langmore, Nucl. Acids Res. 27, 3676 (1999)

[85] C.T. Middleton, K. de La Harpe, C. Su, Y.K. Law, C.E. Crespo-Hernandez, and B. Kohler, Annu. Rev. Phys. Chem. 60, 217 (2009).

[86] Fecko CJ1, Munson KM, Saunders A, Sun G, Begley TP, Lis JT, Webb WW, Photochem Photobiol. 83,1394 (2007).

[87] E.I. Budowsky, M.S. Axentyeva, G.G. Abdurashidova, N.A. Simukova, and L.B. Rubin, Eur. J. Biochem. 159, 95 (1986)

[88] A. Reuther, D.N. Nikogosyan, and A. Laubereau, J. Phys. Chem. 100, 5570 (1996).

[89] E.V. Khoroshilova and D.N. Nikogosyan, J. Photochem. Photobiol. B 5, 413 (1990).

[90] R. Le Harzic, I. Riemann, K. Ko?nig, C. Wüllner, and C. Donitzky, J. Appl. Phys. 102, 114701 (2007).

[91] M.J. Solomon and A. Varshavsky, Proc. Natl. Acad. Sci. USA 82, 6470 (1985).

[92] C.T. Middleton, K. de La Harpe, C. Su, Y.K. Law, C.E. Crespo-Herna?ndez, and B. Kohler, Annu. Rev. Phys. Chem. 60, 217 (2009).

[93] F. Santoro, V. Barone, and R. Improta, Proc. Natl. Acad. Sci. USA 104, 9931 (2007).

[94] S. Baluschev, T. Miteva, V. Yakutkin, G. Nelles, A. Yasuda, and G. Wegner, Phys. Rev. Lett. 97, 143903 (2006).

[95] Angelov, D., M. Charra, C. W. Muller, J. Cadet and S. Dimitrov (2003) Solution study of the NF-kB p50-DNA complex by UV laser protein-DNA cross-linking. J. Photochem. Photobiol. 77, 592596.

[96] Angelov, D., V. Y. Stefanovsky, S. I. Dimitrov, V. R. Russanova, 22. E. Keskinova and I. G. Pashev (1988) Protein-DNA crosslinking in reconstituted nucleohistone, nuclei and whole cells by picosecond UV laser irradiation. Nucleic Acids Res. 16, 45254538.

[97] Budowsky, E. I., M. S. Axentyeva, G. G. Abdurashidova, N. A. 23. Simukova and L. B. Rubin (1986) Induction of polynucleotideprotein cross-linkages by ultraviolet irradiation. Eur. J. Biochem. 159, 95101. 24.

[98] Dobrov, E. N., Z. K. Arbieva, E. K. Timofeeva, R. O. Esenaliev, A. A. Oraevsky and D. N. Nikogosyan (1989) UV laser induced RNA-protein crosslinks and RNA chain breaks in tobacco mosaic virus RNA in situ. J. Photochem. Photobiol. 49, 595598.

[99] Wang, W. Y. Photochemistry and Photobiology of the Nucleic Acids; Wiley: New York, 1976.

[100] Morrison, H. Bioorganic Photochemistry; Wiley: New York, 1990.

[101] Shetlar, M. D. Frontiers of Photobiology; International Congress Series, 1993, 1021, 67-72 and references therein.

[102] Sun, G.; Fecko, C. J.; Nicewonger, R. B.; Webb, W. W.; Begley, T. P. Org. Lett. 2006, 8, 6813.

[103] Dreuw, A.; Head-Gordon, M. J. Am. Chem. Soc. 2004, 126, 400716.

[104] J. Chem. Theory Comput. 2007, 3, 976-987

[105] Micciarelli M.; Altucci C.; Della Ventura B; Velotta R.; Tosa V.; Gonzalez Perez A.B.; Perez M.; de Lera A. R.; Bende A.; Phys. Chem. Chem. Phys., doi:10.1039/C3CP50343G. (2013)

[106] Perdew, J. P. and Schmidt, K. (2001) Jacobs ladder of density functional approximations for the exchange- correlation energy. In Density Functional Theory and Its Application to Materials; Van Doren, V., Van Alsenoy, C., and Geerlings, P. (eds.), AIP, Melville, NY, pp. 120.

[107] Frisch, M. J. et al.; Fox, D. J. Gaussian 09, Revision A.02; Gaussian, Inc.: Wallingford, CT, 2009.

[108] J. P. Perdew, K. Burke, and M. Ernzerhof, Generalized gradient approximation made simple, Phys. Rev. Lett., 77 (1996) 3865-68.

[109] Y. Zhao and D. G. Truhlar, A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions, J. Chem. Phys., 125 (2006), 194101: 1-18.

[110] P.J. Stephens, F.J. Devlin, C.F. Chabalowski, M.J. Frisch, J.Phys.Chem. 98 (1994) 11623-11627

[111] Y. Zhao and D. G. Truhlar, The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals, Theor. Chem. Acc., 120 (2008) 215-41.

[112] T. Yanai, D. Tew, and N. Handy, A new hybrid exchange-correlation functional using the Coulomb-attenuating method (CAM-B3LYP), Chem. Phys. Lett., 393 (2004) 51-57.

[113] R. A. Kendall, T. H. Dunning Jr., and R. J. Harrison, Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions, J. Chem. Phys., 96 (1992) 6796-806.

[114] T. H. Dunning Jr and P. J. Hay, in Modern Theoretical Chemistry, ed. H. F. Schaefer III, Plenum, New York, 1976, 20 vol. 3, pp. 128.

[115] Tomasi, J.; Mennucci, B.; Cammi, R. Chem. Rev. 2005, 105, 2999.

[116] http://www.gaussian.com

[117] The Journal of Chem. Phys. 128, 044118 (2008)

[118] Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. J. Comput. Chem. 1993, 14, 13471363.

[119] S. Matsika, J. Phys. Chem. A, 2004, 108, 75847590

[120] Lorentzon, J.; Fulscher, M. P.; Roos, B. O. J. Am. Chem. Soc. 1995, 117, 9265-9273.

[121] M.J. Peach, A.J. Cohen, D. J. Tozer , Phys. Chem. Chem. Phys. 8, 4543 (2004)

[122] Improta R, Barone V.; J Am Chem Soc. 2004 Nov 10;126(44):14320-1

[123] DeFusco A, Ivanic J, Schmidt MW, Gordon MS.; J Phys Chem A. (2011) 12;115(18):4574-82. doi: 10.1021/jp112230f

[124] Barbatti M.; Aquino A. J. A.; Szymczak J. J.; Nachtigallova D.; Hobza P.; Lischka H.; PNAS 107, 21453 (2010)

[125] I. Tavernelli, B. Curchod, U. Rothlisberger, Nonadiabatic molecular dynamics with solvent effects: A LR-TDDFT QM/?MM study of ruthenium (II) tris (bipyridine) in water. Chemical Physics 391 (2011) 101109

[126] Micciarelli, M. Valadan, Della Ventura, B. Di Fabio, G. De Napoli, L. Bonella, S. Rothlisberger, U. Tavernelli, I. Altucci, C. Velotta, R. Photophysics and photochemistry of a DNA-protein crosslinking model: a synergistic approach combining experiments and theory. Submitted on JPC B

[127] Markovitsi, D.; Gustavsson, T.; Banyasz, A. Absorption of UV Radiation by DNA: Spatial and Temporal Features. Mutat. Res. 2010, 704, 2128.

[128] Gustavsson, T.; Banyasz, A.; Improta, R.; Markovitsi, D. Femtosecond Fluorescence Studies of DNA/RNA Constituents. J. Phys. Conf. Ser. 2011, 261, 012009.

[129] Siewertsen, R.; Renth, F.; Temps, F.; Snnichsen, F. Parallel Ultrafast E-C Ring Closure and E-Z Isomerisation in a Photochromic Furylfulgide Studied by Femtosecond Time-Resolved Spectroscopy. Phys. Chem. Chem. Phys. 2009, 11, 59525961.

[130] Improta, R.; Barone, V. Absorption and Fluorescence Spectra of Uracil in the Gas Phase and in Aqueous Solution: a TD-DFT Quantum Mechanical Study. J. Am. Chem. Soc. 2004, 126, 1432014321.

[131] Dreuw, A.; Head-Gordon, M. Failure of Time- Dependent Density Functional Theory for Long- Range Charge-Transfer Excited States: The Zincbacteriochlorin-Bacteriochlorin and Bacteriochlorophyll-Spheroidene Complexes. J. Am. Chem. Soc. 2004, 126, 40074016.

[132] Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; et Al. Gaussian 09, Revision A.1; Gaussian Inc.: Wallingford, CT, 2009.

[133] Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions,

[134] Kendall, R. A.; Dunning, T. H.; Harrison, R. J. Electron Affinities of the First-Row Atoms Revisited. Systematic Basis Sets and Wave Functions. J. Chem. Phys. 1992, 96, 6796.

[135] Scalmani, G.; Frisch, M. J.; Mennucci, B.; Tomasi, J.; Cammi, R.; Barone, V. Geometries and Properties of Excited States in the Gas Phase and in Solution: Theory and Application of a Time-Dependent Density Functional Theory Polarizable Continuum Model. J. Chem. Phys. 2006, 124, 94107.

[136] Kistler, K. A.; Matsika, S. Three-State Conical Intersections in Cytosine and Pyrimidinone Bases. J. Chem. Phys. 2008, 128, 215102.

[137] Micciarelli, M.; Altucci, C.; Della Ventura, B.; Velotta, R.; To?a, V.; Prez, A. B. G.; Rodrguez, M. P.; de Lera, A. R.; Bende, A. Low-Lying Excited-States of 5- Benzyluracil. Phys. Chem. Chem. Phys. 2013, 15, 7161 7173.

[138] DeFusco, A.; Ivanic, J.; Schmidt, M. W.; Gordon, M. S. Solvent-Induced Shifts in Electronic Spectra of Uracil. J. Phys. Chem. A 2011, 115, 45744582.

[139] Barbatti, M.; Aquino, A. J. a; Szymczak, J. J.; Nachtigallov, D.; Hobza, P.; Lischka, H. Relaxation Mechanisms of UV-Photoexcited DNA and RNA Nucleobases. Proc. Natl. Acad. Sci. U. S. A. 2010, 107, 2145321458.

[140] Thompson, A. L.; Martnez, T. J. Time-Resolved Photoelectron Spectroscopy from First Principles: Excited State Dynamics of Benzene. Faraday Discuss. 2011, 150, 293311.

[141] Valeur, B. Molecular Fluorescence: Principles and Applications; Wiley-VCH Verlag GmbH: Weinheim, 2001; pp. 353355.

[142] Zhou, P.; Song, P.; Liu, J.; Han, K.; He, G. Experimental and Theoretical Study of the Rotational Reorientation Dynamics of 7- Aminocoumarin Derivatives in Polar Solvents: Hydrogen-Bonding Effects. Phys. Chem. Chem. Phys. 2009, 11, 94409449.

[143] Kurnikova, M. G.; Balabai, N.; Waldeck, D. H.; Coalson, R. D. Rotational Relaxation in Polar Solvents . Molecular Dynamics Study of Solute - Solvent Interaction. 1998, 7863, 61216130.

[144] Nachtigallov, D.; Lischka, H.; Szymczak, J. J.; Barbatti, M.; Hobza, P.; Gengeliczki, Z.; Pino, G.; Callahan, M. P.; de Vries, M. S. The Effect of C5 Substitution on the Photochemistry of Uracil. Phys. Chem. Chem. Phys. 2010, 12, 49244933.

[145] Gustavsson, T.; Bnysz, A.; Lazzarotto, E.; Markovitsi, D.; Scalmani, G.; Frisch, M. J.; Barone, V.; Improta, R. Singlet Excited-State Behavior of Uracil and Thymine in Aqueous Solution: a Combined Experimental and Computational Study of 11 Uracil Derivatives. J. Am. Chem. Soc. 2006, 128, 607619.