TESIS DOCTORAL

VNIVERSITAT DE VALÈNCIA
INSTITUTO DE CIENCIA MOLECULAR

Theoretical Research on the Photophysics and Photochemistry of Water Aggregates and DNA/RNA Nucleobases

TESIS DOCTORAL PRESENTADA POR

Javier Segarra Martí

Valencia 2013

Programa de doctorado en Química Teórica
y Modelización Computacional

Dirigida por:
Dra. Manuela Merchán
Dr. Daniel Roca Sanjuán

Dña. Manuela Merchán, Catedrática del Departamento de Química Física y del Instituto de Ciencia Molecular de la Universitat de València, y D. Daniel Roca Sanjuán, Investigador Contratado del Instituto de Ciencia Molecular de la Universitat de València, certifican que la memoria presentada por el Licenciado Javier Segarra Martí con el título **Theoretical Research on the Photophysics and Photochemistry of Water Aggregates and DNA/RNA Nucleobases** corresponde a su Tesis Doctoral y ha sido realizada bajo su dirección, autorizando mediante este escrito la presentación de la misma.

En Valencia, Octubre de 2013

Dra. Manuela Merchán

Dr. Daniel Roca Sanjuán

Doctorando: Javier Segarra Martí

Agradecimientos

Me gustaría que estas líneas sirvieran para expresar mi más profundo agradecimiento a todas aquellas personas que han sido partícipes, ya sea de una forma más directa o indirecta, en la realización del presente trabajo.

A mi directora de Tesis, Manuela, por la orientación, el seguimiento y la supervisión realizada durante toda la Tesis, y sobre todo por creer en mí y en el presente proyecto. Por tener la paciencia de enseñarme paso a paso gran parte de su extenso conocimiento en química teórica y hacerlo siempre de una forma clara y sencilla, así como por estar siempre disponible para prestar consejo ya sea en lo profesional como en lo personal, muchas gracias por todo. A mi codirector de Tesis, Dani, por estar siempre presente más allá de la distancia, por la ayuda en las cuestiones más técnicas y del día a día y por el tremendo trabajo realizado junto con Manuela a la hora de corregir el presente manuscrito.

A todo el grupo QCEXVAL, y en especial a Angelo, Vicenta y Toni, con los cuales he tenido el placer de compartir despacho y muchas anécdotas, así como también de aprender de cada uno de ellos durante estos años. A Vicente, por ayudarme con las máquinas y los diferentes programas, sin los cuales nada de esto hubiese sido posible. A toda la Unidad de Química Teórica del Instituto de Ciencia Molecular, y en particular a Juan y a Joaquín, por todos los buenos ratos pasados tanto en el Instituto como fuera de él.

A los profesores Marco Garavelli y Michael Thoss, haciendo extensivo el agradecimiento a sus respectivos grupos, por recibirme con los brazos abiertos y por ayudarme en todo lo que necesité durante mis respectivas estancias con ellos, así como por brindarme la oportunidad de conocer nuevas culturas y hacer nuevos amigos tanto en Bologna como en Erlangen.

A mis amigos, tanto al grupo originario de Massamagrell como a todos aquellos que he conocido por el camino, muchas gracias por aguantarme durante todo este tiempo y por estar ahí cuando necesité un amigo con el que hablar o unas cervezas que beber.

Por último, me gustaría agradecer a toda mi familia, y en especial a mi hermano y a mi madre, por creer siempre en mí y por estar siempre a mi lado en todo momento más allá de la dificultad o la distancia. Sin ellos nada de esto hubiese sido posible.

Contents

Part I

Introduction

Chapter 1

Scope

> You can't trust water: Even a straight stick turns crooked in it.
>
> *W.C. Fields*

Over the last decade we have witnessed an unprecedented increase in photophysical and photochemical studies focused in biomolecules, both from a theoretical and experimental standpoints. This is mainly due to the technological advances provided in both fronts, which allow for a more accurate and intensive study of the photochemical processes taking place at the molecular level.

Upon light irradiation, a molecule possesses an excess in energy in an electronic excited state, which can derive in a series of scenarios: in a *photochemically productive manner*, evolving towards the formation of photoproducts and subsequently emitting the excess energy as radiation, or in an *unproductive way*, relaxing the excited state while dissipating its excess energy and reaching a

crossing with the ground state that brings the molecule to its starting point and is thus considered **photostable**.[1, 2]

The photoreactivity undergone by the populated excited states can then, depending on the given system and its intricate characteristics, be beneficial or harmful from a biological standpoint. In typical biological chromophores, such as the DNA/RNA nucleobases, photochemically productive pathways between nucleobases may evolve in harmful mutagenic and carcinogenic photoproducts,[3] even though the own monomers have a high degree of photostability.[4–7]

In 2003, the QCEXVAL group (*Quantum Chemistry for the Excited State of Valencia*) belonging to the Instituto de Ciencia Molecular from the University of Valencia, in which this Thesis is ascribed, pioneered the introduction and use of highly correlated *ab initio* quantum chemical methods for the study of the non-adiabatic photochemical pathways present in DNA/RNA systems, in which high-level computational strategies had not been previously reported.[8] Ever since, the QCEXVAL group has deepened the understanding on many systems within the photochemical and photobiological realms.

The present Thesis prospects a step further along the path towards the proper characterization of the main photophysical and photochemical events occurring in solvated biomolecules, hence focusing our main efforts on the study of the intrinsic properties related to water and its structural motifs in biological systems, bearing in mind that water is the main and most common biomolecule present in life. Further work along the DNA/RNA photochemical pathways has also been addressed, building upon the previous knowledge generated over the last decade in the QCEXVAL group and hinting towards an interplay between the photophysical and photochemical routes investigated for water and the decay channels already characterized in the DNA/RNA molecules.

1.1. Water, water, everywhere!

Water has been by far the most studied substance on Earth and yet plenty of mysteries about its properties and structural characteristics prevail.[9–15] It has always been assumed to play a crucial role in life, so much that its very presence in planets and asteroids is linked with the possibility of finding extraterrestrial life. Water was firstly characterized at the end of the nineteenth century by Henry Cavendish and Antoine Laurent Lavoisier,[16] and ever since it has drawn a massive interest from the research community. Plenty of scientists, among them the Nobel laureate Linus Pauling, tried to tackle the different anomalous properties present in water as well as to find possible clues concerning its structure, which so far remains inconclusive. A large amount of biological and industrial processes, such as ionization and solvation, protein folding mechanisms, and several other biomolecular functions, are attributed to its unknown structural motifs.[13, 14, 17, 18] Many are the scientists currently working on these topics related to the structuring properties of water, which usually invoke a plethora of different models to provide answers for the several anomalies found. The problem is then complex and not as straightforward as previously thought, and therefore an underlying question remains to be answered: **What do we mean by water?**

Water has baffled scientists over the years due to its intrinsic complexity. It has three different phases, commonly known as gas, liquid, and solid, yet plenty of frontier zones keep emerging blurring this ideal picture of water.[19–21] Several are the properties changing from one phase to the other, yet a special focus shall be put on the spectroscopic fingerprints assigned to each of them. In the gas phase, water holds an absorption energy of 7.447 eV,[22] which can be related to the property of a single water molecule, due to the large space given between the molecules. Yet, in the liquid state, water has been reported to go to higher energies up to ∼8.2 eV,[23–25] whereas the absorption energy of its solid form has been estimated at ∼8.7 eV.[26] These large changes in the absorption spectrum are indeed arisen due to different conformations and

intermolecular interactions held at the different phases, although plenty of difficulties lie ahead of the explanation.

In the gas phase, water suffers from *Valence-Rydberg mixing* or *Rydbergization* of its orbitals, which represents a mixture between the valence and the Rydberg orbitals of the molecule and provides plenty of problems to be correctly treated computationally. This fact narrows down the theoretical strategies to be used.[27] To provide a proper treatment of the system, highly correlated methods have to be employed thus hampering the extension of the model to large molecular systems due to computational demands. This enhances the difficulty of treating water at the proper computational level in order to yield predictive or at least qualitatively good results. As a way to tackle all the problems inherent of working with condensed phases (large systems), one can study aggregates of water known as *clusters*,[28–31] which are of importance in fields like astrochemistry. Making use of these *clusters*, one can add one molecule at a time and study the perturbations generated in the surrounding water molecules to provide an explanation of their intermolecular interactions. These techniques have been vastly used to ascertain the intricacies of the *hydrogen bond*,[32] a type of interaction that still puzzles the research community. As will be shown in the chapter of results and discussion of the present Thesis, such a treatment does not lead towards the excitation energy value given for liquid water. This might be due to the fact that the structural motifs given in the *clusters* probably do not completely replicate the ones found in liquid water, so the question arises: What is the structure of liquid water?

The current *generally accepted* model of liquid water is the nearly-tetrahedrally coordinated water introduced by Stillinger in 1980.[33] The model takes into account small variations from the tetrahedral structure of ice based on melting experiments and interprets them as dynamically changing motifs with more flexibility as compared to its solid phase.[34] This is what has been commonly coined as **bulk water**, which is the water present in any solution containing nothing but water itself. The structure therefore presents plenty of problems upon its characterization, as the dynamical equilibrium with its surroundings makes it really hard to measure experimentally. Theoretical estimates based

on molecular dynamics predict around 3.2-3.4 hydrogen-bonding coordination on each water as an average,[35] which is consistent with the model previously presented. This model appears to be based on solid grounds, yet it has been recurrently challenged. A series of theoretical models have recently proposed a more linear approach, in which two of the hydrogen bonds would be stronger than the rest, giving rise to chain-like structural motifs.[36, 37] It can therefore be envisaged, that the assumption of the correct structure for liquid water is not straightforward and that the topic is far from being settled. Even the most common processes known to water, such as its *autoionization process*,[38] broadly taught in high schools all over the world, has been under extensive scrutiny ever since its proposal, and there is still no clear answer to its concrete mechanism at the molecular level.[39, 40]

In the light of what has already been outlined, it is not easy to elucidate the factor responsible for the excitation energy shifting towards ∼8.2 eV, or if the energy value corresponds to the same type of state responsible for the excitation energy given in the gas phase or to another type of intermolecular excitation mechanism. What appears to be more troubling, a signal around ∼4.6 eV has been recurrently recorded over the years in the liquid phase,[21, 41–43] the origin of which is still under debate. This signal may be specially relevant, as it appears in the energy range in which the majority of the biochromophores absorb light. It could therefore have crucial effects on their photophysical and photochemical routes by means of facilitating the release of the excess energy upon light absorption or by hampering certain photochemical decay paths. The low-energy absorption signal has been mostly registered in superficial situations, in which water has the ability and capability to structure and further extend itself over a long range. This type of rearrangement has been called **ordered water**, and it usually appears in contact with charged surfaces such as those found in biomolecules or in any other charged environments as well as in metallic surfaces.[44, 45] The structure of the so-called ordered water does therefore look more promising to unveil the underlying solvated photochemical deactivation routes of biological systems and therefore provides new insights into the mechanisms governing solvation. To this end, an effort to give a molecular counterpart to the ∼4.6 eV signal has been attempted in

the present work. In order to further investigate ordered water, a series of models has been built to elucidate the main interactions that give rise to the low-lying energetic absorption. The results and conclusions achieved has been summarized in the results section.

The last phase to be addressed is the solid. Solid water is commonly known as ice, and it holds more mysteries than one would expect. Normal ice (hexagonal ice, I_h) is just one of the fifteen currently existing types of ice. Its structure has already been described above as a tetrahedrally coordinated water system in which the water molecules form hexagonal layers. The constriction of the hydrogen bonded system allows an accurate characterization of its structure, which has been known since long ago. Even then, ice is not exempt of its own controversy: how many water molecules do actually form an ice layer? Due to its similarities with bulk water as well as to the ordered water layers present over surfaces, it is not straightforward to actually distinguish between ice and liquid water. It was not until very recently, that the answer was achieved.[46] The intrinsic properties of the solid start showing up only after 275 water molecules are present in the layer and it totally predominates over the liquid at 475 or more.

It can therefore be concluded that water is not as flawlessly characterized as many people might think. Its intrinsic structural motifs and reaction processes at the molecular level keep humbling scientists all over the world due to its tremendous complexity and a deeper knowledge of its underlying mechanisms would facilitate a proper description of the solvation processes actually occurring *in vivo*. To this end, the QCEXVAL group has focused on the task to elucidate the molecular fingerprints responsible for the interesting ∼4.6 eV fingerprint in water and its effects towards unraveling the real role of water in photoinduced biological and industrial processes.

1.2. DNA/RNA and the genetic code

DNA/RNA has a significant advantage as compared to water when it refers to the determination of their photophysical and photochemical properties: it has a well-defined structure.[47] The knowledge of its structural motifs provides an initial guess as to how the DNA/RNA is placed *in vivo*. As the DNA/RNA chains are comprised by a massive amount of atoms, a general reductionist model states that only those molecules that absorb light prominently are to be considered in a photophysical/photochemical model. Basing our models on that assumption, the DNA/RNA nucleobases are the chromophores that mainly absorb light in the DNA chain, and therefore they are the ones to be studied. The nucleobases can be held accountable of the majority of mutations and photochemical processes happening in the DNA chain upon light irradiation and have therefore been considered as the model study over the years.[4–8, 48]

As mentioned above, the QCEXVAL group main research line has focused over the last ten years on the different processes undergone by the nucleobases in the DNA/RNA chain. Several processes can be accessed by the nucleobases in order to release their excess energy upon light absorption: i) a single nucleobase can liberate the excess energy through a radiationless decay towards the ground state making it photostable;[4–8] ii) in a *Watson-Crick base pair*, which corresponds to each pair of hydrogen-bonded present in the DNA double helix, the decay can take place through excited-state hydrogen transfer processes between the bases of the pair;[49, 50] and iii) in π-stacked nucleobases, the stacking interaction, which is responsible for the red-shifted fluorescence spectrum registered experimentally for DNA, can result in the production of cyclobutane pyrimidine dimers or pyrimidine-pyrimidone (6-4) adducts.[3, 51–54] All these models account for the separated excitation of these subunits describing different types of interactions, whereas in the actual situation they may happen simultaneously and might be competitive with each other. Once having attained a proper description of each of these interactions, ongoing work is currently being developed in this direction in order to as-

certain which are the key factors that prevail upon light irradiation in DNA and therefore obtain a deeper understanding of the photophysical and photochemical properties of DNA upon light irradiation, knowledge of paramount importance for biomedical purposes.

The possibility of photoinduced ionization processes has also been addressed over the years. Such events are of importance in DNA/RNA as they generate charged species, which can contribute to the charge-transfer mechanisms governing the double helix chain.[55] In this case, the DNA/RNA nucleobases are usually taken as the model,[56, 57] but the effects of the sugar and phosphate groups had to be assessed in order to check the reliability of the reduced model to replicate the photoelectron spectroscopic fingerprints registered in the experiments.[58, 59] Once being irradiated with a photoionized light, the loose electrons generated can initiate harmful events like DNA-strand breaks,[60] and to this end further work in this area has been performed here. The *dissociative electron attachment* of uracil has been studied and the conclusions obtained can be seen in the results section.

Finally, due to the advent of new theoretical tools and massively increased computational power, the findings reached over the last decade in the QCEX-VAL group have been revisited with improved and more correlated computational strategies to ascertain the decay paths already predicted and to validate the computational strategy previously used. The use of new computational schemes to further investigate the behavior of the DNA/RNA chromophores is therefore ongoing, in order to provide new insights and try to cease the endless debate coming from approximated approaches and give a definite answer about the deactivation processes occurring in the nucleobases upon light irradiation.

Chapter 2

Objectives

> First, have a definite, clear practical ideal; a goal, an objective. Second, have the necessary means to achieve your ends; wisdom, money, materials, and methods. Third, adjust all your means to that end.
>
> *Aristotle*

Along the full extent of this Thesis, a series of goals has been devised to be tackled. The work focuses mainly on the different photoinduced processes occurring in water, in order to provide a deeper understanding of the solvent surrounding DNA/RNA as well as to treat water itself as a biomolecule. To this end, a proper theoretical description of water has to be given by means of highly correlated *ab initio* quantum chemical methods, which is something that has been missing so far in the literature and has been therefore addressed.

Several water systems ranging from water clusters to more elaborate ordered-water-like models have been designed to investigate their underlying photophysical and photochemical properties and its trend through the use of larger models in order to unveil the common key features responsible for their unique properties.

Water clusters are studied through their photoelectron and UV spectra. Assessing the evolution of their relative fingerprints holds an answer to the influence of the hydrogen bonding network surrounding water and provides valuable estimates for future reference. The study of the dimer, trimer, tetramer, and pentamer clusters and the four hexamer conformers provides a unique way of tying the properties of the gas- and liquid-phase in water, as the smaller clusters are present in the gas phase whereas the more hydrogen-bonded coordinate hexamers relate closely to the 3.2-3.4 hydrogen bond coordination expected for bulk water. The photoelectron spectra shall provide useful information about photo-oxidation processes *in vivo*, whereas the UV spectra shall give an answer to the nature of the excited state of the gas phase and the one responsible for the absorption of bulk water.

The water dimer is extensively studied as it can be considered the smallest building block of water. In this sense, the photophysical and photochemical properties of the gas-phase dimer are thoroughly evaluated and compared with a water dimer that holds a π-stacking interaction as its main feature. Their intrinsic decay channels are compared and their spectroscopical fingerprints used to build upon them to modelize bulk and ordered water, respectively. An hexameric model of water, built upon the knowledge attained from the π-stacking water dimer is devised in order to provide a more realistic molecular counterpart yielding the essentials of the $\sim$4.6 eV spectroscopical fingerprint.

Finally, the intricacies of the *dissociative electron attachment* in uracil are studied as a possible way of breaking DNA/RNA strands by collecting the loose electrons left after ionizing light. All photochemical pathways found in the nucleobases are revised and checked against more correlated methods available nowadays thanks to the improvement in computational power. Explicitly, a comparative study of the decay pathways is undertaken with the CASPT2//CASSCF and CASPT2//CASPT2 protocols.

Chapter 3

Theoretical Background

> With four parameters I can fit an elephant, and with five I can make him wiggle his trunk.
>
> *John Von Neumann*

Many different theoretical methods exist to describe physical and chemical problems. They can be roughly placed within three big groups: classical, quantum, and hybrid methods. Classical methods are based on classical mechanics (Molecular Mechanics, MM), and therefore will be the cheapest, yet the less precise methods available. MM methods can be applied to model huge systems, such as whole nucleic acids or proteins, but are not recommended to describe reactivity or problems in which large changes in the distribution of the electronic density takes place. Quantum methods, on the other hand, rely on the foundations of quantum mechanics (QM), providing a higher accuracy and are therefore able to describe a wider variety of phenomena. The drawbacks are however on the computational requirements, which restrict their applicability to rather small systems. Despite much effort has been focused

on the development of quantum-chemical approximations aimed to deal with large systems, the combination of computational methodologies with different accuracy and computational cost seems to be the most appropriate approach nowadays to treat complex large systems. Hybrid methods combine both previous methodologies (QM/MM) in order to accurately describe small systems embedded in larger environments, hence taking into account surrounding effects.

High-level QM methodologies have been mainly employed in the present Thesis. Therefore, we shall focus on the relevant QM approaches related to the work carried out.

QM methods describe the physical and chemical problems by making use of the *Quantum Mechanics* formulation. In most of the cases in Quantum Chemistry, we are interested in the stationary electronic solutions of the time-independent Schrödinger equation:

$$\hat{H}\Psi = E\Psi \tag{3.1}$$

where $\hat{H}$ is the Hamiltonian operator described as:

$$\hat{H} = \sum_{i=1}^{N} \frac{1}{2}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>1}^{N} \frac{1}{r_{ij}} + \sum_{A=1}^{M}\sum_{B>A}^{M} \frac{Z_A Z_B}{R_{AB}} \tag{3.2}$$

This formulation takes into account several approximations to solve the electronic structure problem, such as the Born-Oppenheimer approximation, which allows the separation of the nuclear and electronic movements, and the neglect of relativistic effects. The Hamiltonian operator holds the physical meaning of the problem and is formed by the addition of different contributions: the kinetic energy, the electron-nucleus attraction, the electron-electron repulsion, and the nucleus-nucleus repulsion. The last term remain constant for a given nuclear configuration under the *Born-Oppenheimer approximation* and can therefore be removed and added afterwards as its value will not vary. The electronic Hamiltonian also bears the meaning for the so-called electronic correlation, which is related to the impossibility for two electrons to be found at the same given elemental volume. The energy obtained as solution of eq. 3.1,

E, corresponds to the total energy at a given fixed nuclei arrangement and provides a potential for the nuclei movement. Slater determinants are the simplest examples of many-electron wave functions fulfilling the identity and antisymmetric principles. They are built by orthogonal one-electron spin-orbitals (χ), which are the product between a spatial orbital and a spin function:

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \begin{pmatrix} \chi_1(\mathbf{x}_1) & \chi_2(\mathbf{x}_1) & \cdots & \chi_N(\mathbf{x}_1) \\ \chi_1(\mathbf{x}_2) & \chi_2(\mathbf{x}_2) & \cdots & \chi_N(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(\mathbf{x}_N) & \chi_2(\mathbf{x}_N) & \cdots & \chi_N(\mathbf{x}_N) \end{pmatrix} \tag{3.3}$$

3.1. Single Reference Methods

In order to solve eq. 3.1 for real systems, *ab initio* methodologies are normally used, most of them starting from the Hartree-Fock (HF) method. HF was firstly introduced in 1930 and makes use of the *variational principle* in order to find the *best* determinant. Let us recall that the variational principle implies a continuous improvement of a given type of wave function causing progressively the lowering of the energy in the system towards the exact value, which would be achieved by including the full electronic correlation. The equations of the HF method are not linear and require an iterative method, also known as the Self-Consistent Field (SCF). In the solution, a series of basis set spatial functions, which will be introduced later in more detail, are required to give rise to the matrix equations of the HF theory: the Roothan, Pople-Nesbet, and Roothan-Hartree-Fock equations for closed-shell, open-shell, and restricted-open-shell systems, respectively. In order to describe open-shell problems, the ROHF and UHF methods were developed. ROHF requires the two closed-shell electrons to have the same spatial function and the ROHF functions are therefore eigenfunctions of the spin operator S^2. On the other hand, the UHF method does not have to fulfill this restriction and can be applied more easily. Nonetheless, the functions obtained are not eigenfunctions of the spin operator; they are mixed with higher-spin functions, in what is known as

spin contamination. Projection techniques have been developed in order to reduce spin contamination. These procedures are mainly applied to the UHF method (PHF), and also more correlated methods based on the UHF reference function, such as the UMP2 which will be described below in more detail.

3.1.1. Configuration Interaction (CI)

The Hartree-Fock approach does not provide enough precision to describe many of the phenomena under study as it lacks the electronic correlation, which can be defined at this point as the difference between the exact energy obtained within the Born-Oppenheimer approximation and the value obtained with the HF method within a complete basis set. Further improvements are therefore required to include to some extent the electron correlation energy. The first and more intuitive idea in this direction is to consider in the wave function of eq. 3.1 not just a Slater determinant, as shown in eq. 3.3, but a linear combination of Slater determinants, which improve the flexibility of the method. Hence, the wave function is written as:

$$\Psi = a_0\Phi_{HF} + \sum_{i=1} a_i\Phi_i \tag{3.4}$$

where a_i corresponds to the coefficient of a certain determinant within the expansion. The terms on this expansion are Slater determinants where some of the spin-orbitals have been changed by spin-orbitals that were not occupied in the HF wave function. Since the HF solution gives rise to N/2 occupied molecular orbitals (MOs) and Q-N/2 virtual orbitals, where Q denote the basis functions, the expansion in eq. 3.4 results in a huge variety of determinants, which can be classified according to the number of interchanged spin-orbitals compared to the HF wave function(singly, doubly, triply excited determinants, and so on). Within this treatment, the correlation energy can be accounted for, but in practice it has to be truncated at some point as the expansion grows considerably fast.

The *Configuration Interaction* (CI) method is therefore based on the variational principle and uses a linear combination of Slater determinants to form its wave function, as described above. In this method, the a_i coefficients from eq. 3.4 are required to minimize the energy. The method is called *Full Configuration Interaction* when all possible excited determinants (singly, doubly, triply excited determinants, etc.) are included in the CI wave function and gives rise to the exact solution within the basis set employed. Normally, due to the fact that the number of determinants and coefficients grows very fast as the number of electrons increase, this expansion shall be truncated obtaining other levels of theory such as *Doubly excited Configuration Interaction* (DCI), *Singly and Doubly excited Configuration Interaction* (SDCI), *Singly, Doubly, and Tryply excited Configuration Interaction* (SDTCI), etc.[1] This process introduces a big problem within the method, which is the fact that it does not escalate properly with the size of the system or that it is not **size-extensive**, meaning that the system gives a worse estimation of the correlation energy whilst increasing the size of the molecule and therefore makes the comparison between systems of different size rather difficult.

3.1.2. Møller-Plesset Second-Order Perturbation Theory (MP2)

Another way of improving upon the Hartree-Fock result, often described as Post-Hartree-Fock methods, would be applying Many-Body Perturbation Theory (MBPT), and more specifically the ones based on the Møller-Plesset (MP) Perturbation Theory up to different orders (MP2, MP3, etc.). This method uses the single-configuration HF process to obtain a set of spin-orbitals (χ_i). Then, making use of an unperturbed Hamiltonian H^0, perturbation theory is used to determine the amplitudes for the different *Configuration State Functions* (CSFs). The amplitude of the reference is taken as unity, whereas the rest are computed by Rayleigh-Schrödinger perturbation using $H - H^0$ as the

[1] *Singly excited Configuration Interaction* (SCI) has been omitted as it does not hold any improvement over the HF reference for the description of its ground state due to the application of *Brillouin's Theorem*; further details are given elsewhere.[61]

perturbation. Normally, a HF single determinant is used as the reference function. Once the reference is provided, the wave function Ψ and the energy E are determined in an order-by-order manner. This means, the first-order wave function correction Ψ^1:

$$\Psi^1 = \sum_{i<j(occ)} \sum_{m<n(virt)} [\langle \chi_i \chi_j | \chi_m \chi_n \rangle - \langle \chi_i \chi_j | \chi_n \chi_m \rangle] \\ [\epsilon_m - \epsilon_i + \epsilon_n - \epsilon_j]^{-1} |\psi_{i,j}^{m,n}\rangle \quad (3.5)$$

is to be determined before computing the second-order correction on the energy. In Ψ^1, the SCF orbital energies are denoted ϵ_k and $\psi_{i,j}^{m,n}$ represents a CSF that is doubly excited (where χ_i and χ_j are replaced by χ_m and χ_n) relative to the zeroth-order wave function Φ. The reason that singly excited determinants are not accounted for is related to the *Brillouin theorem*, which is covered in detail elsewhere.[1] Once the wave function is known up to first order, the energy E can be computed through second-order, yielding:

$$E = \langle \Phi | H^0 + V\Phi + \Psi^1 \rangle = E_{SCF} - \\ \sum_{i<j(occ)} \sum_{m<n(virt)} [\langle \chi_i \chi_j | \chi_m \chi_n \rangle - \\ \langle \chi_i \chi_j | \chi_n \chi_m \rangle]^2 [\epsilon_m - \epsilon_i + \epsilon_n - \epsilon_j]^{-1} \quad (3.6)$$

Ψ and E are expressed in terms of two-electron integrals $\langle \chi_i \chi_j | \chi_m \chi_n \rangle$ coupling the virtual spin-orbitals χ_m and χ_n to the occupied spin-electrons that were excited χ_i and χ_j as well as the orbital energy differences associated to each excitation. For higher-order corrections, contributions from CSFs that are singly, triply, etc. excited relative to Φ are also relevant. This specific case would be representative of the *Second-Order Møller-Plesset Perturbation Theory*, which covers around 90% of the correlation energy, while requiring a relatively low computational cost. One has to bear in mind as well that the perturbations are bound to be small, otherwise the perturbation treatment might not converge. On the other hand, perturbation theory guarantees size-consistency. As in the HF method, MP may suffer from spin contamination problems in open-shell systems. Similar projected techniques as in the HF procedure have been developed giving rise to the *Projected Møller-Plesset*

(PMP) methods. These methods provide the tools needed to analyze the spin contamination problem and usually give an estimate as to which degree the method might be trusted in such open-shell systems.

3.1.3. Coupled Cluster (CC)

The last of the single-reference methods described is the *Coupled Cluster* method. It is a Post-Hartree-Fock method initially developed in the 1960s. It can be seen as a way to implement the size-extensitivy to the CI approach discussed before. This can be achieved through the *Ansatz* formulation, which converts it in an exponential problem, $\Psi = exp(T)\Phi$. This type of formulation allows the truncation at any order, yet it yields a size-extensive method, enabling its safe use in systems of different size, but at a rather large computational cost. The amplitudes T can then be expressed in terms of operators that achieve spin-orbital excitations as follows:

$$T = \sum_{i,m} t_i^m m^+ i + \sum_{i,j,m,n} t_{i,j}^{m,n} m^+ n^+ j i + \cdots \tag{3.7}$$

Coupled Cluster Singles and Doubles (CCSD) is therefore normally the only generally applicable CC method, obtained by truncating the T series at the second term. As in the MP method and other methods based on a HF wave function, it may also suffer from spin contamination in open-shell systems. However, the infinite nature of this method is substantially better at reducing spin contamination. In CC theory, it is often common to start with an initial guess formed by neglecting all the terms that are nonlinear in the t amplitudes and ignoring factors that couple different doubly excited CSFs, which provides amplitudes analogous to those obtained within *Møller-Plesset Perturbation Theory* at the first-order wave function(eq. 3.5):

$$t_{i,j}^{m,n} = -[\langle \chi_i \chi_j | \chi_m \chi_n \rangle - [\langle \chi_i \chi_j | \chi_n \chi_m \rangle][\epsilon_m - \epsilon_i + \epsilon_n - \epsilon_j]^{-1} \tag{3.8}$$

The T_1,[62] based on the T_1 amplitudes ($\sum_{i,m} t_i^m m^+ i$), is a diagnostic that can be consulted in order to check the adequacy of the calculation. Larger

T_1 values than 0.02 imply the necessity of a more correlated approach within the method, and higher-order amplitudes (triple, quadruple, etc.) must be used to reach the energy convergence. In order to check the adequacy of the single reference used for the calculation, the largest T_2 amplitude can be checked, hinting towards multireference behavior at values over 0.2. Since CC methods are computationally very demanding, an alternative to CCSDT is CCSD(T), where the triple-type contributions are evaluated by the less expensive perturbation theory and subsequently added to the CCSD result. The CC methodology is a single reference based approach, like the MP, and its application will be correct as long as the problem can be described properly by a single configuration. In other situations, the multireference approach, *Multireference Coupled Cluster* (MRCC) shall be employed, but its computational cost is extremely high.

Regarding the accuracy of the methods analyzed so far, the following order is often observed with a medium-size basis set:

$$HF << MP2 < CISD < CCSD < CCSD(T)$$

3.2. Multireference Methods

Whereas single reference methods are by far more popular, they often fail to describe situations where **non-dynamical electron correlation** are important, as happens in the description of bond stretching and dissociation, electronic excited states and many open-shell cases. *Multireference Methods* are capable of coping with such problems, even though their successful application requires a more careful analysis and understanding of the details compared to the single reference case, especially in the selection of appropriate reference spaces. A multireference treatment usually starts with a multiconfiguration SCF (MCSCF) calculation, which provides the non-dynamical electron correlation by including all the possible (and most important, all the needed)

CSFs to give a proper description of the system. The remaining contributions, the so-called **dynamical electron correlation**, can then be included mainly in two ways: Through the multireference expansion of the CI method, *Multireference Configuration Interaction* (MRCI), or through the application of perturbation theory on top of the MCSCF-type calculation.
The MCSCF wave function is expressed in the same form as the CI one, and depends also on the variational principle, but both CI coefficients and the molecular orbitals obtained within the previous SCF procedure are optimized. The MCSCF wave function does not have any spin contamination problems since the functions used in the expansion (CSFs) are pure spin states.

In this section special emphasis shall be put on a specific MCSCF method, the *Complete Active Space Self Consistent Field* (CASSCF),[63] and the second-order perturbation treatment on top of it known as *Complete Active Space Second-Order Perturbation Theory* (CASPT2),[64–66] methods that have been extensively used during this Thesis.

3.2.1. Complete Active Space Self-Consistent Field (CASSCF)

CASSCF is certainly the most popular variation of the MCSCF. It was introduced in the 1980s by Björn O. Roos and coworkers,[63] using as a central idea the previously devised procedure of *Direct Configuration Interaction* developed by the same author in order to provide efficient solutions for the CI problem.[67] The method performs an optimization of the multiconfigurational function obtained through a CI procedure over a certain set of MOs and an optimization over the MOs as well, with its corresponding unitary operator for the multiconfigurational function, $exp(\hat{S})$:

$$\hat{S} = \sum_{K \neq 0} S_{K0} \Big(|K\rangle\langle 0| - |0\rangle\langle K| \Big) \tag{3.9}$$

and a unitary operator which describes the orbital rotations, $exp(\hat{T})$:

$$\hat{T} = \sum_{p>q} T_{pq}(\hat{E}_{pq} - \hat{E}_{qp}) \tag{3.10}$$

A general variation of the wave function can then be written as:

$$|0'\rangle = exp(\hat{T})exp(\hat{S})|0\rangle \tag{3.11}$$

The transformed energy expression can be expressed as:

$$E' = \langle 0|exp(-\hat{S})exp(-\hat{T})\hat{H}exp(\hat{T})exp(\hat{S})|0\rangle \tag{3.12}$$

where the conditions to obtain a **stationary value of the energy** as a function of **S** and **T** would be given by:

$$\langle 0|[\hat{H}, \hat{S}]|0\rangle = 0 \tag{3.13a}$$

$$\langle 0|[\hat{H}, \hat{T}]|0\rangle = 0 \tag{3.13b}$$

Assuming the first-order derivatives of these expressions to be zero, a linear system of equations can be built and solved through several procedures detailed elsewhere,[68] optimizing both the multiconfigurational wave function and orbital parameters, related to the operators **S** and **T**, respectively, while obtaining the stationary value of the energy as stated above.

CASSCF is a versatile procedure that allows to generally treat all kinds of systems in a balanced manner yielding a correct description of the nature of the excited states, as well as open-shell systems. Within this method, three different sets of MOs are to be discerned: the inactive, the active , and the secondary spaces. While inactive and secondary MOs are fully occupied and unoccupied, respectively, a full list of CSFs with the required space and symmetry is constructed in what is known as the *Complete Active Space* (CAS) space. The definition of this space affects directly the results obtained. Therefore, the selection of the active MOs is a crucial aspect to consider in this method. The choice must be made depending on the properties or processes under study and must include all orbitals relevant to describe with enough

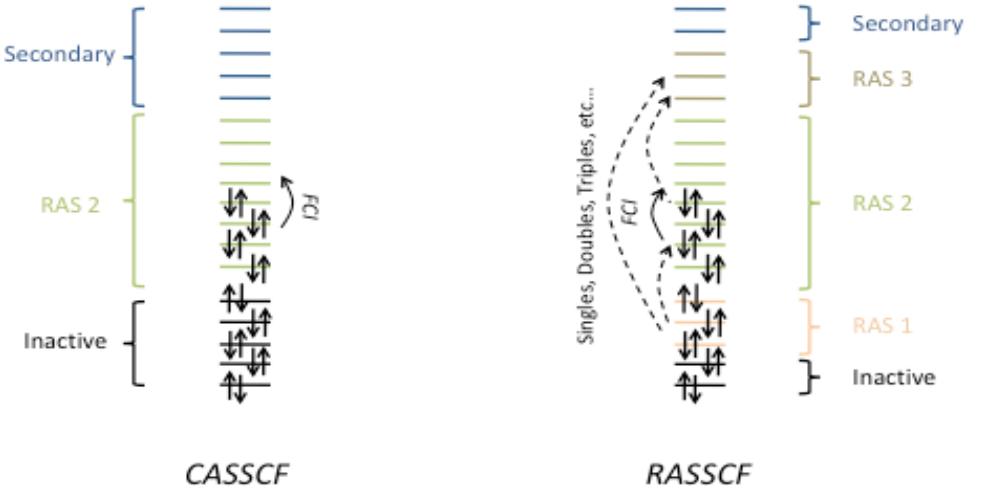

Figure 3.1: *CASSCF* and *RASSCF* schemes with their related possibilities of building CSFs within their different subspaces.

flexibility the problem at hand. Within the active space orbitals, a FCI procedure is performed. This improves the wave function by means of taking care of the non-dynamical correlation, which is a part of the electronic correlation related to near-degeneracies in the energy of different arrangements of electrons in the MOs. The main drawback of the method is the limited size of the active space that might be used, provided by the size of the complete CI expansion K given by the so-called *Weyl's Formula*:

$$K(n,N,S) = \frac{2S+1}{n+1}\binom{n+1}{\frac{1}{2}N-S}\binom{n+1}{\frac{1}{2}N+S+1} \tag{3.14}$$

where n is the number of MOs, N the number of electrons, and S is the spin quantum number. The current implementation in specialized quantum chemical codes, like *MOLCAS*,[69] allows active spaces of around 16 to 18 electrons distributed in a similar number of orbitals, even though such large active spaces might not be practical depending on the basis set employed and the total number of electrons of a given system. A way to partially solve this issue would be to use the *Restricted Active Space Self Consistent Field* (RASSCF) method, in which two extra sets of MOs are defined therefore

yielding: the inactive, the active space comprised in the $RAS1$, $RAS2$, and $RAS3$ subspaces, and the secondary orbitals. Within this method, $RAS2$ is a subspace in which the FCI procedure is performed as in the CASSCF, whereas $RAS1$ and $RAS3$ are orbitals in which a limited number of *holes* and *electrons* are allowed, respectively. The number of *holes* and *electrons* are subsequently related to the possibilities for building the excited spin-orbitals which give rise to the different CSFs that will define the reference wave function. A schematic representation of the different subspaces is given in Figure 3.1 comparing the CASSCF and RASSCF methodologies.

RASSCF can also be used to build upon a certain CASSCF calculation, so that extra orbitals might be correlated in order to refine the active space and thus yield a better description for a given system. The drawbacks from this methodology with respect to the CASSCF are a harder convergence, as the implementation of the optimization procedures for the several subspaces and its restrictions are rather challenging, and the lack of analytical state-averaged gradients, like those provided for the CASSCF,[70] which allow the study of the evolution of the excited states, one of the main features of these methods compared to their single reference counterparts.

3.2.2. Complete Active Space Second-Order Perturbation Theory (CASPT2)

Once a proper CASSCF reference that describes all the features present in the system under study with enough flexibility is obtained, perturbation theory can be used to add the lacking dynamical correlation on top of the CASSCF reference in what is called the *Complete Active Space Second-Order Perturbation Theory*. The method provides a second-order correction to the energy values of the states obtained at the CASSCF level. Such an approach has the virtue of being nearly size-extensive and ought to be computationally more efficient than the MRCI approach introduced above. Møller-Plesset perturbation theory is used on top of the CASSCF wave function in a similar manner

as it was done in the MP2 method, but for a few details. In this case, the CASSCF wave function has to be defined as an eigenfunction of the zeroth-order Hamiltonian to be used in the perturbation procedure. This can be done by defining the interacting space in the following manner:

$$\hat{E}_{pq}\hat{E}_{rs}|CASSCF\rangle \tag{3.15}$$

which is an internally contracted configuration space, doubly excited respect to the CASSCF reference $|CASSCF\rangle$. The functions are linear combinations that span the whole configuration space that interacts with the reference function. Labeling the compound index $pqrs$ as μ or ν, the first-order equation yields:

$$\sum_{\nu}[H^{(0)}_{\mu\nu} - E_0 S_{\mu\nu}]C_\nu = -V_{0\mu} \tag{3.16}$$

where $H^{(0)}_{\mu\nu}$ denotes the matrix elements of the zeroth-order Hamiltonian, chosen as a one-electron operator like in the MP2 method.[2] $S_{\mu\nu}$ is an overlap matrix, as the excited functions are not generally orthogonal to each other, and $V_{0\mu}$ represents the interaction between the excited function and $|CASSCF\rangle$. The main difference between the CASPT2 and MP2 formulations is the more complicated structure of the matrix elements of the zeroth-order Hamiltonian. Whereas in MP2 it is a simple sum of orbital energies, in CASPT2 consists of a Fock operator $\hat{\mathbf{F}}$ with up to fourth-order density matrices of $|CASSCF\rangle$.[64] The zeroth-order Hamiltonian is then written as a sum of projections of $\hat{\mathbf{F}}$ onto $|CASSCF\rangle$:

$$\hat{\mathbf{H}}_0 = \hat{\mathbf{P}}_{|CASSCF\rangle}\hat{\mathbf{F}}\hat{\mathbf{P}}_{|CASSCF\rangle} + \hat{\mathbf{P}}_{\hat{E}_{pq}\hat{E}_{rs}|CASSCF\rangle}\hat{\mathbf{F}}\hat{\mathbf{P}}_{\hat{E}_{pq}\hat{E}_{rs}|CASSCF\rangle} + \hat{\mathbf{P}}_X\hat{\mathbf{F}}\hat{\mathbf{P}}_X \tag{3.17}$$

where the first term represents the projection onto $|CASSCF\rangle$, the second over the internally contracted configuration space (eq. 3.15) and the third term

[2] This yields the main difference between the CASPT and the NEVPT methods, developed by Malrieu and coworkers, in which one-electron operators are used for the inactive and secondary spaces, whereas full two-electron Hamilton operators are used in the active space.[71]

over the remaining configuration space that does not interact with $|CASSCF\rangle$. $\hat{\mathbf{F}}$ represents the generalized Fock operator:

$$\hat{\mathbf{F}} = \sum_{p,q} f_{pq} \hat{\mathbf{E}}_{pq} \tag{3.18}$$

in which:

$$f_{pq} = h_{pq} + \sum_{r,s} D_{rs} \Big[(pq|rs) - \frac{1}{2}(pr|qs) \Big] \tag{3.19}$$

f_{pq} has the property that when the orbital p is doubly occupied (*inactive space*) $f_{pp} = -IP_p$ is the ionization potential and when its empty (*secondary space*) $f_{pp} = -EA_p$ yields an estimate of the electron affinity.[3] In *active orbitals*, the value is between those two extremes yielding $f_{pp} = -\frac{1}{2}(IP_p + EA_p)$ for an *occupation number* of one. In this framework, open-shell systems are favored leading towards an underestimation of binding energies in an homolitic dissociation. To solve this issue, a modified zeroth-order Hamiltonian can be used, which removes this systematic error. It can be built from an interpolation between the two extreme cases introduced above, giving rise to:

$$F_{pp} = -\frac{1}{2}\Big(D_{pp}(IP)_p + (2 - D_{pp})(EA)_p \Big) \tag{3.20}$$

where the formula provides the correct results for $D_{pp} = 0$ and 2, as well as for the singly occupied open-shell. Assuming now that upon an excitation into an active orbital, its energy is replaced by $-EA$, which can be achieved by shifting a value $\sigma_p^{(EA)}$ to eq. 3.20 where:

$$\sigma_p^{(EA)} = \frac{1}{2} D_{pp} \Big((IP)_p - (EA)_p \Big) \tag{3.21}$$

On the other hand, exciting out of this orbital would give an energy value of $-IP$, requiring a shift $\sigma_p^{(IP)}$ where:

$$\sigma_p^{(IP)} = -\frac{1}{2}(2 - D_{pp}) \Big((IP)_p - (EA)_p \Big) \tag{3.22}$$

[3] This stems out from *Koopman's Theorem*, which gives an estimate of the ionization potential depending on the orbital energies yielded by the Fock operator in Hartree-Fock theory, details are given elsewhere.[61]

As the definition of $(IP)_p$ and $(EA)_p$ is not straightforward, an average shift parameter ϵ is added to both shifting values yielding:

$$\sigma_p^{(EA)} = \frac{1}{2} D_{pp} \epsilon \tag{3.23a}$$

$$\sigma_p^{(IP)} = -\frac{1}{2}(2 - D_{pp})\epsilon \tag{3.23b}$$

This procedure is known as **the IPEA shift**, introduced by Ghigo *et al.* in 2004.[72] The value of the average shift parameter ϵ was shown to be optimal at 0.25 a.u., after calibration computations over 49 diatomic molecules.

Whereas the choice of the zeroth-order Hamiltonian might represent a possible problem, a bigger concern is the presence of **intruder states**, that consist of regions of the potential energy surface in which, due to some negative denominators often common for excited states, the energy falls into a singularity and the perturbation series diverge. A general and unbiased solution to solve this problem is the addition of a *level shift correction* giving rise to the *Level-Shift CASPT2* (LS-CASPT2) method:[73]

$$(\hat{H}_0 - E_0 + \epsilon)\bar{\psi}_1 = -(\hat{H}_1 - E_1)\psi_0 \tag{3.24a}$$

$$\bar{E}_2 = -\sum_{i=1}^{M} \frac{|\langle \Phi_i | \hat{H}_i \psi_0 \rangle|^2}{\epsilon_i - E_0 + \epsilon} = -\sum_{i=1}^{M} \frac{|V_i|^2}{\Delta_i + \epsilon} \tag{3.24b}$$

where the singularity comes from the Δ_i term as it vanishes and the shift parameter ϵ therefore prevents the divergence in the perturbation series. This problem arises in perturbation techniques such as the MP2 as well (see eq. 3.6), and is not exclusive of the CASPT2 method. Nowadays, the so-called *Imaginary Denominator Shift*, a technique based on the same grounds, is preferred as it smoothens the regions of the potential energy surface where the singularities would be found and can be used in a more general manner as compared to the regular shift.[74]

The last related technique is the *Multi-State CASPT2* (MS-CASPT2) method. Developed in the late 1990s by Roos and coworkers,[75] it represents an extension of the CASPT2 method for the perturbation treatment of situations

that require two or more reference states to give a proper description of a given problem, as compared to the standard CASPT2 where only an internally contracted CASSCF wave function $\hat{E}_{pq}\hat{E}_{rs}|CASSCF\rangle$ is used as a reference. Within this approach, an effective Hamiltonian matrix is built where the CASPT2 energies and the off-diagonal terms take into account couplings up to second order. Thereby, starting with a set of N orthogonal CASSCF wave functions $\Phi_i (i = 1, N)$, N single-state CASPT2 computations are performed. In order to build the matrix representation of the Hamiltonian with the first order corrections in the wave function, $\Psi_i = \Phi_i + \Psi_i^{(1)}$, the following matrices are defined:[76]

$$S_{ij} = \langle \Psi_i | \Psi_j \rangle = \langle \Phi_i + \Psi_i^{(1)} | \Phi_j + \Psi_j^{(1)} \rangle = \delta_{ij} + s_{ij} \tag{3.25a}$$

$$\langle \Phi_i | \hat{H} | \Phi_j \rangle = \delta_{ij} E_i \tag{3.25b}$$

$$\langle \Phi_i | \hat{H} | \Psi_j^{(1)} \rangle = e_{ij} \tag{3.25c}$$

Since $\langle \Phi_i | \Phi_j \rangle = \delta_{ij}$ and $\langle \Phi_i | \Psi_j^{(1)} \rangle = 0$, but $\langle \Psi_i^{(1)} | \Psi_j^{(1)} \rangle = s_{ij}$ we can conclude that the two CASPT2 related wave functions are not orthogonal. On the other hand, the CASSCF energies for the i state are represented by E_i whereas the e_{ij} elements represent the CASPT2 correlation energies. The Hamiltonian can then be expressed as the sum of a zeroth-order contribution and an extra Hamiltonian term taking care of the remaining effects given for each state:

$$\hat{H} = \hat{H}_i^0 + \hat{H}_i^{'} \tag{3.26}$$

The following expressions can be fulfilled up to second order:

$$\langle \Psi_i^{(1)} | \hat{H} | \Psi_j^{(1)} \rangle \approx \langle \Psi_i^{(1)} | \hat{H}_i^0 | \Psi_j^{(1)} \rangle \approx \langle \Psi_i^{(1)} | \hat{H}_j^0 | \Psi_j^{(1)} \rangle \tag{3.27}$$

The matrix representation of the Hamiltonian is not symmetric, being $H_{12} \neq H_{21}$. The matrix can be made symmetric by taking the average value of the off-diagonal terms represented in eq. 3.27, as they are expected to be similar:

$$\langle \Psi_i^{(1)} | \hat{H} | \Psi_j^{(1)} \rangle = \frac{1}{2} \big(\langle \Psi_i^{(1)} | \hat{H}_i^0 | \Psi_j^{(1)} \rangle + \langle \Psi_i^{(1)} | \hat{H}_j^0 | \Psi_j^{(1)} \rangle \big) \tag{3.28}$$

The matrix elements including terms up to second order take the general form:

$$\hat{H}_{ij} = \langle \Psi_i | \hat{H} | \Psi_j \rangle = \delta_{ij} E_i + \frac{1}{2}(e_{ij} + e_{ji}) + \frac{1}{2}(E_i^{(0)} + E_j^{(0)}) S_{ij} \tag{3.29}$$

Solving its corresponding secular equation, the MS-CASPT2 wave functions and energies can be obtained. The MS-CASPT2 wave function can be finally written as:

$$\Psi_p = \sum_i C_{pi} |i\rangle + \Psi_p^{(1)} = |i_p\rangle + \Psi_p^{(1)} \tag{3.30}$$

where $|i\rangle$ are the CASSCF reference functions and $\Psi_p^{(1)}$ is the first-order corrected wave function for state p. Therefore, $|i_p\rangle$ is formed by the linear combination of the CAS states involved in the MS-CASPT2, being the model state considered as the new reference function for state p. $|i_p\rangle$ are the so-called *perturbation modified* CAS functions, used in the computation of properties and expectation values at the MS-CASPT2 level.

For the proper use of the MS-CASPT2, eq. 3.27 has to be fulfilled, meaning that the asymmetric effective Hamiltonian matrix should have small and similar off-diagonal elements. Otherwise, the average process carried out in eq. 3.28, $(H_{12} + H_{21})/2$, may lead to unphysical results both in the MS-CASPT2 energies and eigenfunctions. The condition $H_{12} \cong H_{21}$ can be achieved by enlarging the active space, thus redefining the zeroth-order Hamiltonian. This method is also usually invoked to treat crossing points between *Potential Energy Surfaces* (PES) for small molecular systems, as it allows the treatment of two or more reference states as compared to the standard CASPT2, thus yielding the proper flexibility to describe such crossings. Large active spaces are recommended for the use of the MS-CASPT2, specially for the simultaneous treatment of valence and Rydberg states, where this method has proven to be very useful, as well as for crossing regions in which two states of different nature are treated simultaneously. Usually, one of the states has a covalent character whereas the other is zwitterionic.[4] The effect of dynamic correlation is usually more pronounced for zwitterionic states than for covalent states. Therefore,

[4]Those are described by hole-hole (covalent) and hole-pair (ionic) structures in *Valence Bond Theory*, described in depth elsewere.[2]

with a moderate (small) active space, the off-diagonal terms become very different as the covalent state is described differently than the zwitterionic state. Active spaces comprising MOs beyond the valence shell might be required to make $H_{12} \cong H_{21}$. In addition, the 2x2 effective Hamiltonian reads:

$$H^{eff} = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} \approx \begin{pmatrix} E_1^{PT2} & \Delta \\ \Delta & E_2^{PT2} \end{pmatrix} \tag{3.31}$$

where $E_1^{PT2} = E_1 + e_{11}$ and $E_2^{PT2} = E_2 + e_{22}$ are the CASPT2 energies in the diagonal of the matrix and $\Delta = (H_{12} + H_{21})/2$ are the averaged off-diagonal terms. If two states become degenerate at the CASPT2 level, $E_1^{PT2} = E_2^{PT2} = E$, the MS-CASPT2 energies and wave functions are:

$$E_{\pm} = E \pm \Delta \tag{3.32a}$$

$$\Psi_{\pm} = \frac{1}{\sqrt{2}}\left(\Psi_1 \pm \Psi_2\right) \tag{3.32b}$$

As the off-diagonal term vanishes, the MS-CASPT2 and CASPT2 solutions are equivalent, as expected in surface crossings. Therefore, by providing enough flexibility to the active space, one has to make sure that the condition in eq. 3.27 is fulfilled and that $\Delta \leq 2$ kcal/mol, **condition to define a crossing**. As a conclusion, in order to provide right estimates for surface crossings at the CASPT2 and MS-CASPT2 level, a larger active space has to be used as compared to the CASSCF.

CASPT2 is a powerful method that provides the right treatment for excited states at a moderate computational cost. The CASSCF/CASPT2 protocol is therefore one of the most common procedures to study the photophysics and photochemistry of a given system from a theoretical standpoint. Generally, the potential energy surfaces are mapped at the CASSCF level and the CASPT2 correction is applied on top of those geometries to provide the right estimate for the energies. Even though it is possible to map the surfaces directly at the CASPT2 level, analytical gradients are not available at the moment for the internally contracted form of the CASPT2, which is the one shown in this section, and such a task becomes cumbersome as it requires the use of numerical gradients. Analytical gradients are available for the uncontracted

CASPT2 in some quantum chemistry packages, but its computational cost is extremely high severely limiting its applications to real systems and reducing the size of the active space that shall be employed to around 8 electrons in as many orbitals.

3.3. Basis Sets

All *ab initio* methods stated above make use of one-electron basis set functions in order to form the molecular orbitals through a linear combination of such functions in what is known as *Molecular Orbitals as Linear Combination of Atomic Orbitals* (MO-LCAO). These basis set functions are atom-centered and span most part of the theoretically complete Hilbert Space, which implies an infinite number of functions. Since the basis sets are truncated and therefore do not span the whole space, larger basis sets will usually produce better results due to a better description and a higher flexibility, even though the magnitude of the property of interest (e.g., ionization potentials, electronic transitions, etc.) may rapidly converge leading to significantly good results with relatively less-demanding basis sets.

Two types of functions are mainly used to describe the one-electron basis functions in electronic structure calculations:

- *Slater Type Orbitals (STO)*
  which present an exponential dependence as the distance between nuclei and electrons vary. These are more efficient and accurate at representing the MOs.
- *Gaussian Type Orbitals (GTO)*
  which present an exponential dependence on the quadratic distance between nuclei and electrons. These functions provide a worse description of the MOs but at the same time present a better computational eval-

uation of the two-electron integrals and are, therefore, extensively used in electronic structure calculations.

The simplest basis set that can be built with such functions is known as the *Minimum Basis*, which implies the lowest number of functions to describe the occupied atomic orbitals. One could think of the oxygen atom as an example, where two s functions ($1s$ and $2s$) and three p functions ($2p_x, 2p_y$, and $2p_z$) are needed at least. These basis sets can be enhanced adding extra functions on each atomic orbital, as in the case of a *Double Zeta* (DZ) basis set, therefore needing four s functions ($1s, 1s', 2s, 2s'$) and six p functions ($2p_x, 2p'_x, 2p_y, 2p'_y, 2p_z, 2p'_z$) to describe the oxygen atom. Further additions can be made to obtain *Triple Zeta* (TZ), QZ, etc... increasing the flexibility progressively. This kind of procedure leads to large basis sets that will translate in higher computational demands. To prevent that, another type of basis sets was developed by increasing only the number of the valence orbitals, which are the most relevant in chemical problems, thus leaving the description of the core orbitals to its simplest form, meaning that the oxygen would have three s ($1s, 2s$, and $2s'$) and six p ($2p_x, 2p'_x, 2p_y, 2p'_y, 2p_z, 2p'_z$) functions for a *Valence Double Zeta* (VDZ) basis set.

Another way to lower the computational requirements is the formation of *Contracted Gaussian Type Orbitals* (CGTO), being built as fixed linear combinations of the so-called primitive functions. This procedure allows decreasing the number of variational parameters to be optimized, hence increasing its computational efficiency. There are two different types of contractions: the segmented and general contractions. In the first case, each primitive function contributes to the formation of just a given contracted function. The general scheme implies that all primitive functions of a certain atom and a certain angular momentum describe all contracted functions of that same momentum with different contractions coefficients, hence providing more flexibility as compared to the segmented procedure.

The basis sets used in this work are the *Pople-type*, *Dunning-type*, and the *Atomic Natural Orbital-type* (ANO) basis sets. The *Pople-type* basis sets have been developed by the Nobel laureate J. A. Pople and coworkers, as a sum of Gaussian-type orbitals to represent Slater-type atomic orbitals.[77, 78] They belong to the segmented scheme and provide a computationally cheap approach of a given problem by yielding qualitatively good results. The one used in this work is the $6{-}31G^{*}$, which is a valence double-zeta polarized basis set. The *Dunning-type* basis sets also belong to the segmented scheme and have been developed over the years by Dunning and coworkers.[79, 80] The basis sets are named *aug-cc-pVXZ* where *aug* represents the addition of diffuse functions, $cc{-}p$ that they are correlation-consistent and polarized, which means that the coefficients within the basis set have been optimized with post-Hartree-Fock computations, and VXZ represents that they are valence X-zeta basis sets. The *Atomic Natural Orbital-type*, on the other hand, belongs to the generally contracted scheme. Originally introduced by Taylor and coworkers, and later on improved by Roos and coworkers,[81–83] it provides a large and flexible basis set that allows precise calculations for the excited states, where the flexibility of the basis set becomes crucial. The ANO-type basis sets have been optimized through several SDCI calculations on different systems and computing many properties, to yield the proper parameters, and it is comprised of a rather large amount of primitive functions, which makes it a rather expensive basis set. Usually named *ANO-large* (ANO-L),[81–83] it can be simplified to obtain a smaller basis set that holds most of its properties and yields precise results, named *ANO-small* (ANO-S).[84] Within these basis sets, a certain contraction scheme has to be specified, in which an oxygen atom with a contraction of $2s1p$ would be the minimal basis, a double-zeta would be $3s2p$, and a double-zeta polarized would therefore be $3s2p1d$. ANO basis sets have been extensively used along the years for their correct characterization of the excited states and have also been intensively used in this work.

A way to improve the performance of large basis sets has been recently developed based on the *Cholesky Decomposition* technique.[85, 86] It allows reducing the number of primitive functions within a certain basis set and therefore accelerates the whole calculation eliminating or reducing as much as possible

the two-electron integral storage, thus yielding faster computations with a very similar level of accuracy. The process works very similarly to other more popular schemes, like the *Resolution of Identity* or *Density Fitting*, being equivalent in their treatment in some of the cases.[85, 86]

As has already been pointed out in this section, large and diffuse basis sets are needed to yield a proper description of the excited states. The description of the atoms might be enhanced because of the functions centered on the neighboring atoms, a phenomenon known as **Basis Set Superposition Error** (BSSE). In order to correct it, the *Counterpoise* correction has been used.[87] A deeper description of this technique is given in Appendix A.

3.4. Geometry Optimization Procedures

Geometry optimization is a key component of any computational chemistry study focused on the structure and/or reactivity of a given system. The algorithms are normally based on energy derivatives, and map the *Potential Energy Surface* (PES) of a given molecule along a certain region, depending on the state to be optimized, and the restrictions imposed within the optimization procedure. The notion of molecular structure and PESs are outcomes of the *Born-Oppenheimer approximation*, that allows separating the motion of the electrons from the motion of the nuclei. As the nuclei can therefore be considered as frozen, the electronic calculation on several different structures provides the description of the molecular energy as a parametric function of the position of the nuclei, thereby yielding a PES. Within the PES, several points of interest shall be discerned.

3.4.1. Minima on the Potential Energy Surface

Minima on the PES are defined as stationary points where the first derivative of the potential energy with respect to each geometry parameter vanishes and the second derivatives are greater than zero:[5]

$$\frac{\partial \mathbf{E}}{\partial \mathbf{q}_1} = \frac{\partial \mathbf{E}}{\partial \mathbf{q}_2} = \cdots = 0 \tag{3.33a}$$

$$\frac{\partial^2 \mathbf{E}}{\partial \mathbf{q}_1^2}, \frac{\partial^2 \mathbf{E}}{\partial \mathbf{q}_2^2}, \cdots > 0 \tag{3.33b}$$

These points are extremely important as they provide stationary structures where the molecule is most stable. A particularly important structure in photochemical studies is the so-called *Franck-Condon* structure (FC), which is defined as the minimum in the ground state from which the excitation process takes place. In order to compute these stationary points, a quadratic approximation to the PES is usually built as:

$$E(\mathbf{q}) = E(\mathbf{q}_0) + \mathbf{g}_0^T \Delta\mathbf{x} + \frac{1}{2}\Delta\mathbf{x}^T \mathbf{H}_0 \Delta\mathbf{x} \tag{3.34}$$

where $\mathbf{g}_0^T$ is the gradient ($\frac{\partial \mathbf{E}}{\partial \mathbf{x}}$) at x_0, $\mathbf{H}_0$ is the Hessian ($\frac{\partial^2 \mathbf{E}}{\partial \mathbf{x}^2}$) at x_0, and $\Delta\mathbf{x} = \mathbf{x} - \mathbf{x}_0$. The eigenvectors of the mass-weighted Hessian in Cartesian coordinates correspond to the normal modes of vibration.[88] Once obtaining the quadratic approximation to the PES, *Newton* and *quasi-Newton* methods to solve nonlinear equations can be used in order to provide estimates for the gradient and Hessian and subsequently solve the optimization procedure. At each step, Newton methods require the explicit calculation of the Hessian, which is quite expensive. Quasi-Newton methods, on the other hand, start with an inexpensive approximation of the Hessian and the difference between the calculated change predicted with the approximate Hessian is used to improve the Hessian at each step in the optimization:

$$\mathbf{H}^{new} = \mathbf{H}^{old} + \Delta\mathbf{H} \tag{3.35}$$

[5] On the other hand, maxima are defined as those stationary points where eq. 3.33a holds true and all the second derivatives in eq. 3.33b are lower than zero.

The algorithm mainly used in current quantum chemical codes, and the one mainly employed within this work, is the *Broyden-Fletcher-Goldfarb-Shanno update* (BFGS),[89] which uses a more complex expression:

$$\Delta\mathbf{H}^{BFGS} = \frac{\Delta\mathbf{g}\Delta\mathbf{g}^T}{\Delta\mathbf{g}^T\Delta\mathbf{x}} - \frac{\mathbf{H}^{old}\Delta\mathbf{x}\Delta\mathbf{x}^T\mathbf{H}^{old}}{\Delta\mathbf{x}^T\mathbf{H}^{old}\Delta\mathbf{x}} \tag{3.36}$$

but that represents the same type of quasi-Newton scheme, in which the Hessian is updated at each step and built upon the previous one.
Different methods and techniques are available for optimization, further details can be read elsewhere.[90, 91]

3.4.2. Saddle points on the Potential Energy Surface

First-order saddle points on the PES represent the *Transition States* (TS), structures between two different minima. Mathematically they differ from the minima in that even though they are both stationary points (eq. 3.33a), in a minimum $\frac{\partial^2\mathbf{E}}{\partial\mathbf{q}_i^2} > 0$ in all directions $\mathbf{q}_i$, whereas in a first-order saddle point $\frac{\partial^2\mathbf{E}}{\partial\mathbf{q}_i^2} < 0$ for the reaction coordinate $\mathbf{q}_i$ which ties two different minima and $\frac{\partial^2\mathbf{E}}{\partial\mathbf{q}_j^2} > 0$ in all other directions ($j \neq i$). For a TS, the PES is therefore a maximum in one direction and a minimum in all other perpendicular directions. Thereby, a TS is characterized by a zero gradient and a Hessian that has one negative eigenvalue yielding one imaginary vibrational frequency. The vibrational mode corresponding to the imaginary frequency is also known as the *transition vector*, which is tangent to the reaction path in mass-weighted coordinates yielded by techniques such as the *Intrinsic Reaction Coordinate* (IRC).[92]

3.4.3. Crossings on the Potential Energy Surface

Crossings on the PES represent *Conical Intersections* or *Singlet-Triplet Crossings* structures, considered in more detail in the next chapter. They arise from the crossing of two different PESs with the same spatial symmetry and equal or different spin multiplicity, giving rise to *non-adiabatic processes*, i.e., given in regions where the *Born-Oppenheimer Approximation* does not hold and a series of non-adiabatic couplings between the different PESs have to be taken into account. From a computational standpoint, the characterization of the crossings is complicated as it implies the correct description of the *seam*, which is a point or a series of points along the PESs characterized by $m-2$ degeneracies between the PESs involved in the crossing, where m denotes the internal degrees of freedom of the system ($m = 3N-6$ for nonlinear and $m = 3N-5$ for linear systems).

The easiest way to tackle the problem consists on computing the *Minimum Energy Conical Intersection* (MECI), as one can simply search for the lowest-energy crossing point fulfilling the non-adiabaticity rules. Calculation of non-adiabatic terms, the so-called *branching space*, is complex. Time-consuming projected-gradient [93] or *Barry's phase* algorithms have to be used to characterize the CI nature. In many cases though, it is safe enough just to compute the *Minimum Energy Crossing Point* (MECP), as the obtained solution already clarifies the presence of a crossing region accurately enough. Even when the MECI or MECP is found, nothing guarantees that this is going to be the relevant structure from a photochemical standpoint, which is what entails us in this work. It can happen that the MECI is placed in a region that cannot be easily accessed from the main decay paths of the molecule [6]. Therefore, the accessibility of the region becomes crucial for its photochemical behavior. In this sense, the first crossing region found along the *Minimum Energy Path (MEP)* and the MECP (or MECI) may diverge in structure, giving rise to two different regions where the PESs can cross. It can happen that the MECI (or MECP) stands lower in energy than the first accessible crossing found along the MEP, but the photochemical decay path is driven to the *easiest accessible*

crossing that might be found along the way. Therefore, it is the singularity relevant from a photochemical point of view.

3.4.4. Linear Interpolation of Internal Coordinates

Linear Interpolation of Internal Coordinates (LIIC) are predefined paths in which each structure is generated from the previous one by adding linearly to each internal coordinate the geometrical change of the initial and the final structures divided by the number of steps. This provides a connected path, but gives rise normally to rather high energy barriers, which can be considered as **upper bounds to the actual barrier**. It is a procedure that shall be avoided as much as possible in general treatments as it usually gives rise to large barriers that do not provide the right estimate. LIIC treatments are still preferred to more popular approaches like the *Relaxed Scans*, based on a series of constrained optimizations along a certain chosen reaction coordinate, which often produce unconnected and useless paths as there is no guarantee that the lowest-energy value for each of the selected frozen-coordinate structures belong to a connected path.

LIICs are usually performed between the molecular geometry at the FC region and a TS or minimum along the ground or the excited state, providing a PES that allows the characterization of the process along the path tying both structures.

3.4.5. Minimum Energy Path

Minimum Energy Paths (MEPs) follow the negative component of the gradient discussed before. When mass-weighted coordinates are used, the MEP coordinate corresponds to the IRC:

$$\mathbf{R}(q) = \left(\sqrt{m_1}\mathbf{r}_1(\mathbf{q}), \sqrt{m_2}\mathbf{r}_2(\mathbf{q}), \cdots, \sqrt{m_1}\mathbf{r}_1(\mathbf{q})\right) \tag{3.37}$$

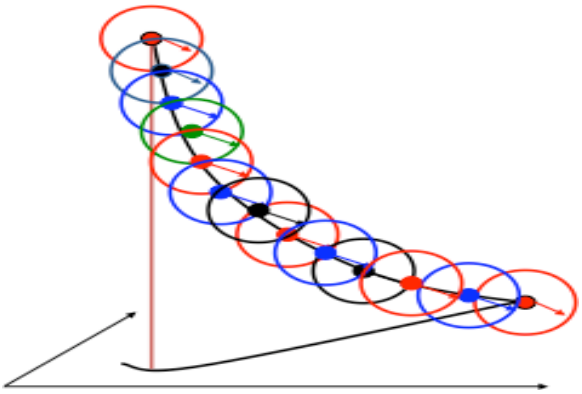

Figure 3.2: Scheme of the hyperspherical constrained optimization along the minimum energy path in its different converged points.

MEPs start either at the excited state energy of the FC geometry, thus simulating the absorption and subsequent evolution of the system, or at TS structures. The algorithms used in the MOLCAS package are taken,[94] as it is the main program used to compute MEPs.

The algorithm implemented in MOLCAS for the MEPs is based on the *Projected Constrained Optimization* method of Anglada and Bofill, described elsewhere.[95] It follows a *Steepest Descent Path* optimization based on the Müller-Brown approach [96] and performs a series of constrained geometry optimizations, each requiring the minimization of the potential energy on a hyperspherical cross section of the PES centered on a geometry and characterized by a given radius:

$$r_1 = \frac{\sqrt{(\mathbf{R}(\mathbf{q}) - \mathbf{R}(\mathbf{q}_{ref}))^2} - R}{\sqrt{M_{tot}}} \tag{3.38}$$

where $\mathbf{R}$ is the radius of the hypersphere and M_{tot} is the total mass of the system.[94]

A schematic picture of the process can be seen in Figure 3.2. Once the first structure within the first hyperspherical cross section is optimized, it is taken

as a reference and another hypersphere with a given radius is built in order to be optimized yielding so the next structure of the path until the bottom of the energy surface is reached. The algorithm therefore provides the descent trend till a barrier is found along the path, where it stops, and yields the perfect computational technique to study the evolution of a system upon light irradiation and its underlying photophysical or photochemical features, as it shall be covered in the next chapter. In the cases where there is no barrier along the descending path, the upper PES will cross with the lower PES, which usually appears in all sorts of photochemical schemes being tremendously important to rationalize the fate of the excited state and its subsequent behavior, as has already been noted above.

3.4.6. Differential Correlation

Another technical concern particularly important for this work arises from the CASPT2//CASSCF protocol. The cheaper CASSCF method is used to determine the geometries and subsequently CASPT2 calculations on top of those are carried out to correct the dynamic electron correlation contributions yielding the right estimates for the energies. This protocol is widely used in up-to-date theoretical photochemistry and presents a possible drawback which has to be addressed: what if both levels of theory do not provide the same topology and topography for the PES? This is due to the **differential correlation effects**, which may occur where the hypersurfaces estimated at the CASSCF and CASPT2 levels are not parallel. It will depend on the relative importance of non-dynamical and dynamical electron correlation contributions in the distinct regions of the PES, which normally can be related to the nature of the electronic state involved.

A schematic picture is shown in Figure 3.3. As can be seen, the dynamic and nondynamic correlation might differ in a certain state giving rise to a complete different topology of the PES. In such cases, the most correlated method, CASPT2 in this case, shall be used to map the PES. Otherwise,

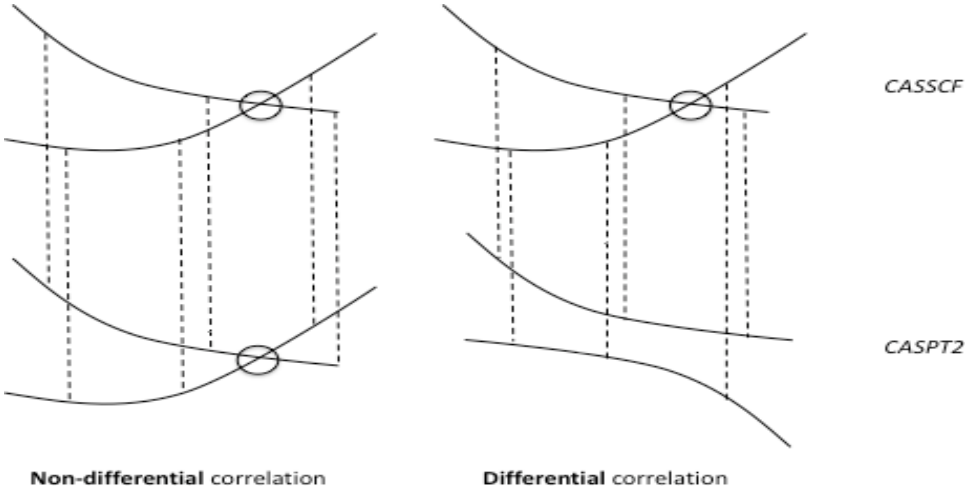

Figure 3.3: Different correlation effects on the CASPT2//CASSCF protocol.

the unbalanced contributions of the dynamical correlation may yield crossings that do not hold any physical sense, as the crossing might be placed in other geometry or not even exist.

3.5. *Ab initio* Molecular Dynamics

Ab initio Molecular Dynamics (AIMD) can then be used to compute an approximation of the time-dependent Schödinger equation to provide an estimate of the time evolution involved in a certain process. Molecular dynamics deals with the movement of atoms within a certain time range. AIMD performs a MD simulation based upon classical dynamics, but in this case the parameters of the equation are optimized at each step with an *ab initio* method, CASSCF in this case as one needs a proper description of the PES. The *velocity-Verlet* algorithm is then used to provide the quasi-classical trajectories in the follow-

ing manner:[97]

$$x(t+\tau) = x(t) + v(t)\tau + \frac{1}{2}\tau^2 \frac{f(t)}{m} \tag{3.39a}$$

$$v(t+\tau) = v(t) + \frac{f(t) + f(t+\tau)}{2m}\tau \tag{3.39b}$$

where τ is a small time increment, m is the mass of the system, and $f(t)$ is the force acting on the particle at a given time t, obtained computing its gradient. The requirement of the gradient in order to obtain the forces restricts the direct application of the CASPT2 method, as it becomes computationally too expensive. Within this scheme, one trajectory can be monitored and followed obtaining the related information. In order to give a good statistical description of the system, not one but a large amount of trajectories are to be computed and its mean value usually yields an appropriate answer for the problem at hand, increasing in accuracy as we increase the number of trajectories and therefore requiring a sampling method for the task. The sampling method used in this work is based on the Wigner distribution of the quantum harmonic oscilator [98] and implemented in the $NEWTON{-}X$ program package,[99] which makes use of the normal modes of the molecule previously computed. In order to allow the population transfer from one state to the other, a *surface hopping* scheme has to be applied. In this work, the *surface hopping* algorithm used is based on the approach widely applied by Olivucci and coworkers, in which the CSF coefficients describing the upper PES in each time step are compared to those of the lower PES (ground state) of the next time step, and vice versa. Large increases in their scalar product indicate a region featuring non-adiabatic couplings and hop probabilities. In practice, the hopping is performed when the scalar product between both CSF sets reach the 0.5 value.[100]

Chapter 4

Photophysics and Photochemistry

> Everything should be made as simple as possible, but not simpler.
>
> *Albert Einstein*

Photoinduced phenomena are of paramount importance in life. They rule the photophysical and photochemical processes, which are those derived from the absorpion of *photos*, which means light in greek. After absorption, a richness of different processes can be described, which are ascribed to either the photophysical or photochemical realms. In this chapter a brief introduction about these processes is given.

4.1. Theoretical Spectroscopy

Spectroscopy deals with the interaction between electromagnetic radiation and matter. Electron Spectroscopy is a very useful technique vastly used in physics and chemistry.[1, 2] It studies the promotion of electrons from one energy level to another due to their interaction with an incoming radiation with a given frequency. When the electron goes from a higher energy level to a lower one, the exceeding energy is emitted at a frequency determined by the same difference in energy described by the aforementioned energy levels. This is the so-called *Bohr resonance condition* and it is expressed as follows:

$$\Delta E = h\nu \tag{4.1}$$

The lowest energy level of a given molecule corresponds to its most stable arrangement and it is called the **ground state**. The higher states given above it are known as **excited states**, which are known to release their excess energy through three mechanisms: *radiative transition*, which emits the excess energy in the form of radiation, *nonradiative transition*, in which the energy is dissipated as heat along the states of the molecule or transferred from a molecule to another, and a *chemical reaction*.

A semiclassical approach is employed to tackle the interaction between radiation and matter, where matter is considered as a quantized system with its rotational, vibrational, and electronic levels, while the electromagnetic radiation behaves as a classical wave. Every electronic state has several vibrational states close in energy associated to interatomic vibrations, and every vibrational state has rotational ones even closer in energy provided by the rotational motions of the molecules around their mass centers. Upon light irradiation, electromagnetic radiation interacts with matter by distributing a certain amount of energy quanta among the different degrees of freedom or energy levels available in the molecule. From a photochemical point of view, the states of interest are the vibrational and electronic ones, as they are related to the main spectroscopic features given in most biological and chemical processes.

Prior to light absorption, all molecules are in their respective ground states. However, there is a vibrational threshold energy known as *Zero Point Energy* (ZPE), which is associated to the minimum energy vibration level accessible of a molecule. Higher-lying vibrational states can be accessed through thermal population or electromagnetic radiation.

The ground state absorbs the incoming irradiation, whose energy is used to either change the charge distribution of the molecule or relocate some of its electrons. The molecule is thus polarized due to an arisen electric field that yields a *Transition Dipole Moment* (TDM):[2]

$$TDM_{n\to m} = \sum_{\alpha=x,y,z} \langle\psi_m|\hat{\mu}_\alpha|\psi_n\rangle \tag{4.2}$$

in which $\hat{\mu}_\alpha$ is the dipole moment operator for component α. The square of the transition dipole moment estimates the probability of the transition $n \to m$, and thus its band intensity through *Fermi's Golden Rule*. When the dipole is zero, the one-photon transition is forbidden. The selection rules arise from Group Theory. A transition is allowed if and only if the direct product of the irreducible representations of the components, $\Gamma(\psi_m)\otimes\Gamma(\hat{\mu}_\alpha)\otimes\Gamma(\psi_n)$, is equal to the totally symmetric irreducible representation of the point group where the molecule belongs to.

When a transition is the most likely to take place, the intensity reaches its maximum value. The **oscillator strength** (f) provides an estimate of the intensity of the spectral band. It represents the probability of a system to access from one state to a higher-lying excited state due to the absorption of one photon. It is related to the transition dipole moment as shown below:

$$f_{n\to m} = \frac{2}{3}\Delta E_{nm}|TDM_{n\to m}|^2 \tag{4.3}$$

where n represents the initial and m the final state, expressing ΔE and $TDM_{n\to m}$ in atomic units. According to the **Franck-Condon** (FC) principle, the transition most likely to take place corresponds to the one without relative changes in the position of the nuclei of the molecule, as the absorption

of a photon is a very fast process compared to nuclear motion. Based on this principle, the most stable ground state geometry shall be labeled the *Franck-Condon region* (FC) as it will be the geometry in which the system initially absorbs light.

4.2. Potential Energy Landscapes

The excited states populated upon irradiation have a limited lifetime, as the molecule is unstable as compared to its ground state. This drives the molecule towards the search of relaxation channels in order to be deactivated. Energy relaxation occurs mediated by several processes that differ in both their timescale and their probability. The absorption is an ultrafast process in the femtosecond timescale, which implies electron motion as there is no time for nuclear reorganization. After absorption between states of the same spin multiplicity, most of the molecules are in a vibrational excited state (or vibrationally hot state). The first potential step undertaken by the molecule is the relaxation of its vibrational energy back to the ZPE, which is the lowest-lying vibrational level available in any given electronic state. The vibrational decay process is the *Intramolecular Vibrational Relaxation* (IVR), where the energy disipates through vibrational motions in the molecule. From this point the molecule goes through a minimum energy path from the initial structure in the FC region towards the minimum of the excited state. This path can be crossed by potential energy surfaces with the same (*Internal Conversion*, IC) or different spin multiplicity (*Intersystem Crossing*, ISC). In the former, the existence of an effective crossing between the surfaces is accounted for in the form of either *Conical Intersections* (CI) or *Avoided Crossings* (AC). When the molecule reaches the bottom of the potential energy surface and is trapped in the minimum of the excited state, another process takes place in the form of *radiative emission* to other lower-lying states. The emission can be produced from states with equal or different spin multiplicity yielding the *fluorescence* and *phosphorescence*, respectively. Ionization processes are also possible and direct access from the ground state to a doublet state (cation

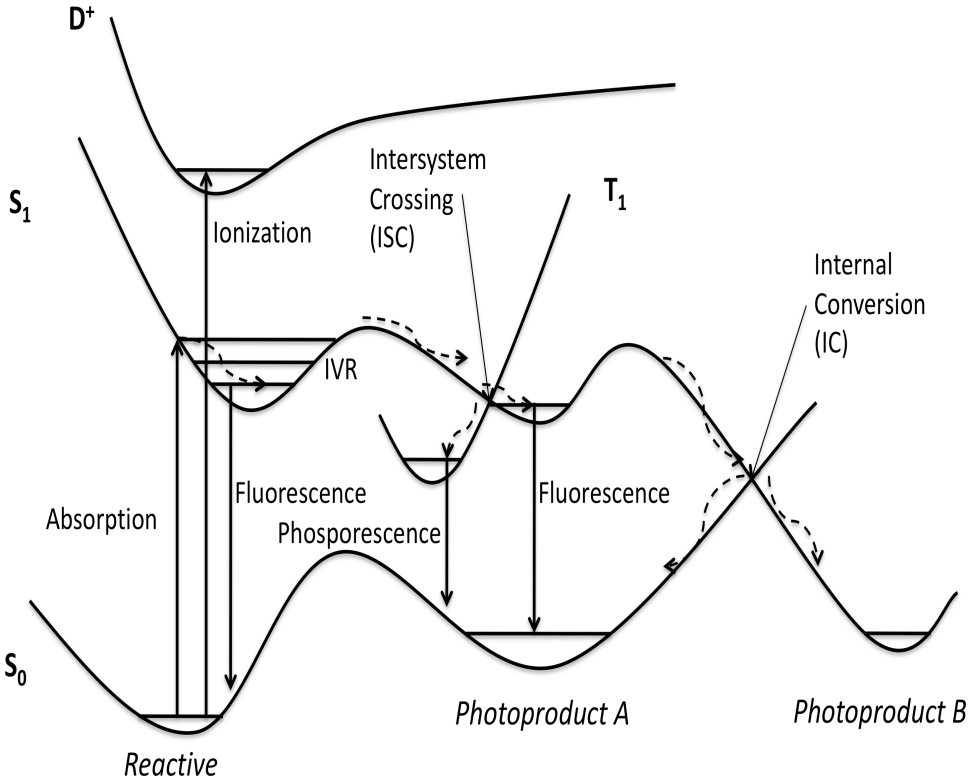

Figure 4.1: Main intramolecular photochemical pathways.

or anion) is also possible, giving rise to several properties which have been covered in detail in chapter 7. A schematic view of all processes described can be seen in Figure 4.1. *Kasha's rule* states that emission occur from the lowest-lying energetically-wise excited state to the ground state. The emitted radiation has a longer wavelength than the incident radiation on the molecule upon absorption. This phenomenon is coined as *Stoke's shift* and it is due to the internal relaxation exerted in the excited state that drives nuclear motion towards the minimum of the excited state,[101] where the difference in energy between the excited and the ground states is smaller. The lifetimes of emissive processes are relatively longer than those found for nonradiative processes such as surface crossings. Electronic transitions between states of different multi-

plicity are forbidden, unless the *spin-orbit coupling* (SOC) propels the mixture between singlet and triplet states.

The ratio between the absorbed and emitted radiation is known as *quantum yield* (Φ). Hence, systems that emit eminently hold a quantum yield value close to one, whereas low values of this magnitude represent the presence of nonradiative decay channels or the presence of quenchers. Nonradiative phenomena occur when a funnel is found between potential energy surfaces of the same spin multiplicity (IC) or different (ISC), through a hyperline or a hyperplane, respectively. These processes are faster as the states are nearly degenerate in energy.

4.2.1. Photophysical and Photochemical Processes

Relaxation processes from the excited state might yield the so-called **photochemical reactions**, which are those in which a new ground-state minimum is reached by the system after deactivation. Adiabatic and non-adiabatic photochemical processes are likely to occur through relaxation routes. Adiabatic photochemistry is characterized by a reaction taking place in one PES, whereas non-adiabatic photochemistry implies the formation of a photoproduct after population transfer between two PESs through a crossing. Figure 4.2 depicts both types of photochemical reactions, as well as the photophysical ones. Upon energy absorption, the reactive R is promoted to an electronic excited state R^*. The reaction $R^* \rightarrow P$ in the right-hand represents an adiabatic process leading to an emitting feature, whereas the left side of the figure depicts a non-adiabatic process, since there is a surface crossing mediating a nonradiative decay back to the ground state. **Photophysical processes** are concerned in the central part of the figure. The emission brings the molecule to the initial structure, thus no chemical change is given. Modern photochemistry relies on the efficiency of nonradiative decays between different electronic states involved in IC or ISC processes, associated to the presence of CIs and crossing hyperplanes that behave as funnels where the probability for nonadiabatic

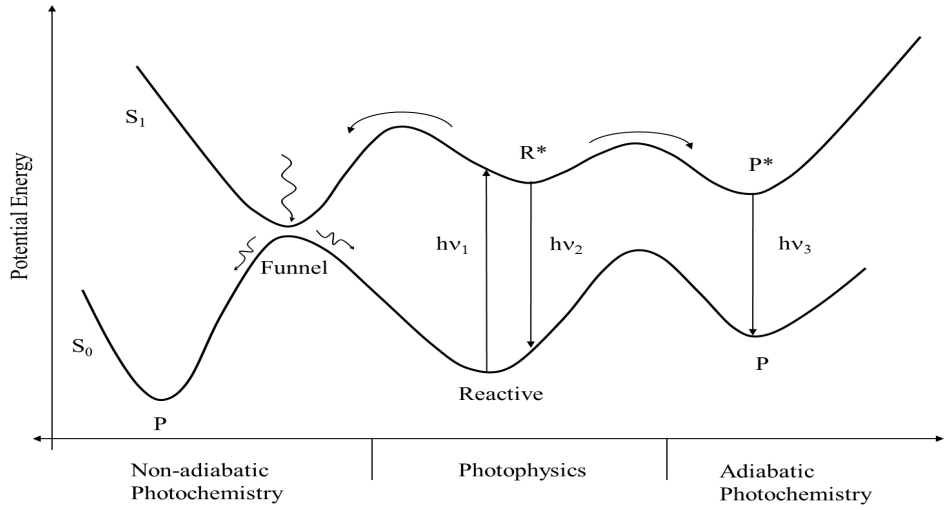

Figure 4.2: Photophysical and photochemical processes.

jumps is high. A crossing seam occurs between two states of the same spin multiplicity when their PESs intersect along a (F-2)-dimensional hyperline as the energy is plotted against the F nuclear coordinates, where F denotes the internal degrees of freedom (3N-6, 3N-5 if linear). A summary can be seen in Figure 4.3. In any point of the (F-2)-dimensional intersection space the energies of the two states are the same. The degeneracy is lifted along the two remaining linearly independent coordinates, x_1 and x_2, wich span the branching subspace corresponding to the gradient difference vector and the non-adiabatic coupling vector, respectively:[2]

$$x_1 = \frac{\partial(E_1 - E_2)}{\partial Q} \tag{4.4a}$$

$$x_2 = \langle \psi_1 | \frac{\partial \psi_2}{\partial Q} \rangle \tag{4.4b}$$

State crossings of different spin multiplicity yield a (F-1)-dimensional set of points in which intersystem crossing processes may occur funneling the population to a state of different multiplicity. The non-adiabatic coupling vector

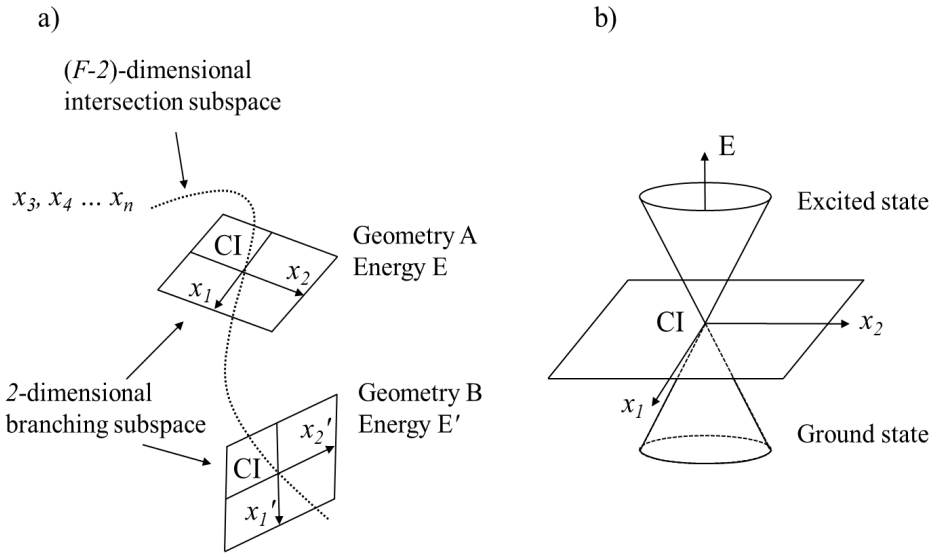

Figure 4.3: Hyperline resulting from the intersection of two PESs (a) and the energy representation of the excited and ground states against the x_1 and x_2 coordinates of the branching space (b).

vanishes here, therefore one is dealing with the one-dimensional branching space.

4.2.2. Spin-Orbit Coupling

Spin-orbit coupling (SOC) terms in the non-relativistic Hamiltonian account for the interaction between the magnetic spin momentum of the electron ($\overrightarrow{S}$) and the magnetic momentum due to the electronic orbital movement ($\overrightarrow{L}$). The spin selection rules are a consequence of the electric dipole moment operator which does not depend on the spin, hence allowing only transitions of the same mulitplicity. Singlet-triplet absorptions are hardly observable in the spectra of

any molecule aside from the heavy atoms. The spin-orbit Hamiltonian reads:

$$\hat{H}_{SO} = \frac{e^2}{2m_e^2c^2}\sum_j\sum_A \frac{Z_A}{|\mathbf{r_j^A}|^3}\mathbf{l_j^A}\cdot\hat{\mathbf{s}}_j \tag{4.5}$$

where $\mathbf{l_j^A}$ is the orbital angular momentum operator of electron j and $\mathbf{r_j^A}$ is the vector that points from nucleus A to electron j.[2] Eq. 4.5 shows the dependence on the atomic number Z_A, which explains why these processes are more probable in heavy atoms. The $\hat{H}_{SO}$ represents a small term inside the electronic Hamiltonian where singlet excited states will have a low contribution to the description of triplet states. These slight differences make the singlet-triplet momentum different from zero and enhance the possibility of characterizing transitions $S_i \rightarrow T_f$. The length of the SOC vector yields:

$$SOC_{fi} = \sqrt{\sum_u |\langle T_{f,u}|\hat{H}_{SO}|S_i\rangle|^2} \tag{4.6}$$

where $u = x, y, z$. Spin-orbit couplings provide the means to evaluate the efficiency of an intersystem crossing process. Hence, exceedingly small SOC values will yield low ISC rates. On the other hand, taking into account that the TDM describing the transition is dependent on the energy difference between states:

$$\begin{aligned} TDM_l(T^k) = \langle S|\hat{r}^l|T^k\rangle = \sum_n \frac{\langle S^0|\hat{r}^l|S_n^0\rangle\langle S_n^0|\hat{H}_{SO}^k|T^{k,0}\rangle}{E(T^0)-E(T_m^0)} \\ + \sum_m \frac{\langle S^0|\hat{H}_{SO}^k|T_m^{k,0}\rangle\langle T_m^{k,0}|\hat{r}^l|T^{k,0}\rangle}{E(S^0)-E(T_m^0)} \end{aligned} \tag{4.7}$$

where k labels the magnetic spin sublevels, $l \in (x, y, z)$. $T^{k,0}$ and T^k denote the zeroth- and first-order wave functions of perturbation theory, respectively, and $\hat{H}_{SO}^k$ is the kth component of the SOC operator. Thus, it can be clearly seen that an exceedingly small difference in energy favors ISC processes.

Part II

Results & Discussion

Chapter 5

First steps into the Photochemistry of Water: The Water Dimer

> Water, the Hub of Life. Water is its mater and matrix, mother and medium. Water is the most extraordinary substance! Practically all its properties are anomalous, which enabled life to use it as building material for its machinery. Life is water dancing to the tune of solids.
>
> *Albert Szent-Gyorgyi*

Water has very unique and anomalous properties, mainly arisen due to its structural motifs.

As mentioned above, the structure governing the liquid state of water is far from being understood and therefore all avenues of inquiry shall be pursued in order to properly characterize the liquid and its underlying photophysical and photochemical mechanisms. As a way to tackle this problem, the use of the water dimer as a model system has been proposed. This is due to the fact that, for most of the researchers working in the field, it constitutes the smallest building block of liquid water and can be used to represent both types of intermolecular interactions that can be found in water: i) through hydrogen bonding, giving rise to the *conventional water dimer*, a water cluster appearing in the gas phase characterized in the late seventies by Dyke and coworkers,[102] or ii) through non-covalent interactions giving rise to the π-*stacked water dimer*, that shall be used in order to rationalize the energetically low-lying spectroscopic fingerprint found in ordered water. Both approaches are compared and a number of conclusions have been attained, propelling forward the knowledge of both intermolecular interactions and providing clues as to how water molecules might be actually organized based on the spectroscopic fingerprints of its smallest subunits.

The Conventional Water Dimer

In the present Thesis, the gas-phase water dimer has been extensively studied regarding its hydrogen bonded character as well as its photochemical decay paths upon light absorption. The water dimer was characterized by Dycke and co-workers,[102] and ever since it has been postulated to play a crucial role in a plethora of processes, including ice nucleation,[103] and radiation filtering in the atmosphere,[104–107] among others.

Computational Details A well-established quantum-chemical *ab initio* method, namely the CASPT2//CASSCF, has been used for this study as implemented in the MOLCAS-7 software package.[69] The basis set of Atomic Natural Orbital (ANO) type with the primitive set O(14s9p4d3f)/H(8s4p3d),

the ANO-L set, contracted to O[5s4p1d]/H[2s1p] has been used throughout. It represents a quadruple- and double-ζ plus polarization scheme for the oxygen and hydrogen atoms, respectively. The choice has been based on the results of the previous calibration calculations reported earlier on the study of the valence and Rydberg character of the excited states of the water molecule.[27] Multiconfigurational wave functions have been initially determined at the CASSCF level. In all the calculations, the oxygen 1s electrons have been treated as inactive and all the valence orbitals have been included in the active space. The full valence space comprises eight electrons and six orbitals for each water molecule considered. The dimer has therefore been described by the CASSCF(16,12) wave function, employing five roots in the averaging procedure. It represents for two water molecules in a singlet state 70785 configuration state functions with no spatial symmetry constraints. Since wave function symmetry breaking is a prerequisite to describe correctly the asymptotic limit for the lowest singlet electronic state of the moieties considered, no symmetry constraints have been imposed (C_1 symmetry), even though the global molecular system might display a given spatial symmetry. On the other hand, the corresponding results including the second-order corrections shall be labeled as CASPT2(16,12). The oxygen 1s electrons have been kept frozen in the perturbation step. In order to minimize weakly interacting intruder states, the imaginary level-shift technique, with a value of 0.2 au, has been employed. Optimized structures and MEPs have then been obtained at the CASSCF level, adding the necessary dynamic correlation through the CASPT2 energy calculations on top of the geometries obtained, with the zeroth-order Hamiltonian as originally implemented in MOLCAS.[65] MEPs have been extensively used in this work, making use of mass weighted coordinates. An estimate of a given photochemically relevant conical intersection (CI) has been performed by determining the Crossing Point along the corresponding MEP.

Born-Oppenheimer molecular dynamics calculations have been performed on the excited-state surface of the water dimer with the MOLCAS-7 program. Newton equations have been solved by the velocity-Verlet algorithm with a 1 fs time step.[97] Analytical gradients have been computed on-the-fly at the same level of theory employed in the calculations of the electronic struc-

ture, that is, at the CASSCF(16,12)/ANO-L 541/21 level of computation and no symmetry constraints. A set of 50 initial conditions have been generated with the NEWTON-X program,[99] according to the Wigner distribution for the quantum harmonic oscillator,[98] and using the normal modes computed at the MP2/aug-cc-pVDZ level of theory with the GAUSSIAN-09 software package.[108] The surface hopping algorithm implemented in the MOLCAS package [100] has been used to enable the non-adiabatic process from the excited state to the ground state. All trajectories have been run until the complete dissociation of the OH bond. At the geometries obtained in the simulations, the energies have been subsequently computed at the CASPT2(16,12)/ANO-L 541/21 level of theory.

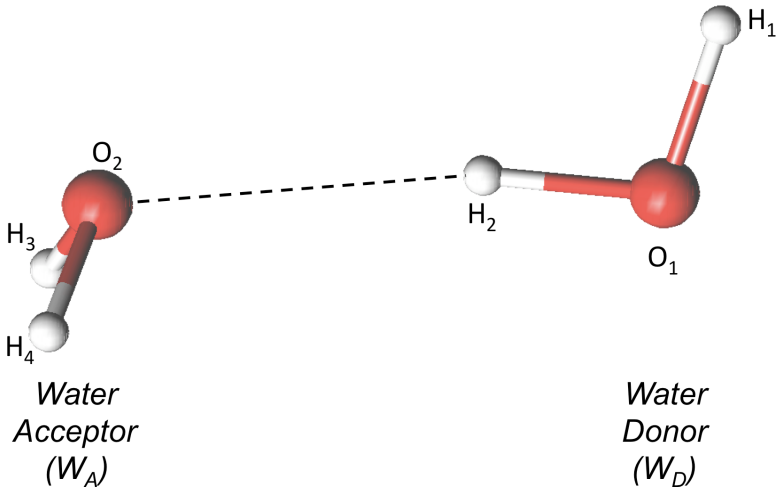

Figure 5.1: Gas-phase water dimer structure and labeling.

Results and Discussion The water dimer structure is depicted in Figure 5.1. Both water molecules are tied by a hydrogen bond taking place between the H_2 and O_2 atoms. The presence of the hydrogen bond alters the photophysical and photochemical properties of the water molecule, giving rise to slight differences between the dimer and the monomer. In this sense, two different types of water can be found in the dimer: one that *gives* the hydrogen

bond (water donor, W_D), and one that *receives* it (water acceptor, W_A). The effect of the hydrogen bond on the absorption energy appears to be rather small, with a minor red-shift in the W_D and a relative blue shift in the W_A as compared to its monomer counterpart.[27] The results here obtained have been assessed comparing with other high level quantum chemical *ab initio* methods as well as with the experimental data given for the monomer,[22] yielding comparable results.

Table 5.1: Vertical electronic transitions (ΔE; in eV) and oscillator strength in the singlet manifold of the water dimer.

		CASPT2(16,12)[a]		EOM-CCSD[b]		Exp[c]
State		ΔE	f	ΔE	f	ΔE
S_1	$lp_D\sigma^*_D$	7.36	0.018	7.58	0.052	7.447
S_2	$lp_A\sigma^*_A$	7.78	0.061	8.02	0.061	
S_3	$lp_D\sigma^*_A$	8.96	0.00005	8.93	0.000	
S_4	$\sigma_D\sigma^*_A$	9.46	0.045	-	-	

[a]Reference [109]
[b]Reference [110]
[c]Reference [22]

The ground-state equilibrium structure of the water dimer has been optimized at the CASSCF(16,12)/ANO-L 541/21 level. At this geometry, the absorption spectrum has been determined with the CASPT2 method. Table 5.1 summarizes the energies of the lowest-lying singlet-singlet transitions (ΔE), the oscillator strength (f), and the nature of the excited state. For the sake of comparison, the experimental value obtained by Mota *et al.*[22] for the maximum of the lowest absorption band of the water molecule, as well as the computed EOM-CCSD/aug-cc-pVTZ results for the dimer reported by Cabral do Couto *et al.*,[110] have also been included. As can be seen by inspection of Table 5.1, the CASPT2//CASSCF approach gives rise to absorption energies in agreement with the EOM-CCSD findings, although the ground-state CASSCF optimized geometry overestimates the hydrogen-bond distances due to the lack of dynamical correlation.

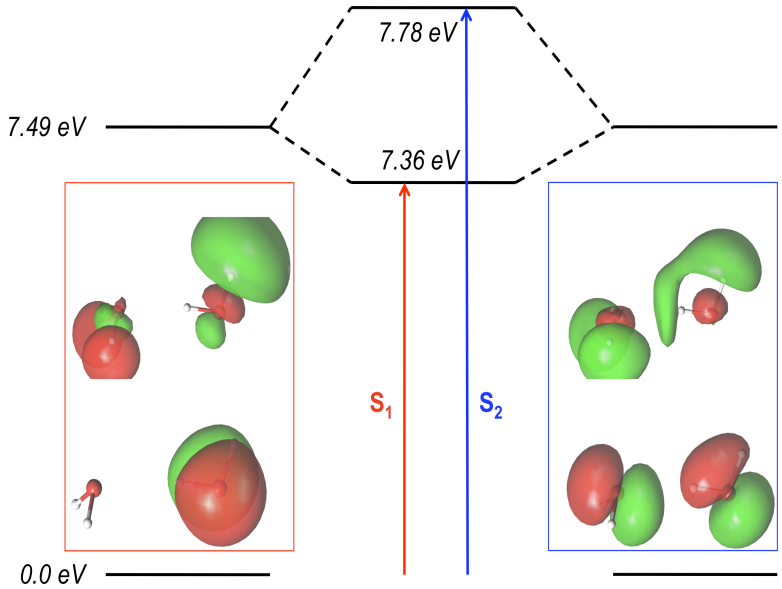

Figure 5.2: Diagram of the lowest-lying excited states of the water dimer. Red lines denote the S_1 excited state and blue lines the S_2 excited state, which feature $lp_D\sigma^*_D$ and $lp_A\sigma^*_A$ nature, respectively (see text). The singly occupied CASSCF Natural Orbitals at the Franck-Condon region are depicted for S_1 and S_2. The lowest vertical transition energies are also included.

The excitations on the water dimer are described by the electronic promotion of the *highest occupied molecular orbital*, (HOMO)-type orbital (lp), to the *lowest unoccupied molecular orbital*, (LUMO)-type orbital, localized on each water molecule (σ^*). The orbitals can be seen in Figure 5.2, together with the estimates of the energies obtained for the W_D (red) and W_A (blue), and the comparison to the theoretical value obtained at the same level of theory for the monomer.[27] These orbitals correspond to the $1b_1$ and $4a_1$ irreducible representations of the C_{2v} symmetry point group and are involved in the description of the lowest-energy excited state of valence nature in the water molecule. We can therefore conclude that the effect of the hydrogen bond in the absorption spectra of the dimer is rather small, and the essentials provided by the

monomer are maintained by the union of both monomeric units through the hydrogen bond.

In particular, S_1 is described by the one-electron promotion corresponding to the $lp_D\sigma^*_D$ electronic transition of the W_D water, whereas S_2 has the same characteristics centered in the W_A moiety ($lp_A\sigma^*_A$). At higher energies, S_3 shows charge transfer character between the W_D and W_A monomers ($lp_D\sigma^*_A$). The oscillator strength (f) of the $S_0 \rightarrow S_3$ transition is small, which points out to a low probability for its population during light irradiation. The remaining state computed, S_4, appears 2 eV above the lowest-lying excited state and has also a charge transfer nature involving sigma orbitals. Regarding the ordering of the states, two different scenarios can be found in the literature: i) the lowest-lying excited state has charge transfer nature [111–115] or ii) it is a locally excited state in one of the water molecules.[110] According to the present analysis, the highest-level quantum chemical methods, that is, the CASPT2//CASSCF and EOM-CCSD,[110] agree with the ordering and the nature of the states (see Table 5.1), thus highlighting the misleading results of a charge-transfer state as the lowest-lying excited state in the water dimer obtained by several authors in the literature.[111–115]

The lowest vertical absorption energy of the water dimer shows a small red-shift of around a tenth of an eV relative to the water monomer. This has also been previously noted both theoretically and experimentally.[110, 116, 117] A larger blue-shift of around 0.3 eV is found for the state related to the excitation localized on the W_A molecule (S_2), in agreement with previous theoretical studies.[115, 116] In the light of the present analysis, we can predict that $lp_D\sigma^*_D$ and $lp_A\sigma^*_A$ are the states which will be populated during the light irradiation of the water dimer. Therefore, the photochemistry of the system will be mainly driven by these excited states.

Once absorbing the light, the excess energy is relaxed via decay paths towards the photodissociation features given also for the water monomer.[118] The evolution of both excited states yield dissociative PEHs, which is reasonable

considering the fact that the states imply an electron promotion from a lp molecular orbital to the antibonding σ^* orbital of the OH bond of the same water molecule that is not involved in the hydrogen-bonding network. This result is in contrast to some previous theoretical studies,[111] and supports other works in which the photodissociative character of the lowest-lying states is shown.[119, 120] Regarding the nature of the dissociation process of the water dimer, it is also worth mentioning that the data obtained here must not be confounded with the dissociation energies of the water dimer reported in other theoretical and experimental studies,[121–123] as it refers to different magnitudes. Whereas those studies elucidate the thermal energy needed to dissociate the water dimer in two separate water molecules, breaking the hydrogen bond present between the H_2-O_2 atoms, our results emphasize on the localized excitation on the water molecules of the dimer that undergo photodissociation processes analogous to those already observed for the water monomer. The dissociations occur in this case at the O_1-H_1 and either the O_2-H_3 or O_2-H_4 bonds for the W_D and the W_A, respectively.

The excitation on the two lowest-lying excited states are centered in the W_D and W_A and yield dissociative pathways, meaning they do not show a minimum, therefore leading towards the photodissociation of the *dangling* hydrogen atom (the one that does not participate in the hydrogen bonding) producing the hydrogen-bonded $H_2O\cdots HO^{\cdot}$ species plus a hydrogen ($H^{\cdot}$) atom.

Figure 5.3 depicts the evolution of the S_1 state, $lp_D\sigma^*_D$, along its MEP, starting at 7.36 eV as shown in Table 5.1 and leading to a conical intersection (CI), namely $(lp_D\sigma^*_D/S_0)_{CI}$, and its dissociation at 5.15 eV. This value is in agreement with the experimental value found for the water monomer reported in Herzberg,[118] at 5.11 eV, thus highlighting the photochemistry of the water dimer as indeed a combination of two water molecules where the hydrogen bond only plays a minor role.

To our knowledge, the evolution of the second excited state ($lp_A\sigma^*_A$) of the water dimer after light irradiation has not been so far investigated. This state

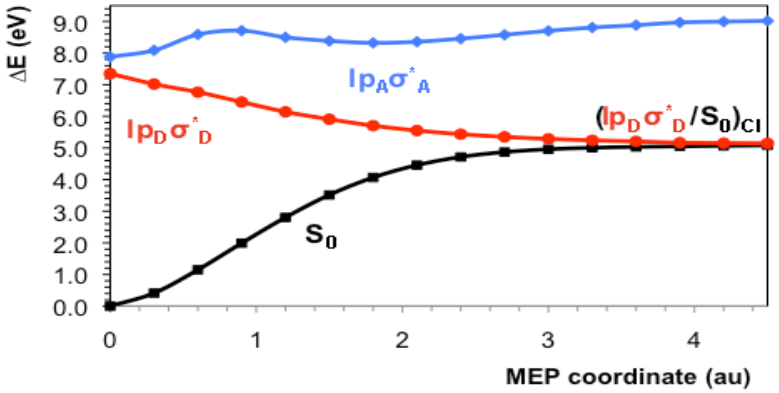

Figure 5.3: CASPT2 energies of the ground (S_0) and lowest-lying singlet excited states ($lp_D\sigma^*_D$ and $lp_A\sigma^*_A$) of the water dimer from the FC geometry and along the S_1 MEP.

requires an initial excitation (7.78 eV) slightly higher than ($lp_D\sigma^*_D$) and might exhibit a higher population compared to the latter, as its oscillator strength is around three times larger (see Table 5.1). Therefore, the contribution of $lp_A\sigma^*_A$ in the photochemistry of the system cannot be discarded. Following Kashas rule, excited states higher than S_1 are expected to relax in an ultrafast fashion to the lowest-lying excited state. In order to verify it and determine the photochemical evolution of the state localized in the W_A water monomer, we have performed a MEP computation on the S_2 (see Figure 5.4).

At the CASPT2 level of theory, the CI between both states, namely the $(lp_A\sigma^*_A/lp_D\sigma^*_D)_{CI}$, is found at 7.52 eV. Therefore, a non-adiabatic process is predicted along the photochemical evolution of the $lp_A\sigma^*_A$ state, which will transfer the reactivity to the S_1 PEH. From the FC region to the $(lp_A\sigma^*_A/lp_D\sigma^*_D)_{CI}$ crossing point the system shows a small increase in either the O_2-H_3 and O_2-H_4 bonds, corresponding to two equivalent degenerate crossing regions, as the populated σ^* orbital in the excitation represents a plus-linear combination of both σ^* orbitals for the O_2-H_3 and O_2-H_4 bonds (cf. Figure 5.2).

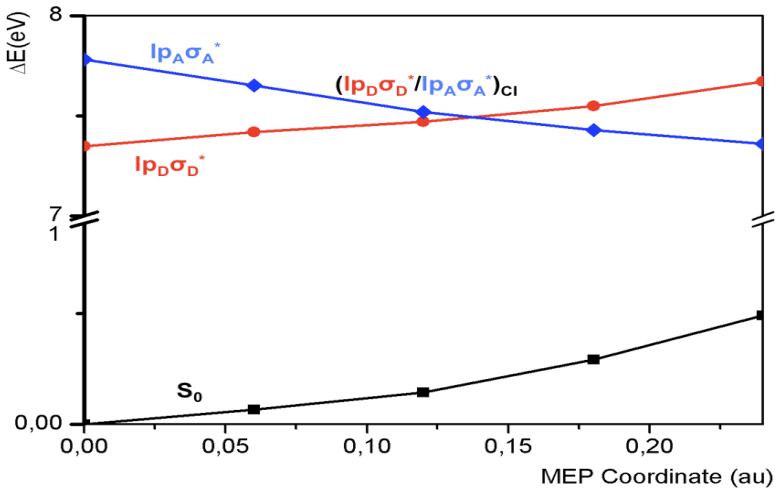

Figure 5.4: CASPT2 energies of the ground (S_0) and lowest-lying singlet excited states ($lp_D\sigma_D^*$ and $lp_A\sigma_A^*$) of the water dimer from the FC geometry and along the S_2 MEP.

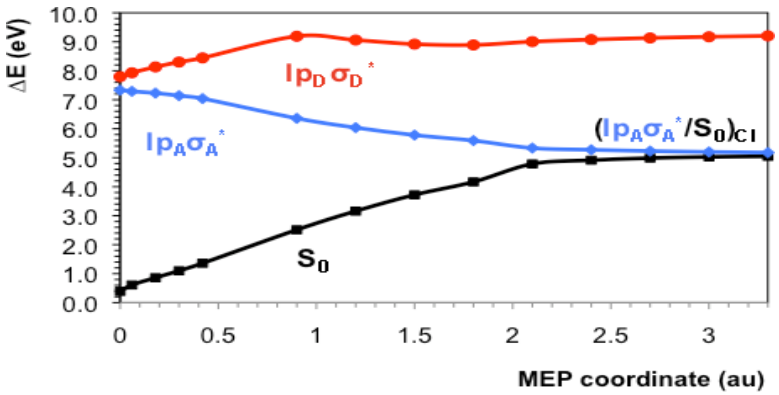

Figure 5.5: CASPT2 energies of the ground (S_0) and lowest-lying singlet excited states ($lp_D\sigma_D^*$ and $lp_A\sigma_A^*$) of the water dimer along the S_1 MEP from the CASSCF optimized structure determined along the S_2 MEP from the FC geometry (see Figure 5.4).

Figure 5.5 shows the decay of the S_1 excited state after the non-adiabatic crossing mediated by the region of close degeneracy between the S_1 and S_2 states reached at the end of the S_2-MEP. The water dimer continues decaying toward the CI with the ground state, $(\mathrm{lp}_A\sigma_A^*/\mathrm{S}_0)_{CI}$, while the elongation of either of the O_2-H_3 or O_2-H_4 bonds becomes more pronounced. The dissociation occurs in the W_A molecule, as it is the molecule in which the $\mathrm{lp}_A\sigma_A^*$ state is centered. The $(\mathrm{lp}_A\sigma_A^*/\mathrm{S}_0)_{CI}$ crossing point corresponds to the dissociation limit of these bonds, as was analogously found for the $\mathrm{lp}_D\sigma_D^*$ state. In this case, the $(\mathrm{lp}_A\sigma_A^*/\mathrm{S}_0)_{CI}$ point is found at 5.18 eV, which is in agreement with the experimental data compiled by Herzberg.[118]

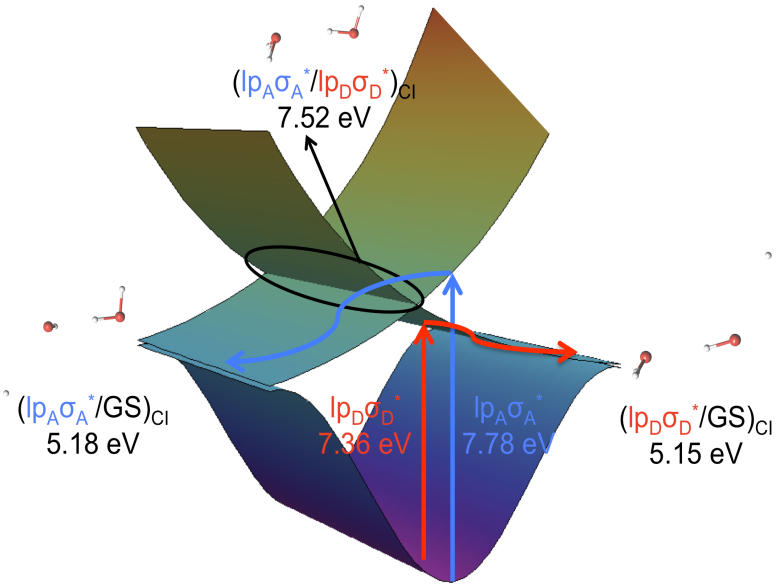

Figure 5.6: Scheme of the photochemical processes held in the water dimer after UV light irradiation.

A schematic view of the process can be found in Figure 5.6, where the direct excitation of the lowest-energy excited state (red) localized in the W_D at 7.36 eV evolves towards the photodissociation feature centered in its *dangling* O_1-H_1 bond and the absorption on the second excited state (blue) localized in the W_A at 7.78 eV develops in the photodissociation process given in either the

O_2-H_3 or the O_2-H_4 bonds as they are both equivalent. The present findings do not support an electron-driven proton transfer mechanism in which the $H_3O\cdots HO$ complex is formed in the photodissociation. Such model was suggested in previous theoretical studies in which the level of theory employed was not able to describe the correct nature of the lowest-lying excited state.[111–115] The electron-driven proton transfer determined in these calculation was a consequence of the charge transfer nature of the S_1 state. Instead, the mechanism established here corresponds to a homolytic photodissociation resulting in the production of the hydrogen-bond complex $H_2O\cdots OH$ plus a hydrogen atom.

The decay paths of the water dimer shown in Figure 5.6 point out to the water dimer as a potentially important UV-light filtering species in the atmosphere, where the concentration of this cluster has been suggested to be high.[104–106] The flat topology of the PEHs pictured can also hint at the production of $OH(X_2\Pi)$,[120] since flat electronic surfaces favor the dissociation into vibrationally cold molecules,[119] which might be relevant in atmospheric chemistry.

AIMD computations have been performed on the lowest-excited state S_1 of the dimer. As mentioned above, a slight elongation of the O_2-H_3 and O_2-H_4 bonds is enough to bring the $lp_A\sigma^*_A$ state below the energy of the $lp_D\sigma^*_D$ state. In fact, from the set of 50 initial conditions obtained by matching a Wigner distribution, 34 structures have an inversion on the nature of the two lowest-lying excited states with respect to the ordering at the CASSCF optimized structure, whereas the rest of geometries keep the same ordering, that is, a $lp_D\sigma^*_D$ and $lp_A\sigma^*_A$ character for the S_1 and S_2 states, respectively. This ratio can be explained by the fact that the plus-linear combination of σ^* orbital for S_2 (see Figure 5.2) provides twice as many probabilities to be found as the lowest-lying excited state after the Wigner distribution sampling. Therefore, the photodynamics of both the $lp_D\sigma^*_D$ and $lp_A\sigma^*_A$ states has been modeled in the present study (16 of those trajectories have been carried out on the locally excited state centered on the W_D, while 34 represented the W_A). All the dynamics simulations have given rise to the dissociation of the water molecule hosting the excitation.

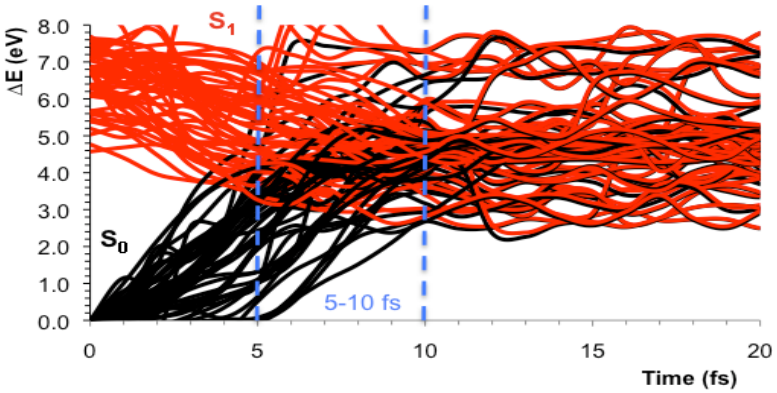

Figure 5.7: CASPT2 energies of the ground (S_0) and lowest-lying excited state (S_1) of the water dimer along the CASSCF *ab initio* molecular dynamics simulations performed on S_1.

In all trajectories the dissociative character of the $lp_D\sigma^*_D$ and $lp_A\sigma^*_A$ states has been confirmed and the photodissociation of the locally excited water molecule takes place. Figure 5.7 displays the CASPT2 energies of the ground (S_0) and low-lying excited state (S_1) vs. time for all the trajectories. An ultrafast process is predicted for the photoinduced homolytic cleavage of the OH bonds, with an estimation of the time for dissociation of around 5-10 fs after light irradiation. Similar time results were found in previous computational photodynamics studies and ground-state quantum dynamics works,[113, 115] whereas a much larger time scale was proposed in more recent photochemical study of the dynamics of the water system.[114] However, these excited-state dynamics were carried out in a charge-transfer excited state which is not present among the computed lowest-lying excited states. We assign the present computed photodissociation time of 5-10 fs to the locally excited state and the $H_2O\cdots HO + H$ dissociation limit. It is worth noting that the experimental value estimated for the photodissociation of the liquid water is also around 10 fs.[124] The agreement provides further evidence of the major role of the intrinsic properties of the water molecule and minor role of intermolecular interactions in order to rationalize the photochemistry of water.

Conclusions High-level *ab initio* CASPT2//CASSCF results are addressed to elucidate the photochemical mechanism operating in the gas-phase water dimer. A small red-shift and a larger blue-shift relative to a single water molecule have been found for the lowest-lying electronic transitions. The electronic excited states S_1 and S_2 are characterized by local excitations centered in the W_D ($lp_D\sigma^*_D$ state) and W_A ($lp_A\sigma^*_A$ state), respectively. Determination of the nature of the state is in contrast to previous computational studies on the photoinduced chemistry of the dimer, where a charge-transfer lowest-lying state was predicted.[111–115] The decay channels of the S_1 and S_2 excited states have been studied, having both of them dissociative character. S_1 leads directly to the photodissociation, exhibiting a non-adiabatic relaxation to the ground state, via a CI crossing, at energies consistent with the available experimental evidence on the dissociation of the water molecule. S_2 yields an analogous photodissociation showing two non-adiabatic processes via CIs first with S_1 and subsequently with the ground state. The dissociative nature of the lowest-lying S_1 and S_2 excited states have been predicted by AIMD dynamics simulations, obtaining a photodissociation time of 5-10 fs.

Based on the present study, the mechanism proposed for the photodissociation of the water dimer corresponds to the production of $H_2O\cdots HO$ species plus an H atom and not the $H_3O\cdots HO$ species suggested in previous studies.[111–115] Locally excited states ($lp_D\sigma^*_D$ and $lp_A\sigma^*_A$) are determined as the driving states in the photochemical events of the dimer, thus highlighting the importance of the electronic properties of the water molecule itself in the photochemical processes over intermolecular interactions held by the water dimer.

It can therefore be concluded that the hydrogen-bonded dimer stands closer to the fingerprint given for bulk water, specially in the case of the W_A in which the excitation energy is blue-shifted as compared to its monomeric counterpart. This blue shift is not large enough as to reach the ∼8.2 eV given for bulk water, but it can be envisaged that a higher donor-type related structures bearing a higher hydrogen-bonded coordination might further approach this value. What can be already discarded is the role of hydrogen bonding in the ∼4.6 eV fingerprint registered for ordered water, as the red-shift attained by

the W_D is exceedingly small and the effects when the water molecule acts as a hydrogen bond acceptor are more pronounced towards the blue-shift, thus making it hard to see how the signal can go down in energy so much by means of this type of interaction.

The π-Stacked Water Dimer

The π-stacked water dimer represents the first theoretical model, as far as we know, that proposes the non-covalent π-stacking interactions in water molecules. It aims to provide a molecular counterpart of the $\sim$4.6 eV ($\sim$270 nm) fingerprint registered in the absorption spectrum of ordered water and the associated fluorescence.[21, 41–43] The model is built based on the experience gathered over the years in the QCEXVAL group by treating the red-shifted excimeric fluorescence present in the DNA.[3, 51, 52] Experience has shown us that excimer-type interactions, due to π-stacked molecular configurations, produce very different results as compared to its monomeric counterparts. The excited state usually delocalizes between the monomers, yielding a completely different picture on the photophysical and photochemical pathways governing the processes undergone upon light absorption.

Computational Details The complete-active-space self-consistent-field second-order perturbation theory (CASPT2),[65] together with the basis set of Atomic Natural Orbital (ANO) type with the primitive set Li,O(14s9p4d3f)/ H(8s4p3d), the ANO-L set, contracted to O[5s4p1d]/Li[5s4p1d]/H[2s1p] has been used throughout. It represents a quadruple-ζ plus polarization scheme for the oxygen and lithium atoms and a double-ζ plus polarization scheme for the hydrogen atoms. The choice has been based on the results of the previous calibration calculations reported earlier on the study of the valence and Rydberg character of the excited states of the water molecule.[27] Employing that contraction in the basis set, the vertical excitation energy has been computed to be 7.49 eV at the CASPT2 level, consistent with the band maximum

displayed in the high resolution absorption spectrum at 7.447 eV.[22] Thus, the selected basis set represents a compromise between accuracy and demands of computational resources. Unless otherwise stated, computations have been carried out at the experimental equilibrium geometry of the ground state reported by Herzberg (r(OH)= 0.957 Å and <HOH=104.5°).[118] Geometries for the hydrated π-stacked water molecules are based on the experimental structures obtained by Odutola and Dycke for the water dimer.[102]

Multiconfigurational wave functions have been initially determined at the CASSCF level. In all the calculations, the oxygen 1s electrons have been treated as inactive and all the valence orbitals have been included in the active space. The full valence space comprises eight electrons and six orbitals for each water molecule considered. The dimer has therefore been described by the CASSCF(16,12) wave function, employing four roots in the averaging procedure. On the other hand, the corresponding results including the second-order corrections shall be labeled as CASPT2(16,12). The oxygen 1s electrons have been kept frozen in the perturbation step. In order to minimize weakly interacting intruder states, the imaginary level-shift technique, with a shift of 0.2 au, has been employed. Dynamic electron correlation taken into account through the CASPT2 method becomes crucial to give a realistic description of the interaction between two water molecules in the excited state, which is at the CASSCF level basically repulsive. Since calibration calculations have shown that the π-stacked interaction can be properly described by two active electrons distributed into two active orbitals (HOMO-, LUMO-type) of the two water molecules involved,[125] computations for the fluorescence were carried out using this active space. Thus, the excimer has been optimized at the CASPT2(8,8) level making use of numerical gradients.

External electric fields have been introduced through finite-field perturbation theory as implemented in the MOLCAS package. All computations have been carried out by using the MOLCAS 7.4 quantum-chemical software.[69] Basis set superposition error (BSSE) has been explored by using the counterpoise correction.[87]

Results and Discussion In the quest for the basic molecular moiety consistent with the recorded experimental fluorescence spectra of structured water, a number of preliminary facts are worth considering. From an experimental point of view, it is known that the spectroscopic features observed in the ordered zone (free-solute zone) adjacent to hydrophilic surfaces are similar to those of the aqueous region around the solutes, implying that the phenomenon is due to water itself. The fluorescence of NaCl, KCl, and LiCl solutions have the maximum of emission centered around 480-490 nm (2.58-2.53 eV), giving rise to the highest intensity for the latter. For saturated solutions of LiCl, Chai *et al.*[42] also analyzed the fluorescence dependence. Interestingly, exciting at ∼310 nm (4.00 eV), the recorded emission was faintly visible. However, the observed peak intensity considerably increased at 250 nm (4.96 eV) or 270 nm (4.59 eV). Under these conditions, blue fluorescence λ_{max} ∼483 nm (2.57 eV) was clearly visible to the naked eye.[42] Therefore, the relaxed excited state source of the fluorescent water-based system is expected to be placed above 4 eV having a vertical emission around 2.6 eV. A key question to be addressed seems to be: what is the smallest molecular unit capable of describing such a blue fluorescence feature?

The vertical excitation energies of the gas-phase conventional water dimer have been found slightly red- and blue-shifted for the water donor (W_D) and the water acceptor (W_A), respectively, with respect to the single water molecule.[109] It is also the case for larger water clusters.[126] That is, the presence of hydrogen bonds in the ground-state equilibrium structures of the clusters does not have a pronounced effect on the low-lying vertical electronic transitions. Furthermore, the main photoinduced chemistry of the water dimer can be described as a process occurring between two single molecules of water, with the effect of the hydrogen bonds playing a minor role.[109] The lowest decay routes of the dimer correspond to the photodissociation processes, one for each water molecule (W_D, W_A), leading to the hydrated hydroxyl radical $OH\cdots H_2O$ and the hydrogen atom (H) at 5.1-5.2 eV, i.e. the well-known photolysis of water recorded experimentally.[118] The emissive feature we are looking for cannot therefore be related to the H-bonded dimer itself.

The characteristics expected from the emissive moiety, together with the considerations mentioned above, lead us to explore the hypothesis of excimer-type fluorescence formed by two relaxed π-interacting side-hydrated dimers. Excimers are well-defined excited states characterized by a repulsive interaction in the ground state and play a crucial role in many biological processes.[3, 51, 52, 127, 128] In the study of biopolymers, including DNA, excimers (excited dimers) are widely invoked in order to rationalize a number of spectroscopic features which are not present in the monomers.[128] In this manner, the appearance of long-lived emissive states observed in the fluorescence spectra of DNA and related oligomers could successfully be explained, both experimentally[127] and theoretically.[3, 51] From a computational standpoint, as described elsewhere,[3] excimers are particularly difficult to handle correctly. In general, in order to describe an excimer properly, highly accurate quantum-chemical methods are required.

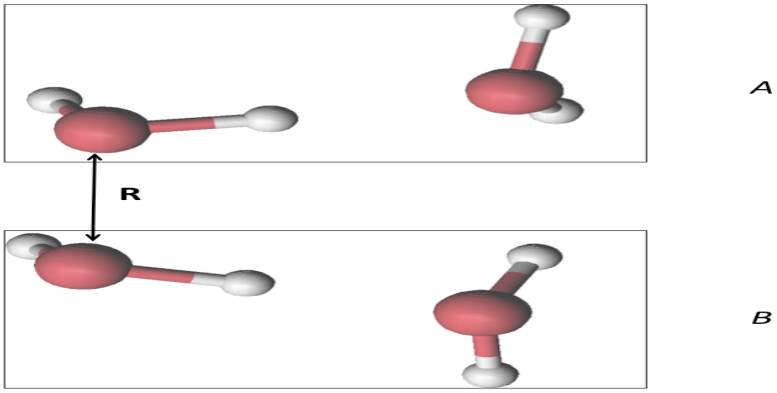

Figure 5.8: Optimized geometry for the lowest excited state of the tetramer formed by two relaxed π-interacting side-hydrated dimers at the CASPT2(8,8) level of theory.

The equilibrium geometry for the excimer computed at the CASPT2(8,8) level is displayed in Figure 5.8. The excimer is formed by two relaxed π-interacting side-hydrated dimers. As expected, an elongation of the O-H bonds in the π-stacked water molecules (1.034 Å) with respect to the ground state (0.957 Å) has occurred. The excimer can basically be seen as two π-stacked water

molecules in which each water is hydrated by an additional molecule. The wave function of the excimer can be clearly described by the one-electron promotion from the Natural Orbital (NO) NO-plus-HOMOs type, the out-of-phase plus linear combination of $1b_1$ molecular orbitals (MOs) of the two monomers belonging to the side-hydrated dimer (A and B in Figure 5.8), to the NO-plus-LUMOs type, involving the $4a_1$ MOs in the C_{2v} point group of the water molecules (see Figure 5.9).

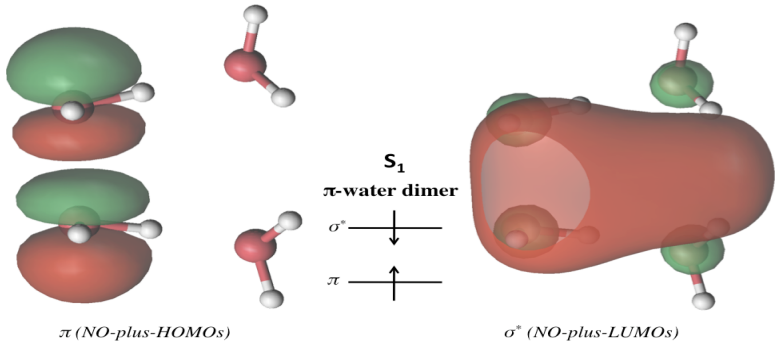

Figure 5.9: Main CASSCF(8,8)/ANO-L O[5s4p1d]/H[2s1p] natural orbitals (NOs) involved in the description of the S_1 state at the equilibrium distance $\mathbf{R}_{min}$ of two relaxed π-interacting side-hydrated dimers.

Employing the CASPT2 optimized geometry of the excited state, the vertical emission is computed to be at 3.06 eV (∼405 nm), clearly about half an eV blue-shifted with respect to the experimental data. In order to get additional insight on the fluorescent feature, a displacement between the π-stacked side-hydrated dimers has been carried out, keeping the intra-dimer geometry fixed.

As can be seen from Figure 5.10, the potential energy built with respect to the intermolecular oxygen-oxygen distance R(O-O) of the respective subunits of the excimer is very shallow. Because of the repulsiveness of the ground state, a small decrease in R(O-O) causes a pronounced decrease in the predicted vertical fluorescence, which covers the energy range 2.47-3.06 eV corresponding to the R(O-O) interval 1.92-2.02 Å. The results are consistent with the

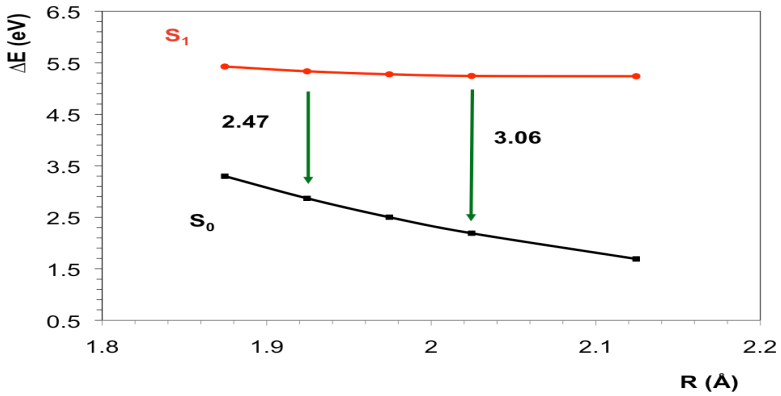

Figure 5.10: CASPT2(8,8)/ANO-L O[5s4p1d]/H[2s1p] potential energy curves built with respect to the intermolecular distance **R** (see Figure 5.8) of two relaxed π-interacting side-hydrated dimers.

experimental data, although the actual observed feature would depend on the intricacies of a given system. Excimers of larger molecular size could equally well describe the fluorescent species because the electronic transition is mainly localized in the π-stacked water molecule belonging to distinct side-hydrated dimers (A and B in Figure 5.8) and the additional water molecule just plays a hydration role. The excitation does not further delocalize, thus pointing towards the excimer formed by two relaxed π-interacting side-hydrated dimers as the basic molecular unit responsible for the fluorescence observed in structured water.

Taking into consideration the mentioned theoretical background, the smallest unique moiety responsible for the observed absorption at 4.59 eV [42] has to most probably be related with the absorption of two π-stacked water molecules. The *anti*- and *syn*-type interactions depicted in Figure 5.11 can be therefore seen as models of the bimolecular interaction taking actually place in structured water. To our knowledge, how the electronic spectrum of a water molecule is affected by the presence of another π-stacked water molecule has not been so far analyzed. Thereby, we have computed the vertical transition

for the lowest excited state of the π-stacked *anti* water dimer in order to elucidate whether it could be related to the spectroscopic fingerprint characteristic of structured water observed at 4.59 eV.[42]

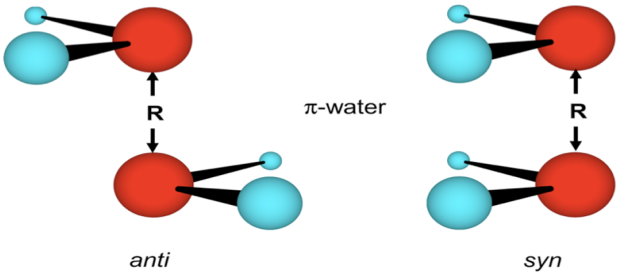

Figure 5.11: Interaction of two π-stacked water molecules in the *anti* and *syn* orientation separated by the intermolecular distance **R**.

Table 5.2: CASPT2 vertical transition energies (ΔE) for the $S_0 \rightarrow S_1$ π-stacked water homodimer in the *anti* and *syn* orientations.

R(Å)	*anti* ΔE(eV)	*syn* ΔE(eV)	Exp.
50.00	7.39	7.35	7.447[a]
3.00	7.04	6.93	
2.10	4.97	4.66	
2.05	4.73	**4.41**[c]	4.59[b]
2.00	**4.46**[c]	4.26	4.59[b]
1.95	4.17	3.85	

[a]Reference [22]
[b]Reference [42]
[c]Reference [129]

Table 5.2 lists the results obtained for the $S_0 \rightarrow S_1$ transition energy at distinct intermolecular separations (**R**) of the water dimer. As expected, the

absorption calculated at 50.00 Å corresponds to the monomer absorption and it is here computed at 7.39 eV, in agreement with previous findings and experimental evidence.[22, 27] More interestingly, in the interval of **R** ranging between 2.00-2.10 Å, the electronic $S_0 \rightarrow S_1$ transition energy decreases to 4.46-4.97 eV (*anti*) and 4.26-4.66 eV (*syn*), which closely resemble the absorption maximum recorded experimentally for structured water.

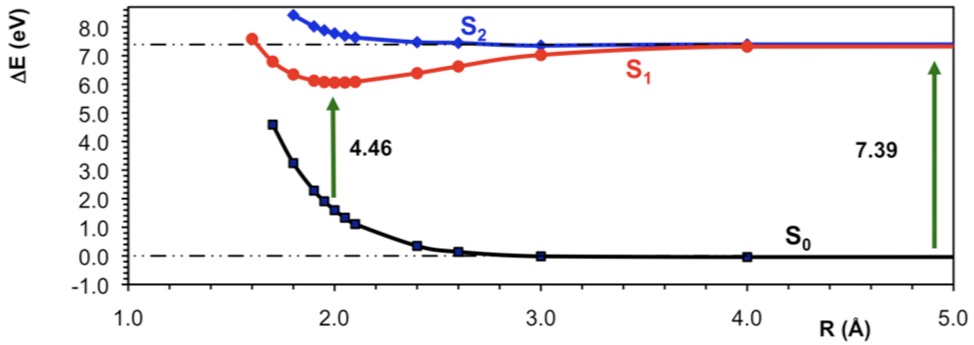

Figure 5.12: CASPT2(16,12)/ANO-L O[5s4p1d]/H[2s1p] potential energy curves built with respect to the intermolecular distance **R** (see Figure 5.11) of two *anti* π-stacked water molecules.

As can be noted in Figure 5.12, where the potential energy curves along the intermolecular distance (**R**) are shown, the equilibrium intermolecular distance for the S_1 state ($\mathbf{R}_{min}$) is precisely placed at 2.00 Å for the anti arrangement. It can be seen that the lowest excited state becomes clearly bound by more than 1 eV. Furthermore, once the BSSE has been applied, the *anti* S_1 state has a binding energy of 1.20 eV. The BSSE correction (0.13 eV) is relatively small in the π-water dimer, both in the *anti* and *syn* arrangements, as compared to other systems.[3, 51]

At the equilibrium geometry of the S_1 state, the computed oscillator strength for the $S_0 \rightarrow S_1$ vertical transition is similar (0.039) to that obtained for the monomer.[27] The S_1 state is mainly described by the one-electron promotion

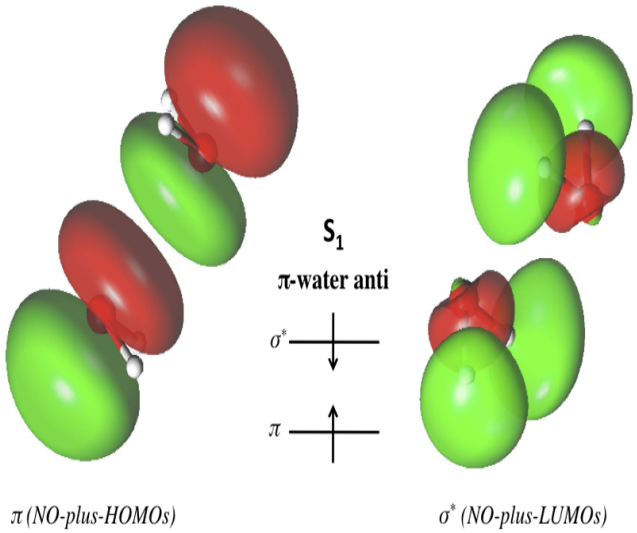

Figure 5.13: Main CASSCF(16,12)/ANO-L O[5s4p1d]/H[2s1p] natural orbitals (NOs) involved in the description of the S_1 state at the equilibrium distance $\mathbf{R}_{min}$ of two *anti* π-stacked water molecules.

from the Natural Orbital NO-plus-HOMOs type, the out-of-phase plus linear combination of $1b_1$ MOs of the two monomers, to the NO-plus-LUMOs type, involving the $4a_1$ MOs under the C_{2v} point group of the two water molecules (see Figure 5.13). Consequently, it is predicted that upon light irradiation, the π antibonding oxygen-oxygen interaction is released and the stretching of the O-H bonds is favored, in accordance with the experimental evidence.[42]

In order to study π-stacked configurations, a PES governed by the stacking coordinate is usually built. In our case, the stacking coordination is defined as the O-O distance between the water molecules, yielding the curve given in Figure 5.12. As can be seen, at *infinite* distance between the monomers (50.00 Å), the excitation energy would be that of the monomer as there is no interaction happening between them. As we bring them together, a whole different

picture arises though. The orbitals represented in Figure 5.13 start interacting with each other yielding a linear combination of the orbitals describing the monomers, and the initially localized orbitals delocalize over both molecules. The energy profile of the lowest-energy excited state rapidly drops till a minimum placed vertically at ∼4.6 eV with an inter-molecular O-O distance of ∼2 Å.

A closer look to the orbitals involved in the excitation and depicted in Figure 5.13 provide a broader analysis of the absorption process. Upon light absorption, the excitation is focused in the π-π interaction occurring between the water molecules. This delocalization occurring in the molecules helps rationalizing why they are driven to such a close arrangement and it can also be related to the growing size of the ordered water regime under light irradiation, facilitating the energy release through the stretching of the O-H bonds as has been observed experimentally.[42] The fluorescence observed experimentally at ∼2.57 eV (∼483 nm) can be rationalized by invoking a widely and previously introduced model based on excimeric interactions. This is accomplished by extending the model through the hydration of both π-stacked water molecules, and then optimizing to its minimum in their excited state in order to follow the energetic decay till it reaches the bottom from where they will fluoresce. This yields a convincing explanation of the fluorescence present in the ordered water that resembles those previously found for DNA.[3, 51, 52]

In order to get further insight into the nature of the absorption feature, the study has been additionally extended by considering two π-stacked top-hydrated water molecules. As shown in Figure 5.14, the trends of the potential energy curves and the computed equilibrium distance obtained for S_1 (∼2.0 Å) are quite similar as those obtained for just two π-stacked water molecules. It is however worth mentioning that the binding energy decreases to 0.83 eV. The finding gives support to the view on the prevalence of the π stacking interaction of isolated water molecules in the phase characterized by a strong absorption of light at ∼270 nm. The spectroscopic feature is also a characteristic of water in contact with various charged molecules. In order to check the stability of the whole π-stacked system in a charged environment, an ad-

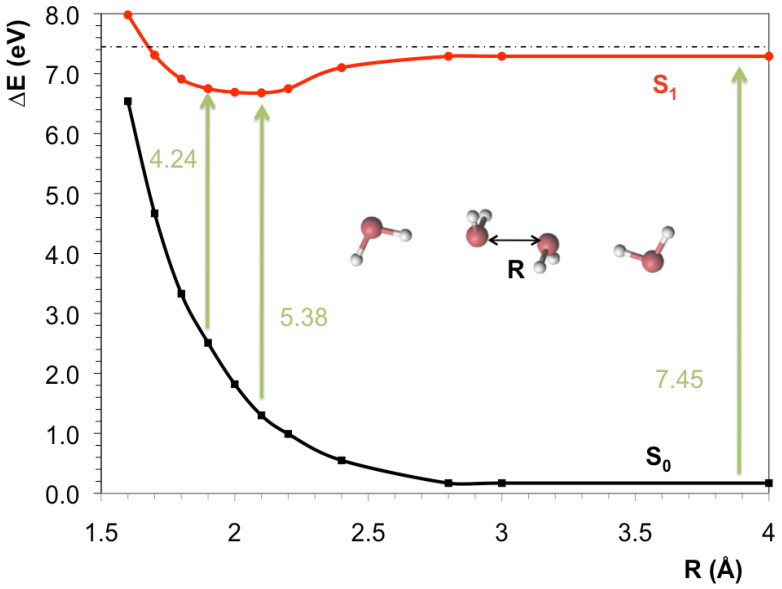

Figure 5.14: CASPT2(8,8)/ANO-L O[5s4p1d]/H[2s1p] potential energy curves built with respect to the intermolecular distance **R** (in the Figure) of two *anti* π-stacked top-hydrated water molecules.

ditional analysis has been conducted under the influence of positive lithium ions.[130]

The π-stacked water dimer provides a good model to test the adequacy of the π-stacking interactions to partially understand the experimental signal registered, but a number of the intrinsic properties characterized experimentally to this ordered water escape the extent of the model. In order to provide a more complex model, electrostatic interactions have been considered, showing that the essentials of the interaction remain unchanged due to the effect of external electric fields or charged environments. Slight quantitative changes are produced, but not qualitative differences.

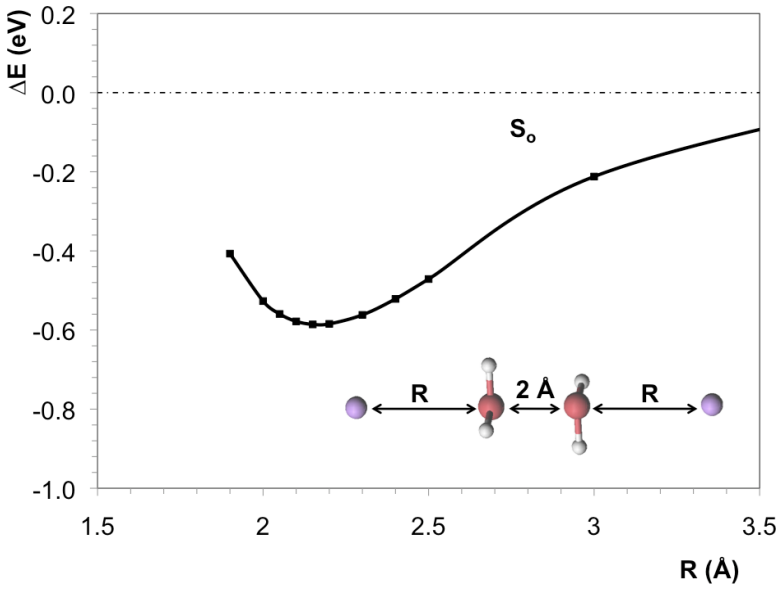

Figure 5.15: CASPT2(16,12)/ANO-L O[5s4p1d]/Li[5s4p1d]/H[2s1p] potential energy curves built with respect to the intermolecular distance **R** (in the Figure) of two *anti* π-stacked water molecules in the presence of two lithium cations.

As shown in Figure 5.15, when two π-stacked water molecules separated by 2 Å are in the presence of two lithium cations in a top-like arrangement for each water molecule, the full system in the ground state is stabilized by around 0.6 eV, reaching a minimum at about 2.15 Å, in accordance with the experimental evidence.[131] Such behavior has also been found experimentally, where solvated lithium solutions have shown a distinctive structure in the form of strings,[132–134] and the negatively charged counterions are pushed out of the solution to the surface.[135] Therefore, the repulsive character of the ground-state of the two π-stacked water molecules can be overcome by the actual interaction with the charged environment. In addition, the potential energy curves with respect to the **R** distance of two water molecules have been rebuilt in the presence of two lithium cations, keeping fixed the interatomic distance Li-Li at 8.0 Å, 15.0, and 20.0 Å.

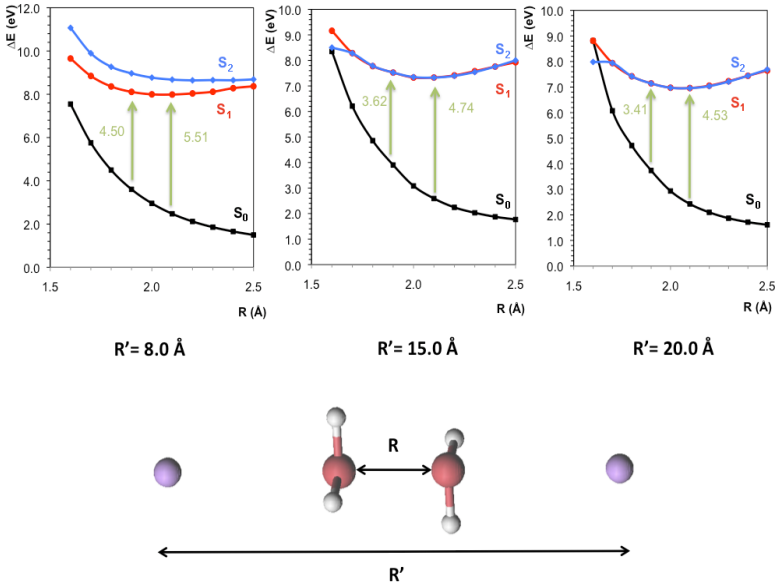

Figure 5.16: CASPT2(16,12)/ANO-L O[5s4p1d]/Li[5s4p1d]/H[2s1p] potential energy curves built with respect to the intermolecular distance **R** of two *anti* π-stacked water molecules in the presence of two lithium cations at a fixed Li-Li interatomic distance $\mathbf{R}'$.

By inspection of the results depicted in Figure 5.16, it is concluded that the confinement of two π-stacked water molecules in a charged environment does not significantly change the essentials of the π interaction obtained in vacuo. From the recent work of Brunet and Joly [136] one can also anticipate that the π-like interaction might also play a fundamental role in the reaction of metals with water,[136–138] as well as in proteins and other biological and hydrophilic environments.[139, 140] Within this framework, the effect of an externally-induced electric field might be of importance and has therefore been analyzed. For relatively weak to moderately strong fields, 0.0001-0.04 au, the respective ordering of the excited states does not vary for both the *anti* and *syn* orientations.

A notable feature of the ordered water region is its negative charge, with local charge separation between the ordered zones and the regions beyond, the latter bearing positive entities. Nevertheless, based on previous works on the photochemistry of water clusters, those negatively or positively photoinduced species are not expected to contribute to the ∼270 nm absorption. As pointed out earlier, the $(H_2O)_5$ cluster represents the smallest cluster exhibiting a well-developed second minimum on the ground-state hypersurface corresponding to the hydrate ion pair $H_3O^+(H_2O)_3OH^-$.[126] The absorption maximum of the cyclic water pentamer was obtained at about 7.5 eV, close to the absorption maximum of water in the gas phase, whereas from the zwitterionic ground state the vertical transition around 6 eV was computed. On the other hand, the biradical form described by Sobolewski and Domcke for the same system has the characteristic absorption spectrum of the hydrated hydronium radical peaking around 1.5 eV.[126] It is also worth mentioning that the lowest vertical electronic transition of the hydroxyl radical hemibonded to water was calculated by Chipman at 4.41 eV (280 nm).[141] The result is here confirmed at the CASPT2(15,12) level, 4.47 eV. The associated oscillator strength for the transition is however exceedingly small. Therefore, it might contribute to the recorded ∼270 nm absorption in a minor extent. Moreover, according to molecular dynamics simulations performed by Vassilev *et al.*[142] on the hydroxyl radical in liquid water, hemibonded water molecules are expected to form part of the first coordination shell of the hydrated OH, although the authors of this work emphasize that further work is required to be totally conclusive in this respect.

Conclusions From a constructive interplay between the available experimental data on structured water and the present theoretical results, several conclusions may be derived. According to the current research, the ∼270 nm (4.59 eV) experimental feature of the absorption spectrum of water observed in the ordered zone of interfacial water and in different solutions of salts, can be related to the presence of π-stacked water dimers at relatively short intermolecular distances (around 2 Å). Employing the ground-state equilibrium geometry of the two water molecules, the lowest π-stacked singlet state of the

dimer, the S_1 state, has the minimum for the intermolecular distance around 2 Å. The vertical transition energy $S_0 \rightarrow S_1$ computed at the equilibrium intermolecular distance obtained for S_1 is calculated in the energy range 4.4-4.5 eV. The lowest π-stacked singlet state is bound by more than 1 eV. Thus, light favors the formation of π-stacked water at short intermolecular distances. It helps to explain the reason why the exclusion zone actually grows upon light irradiation.

The fate of the absorbed energy is predicted to be released as excimer-type fluorescence. The excimer (excited tetramer) can be primarily seen as two relaxed π-interacting dimers. Independently of the real macroscopic arrangement of structured water in a given environment, i.e. forming a network of ice-like layers or consisting of polymer-like chains, the present findings suggest that the underlying microscopic phenomenon is related to π-stacked water molecules. Therefore, the spectroscopic characteristics of structured water determined experimentally together with the current interpretation offered here from high-level quantum-chemical results point out that π-stacked water can be considered as the actual fingerprint of the spectral properties of structured water.

It can therefore be concluded that the fingerprint observed for ordered water can be assigned to π-stacking interactions happening at short intermolecular distances in water. The dimer provides the smallest building block in which the absorption can be explained through the π-stacking interaction and the fluorescence requires an extended model that presents a hydration on both π-stacked molecules. The environmental effects suffered by the ordered water due to charge or external electric fields are deemed to be exceedingly small and thus exert no qualitative changes on the essentials of the interaction. From the overall results, we can concur that a larger model based on the characterized π-stacking interactions shall be built in order to group both absorption and fluorescence into the same building block, a premise that has been tackled over the next chapter.

Conventional vs π-Stacked Water Dimer

A brief comparison of the results obtained between the conventional and the π-stacked water dimers shall be addressed to clarify the key concepts that help unravel the very distinct photophysical and photochemical properties characterizing them.

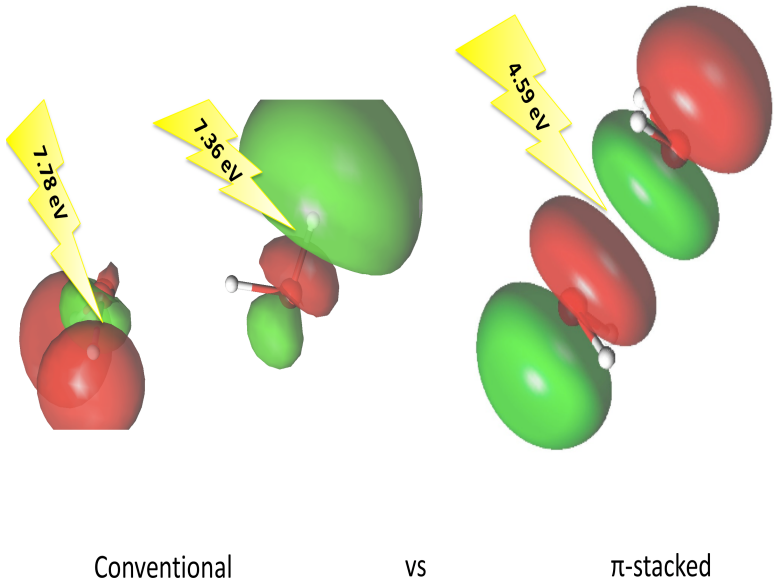

Figure 5.17: Scheme of the localized vs delocalized excitation processes undergone by the conventional and the π-stacked water dimers.

The main distinctive feature is the localization of the excitation, as it appears localized in the monomers of the hydrogen-bonded dimer whereas it is totally delocalized in the π-stacked one. A schematic view of both processes can be seen in Figure 5.17. This delocalization provides a massive energetic stabilization which brings the absorption to about 3 eV lower than what it would be expected in its monomeric constituent parts, whereas the localized excitation given in the hydrogen-bonded dimer yields values in accordance to those observed in separated water molecules. The results stress out the negligible

effect of the hydrogen bond in the excitation energy and highlighting the importance of the π-stacked interactions as a plausible way to account for such a difference in energy between both spectroscopical fingerprints.

Once the excitation becomes localized, it is driven by photophysical and photochemical processes analogous to those previously found for the water monomer. However, when the excitation is delocalized, the photochemical pathways vary completely giving rise to a whole new picture, represented in the low-lying absorption but also in the fluorescence emission features present in ordered water, which are lacking in any other kind of water aggregates.

Chapter 6

Ordering Water: The Hexagonal Ice-like Model for Structured Water

> Water is H_2O, hydrogen two parts, oxygen one, but there is also a third thing that makes water and nobody knows what that is.
>
> *D. H. Lawrence*

In order to give a proper description of the so-called ordered water, a reduced model was presented in the preceding chapter based upon the knowledge of its intrinsic spectroscopical fingerprint and the need of a π-stacking interaction to justify the pronounced difference in energy given between the registered signal at 4.59 eV [42] and the one widely known for liquid water around 8.2 eV.[23–25] As has already been discussed, the model comprising two π-stacked water

molecules yield the essentials of the absorption fingerprint but needs a further extension to tackle the fluorescence feature as well as the environmental effects registered in the vicinities of ordered water.

Structured water, also known as ordered water, has a number of distinct features as compared to bulk water.[13, 21, 42, 132, 143] Evidence that surfaces impact nearby water has been shown by using different experimental techniques, such as absorption spectroscopy,[42] infrared absorption and emission,[144] nuclear magnetic resonance,[143] as well as from viscosity determinations, [145] and optical measurements.[146, 147] On the other hand, a number of structured layers are also clearly identified in carbon nanotubes containing water.[148, 149] We shall focus in the present theoretical contribution on the absorption spectroscopy intrinsic to the presence of structured water.

On the basis of high-level *ab initio* computations, we have previously reported that the spectroscopic absorption observed experimentally at ∼4.59 eV can be rationalized at the molecular level in terms of two π-stacked water molecules,[129] which as far as we know, constitutes a completely novel concept within this context. Nevertheless, that simple scheme does not take into account the experimental fact that in the ordered zone a local charge separation between the ordered region and the regions beyond does exist, being the latter with abundant protons.[148, 150] In this sense, the hexagonal ice-like model recently suggested by Pollack and co-workers looks much more appealing.[21] The honeycomb-model proposed is consistent with the experimental evidence repeatedly offered by different labs leading essentially to the same conclusion: when water absorbs light, the absorbed energy builds structural order and drives charge separation.[21, 148] Of the evidence showing hexagonality, it is worth noting that water hexamers have been observed next to diverse surfaces, including metals,[44, 45] protein subunits,[151] and graphene.[149] Near-surface hexamers have also been found in supercooled water.[152] Therefore, the hexagonal ice-like model for structured water is largely supported experimentally.

A detailed analysis of the hexameric model for structured water at the molecular level may hopefully provide additional insights into the underlying properties responsible for the spectroscopic feature observed at ∼270 nm, which serves indeed as a fingerprint to identify the presence of structured water. Within that spirit the present contribution is offered. As discussed in detail below, predictive high-level *ab initio* results fully support the presence of such a spectroscopic peak in the π-stacked system built from the most plausible basic molecular unit (**BMU** hereafter) of the hexagonal ice-like model, here identified as the negatively charged radical $[H_{19}O_{10}]^{\cdot -}$. Knowledge provided on theoretical grounds may be potentially relevant in the realms of biology, life-sciences, and nanotechnology, where undoubtedly ordered water plays a fundamental role.[21, 45]

Computational Details Ground-state geometry optimizations for the **BMU** system and its π-stacked dimer have been performed at the MP2 level through the program GAUSSIAN 09.[108] At the equilibrium structures so obtained, the corresponding vertical electronic transitions have been subsequently computed by using the CASPT2 approach, a well-established *ab initio* methodology especially suited for that purpose,[66] as implemented in the MOLCAS-7 software package.[69] The basis set 6-31G** has been used throughout, as a compromise between accuracy and computational resources.[6]

Multiconfigurational wave functions have been initially determined at the CASSCF level. In order to have enough flexibility in the employed wave functions to describe the low-lying excited states up to 5 eV, especial attention has been paid to the selection of the active space that has been based on a systematic number of calibration computations. For the **BMU** system, the active space comprises 13 electrons distributed in 12 active orbitals, hereafter denoted as the CASSCF(13,12). Accordingly, the respective results including electron dynamic correlation at the second-order level, will be named as CASPT2(13,12). For the π-stacked dimer, the CASSCF(12,15) wave function has been used, where the second-order findings will be coined as CASPT2(12,15). Details on the wave functions will be next given as ap-

propriate. The oxygen 1s electrons have been kept frozen in the perturbation step. In order to minimize weakly interacting intruder states, the imaginary level-shift technique, with a value of 0.2 au, has been employed. The CASPT2 method with the zeroth-order Hamiltonian as originally implemented has been used.[65]

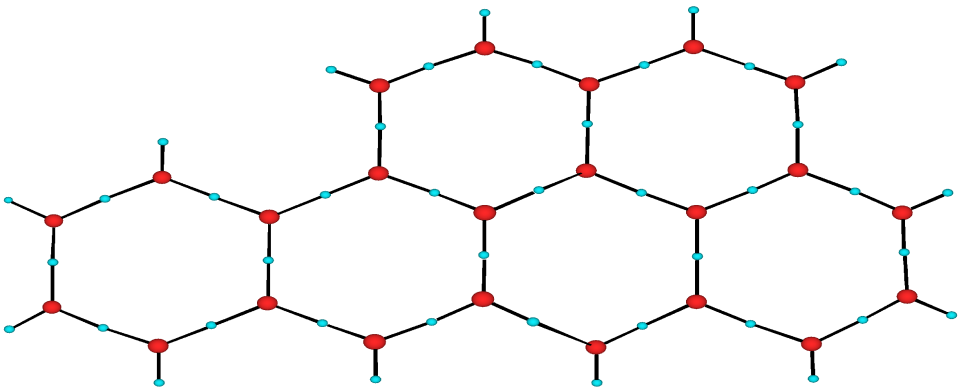

Figure 6.1: Scheme for the ice-like multi-layer model of structured water formed by hexagonal units. The system is negatively charged (see text).

Results and Discussion As a way to provide a larger and more accurate model that contains all the spectroscopical properties reported for ordered water, an effort was attempted to build an extended model holding the previously characterized essential features.[129] In this structure, which can be seen in Figure 6.1, the water molecules are arranged building hexagons linearly, similar to an ice layer in a polymer-like manner. The new model is based upon the evidence piled up over the years on the hexagonal-like arrangement of water next to many surfaces. The proposed model changes slightly as compared with previous suggestions,[21, 153] although all proposals emphasize that the planes themselves are negatively charged.

The net charge of the six-hexagon structural diagram drawn in Figure 6.1 is -5. It can be easily checked that, in general, for N-fused hexagons, the charge

acquired by the global system is 1-N.[153] Therefore, for the hexagonal water model, the basic molecular unit corresponds to the negatively charged radical species $[H_{19}O_{10}]^{\cdot -}$ constituted by two-fused hexagons, previously denoted as **BMU**. In order to fulfill the empirical formula $(H_2O)_n$, for each couple of fused hexagons built, a proton is assumed to be released, which is consistent with the local charge separation observed between the ordered water zone and the regions beyond confirmed by Chai *et al.* using pH measurements.[150] As stated in the introduction, water next to many surfaces is evidently hexagonal,[44] which agrees with the proposed model. It is however worth noting that the hexagonal model proposed here is different from previous suggestions,[21, 153] although the three proposals emphasize that the planes themselves are negatively charged. In 1969, Lippincot *et al.* suggested the hexagonal model of water in the generic form depicted in Figure 6.1,[153] assuming that an appropriate number of hydronium ions, protons, or tetra-coordinate hydrogen-bonded oxygen atoms would be required to maintain the empirical formula $(H_2O)_n$. On the other hand, in the Pollack′s model,[21] proposed forty-four years later, the net charge of the hexagon is -1, which corresponds to the infinite lattice. In our proposal, two-fused hexagons, the **BMU** system, bear the negative charge. Of course, as larger becomes the number of fused hexagons considered (N), the charge per hexagon (1-N)/N gets closer to -1. As shall be seen below, the π-stacking of two **BMUs** leads to the characteristic spectroscopic fingerprint of structured water. The two protons (one for each **BMU** considered) required to fulfill the empirical formula $(H_2O)_n$ are not directly involved in the model. It would be interesting to actually verify whether the larger number of $[H_{19}O_{10}]^{\cdot -}$ plus H^+ systems potentially generated by light irradiation is consistent with the experimental evidence of the growing size of the ordered water region in relation to the decrease of pH measured just beyond it, as it seems to be the case.[150] It is anticipated that the electronic transitions of the π-stacked **BMU** dimer are locally centered in the central O-H-O subunits, being responsible for the unique spectroscopic signature of structured water. According to the conclusions derived from the present computational research, the O-H-O subunit is essentially the key to understand the spectroscopic and photochemical properties of ordered water.

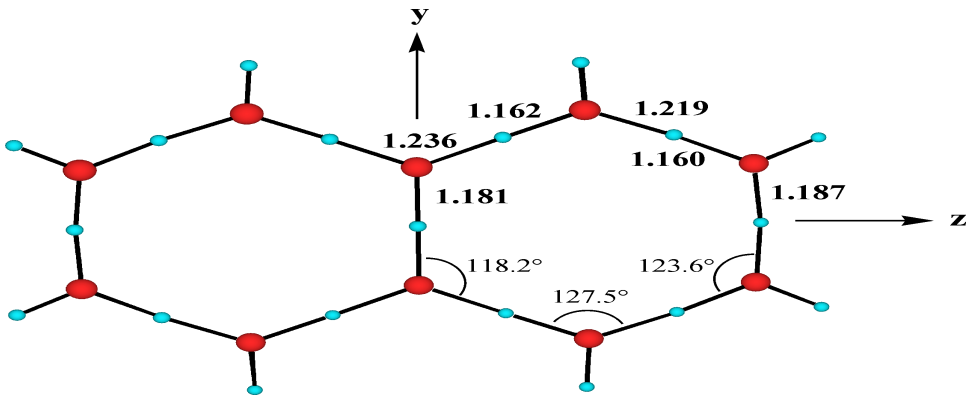

Figure 6.2: Optimized geometrical parameters for the 1^2B_{2u} ground state computed at the MP2(UHF)/6-31G** level. The molecule is placed in the yz plane with the central O-H-O moiety along the y axis (D_{2h} symmetry).

Therefore, we have concluded that the smallest unit found to be of relevance for its spectroscopical features, namely the *Basic Molecular Unit* (**BMU**), is that of two-fused hexagons, depicted in Figure 6.2, and in which the formula reads as $[H_{19}O_{10}]^{\cdot -}$. The O-H distances are elongated as compared to the ones found in the hexameric water cluster,[125] and the fused nature of the **BMU** makes it a negatively charged radical, that fits perfectly into the experimental evidence pointing towards a highly negatively charged region where the ordered water lies being surrounded by a positively charged environment beyond that region.[150]

As a result of the computation, the ground state becomes the 1^2B_{2u} state with the *Singly Occupied Natural Orbital* (SONO) localized on the central O-H-O bond. The SONO is mainly described by the plus linear combination of the p_y orbitals of the central oxygen atoms, where the spin density is localized. The vertical transition energies computed at the CASPT2(13,12)/6-31G** level together with the corresponding SONOs derived from the CASSCF(13,12)/6-31G** computation are compiled in Figure 6.3. The CASSCF(13,12) corresponds to 13 electrons distributed among 12 active molecular orbitals: 1, 2, 1,

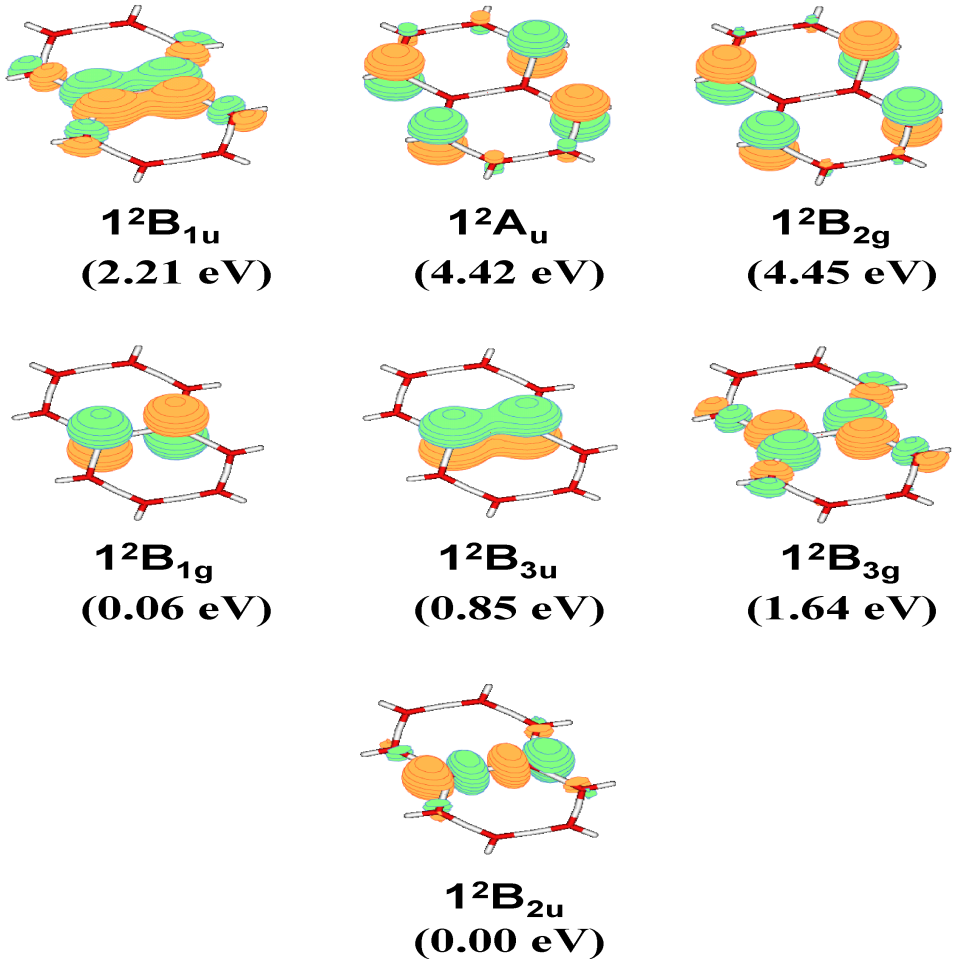

Figure 6.3: CASSCF(13,12) singly occupied natural orbitals (SONOs) for the low-lying electronic states of the basic molecular unit (**BMU**). The isodensity surface value for the orbitals is 0.030. The CASPT2(13,12) vertical electronic transitions (within parentheses) computed at the 1^2B_{2u} ground-state equilibrium geometry (see Figure 6.2) are also included. The 6-31G** basis set has been employed.

2, 1, 2, 1, 2, belonging to the irreducible representations a_g, b_{3u}, b_{2u}, b_{1g}, b_{1u}, b_{2g}, b_{3g}, a_u, respectively. In each symmetry, state-averaged CASSCF(13,12) computations have been carried out involving the four lowest roots. For the sake of completeness, the inactive orbitals employed are also given: 11, 2, 10, 2, 8, 1, 8, 1. The core orbitals (3, 0, 3, 0, 2, 0, 2, 0) have been kept frozen in the CASPT2 step. The seven electronic states obtained correspond to the respective lowest root of each irreducible representation considered.

Six low-lying vertical excited states are characterized for **BMU** below 5 eV. In the near infrared region, the 1^2B_{1g} and 1^2B_{3u} excited states are placed vertically at 0.06 and 0.85 eV, respectively. The respective SONOs are described by the π-antibonding (1^2B_{1g}) and π-bonding (1^2B_{3u}) linear combination of the p_x orbitals of the fused oxygen atoms. In the visible range, the 1^2B_{3g} (1.64 eV) and 1^2B_{1u} (2.21 eV) are next found. Both states have σ-type SONOs involving the central O-H-O subunit. Finally, the 1^2A_u and 1^2B_{2g} states of π character are predicted in the near ultraviolet zone of the electromagnetic spectrum at about 4.4-4.5 eV. Therefore, the **BMU** is revealed as a quite versatile system, capable of producing electronic excited states in an ample energy range, from near infrared to near ultraviolet. In fact, taking into account that the two lowest excited states lie within 1 eV with respect to the ground state, it can be envisaged that the nature of the ground state itself may easily vary depending on the environment and/or external influences. The oscillator strengths computed for the respective dipole-allowed transitions are however small, as it is in the water molecule[27] and its dimers.[109, 129] The intensity for the electronic transition $1^2B_{2u} \rightarrow 1^2B_{1g}$ is particularly predicted to be exceedingly weak. The oscillator strength associated to the $1^2B_{2u} \rightarrow 1^2B_{3u}$ electronic transition is $5{\cdot}10^{-3}$. The remaining electronic transitions from the 1^2B_{2u} ground state have in D_{2h} symmetry dipole-forbidden character. Nevertheless, the 1^2B_{3u} and 1^2A_u excited states may be populated from the 1^2B_{1g}. The oscillator values for the corresponding transitions are computed to be 0.094 and 0.005, respectively. The 1^2B_{2g} state, which lies at 4.45 eV above the ground state, might also be populated from the 1^2B_{3u} state placed vertically at 0.85 eV. The oscillator strength for the $1^2B_{3u} \rightarrow 1^2B_{2g}$ electronic transition is calculated to be $5{\cdot}10^{-3}$. In fact, it is the only chance to

get population directly on the 1^2B_{2g} state because the electronic transitions from the 1^2B_{2u} and 1^2B_{1g} states correspond to dipole-forbidden features.

The **BMU** has therefore been thoroughly studied yielding several spectroscopical signals over the near infrared, the visible, and the near ultraviolet domains showing an impressive flexibility that allows the system to accommodate a plethora of different excitations and subsequent photophysical/photochemical processes. Most of the excitations arise from the O-H-O central subunit, pointing towards this central subunit as the active site of the **BMU**. However, the oscillator strength associated to the electronic transitions given in this system is moderately small so the probabilities of population upon light irradiation are relatively low.

Given the slight chances of populating those states, and taking into account the relevance of π-stacked water, the model was further extended considering two **BMUs** in a π-stacking manner as previously done for the water dimer.[129]

In principle, when two **BMU** systems of spin multiplicity two (doublet states) get relatively close to each other in a parallel orientation, the π-stacked dimer is expected to get stabilized producing electronic states of singlet multiplicity. Therefore, the ground state of the π-stacked dimer formed by two **BMUs** has been initially optimized at the MP2(RHF)/6-31G** level. A conservative estimate of the stabilization gained by the π-stacked **BMU** dimer with respect to the two ground-state isolated **BMU** is larger than two eV. The main optimized geometrical parameters for the **BMU** dimer are displayed in Figure 6.4.

According to the standard convention defined above (see Figure 6.4), in the π-stacked dimer the central hydrogen atoms are located in the z axis, whereas the oxygen atoms of the fused rings are placed in the yz plane. The **BMUs** in the dimer are symmetrically displaced with respect to the inversion center of the full system along the z axis. As compared to the ground-state equilibrium of the **BMU** system, the bond distances in the dimer are in general different,

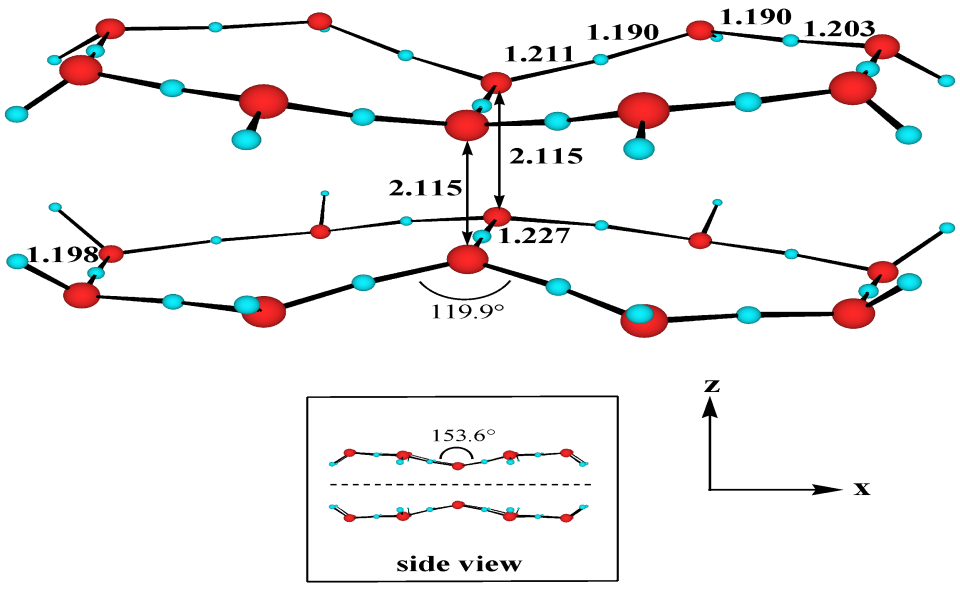

Figure 6.4: Optimized geometrical parameters obtained for the 1^1A_g state of the π-stacked dimer, formed by two symmetrically oriented basic molecular units (**BMUs**), at the MP2(RHF)/6-31G** level.

showing a distinct pattern. In particular, the O-H bond distance of the central subunit has been elongated by 0.046 Å with respect to that of the isolated fused monomer. The inter-atomic distance between the central oxygen atoms belonging to distinct **BMUs** is 2.115 Å, whereas the two central hydrogen atoms are separated by 2.056 Å.

The hexagonal ice-like model provides a pronounced ground-state stabilization, as it is comprised of two **BMUs** that are negatively charged radicals yielding a singlet ground state by pairing the two electrons while forming the $H_{38}O_{20}$ structure. The model also prospects the release of protons out of the zone where the ordered water is extended and points towards the experimental evidence available in this respect that registers the presence of protons in the region beyond ordered water.[150] The loss of protons can be assessed by

the two-proton deficit that is shown in the $H_{38}O_{20}$ stoichiometry, and hints towards a similar pattern for other rearrangements laid upon the same foundations.

At the ground-state geometry of the **BMU** dimer, the vertical electronic transition and the associated oscillator strengths have been determined at the CASPT2(12,15) level, which corresponds to 12 electrons distributed among 15 active molecular orbitals: 1, 1, 3, 2, 2, 1, 3, 2, belonging to the irreducible representations a_g, b_{3u}, b_{2u}, b_{1g}, b_{1u}, b_{2g}, b_{3g}, a_u, respectively. The remaining (inactive) orbitals are distributed as follows: 14, 11, 12, 10, 13, 11, 12, 10, for each irreducible representation of the D_{2h} symmetry group written above. For each symmetry, state-averaged CASSCF(12,15) computations have been carried out involving the four lowest roots, except for the singlet states of A_g symmetry where five roots have been employed. The core orbitals (3, 2, 3, 2, 3, 2, 3, 2) have been kept frozen in the CASPT2 step. The CASSCF(12,15) wave function provides enough flexibility to describe the electronic excitations of the π-stacked dimer up to 5 eV, implying the corresponding orbitals derived from the plus and minus linear combinations of the SONOs of the low-lying excited states of the **BMU** depicted in Figure 6.3.

The canonical orbitals obtained from the closed-shell Hartree-Fock (HF) calculation (14, 11, 14, 11, 14, 11, 13, 11) of the π-stacked dimer together with the related natural orbitals (NOs) coming out from state-averaged CASSCF computation are shown in Figure 6.5. It can be readily noticed that the lowest unoccupied molecular orbital (LUMO) has a negative orbital energy. It is not therefore surprising that the ground-state 1^1A_g is described at the CASSCF(12,15) level by a mixing of two configurations: the closed-shell HF and the doubly excited configuration $(H)^2 \rightarrow (L)^2$ with weights 58% and 36%, respectively. The H and L labels denote the HOMO-like and LUMO-like NOs derived from the five-root state-averaged CASSCF(12,15) wave function involving 717760 configuration state functions in the computation of the A_g states. The square of the coefficients related to the respective spin-adapted configurations corresponds to the weights given. In other words, the single closed-shell HF determinant is a somewhat limited description of the π-stacked

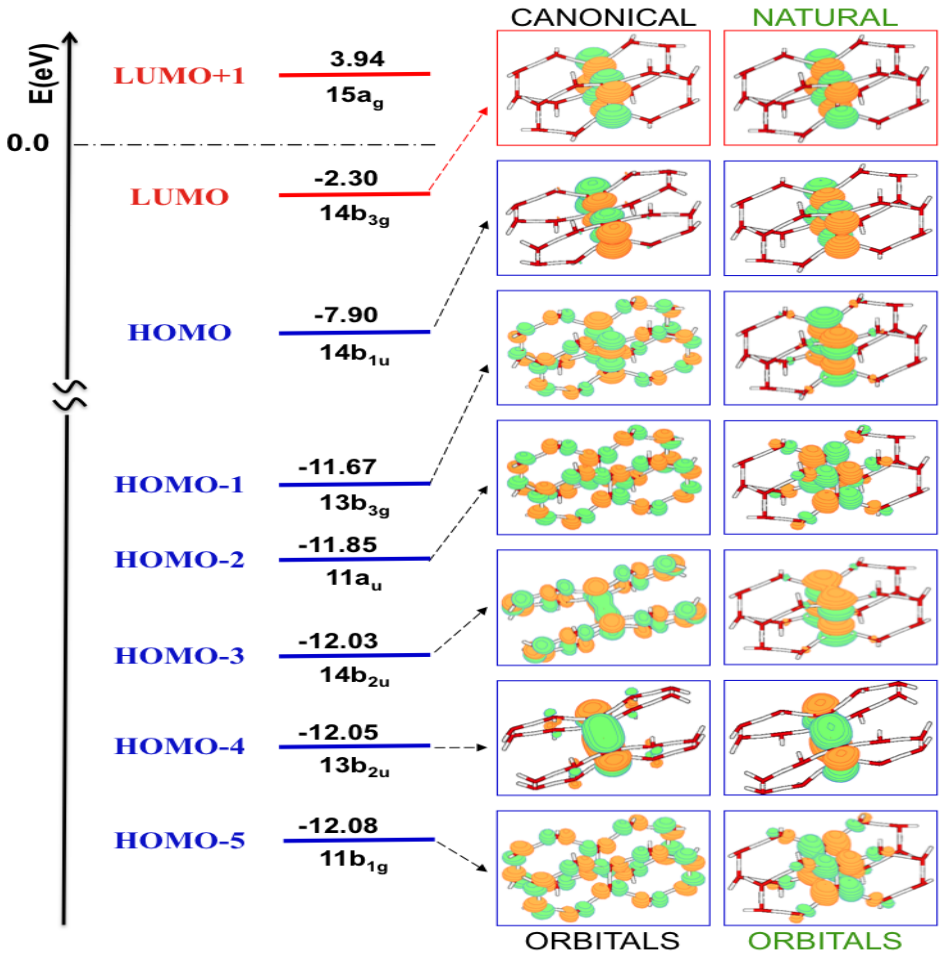

Figure 6.5: Canonical orbitals and the respective orbital energies derived from a closed-shell Hartree-Fock computation obtained at the MP2(RHF)/6-31G** optimized geometry of the ground-state 1^1A_g of the π-stacked dimer. The isodensity surface value for the orbitals is 0.030. For the sake of comparison, the related natural orbitals derived from the five-root state-averaged CASSCF computation of the A_g symmetry are also included.

dimer. According to the CASSCF(12,15) wave function, it can be better seen as a biradical species. Indeed, the occupation numbers for the HOMO-like and LUMO-like NOs of the 1^1A_g state are 1.23 and 0.80, respectively. Since the MP2 computation is based on the closed-shell HF description of the system, we were wondering whether the MP2 geometry for the ground state could be considered appropriate. Consequently, the inter-atomic distance between the central oxygen atoms was re-optimized at the CASPT2(12,15) level. A similar minimum to that already provided from the MP2 computation was obtained. It was therefore concluded that the MP2 wave function, built from doubly excited configurations from the HF reference (see eq. 3.5), is indeed capable to render the essentials of the electronic structure of the ground state of the **BMU** dimer.

Table 6.1 summarizes the energies of the lowest-lying singlet-singlet transitions (ΔE), the related oscillator strengths computed for the respective vertical electronic transitions, and the nature of the excited state employing the optimized geometry of the ground state calculated at the MP2(RHF)/6-31G** level. In order to complement the information on the nature of the low-lying excited states, the occupation numbers of the NOs for each excited state considered are listed in Table 6.2.

By inspection of Table 6.1, it is clear that no electronic transitions are found in the near infrared range (0.02-1.61 eV) for the $H_{38}O_{20}$ π-stacked dimer, unlike it occurs in the **BMU** system. Seven low-lying excited states are computed within the visible interval (1.61-3.22 eV). The lowest singlet excited state appears at 1.74 eV. The associated vertical electronic transition $1^1A_g \rightarrow 1^1B_{2u}$ has a small oscillator strength. The 1^1B_{2u} state has a biradical character with the electrons on the HOMO-like and (HOMO-1)-like NOs, consistent with the doubly excited nature of the state from those NOs to the LUMO-like NO. Whereas the HOMO-like NO is similar to the canonical HOMO, as it is also the case with the LUMO and HOMO-4 types, the (HOMO-1)-like NO is localized on the central O-H-O subunits as compared to the canonical HOMO-1. The same comment can be also applied to the HOMO-2, HOMO-3, and HOMO-5 types. The canonical MOs are simultaneously eigenfunctions of the Fock oper-

Table 6.1: Vertical transition energies (ΔE) computed for the $H_{38}O_{20}$ π-stacked dimer at the CASPT2(12,15)/6-31G** level.

State	ΔE(eV)	ΔE (nm)	Osc. strength	Assignment (Weight)[b]
1^1A_g (GS)				closed-shell HF (58%)
				(H)2 → (L)2 (36%)
1^1B_{2u}	1.74	713	0.0005	H-1, H → L,L (94%)
2^1B_{2u}	2.01	616	0.0090	H → L (88%)
1^1B_{3g}	2.24	554	*forbidden*	H-3, H → L,L (92%)
1^1B_{1g}	2.39	520	*forbidden*	H-2, H → L,L (95%)
2^1A_g	2.72	456	*forbidden*	H-1 → L (71%)
				H-3, H-4 → L, L (16%)
1^1B_{1u}	2.86	433	0.0001	H-3 → L (71%)
				H-4, H-1 → L, L (24%)
1^1B_{3u}	2.95	420	0.0005	H-2 → L (74%)
				H-4, H-5 → L, L (14%)
1^1A_u	3.40	365	*forbidden*	H-5, H → L, L (93%)
3^1A_g	3.47	357	*forbidden*	closed-shell HF (23%)
				$(H)^2 \rightarrow (L)^2$ (44%)
				$(H-4)^2 \rightarrow (L)^2$ (16%)
1^1B_{2g}	3.54	350	*forbidden*	H-5 → L (70%)
				H-4, H-2 → L, L (19%)
2^1B_{1u}	4.58	271[a]	0.2894	H-4 → L (71%)
				H-3, H-1 → L, L (19%)
4^1A_g	4.68	265	*forbidden*	$(H-1)^2 \rightarrow (L)^2$ (53%)
				$(H-3)^2 \rightarrow (L)^2$ (34%)

[a] Experimental value [42]

[b] Weight

ator and symmetry operator of the corresponding irreducible representation of the D_{2h} symmetry group, as it is also the case for the CASSCF wave functions. The orbital optimization procedure leads in the latter towards localization of the active MOs because it increases the exchange integral between the MOs involved, which is concomitant to the decrease of the energy in the variation

CASSCF procedure. As a result, the CASSCF NOs are localized in the O-H-O subunits.

In terms of the SONOs of the **BMU** system, in the π-stacked dimer the plus and minus linear combinations of the two SONO-1^2B_{2u} lead to the HOMO-3 and HOMO-1, respectively. The HOMO-4 and LUMO of the dimer are similarly derived from the two SONO-1^2B_{1g} type. On the other hand, the HOMO comes from the minus linear combination of the two SONO-1^2B_{3u} type. The plus linear combination of the two SONO-1^2B_{3g} yields the HOMO-5, being the corresponding minus linear combination related to the HOMO-2. The remaining combinations are related to electronic excited states in the $H_{38}O_{20}$ π-stacked system above 5 eV. The second-energy excited state 2^1B_{2u} is placed vertically at 2.01 eV. The oscillator strength associated to the transition is about 0.01, being the largest within the visible region. The 2^1B_{2u} state is described mainly by the one-electron promotion from the HOMO- to the LUMO-like NOs. The dipole-forbidden electronic transitions $1^1A_g \rightarrow 1^1B_{3g}$ and $1^1A_g \rightarrow 1^1B_{1g}$ are next predicted at 2.24 and 2.39 eV, respectively. The excited states are described mainly by simultaneous singly excited configurations to the LUMO-like NO. The excited states have a clear biradical character (cf. Table 6.2). The 2^1A_g state computed at 2.72 eV has a more complex wave function, as it is also the situation for the excited states 1^1B_{1u} and 1^1B_{3u}, which lie vertically at 2.86 and 2.95 eV, respectively. The latter two states are associated to weakly dipole-allowed electronic transitions.

Table 6.2: Occupation numbers of the natural orbitals (NOs) derived from the CASSCF(12,15) wave function of the $H_{38}O_{20}$ π-stacked dimer in terms of the topologically equivalent HOMO-like (H) and LUMO-like (L) NOs (see Figure 6.5).

State	LUMO-like	HOMO-like	(H-1)-like	(H-2)-like	(H-3)-like	(H-4)-like	(H-5)-like
1^1A_g	0.80	1.23					
1^1B_{2u}		1.02	1.00				
2^1B_{2u}	1.01	1.00					
1^1B_{3g}		1.02			1.01		
1^1B_{1g}		1.02		1.00			
2^1A_g	1.51		0.94		1.67		
1^1B_{1u}	1.26		1.73		1.25	1.75	
1^1B_{3u}	1.16			1.14		1.85	1.85
1^1A_u		1.04					1.00
3^1A_g	1.44	1.04				1.66	
1^1B_{2g}	1.21		1.79	1.79		1.80	1.19
2^1B_{1u}	1.20		1.78		1.78	1.22	
4^1A_g			0.78		1.28		

In the near ultraviolet region, five electronic transitions are predicted, one of them computed with a relatively large oscillator strength, whereas the remaining are dipole-forbidden features. The 1^1A_u state lies at 3.40 eV above the ground state. It has a biradical character with the unpaired electrons on the HOMO-like and (HOMO-5)-like NOs. In contrast, the 3^1A_g state placed at 3.47 eV has multiconfigurational character, as it also occurs with the nearby 1^1B_{2g} state computed to be at 3.54 eV. The electronic transition $1^1A_g \rightarrow 2^1B_{1u}$ is predicted at 4.58 eV (271 nm) with a computed oscillator strength of 0.29. The finding can be clearly assigned to the experimental band maximum recorded at ∼270 nm (4.59 eV).[42] The theoretical prediction shows a remarkably good match with the observed band maximum. It is not surprising, because the computational strategies employed normally render vertical excitation energies ± 0.2 eV of well-defined experimental data.[66] The 2^1B_{1u} state is described mainly by the (HOMO-4)-like → LUMO-like singly excited configuration with a weight of 71%. The doubly excited configuration from the (HOMO-3)-like and (HOMO-1)-like to the LUMO-like NO has also a non-negligible contribution (19%). Finally, the 4^1A_g state involving doubly excited configurations from the (HOMO-1)-like and (HOMO-3)-like NOs to the LUMO-like NO is vertically found at 4.68 eV.

For the $H_{38}O_{20}$ system, plenty of signals have also been predicted. All excitations appear delocalized in the π-stacked O-H-O bridging subunits of both **BMUs** and thus mirror the behavior already analyzed for the π-stacked water dimer.[129] For the π-stacked **BMUs** a transition is computed to be at ∼4.58 eV with an associated oscillator strength of 0.29, which provides a clear assignment for the ∼4.6 eV experimental fingerprint associated to the presence of ordered water. As compared to the π-stacked water dimer, the oscillator strength is rather large making it more consistent with the experimental evidence and thus highlighting the importance of an extended model that yields a closer representation of the experimental findings. The **BMU** system under study has very unique properties as they are negatively charged anions and the treatment of such systems computationally is regarded as highly challenging. The theoretical method employed has previously been assessed to yield sound results for these cases and has therefore been used.

Conclusions The spectroscopic features of the hexameric ice-like model of structural water have been examined. Firstly, the basic molecular unit (**BMU**) of the model has been identified as the negatively charged radical species $[H_{19}O_{10}]^{\cdot -}$. The results show that it is a versatile unit, having two low-lying excited states in the near infrared region, two in the visible range, and two in the near ultraviolet zone. The main absorption feature is found however for the π-stacked dimer formed by two **BMUs**, the $H_{38}O_{20}$ system, which has a pronounced stabilization in the ground state with respect to the isolated **BMUs**. Most of the states of the π-stacked dimer, including the ground state, have a marked biradical character. In all the cases, the one- and two-electron promotions are located on the central O-H-O subunits of the fused hexagons. A total of twelve singlet→singlet electronic transitions have been computed for the $H_{38}O_{20}$ system, seven in the visible range and five in the near ultraviolet region. In the latter, the electronic transition $1^1A_g \rightarrow 2^1B_{1u}$ of π character is predicted at 4.58 eV (271 nm) with a computed oscillator strength of 0.29 in agreement with the experimental band maximum recorded at ∼270 nm (4.59 eV).[42] The 2^1B_{1u} state is predominantly (71%) described by the singly excited configuration from the plus linear combination of the out-of-phase π orbitals of the oxygen atoms placed in the central O-H-O subunit of the two fused hexagons, (HOMO-4)-like, to the respective minus linear combination (LUMO-like). In sum, according to the present results, it can be safely concluded that the hexameric ice-like model of structured water is fully consistent with the observed feature at ∼270 nm (4.59 eV). The key to understand structured water at the molecular level seems to be related with the ability of water to build π-stacked aggregates (formed at least by two water molecules)[129] under certain conditions, a possibility virtually missing in the literature up to date. In this sense, distinct π-stacked crystalline-like structures can in addition be envisaged to be consistent with the characteristic spectroscopic fingerprint determined experimentally, depending their relative relevance on the actual environmental conditions.

The hexagonal ice-like model has been built based upon what was learnt in the previous chapter from the studies of the different intermolecular interactions present in water. As can be seen, the molecular counterpart of the ∼4.6

eV fingerprint is again associated to the presence of a π-stacking species, in this case a hexameric-like system that prospects the possibility of extending itself over layers as well as stacking them thus providing the aforementioned spectroscopical feature.

The large stabilization effect present in the $H_{38}O_{20}$ provides a sound argument to argue about the possibility of such an species to be present in ordered water. Together with the experimental evidence, it can therefore be concluded that upon light absorption, the excess energy builds structural order and drives charge separation.[129]

The essentials of the interaction are expected to be kept in the presence of an external perturbation due to the charged surroundings or the application of an external electric field, as it has already been shown for the π-stacked water dimer.[129] The mechanism governing the fluorescence, another of the key properties of this ordered water, is currently being under study. Therefore, the model holds the essential features previously reported unveiling hidden key aspects about the photophysical and photochemical processes implied in ordered water.

Chapter 7

Water oxidation properties through its Ionization Potentials

> The formula for water is H_2O. Is the formula for an ice cube H_2O squared?
>
> *Lily Tomlin*

In this chapter, the effect of ionizing radiation in water has been addressed through the systematic study of the smallest existing water clusters aiming towards the extrapolation of their properties onto bigger systems with similar hydrogen-bonded coordination patterns. In particular, the clusters considered are the dimer, trimer, tetramer, and pentamer, as well as the book, cage, ring, and prism conformers of the hexamer. The properties under study are the different magnitudes related to the *ionization potentials* of the water clusters

and the analysis has been carried out through the comparison between clusters whilst increasing their size, as well as their hydrogen bond connectivity.

A deeper knowledge of the ionization process of water presents itself as a very promising path towards the proper characterization of solvation mechanisms related to biochemical and industrial processes of paramount importance. The effect of such radiation bears the quality of removing an electron, and subsequently starting a whole new chain of events such as *dissociative electron attachment* processes that have been considered over the next chapter and that can end up in DNA-strand breaks. To that end, a broader understanding of the oxidation processes related to the cationic water species, as well as the proper assignment of their ionization values yield a unique way of studying the underlying mechanisms responsible for the ionization process. The use of the water clusters as model systems offers a systematic way of studying water's behavior by increasing the system's size as well as its hydrogen-bond connectivity. The study here presented provides a benchmark to validate the computational approach used throughout the whole Thesis.[125]

Ionizing radiation in water clusters lead to the following process:

$$(H_2O)_n + h\nu \rightarrow [(H_2O)_n]^+ + e^- \tag{7.1a}$$

$$\rightarrow (H_2O)_{n-1}H^+ + OH + e^- \tag{7.1b}$$

in which the initially formed cationic species given in eq. 7.1a undergoes an intra-cluster charge redistribution favoring proton transfer processes and the subsequent loss of the OH (eq. 7.1b). Ionization then drives charge redistribution, a process vastly important in the mechanisms towards ordering water.

The present study contemplates the assessment of a series of magnitudes related to the ionization potential in the water clusters, represented in Figure 7.1. As can be seen in the figure, the neutral water clusters (C) absorb ionizing light bringing them to their cationic state (C^+). The vertical transition from the FC region is known as the *Vertical Ionization Potential* (VIP), which yields the $[(H_2O)_n]^+$ cationic species. The excess energy in the cationic state

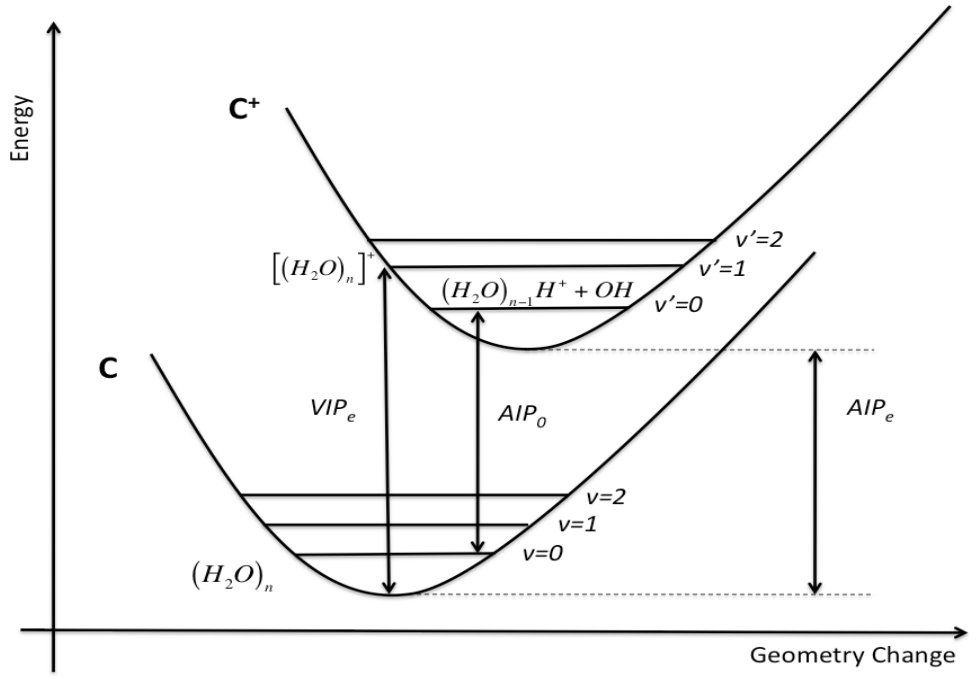

Figure 7.1: IP diagram. Definitions of the theoretical magnitudes related to IP are graphically shown through the electronic and vibrational potential energy levels. Those magnitudes are VIP_e, AIP_e, and AIP_0 (see text).

is released through proton transfer processes until reaching the minimum of the potential energy surface, giving rise to the $(H_2O)_{n-1}H^+ + OH$ species. The difference in energy from the minimum of both wells is given by the *Adiabatic Ionization Potential* (AIP), which can be computed both with the ZPE correction or without it. When it does contain the ZPE correction, the magnitude is related to the 0-0 or T_0 experimental band, and it represents the computed value that actually has an experimental counterpart.

Computational Details Geometry optimizations have been carried out with no symmetry restrictions at different levels of theory. Both the ground

and the lowest cationic states of the water clusters have been optimized at the MP2, CCSD, CASSCF, and CASPT2 levels of theory, making use of numerical gradients in the latter. The minima obtained have then been characterized by computing second-order derivatives at the MP2 level, for the MP2 and CCSD geometries, and at the CASSCF level, for the CASSCF and CASPT2 structures. The MP2 and CCSD calculations have been performed with the aug-cc-pVDZ Dunning basis set and the atomic natural orbital ANO-L basis set contracted to O[4s3p1d]/H[2s1p] (hereafter ANO-L 431/21) has been employed with the CASSCF and CASPT2 methods. ZPE corrections have been included through the second-order derivatives of the energy using the harmonic approximation. Further calculations determining third and fourth-order derivatives have also been performed at the MP2 level to test the adequacy of the harmonic approximation in the water clusters, which were taken into account by Mella *et al.* as highly anharmonic systems.[154] Additional CCSD(T) calculations have been performed at the CCSD optimized geometries with the aug-cc-pVDZ and aug-cc-pVTZ basis sets. At the CASSCF and CASPT2 equilibrium structures, computations with the ANO-L basis set contracted to O[4s3p2d1f]/H[3s2p1d] (hereafter ANO-L 4321/321) have also been carried out. The performance of the IPEA correction within the CASPT2 method has also been analyzed by using two different values: 0.00 and 0.25 au, as done in previous works.[56] The multiconfigurational CASSCF calculations for the geometry optimizations have been carried out with an active space comprising both the topologically equivalent highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) of the single water molecule, that is, the 2p lone pair orbital perpendicular to the plane of the molecule (referred as $1b_1$ in the C_{2v} point group) and the anti-bonding sigma orbital (referred as the $4a_1$ in the C_{2v} point group), meaning two electrons in two orbitals for each water molecule comprising in the cluster. Thus, in order to exert an equivalent treatment on all systems, four electrons distributed into four orbitals, 4-in-4, 6-in-6, 8-in-8, 10-in-10, and 12-in-12 orbitals have been considered in the active space for the dimer, trimer, tetramer, pentamer, and the four hexamers, respectively. The quantum-chemistry packages of software used in the present study are Gaussian-09,[108] employed for the calculations

at the MP2, CCSD, and CCSD(T) levels of theory, and MOLCAS-7,[69] which has been used for the CASSCF and CASPT2 computations.

Results and Discussion To begin with, the clusters are optimized at the FC-equilibrium structure. All geometries obtained can be seen in Figure 7.2. All clusters considered in this study have been previously characterized experimentally.[28–30] In order to study the hydrogen bond properties of the different clusters in the ground state, the interatomic O-O distance has been used to unveil the full extent of the hydrogen bond cooperativity in the different systems and its variation whilst increasing the cluster's size.[31]

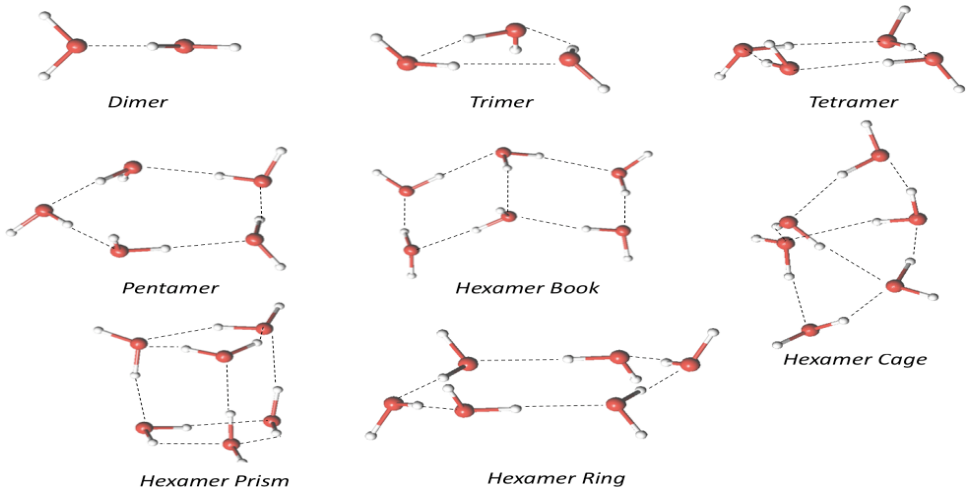

Figure 7.2: Equilibrium structures for the ground state of the neutral water clusters.

Table 7.1 compiles the present computed average O-O distances for the neutral water clusters, together with the corresponding theoretical values and experimental measurements reported in the literature. CCSD provides with the closest results to the experiment, with an average deviation, after considering all the clusters, of 0.01 Å. CASPT2, MP2, and DFT methods follow, with average differences of 0.05 Å. Both MP2 and CASPT2 give rise to sim-

ilar values, whereas CCSD gives the best result, as expected from a single reference problem as the neutral water clusters. Methods that do not include dynamic electron correlation, such as the HF and CASSCF, lead to the largest deviations, with differences up to 0.09 Å. Such behavior reflects, as expected, a better description of the hydrogen-bonded systems at hand with dynamic-correlated methods. From a general point of view, all methods show the same trend for the O-O distances in the different clusters studied. Thus, the dimer has the most separated adjacent water molecules, and the O-O length decreases with the increasing size of the cyclic clusters: trimer > tetramer > pentamer > hexamer ring. Regarding the hexamer clusters, the oxygen distances vary depending on the average hydrogen-bond-wise coordination order of the water molecules in the cluster (cf. Figure 7.2), rising from the ring to prism conformations, in agreement with previous available results.[155]

Reported values on liquid water show an average O-O distance of 2.85 Å,[35] being in close agreement with the distances obtained for the hexamer prism at the CASPT2 and MP2 levels of theory and hexamers cage and prism at the CCSD level. These hexamer clusters are the most coordinated forms within the six-water-members conformations studied and hence the ones that relate closer to the average of 3.2-3.4 hydrogen bonds per water molecule given for bulk water.[35]

It can be concluded that the interatomic O-O distance decreases while increasing the size of the cluster, which implies a stronger cooperative effect in the hydrogen bonding network. The high-level methods behave better than those lacking dynamic correlation, which is deemed important to describe hydrogen bonding interactions.

Table 7.1: Average R(O-O) distances (in Å) of the neutral water clusters obtained by different theoretical methods, together with the corresponding experimental data available in the literature.

System/Method	Exp[a]	HF[b]	B-LYP[b]	CASSCF[c]	MP2[c]	CCSD[c]	CASPT2[c]
$(H_2O)_2$	2.976	3.032	2.939	3.041	2.917	2.944	2.897
$(H_2O)_3$	2.845	2.927	2.808	2.936	2.802	2.836	2.794
$(H_2O)_4$	2.782	2.880	2.743	2.869	2.748	2.785	2.734
$(H_2O)_5$	2.760	2.867	2.727	2.852	2.731	2.768	2.713
$(H_2O)_6^{book}$				2.922	2.772	2.808	2.757
$(H_2O)_6^{cage}$	2.820			2.965	2.812	2.848	2.797
$(H_2O)_6^{prism}$				3.001	2.840	2.875	2.834
$(H_2O)_6^{ring}$	2.759	2.855	2.714	2.840	2.722	2.757	2.697

[a]References [28, 156]
[b]Reference [31]
[c]Present Work

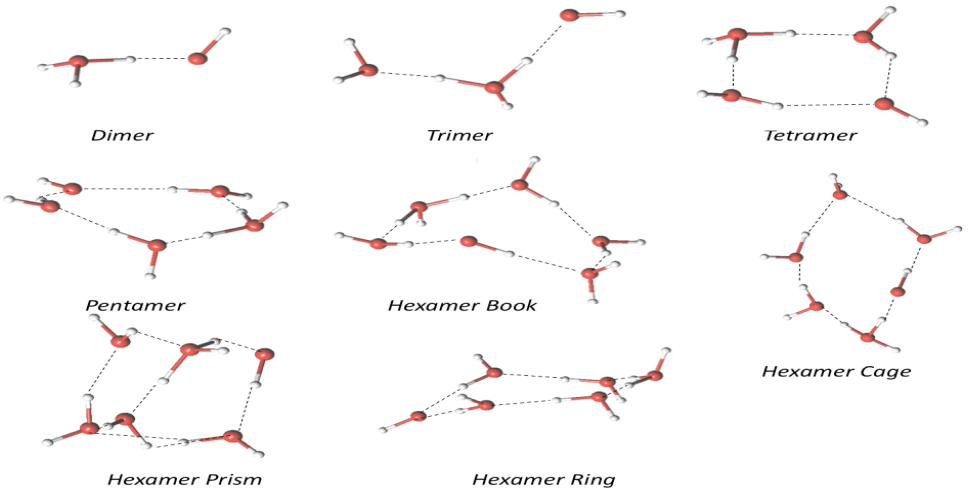

Figure 7.3: Equilibrium structures for the lowest-energy cationic state of the water clusters.

Cationic structures of water clusters have also been optimized at the same levels of theory as those employed for the neutral systems. The results allow analyzing both the performance of the different methods and the structural changes which occur in water clusters after the ionization process. Figure 7.3 displays the obtained distortions. All structures end up as a combination of $(H_2O)_{n-1}H^+$ and OH species, where n denotes the number of water molecules, in agreement with the results reported earlier, in which geometry optimizations or dynamics calculations have been carried out reaching similar structures.[157–159] A common trend is shown in Figure 7.3 for the tetramer, pentamer, and hexamer book clusters: the double proton transfer via adjacent water molecules. The phenomenon has also been noted in other theoretical studies that describe it as a very fast process occurring directly after ionization in a femtosecond timescale.[157] The hexamer ring exhibits a triple proton transfer while the remaining of the clusters, the dimer, hexamer cage, and hexamer prism, do have just one proton transfer.

Table 7.2: Average R(O-O) distances (in Å) of the cationic water clusters obtained by different theoretical methods.

System/Method	CASSCF[a]	MP2[a]	CCSD[a]	CASPT2[a]
$(H_2O)_2$	2.553	2.526	2.545	2.498
$(H_2O)_3$	3.152	3.111	3.123	3.083
$(H_2O)_4$	3.244	2.724	2.742	2.709
$(H_2O)_5$	3.372	2.710	2.371	2.691
$(H_2O)_6^{book}$	2.767	2.686	2.708	2.669
$(H_2O)_6^{cage}$	2.761	2.675	2.698	2.657
$(H_2O)_6^{prism}$	3.113	2.895	2.910	2.845
$(H_2O)_6^{ring}$	3.093	2.709	2.731	2.689

[a]Present Work

Table 7.2 compiles the average values for the intermolecular O-O distances in the cationic water clusters. Due to the lack of experimental data regarding such clusters, CCSD results will be used as reference. MP2 and CASPT2 methods show good agreement with average deviations around 0.02 Å and 0.04 Å, respectively. The description of cationic structures provided by non-correlated methods, such as the CASSCF, is worse than in neutral systems, with an average difference of 0.23 Å. This is especially noticeable in tetramer, pentamer, hexamer prism, and hexamer ring, where differences can reach up to 0.6 Å and the cycle opens up becoming almost broken, as it occurs for the water trimer and hexamer prism by using all methods. Previous theoretical works with molecular dynamics agree with the opening structures for the water trimer,[160] thus backing up the results obtained in the current study. Moreover, the authors also suggest a decrease in the O-O distances whilst increasing the cluster size, a factor attributed previously to the increase of the hydrogen-bond energy.[160] The same general trend in the distances can be seen in Table 7.2, with the exceptions of the dimer and hexamer prism. The correspondence with the hydrogen-bond energies can therefore be discarded, since not all the cationic neither the neutral systems behave accordingly.

Comparison between the most accurate results present in Table 7.1 and 7.2 (MP2, CCSD, and CASPT2) shows a decrease of the average oxygen distances from neutral to cation systems, with exception of the trimer and hexamer prism, which suffer a ring opening as stated above. Changes are more pronounced at the CCSD level, with an average deviation of 0.06 Å, followed by the CASPT2 and MP2 methods, with differences of 0.05 and 0.04 Å, respectively. The cationic water clusters show a similar trend as compared with the neutral forms, having the shortest lengths in the hexamer ring and the largest in the trimer instead of the dimer as it occurs for the neutral forms. Regarding the non-ring hexamers, smaller deformations appear with respect to the neutral structures, likely due to the increase of the hydrogen bonds per molecule that allows these clusters to be less flexible.

Table 7.3 compiles the present computed results for the lowest-energy VIPs of the water clusters. From the careful analysis of the performance of different quantum-chemistry methods and computational strategies, the CCSD(T) methodology with the aug-cc-pVTZ basis set can be selected as reference, since the electronic structure of all the neutral and cation forms are well described by a single reference wave function, no problems due to spin contamination are present within the CCSD(T) calculations (the eigenvalues of the S^2 operator in the cation systems have values close to 0.75), and the T_1 diagnostic is favourable ($T_1 < 0.02$). The scarce experimental data for the IPs of water clusters refer to *appearance energies*, which are associated in the literature to the VIP.[161] The CCSD(T) results have however a large deviation, over one eV in certain cases. The reliability of such experimental technique to measure the vertical ionization properties is therefore predicted to be low, and its values can be seen as a threshold at which the ionization might start, providing thus a range for the ionization.

Table 7.3: Experimental *appearance energies* (in eV) and low-lying vertical ionization potentials (VIP_e, in eV) of water clusters obtained by different theoretical methods.

Method	$(H_2O)_2$	$(H_2O)_3$	$(H_2O)_4$	$(H_2O)_5$	$(H_2O)_6^{book}$	$(H_2O)_6^{cage}$	$(H_2O)_6^{prism}$	$(H_2O)_6^{ring}$
Experimental[a]	11.25	11.15	10.94	10.94		10.93		
BhandHLYP/6-31++G**[b]	10.59	10.77	10.67	10.43		10.46		
LC-ωPBE/6-31++G**[b]	11.06	11.10	10.98	10.87		10.73		
BhandHLYP/aug-cc-pVTZ//B3LYP/cc-pVTZ[c]	11.50	10.88	11.87	11.67			11.27	
B2PLYP/aug-cc-pVTZ//B3LYP/cc-pVTZ[c]	11.81	11.94	11.43	11.85			11.45	
B2GP-PLYP/aug-cc-pVTZ//B3LYP/cc-pVTZ[c]	11.80	11.98	11.76	11.81			11.59	
EOM-IP-CCSD/6-311++G**[d]	11.40							
ADC(3)[e]	11.90	12.40	12.18					
MP2/6-31++G**[d]		12.32	12.10	11.96				
MP2/aug-cc-pVTZ//B3LYP/cc-pVTZ[c]	11.98	12.55	12.06	12.34			11.90	
CCSD/aug-cc-pVTZ//B3LYP/cc-pVTZ[c]	11.71	12.30	12.47					
CCSD(T)/aug-cc-pVTZ//B3LYP/cc-pVTZ[c]	11.81	12.21	12.19					
MP2/aug-cc-pVDZ[f]	11.79	12.37	11.93	12.17	11.69	12.00	11.75	11.58
PMP2/aug-cc-pVDZ[f]	11.74	12.31	11.78	12.12	11.64	11.77	11.69	11.47
CCSD/aug-cc-pVDZ[f]	11.55	12.08	12.33	11.90	11.44	11.77	11.46	12.34
CCSD(T)//CCSD/aug-cc-pVDZ[f]	11.65	12.12	12.13	11.95	11.50	11.73	11.50	12.02
CCSD(T)/aug-cc-pVTZ//CCSD/aug-cc-pVDZ[f]	11.79	12.27	12.27	12.10	11.69	11.99	11.65	12.14
CASPT2//CASSCF/ANO-L 431/21[f]	11.82	12.32	12.27	12.16	11.67	11.94	11.80	12.16
CASPT2/ANO-L 4321/321//CASSCF/ANO-L 431/21[f]	11.87	12.36	12.31	12.22	11.71	11.98	11.87	12.22
CASPT2/ANO-L 431/21[f]	11.66	12.22	12.15	12.08	11.49	11.79	11.63	12.14
CASPT2/ANO-L 4321/321//CASPT2/ANO-L 431/21[f]	11.72	12.26	12.19	12.19	11.53	11.81	11.67	12.09

[a] Reference [161]
[b] Reference [162]
[c] Reference [163]
[d] Reference [164]
[e] Reference [165]
[f] This Work

The overall theoretical methods show a large dispersion of values, mainly due to the DFT results. Several functionals have been used within this method, showing different VIP responses and large sensitivity to the type of basis set, as can be seen from the comparison of the BHandHLYP results with the 6-31++G** and aug-cc-pVTZ basis sets (differences ranging 0.11-1.24 eV). Similar conclusions on the performance of DFT in the determination of ionization potentials and electron affinities were pointed out in previous studies, focused in those cases on DNA nucleobases and nucleotides.[56, 57] The worst DFT results obtained for the VIPs of water clusters correspond to the BHandHLYP/6-31++G** level,[162] which underestimates the VIP by 1.20-1.67 eV, with respect to the CCSD(T) results. The B2GP-PLYP functional gives rise to the best estimates with an average deviation of 0.25 eV.[163] Meanwhile, the only value reported in the literature for the dimer cluster at the EOM-IP-CCSD level is 0.39 eV below the reference.[164] ADC(3), which is a propagator method that evaluates the one-particle Green function according to a third order algebraic diagrammatic construction scheme,[165] provides results in close agreement with CCSD(T) for the water clusters up to tetramer. The MP2 and CC levels of theory have also been employed in the literature to determine the VIP of the small clusters and the hexamer prism (in the last case with the MP2 method), providing a good agreement with our reported VIPs. Deviations are due to the use of DFT optimized geometries in the values from the literature.

The present computed VIPs include several strategies within the MP2, CC, and CASSCF/CASPT2 methods. Figure 7.4 illustrates the analyses performed via six histograms. Both MP2 and Projected MP2 (PMP2) VIP values have been obtained, rising mostly negligible differences except for the tetramer, hexamer cage, and hexamer ring systems, where the differences can go over 0.2 eV, as for the hexamer cage (cf. Figure 7.4a). Such differences seem to come mainly due to spin contamination, as all three systems show a rather big deviation in their S^2 operator value. However, in this case the projected technique enlarge the deviations with respect to the CCSD(T) reference data by decreasing the VIPs. MP2 values fare fairly well in comparison to CCSD and the same basis set (aug-cc-pVDZ), with the exception of the tetramer and

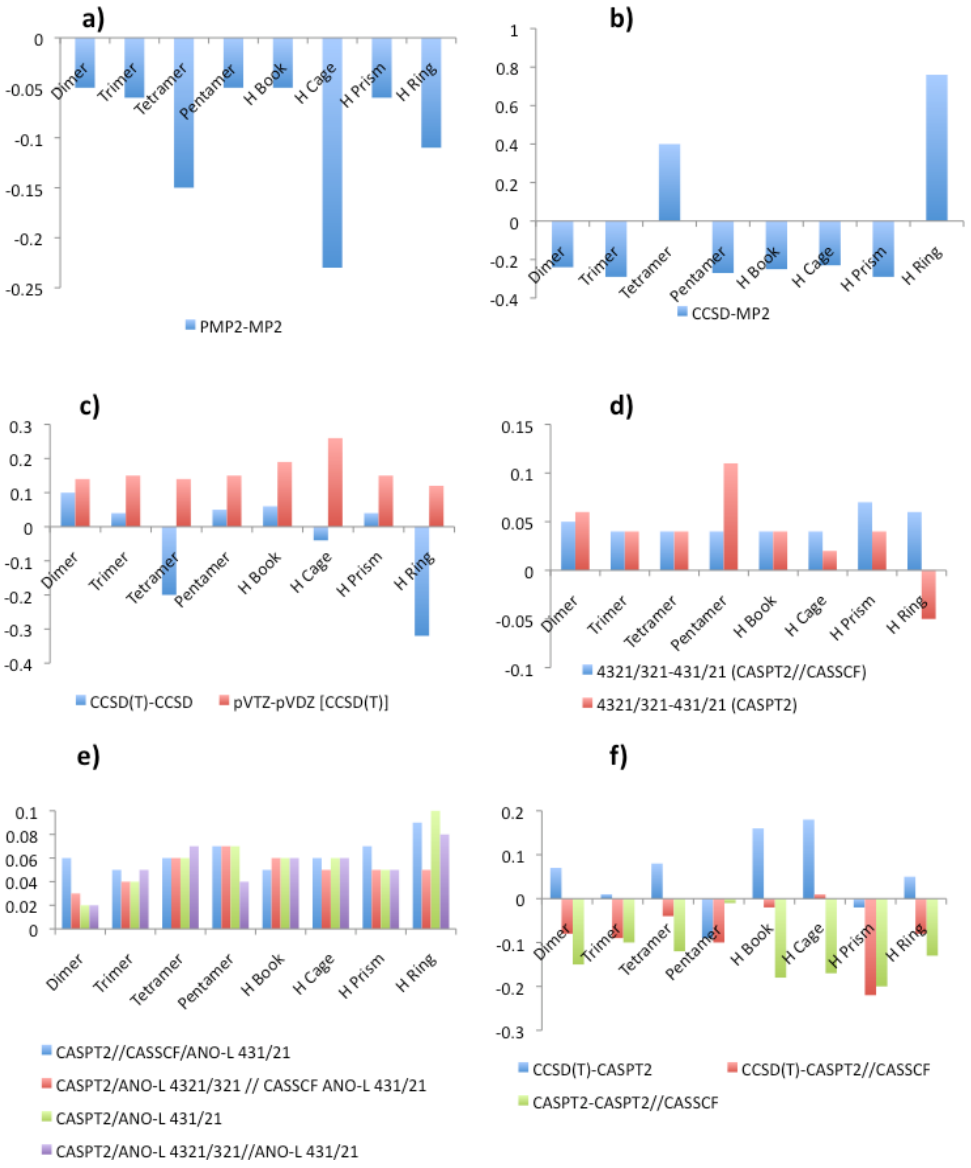

Figure 7.4: Comparative analyses on the performance of the most relevant computational strategies employed in the present work to determine the VIPs of water clusters. All energies are given in eV. See text for further details.

the hexamer ring systems, giving closer VIPs to the CCSD(T), which can be attributed to error cancellation in the MP2 results. Systematic decreases of the VIP of around 0.2 eV are found from MP2 to CCSD, with the exception of the tetramer and hexamer ring, which show the largest energy differences and with opposite sign, being especially important in the latter one (cf. Figure 7.4b). Both systems are also highly affected by the projected values as stated above, which hint to the spin contamination problem as a possible explanation of this large difference. Regarding the CC calculations, both the introduction of the perturbative estimate of triple excitations and the enlargement of the basis set have been analyzed. The first factor brings a small increase of the VIPs below 0.1 eV for almost all the systems, pointing out to the convergence of the dynamical correlation (see Figure 7.4c). The tetramer and hexamer ring are again exceptions, showing negative VIP decreases of -0.20 and -0.32 eV, respectively. Hexamer cage also shows a decrease in the energy, albeit it is in comparison rather small. The basis set factor increases an average of 0.16 eV the VIPs for all systems, being especially relevant for the hexamer cage cluster, in which the energy is improved by 0.26 eV (cf. Figure 7.4c). As for the CASPT2//CASSCF and CASPT2 methodologies, small changes in the VIP are obtained when the basis set is improved from ANO-L 431/21 to ANO-L 4321/321: +0.11 eV in the CASPT2 values for the pentamer and below 0.06 eV in all the other cases (cf. Figure 7.4d). These differences are much smaller than those obtained in the analysis of the aug-cc Dunning basis sets within the CC methodologies, which highlight the better quality of the ANO-type basis sets. In fact, the general contraction scheme of the ANO-type basis sets allows a larger flexibility, which leads to a faster convergence of the property under study. As can be seen from Table 7.3 and Figure 7.4d, the ANO-L 431/21 basis set already provides with good estimates of the VIP values, which is very promising for the study of oxidation processes in water clusters due to the lower computational demands required for this basis set in comparison to ANO-L 4321/321. Within the CASPT2//CASSCF and CASPT2 computational strategies, the use of an IPEA correction with a value of 0.25 instead of 0.00 au (standard CASPT2) give rise to very small differences of 0.06 eV in average, in all the cases below a tenth an eV (cf. Figure 7.4e). Since the IPEA correction

(0.25 au) does not significantly change the results, its use is not really required in studies on oxidation processes in water clusters.

Several conclusions can be drawn from the comparison of the best VIP values within the CCSD(T), CASPT2//CASSCF, and CASPT2 methodologies. The CASPT2 method with the ANO-L 4321/321 basis set perform rather well compared with CCSD(T), giving average energy differences of 0.08 and -0.06 eV for the values coming from the CASSCF and CASPT2 optimized geometries, respectively. Both CASPT2//CASSCF and CASPT2 VIPs deviate from each other an average of 0.14 eV. It is worth noting that even when the systems display a clear single reference character, multireference perturbation theory methods outperform those based on single reference Møller-Plesset perturbation theory. This is probably a consequence of the fact that there is no spin contamination in the CASSCF due to the way it is constructed through pure spin functions.

Results for the AIP property of the water clusters are summarized in Table 7.4 and Figure 7.5. Theoretical and experimental data, both obtained within this work and present in the literature, have been analyzed showing the values a much shorter dispersion as compared to the VIPs. To the best of our knowledge, only experimental measurements for the AIP of the dimer cluster have been reported.[166] The value is in agreement with the theoretical results obtained at the higher levels of theory. Similar to the VIP analysis of theoretical data, the CCSD(T) method with the aug-cc-pVTZ basis set will be used as reference, providing us with the most accurate AIP within the present computed and reported data.

Table 7.4: Experimental and theoretical lowest-lying adiabatic ionization potentials (AIP_0, in eV) of water clusters.

Method	$(H_2O)_2$	$(H_2O)_3$	$(H_2O)_4$	$(H_2O)_5$	$(H_2O)_6^{book}$	$(H_2O)_6^{cage}$	$(H_2O)_6^{prism}$	$(H_2O)_6^{ring}$
Experiment[a]	10.8							
BhandHLYP/aug-cc-pVTZ//B3LYP/cc-pVTZ[b]	10.42	9.66	9.25	8.96			8.75	
B2PLYP/aug-cc-pVTZ//B3LYP/cc-pVTZ[b]	10.75	10.02	9.64	9.34			9.16	
B2GP-PLYP/aug-cc-pVTZ//B3LYP/cc-pVTZ[b]	10.75	10.04	9.66	9.37			9.27	
MP2//UHF/4-31++G**[c]	10.7	9.9	9.7	9.4				9.2
UMP2/6-31++G(1d,1p)[d]			9.57	9.25				
MP2/aug-cc-pVTZ//B3LYP/cc-pVTZ[b]	10.98	10.28	9.89	9.59			9.48	
CCSD/aug-cc-pVTZ//B3LYP/cc-pVTZ[b]	10.69	9.97	9.81					
CCSD(T)/aug-cc-pVTZ//B3LYP/cc-pVTZ[b]	10.78	10.08	9.94					
MP2/aug-cc-pVDZ[e]	10.82	10.07	9.79	9.48	9.35	9.46	9.49	9.31
PMP2/aug-cc-pVDZ[e]	10.77	10.03	9.74	9.43	9.31	9.41	9.44	9.26
CCSD/aug-cc-pVDZ[e,f]	10.56	9.81	9.51	9.21	9.08	9.19	9.23	9.04
CCSD(T)//CCSD/aug-cc-pVDZ[e,f]	10.62	9.88	9.59	9.28	9.17	9.31	9.31	9.09
CCSD(T)/aug-cc-pVTZ//CCSD/aug-cc-pVDZ[e,f]	10.75	10.00	9.71	9.39	9.27	9.38	9.41	9.22
CASPT2//CASSCF/ANO-L 431/21[e]	10.67	9.94	9.69	9.45	9.15	9.30	9.45	9.31
CASPT2/ANO-L 4321/321//CASSCF/ANO-L 431/21[e]	10.77	10.05	9.76	9.54	9.27	9.41	9.54	9.37
CASPT2/ANO-L 431/21[e,g]	10.64	9.92	9.59	9.29	9.03	9.31	9.42	9.12
CASPT2/ANO-L 4321/321//CASPT2/ANO-L 431/21[e,g]	10.77	10.08	9.71	9.42	9.19	9.46	9.57	9.24

[a] Reference [166]
[b] Reference [167]
[c] Reference [168]
[d] Reference [158]
[e] This Work
[f] ZPE computed at the MP2/aug-cc-pVDZ level
[g] ZPE computed at the CASSCF/ANO-L 431/21 level

Several theoretical approaches have been employed in the literature for the calculation of AIPs, going from DFT to perturbation theory and CC methods with different basis sets. The DFT values range again quite highly, obtaining the best values with the B2GP-PLYP functional as in the VIPs.[167] The MP2 methods perform quite similarly with Pople and Dunning basis sets, showing the largest deviations for the values computed at B3LYP optimized geometries.[158, 167, 168] The CCSD(T) values reported in the literature have larger energy differences with respect to those obtained here as compared to the VIPs, reaching a maximum deviation of 0.23 eV in the tetramer. This is probably due to the fact that the AIP property is more sensitive to the method employed for geometry optimizations, which is more accurate in the present study.

Within the computational strategies employed in the present work to determine the AIPs, we analyzed also the MP2 projected technique, the inclusion of triples excitation through perturbation theory in the CCSD method, the CASPT2//CASSCF and CASPT2 approaches, the effect of increasing the basis set quality, and the IPEA correction within the CASSCF/CASPT2 approach (see Figure 7.5). A rather small systematic reduction of the AIPs of around 0.05 eV is found when the PMP2 method is used instead of the non-projected MP2, which is consistent with the fact that no spin contamination is present in this case, in contrast to the VIP calculations (cf. Figure 7.5a). On the other hand, error cancelation similar to the VIP seems to take place within the MP2 method, as can be concluded from the inspection of Figure 7.5b and the fact that the CCSD method brings the AIP values farther away from the CCSD(T) reference, with respect to the MP2 method and the same basis set. Systematic decreases of the AIP are obtained from the MP2 to CCSD results, ranging 0.26-0.28 eV. The inclusion of the triples excitations within the CC methodologies and the enlargement of the basis set account for an average of around two tenths of an eV in total, being equally important, with improvements of 0.11 and 0.09 eV, respectively (cf. Figure 7.5c). Similar conclusions are obtained in the analysis of the enlargement of the ANO-type basis set with the multiconfigurational methods (see Figure 7.5d). As compared to the VIPs, the effect of the basis set on the CASPT2//CASSCF and CASPT2 AIPs

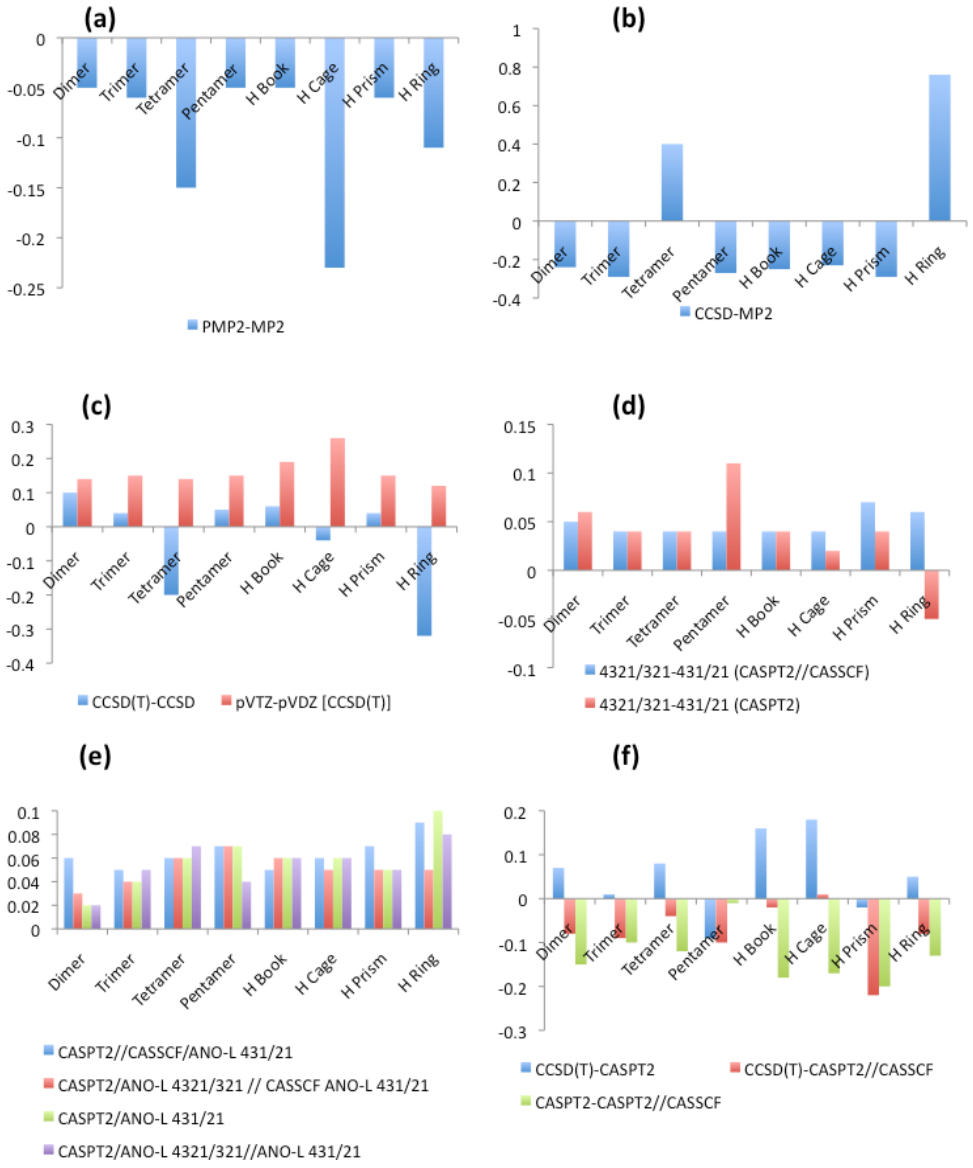

Figure 7.5: Comparative analyses on the performance of the most relevant computational strategies employed in the present work to determine the AIP_0 of water clusters. All energies are given in eV. See text for further details.

appears to be more relevant, giving rise to an increase up to 0.16 eV, being especially important for the CASPT2 values. On the other hand, the IPEA correction is even less relevant than in the case of the VIPs, with the exception of the hexamer ring at the CASPT2 level (cf. Figure 7.5e). AIP improvements are again not general and therefore a value of 0.25 au for the IPEA parameter is neither required for accurate computations of the adiabatic property.

The overall CASPT2//CASSCF/ANO-L 4321/321 and CASPT2/ANO-L 4321 /321 values stand in agreement with the CCSD(T)/aug-cc-pVTZ results, having differences of 0.07 and 0.04 eV, respectively. Therefore, although the CASSCF optimized geometries for cations are less accurate, both the CASPT2 method and the CASPT2//CASSCF approach provide with good estimates of the AIPs, validating the use of CASPT2 methodologies for hydrogen-bonded problems as the one studied within this work and also the application of the CASPT2//CASSCF protocol once the geometries are carefully analyzed.

Table 7.5: Zero-point vibrational energy corrections (eV) of neutral and cationic water clusters and corresponding differences calculated at the MP2 and CASSCF levels of theory with the harmonic (MP2 and CASSCF) and anharmonic ($MP2_{Anharmonic}$) approaches.

System/Method	MP2	$MP2_{Anharmonic}$	CASSCF
$(H_2O)_2$	1.25/1.22/0.03	1.22/1.17/0.05	1.31/1.30/0.01
$(H_2O)_3$	1.97/1.90/0.07	1.93/1.85/0.08	2.02/2.00/0.02
$(H_2O)_4$	2.68/2.59/0.08	2.62/2.53/0.09	2.76/2.70/0.06
$(H_2O)_5$	3.35/3.22/0.13	3.27/3.00/0.27	3.45/3.37/0.08
$(H_2O)_6^{book}$	4.07/3.95/0.12	3.97/3.90/0.07	4.21/4.14/0.07
$(H_2O)_6^{cage}$	4.07/3.96/0.11	3.97/3.82/0.15	4.18/4.12/0.06
$(H_2O)_6^{prism}$	4.08/3.96/0.12	3.98/3.84/0.14	4.18/4.14/0.04
$(H_2O)_6^{ring}$	4.03/3.91/0.12	3.94/3.80/0.14	4.15/4.08/0.07

The AIP values discussed so far correspond to the AIP_0 magnitude, which include the ZPE vibrational correction. Computed ZPE values are given in Table 7.5. Both CASSCF and MP2 methods have been employed with the harmonic

approach to obtain the ZPEs. In addition, the anharmonic approach has been used with the MP2 method in order to analyze the anharmonicity of the water cluster cations and to check the reliability of the harmonic approach in these systems. Within the harmonic approach, both MP2 and CASSCF differences of ZPE between the cation and neutral forms of the clusters increase in general with the size of the cluster up to the pentamer, giving rise to similar results for all the hexamer clusters. CASSCF ZPEs have smaller energy differences, below 0.08 eV, whereas the corresponding MP2 results reach a maximum value of 0.13 eV. The ZPE differences between the cation and neutral structures with the anharmonic method at the MP2 level are fairly similar to the previous MP2 results obtained with the harmonic approach. The largest deviation occurs for the pentamer cluster, in which an increase of 0.14 eV is obtained for the ZPE difference, whereas in all the other cases the corresponding differences average 0.02 eV. Anharmonicity is therefore not a general feature of the water clusters, which is in contrast to previous Monte Carlo studies carried out by Mella *et al.*,[154] and is in agreement with more recent works by Svoboda *et al.* based on path integral molecular dynamics.[162]

Figure 7.6 displays the values obtained in the previous sections for both VIPs and AIPs at the highest levels of theory, that is, the CCSD(T), CASPT2//CAS SCF, and CASPT2 methods with the largest basis set. The experimental data in the liquid water, reported by Mozumder[169] and Winter *et al.*[170] for the vertical property and by Watanabe *et al.*[171] for the AIP, are also shown for the sake of comparison. By inspection of Figure 7.6, the AIP results show a clear trend. The water dimer has the largest value and AIPs become lower in energy with the increasing size of the cluster: dimer > trimer > tetramer > pentamer > hexamers. This behavior implies a progressive evolution of the AIPs and might correspond to the transition from a water molecule (experimental IP of 12.6 eV)[172] to the liquid water, shedding some light on the structure present in the liquid phase. In fact, the experimental value for the AIP of liquid water appears at 9.3 eV (red solid line in Figure 7.6), in the region of the present computed AIPs for the group of hexamer clusters, being especially close to the book and ring conformers. VIPs show similar trends, with the exception of the dimer, which is located below the trimer, tetramer, and

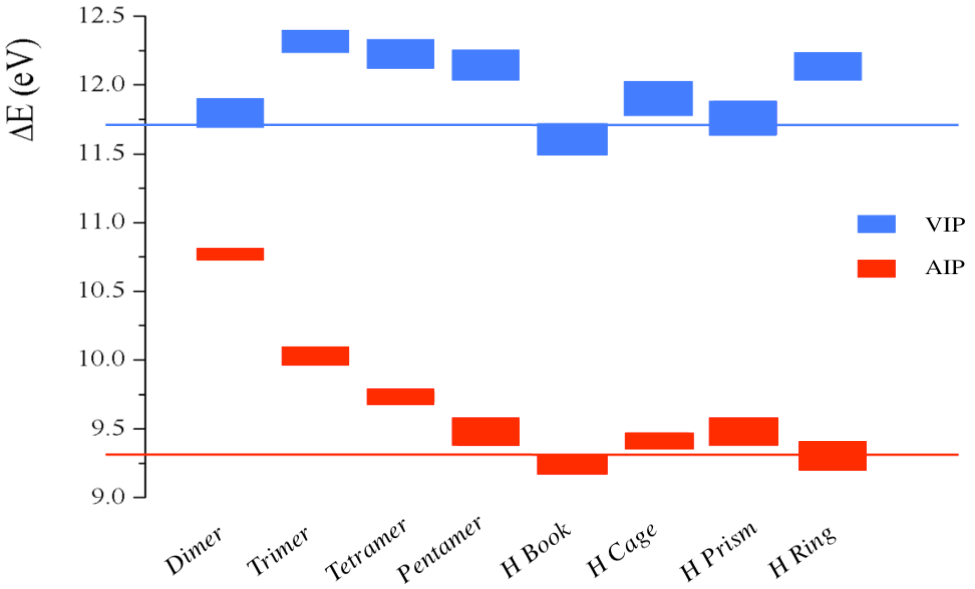

Figure 7.6: Diagram of the VIPs and AIPs obtained within this work and the dispersion shown for the CCSD(T), CASPT2//CASSCF and CASPT2 methodologies with triple-ζ basis sets. Straight lines represent the experimental values of these properties for liquid water.[169–171]

pentamer clusters. The hexamers stand again in agreement with the experimental value for the liquid water, being closer the highest hydrogen-bonded coordinated species, whereas the hexamer ring has the largest difference of around 0.5 eV.

The differences between the VIPs and AIPs for each system are relevant to understand the relaxation process which takes place immediately after the ionization phenomenon. Figure 7.7 displays the energy differences between both IPs at the CCSD(T), CASPT2//CASSCF, and CASPT2 levels of theory. The values are larger than two eV,[158] with the exception of the smallest and the only non-cyclic water cluster, i.e., the dimer, which shows a decrease of the IP from the vertical to the adiabatic property of around one eV. A high

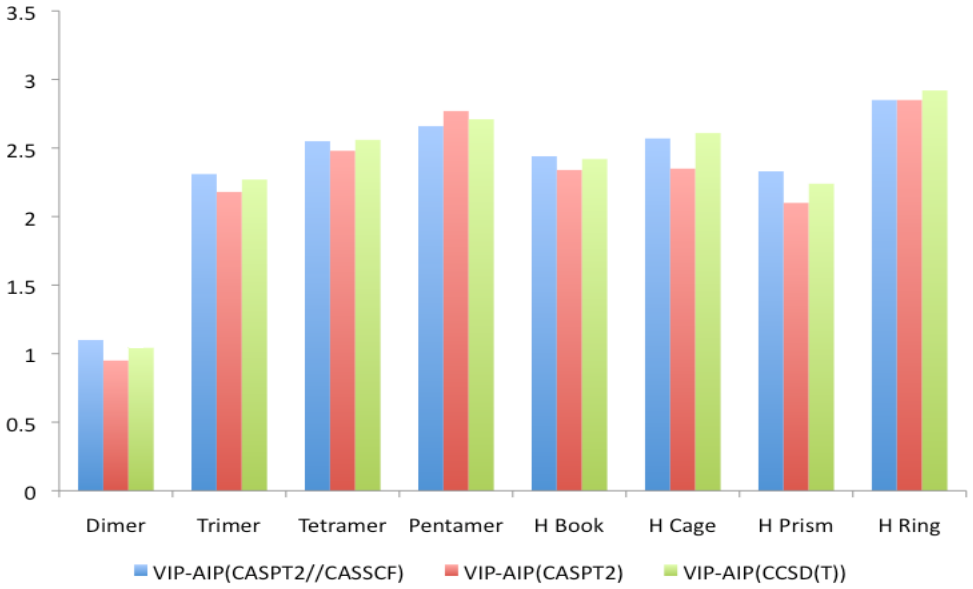

Figure 7.7: Energy differences (in eV) between the water clusters VIP and AIP values obtained at different level of theory: CASPT2/ANO-L 4321/321//CASSCF/ANO-L 321/21 (blue), CASPT2/ANO-L 4321/321//CASPT2/ANO-L 431/21 (red), and CCSD(T)/aug-cc-pVTZ//CCSD/aug-cc-pVDZ.

correlation is found between the VIP-AIP energy differences and the number of proton transfer events which occur in the water cluster cations. Thus, the dimer, trimer, hexamer cage, and hexamer prism, which suffer a single proton transfer in the vibrational relaxation process towards the equilibrium structure of the cation, have the lowest differences between VIP and AIP, whereas the hexamer ring, with a triple proton transfer, show the largest energy difference. The other clusters (tetramer, pentamer, and hexamer book) exhibit a double transfer process and have intermediate values, with the exception of the CCSD(T) and CASPT2//CASSCF results for the hexamer book, which are only slightly below the corresponding values of the hexamer cage and correlates well at the CASPT2 level.

As can be seen, the clusters approach the liquid water value as they grow in size, as well as in hydrogen bonding coordination, providing a reasonable agreement with the liquid fingerprint. It can therefore be envisaged that the properties of the highly coordinated hexamer prism, which holds 3 hydrogen bonds per water molecule, provides the best estimate as compared to the 3.2-3.4 hydrogen bonds in bulk water.[35] The water dimer shows a pretty distinct behavior, mainly due to being the only cluster that is not cyclic. The rest of the clusters behave accordingly and provide a clear trend from the gas-phase or the monomer value towards the liquid phase. The difference in energy between VIPs and AIPs can be used to monitor the number of proton transfer processes occurring within a cluster, as the energy released increases as the number of proton transfers raises.

Conclusions Several *ab initio* methods have been used and benchmarked for the calculation of both the structural and ionization properties of different water clusters, namely: dimer, trimer, tetramer, pentamer, hexamer book, hexamer cage, hexamer prism, and hexamer ring. Geometries for the ground state at the neutral form of the water clusters are compared with experimental values, showing a good agreement with all the dynamic-correlated methods employed. Cationic structures have been compared with the most trusted method, CCSD, showing rather small differences with second-order perturbative methods and larger deviations with non-dynamic-corelated methods such as HF or CASSCF, thus highlighting the importance of such correlation for hydrogen-bonded systems. VIPs and AIPs have been computed and compared with experimental data and with other theoretical methods from the literature in order to obtain accurate values for such properties, validate different computational strategies, and discard odd values observed in experiments. ZPEs have also been calculated with both harmonic and anharmonic approaches in order to validate the use of the harmonic approximation in hydrogen-bonded water aggregates. The CCSD(T) method with the aug-cc-pVTZ basis set, as well as the CASPT2//CASSCF and CASPT2 methods with the ANO-L 4321/321 basis sets, provide us with the highest accurate VIPs and AIPs results, which are located between the experimental IP values of the water

molecule and the liquid phase. The IPs show a general trend, water dimer VIP aside, of decreasing energy whilst increasing the size of the cluster, converging towards the experimental values attributed to the liquid water, which is particularly clear for the adiabatic property. Among the hexamers, the highest values of hydrogen-bond coordination correlate better with the VIP value for the liquid phase and the lowest hydrogen-bond coordinated clusters are close to the AIP value. Finally, proton transfer processes have been analyzed along the vibrational relaxation phenomenon from the neutral to the cationic equilibrium structures, which determine the amount of energy stabilization in the water cluster and hence the difference between VIPs and AIPs.

The oxidation processes in water have therefore been assessed through the systematic study of water clusters. Such processes hold a very important role in nature, as they can be precursors of a plethora of processes, such as dissociative electron attachments that will be discussed over the next chapter. The ionization process drives a charge separation mechanism, similar to those exerted by UV light irradiation in ordered water, which hints towards a possible common mechanism to provide charge separation through cationic species and thus extend the range of the water ordering. The study corresponds to the first work performed in this Thesis chronologically speaking, and thus enabled benchmarking the methods used throughout. The provided high-level results can be used to set up new estimates for future reference and improve the energetic values assigned to the different processes described within this work.

Chapter 8

Picking up loose electrons: Dissociative Electron Attachment Processes in Uracil

> DNA neither cares nor knows. DNA just is. And we dance to its music.
>
> *Richard Dawkins*

Low-energy electrons arise as a harmful secondary reaction capable of generating photochemical damage in the form of DNA/RNA-strand breaks.[60] The generation of low-energy electrons can be due to a plethora of processes within the cellular system after an initial ionization process, as the one described in the previous chapter occurring in water, present in all biological environments. Once the ionization process is completed, a subsequent attachment to one of

the π^* or σ^* orbitals in a nucleobase might take place yielding N-H dissociation processes that can be considered as precursors of DNA/RNA-strand breaks.

Experimental evidence has been piled up over the years supporting dissociations in DNA/RNA caused by electrons in the 0.1-3 eV energy range.[173, 174] The nucleobases can undergo decomposition reactions after the attachment of low-energy electrons named *Dissociative Electron Attachment* (DEA) processes. The current investigation prospects the analysis of the N_1-H and N_3-H dissociation reactions as a result of adding a low-energy electron into an empty π^* orbital. Potential energy curves have been built by elongating the N-H bond towards the dissociation limit.[175]

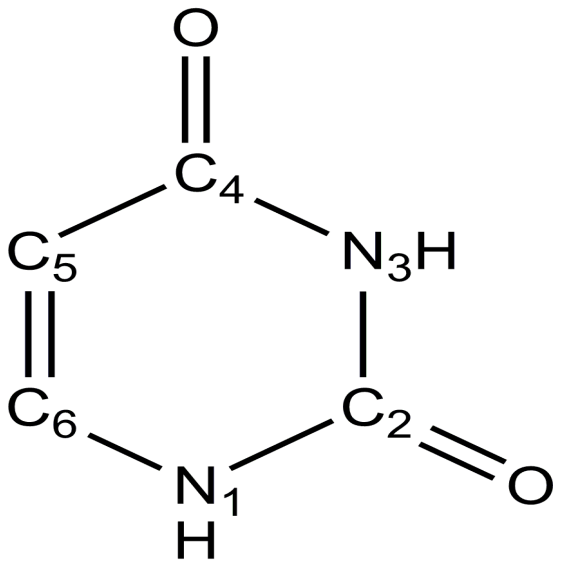

Figure 8.1: Numbering and atom labeling for uracil.

Figure 8.1 shows the numbering and atom labeling for uracil. Both DEA dissociation channels are shown experimentally to be centered at the N-H bonds. The N_1-H is found to be responsible for the lowest-lying experimental signals (< 1.4 eV), whereas N_3-H register the higher ones.[174] The reaction

therefore represents the formation of $(U\text{-}H)^-$ complex leading to a hydrogen atom dissociated as a consequence of the DEA process. Theoretical studies able to provide the proper counterpart to the experimental observations are difficult to carry out, because two different types of anions are present in uracil at low energies: the *Valence Bound* (VB) and *Dipole Bound* (DB) anions. High-level *ab initio* quantum-chemical methods have been used in this work, performing careful analyses of the wave function electron density in order to distinguish between VB or DB anionic states.

Computational Details Characterization of the lowest-lying valence states of the uracil anion and the ground state of the uracil neutral molecule along the N_1-H and N_3-H reactive coordinates is performed in the present study at the CASSCF/CASPT2 level of theory to establish the mechanism for hydrogen loss caused by electron-attachment. Previous benchmark calculations on the determination of the vertical and adiabatic EAs of DNA/RNA nucleobases were considered here to select the most appropriate methodology.[57] These benchmark reference values lead to the characterization of the ground state of the valence anion of uracil as a temporary anion state (resonance), in line with the *Electron Transmission Spectroscopy* (ETS) data, since its energy is higher (vertical and adiabatically) than the energy of the ground state of the neutral molecule and consequently it is unstable with respect to electron detachment. The excited states of the uracil anion have accordingly the same characteristic. Temporary anion states are difficult to treat using conventional quantum chemistry techniques since calculations can tend to put the extra electron into the most diffuse orbital available in order to simulate the neutral molecule plus a free electron. Despite the difficulties, it is also possible to obtain reliable solutions representing the resonance states.[57] It was shown in previous studies of biphenyl and *p*-benzosemiquinone radical anions,[176, 177] that the CASSCF method is able to provide well-localized solutions which can be regarded as a discrete representation of the temporary anion states. Spurious solutions where the extra electron is located far from the molecule in a diffuse orbital can also appear and are not reliable. To distinguish between both types of solutions, in addition to the analysis of the natural orbital with

the unpaired electron, we have determined the spatial extension of the electron density ($<r^2>$) by means of the trace of the second Cartesian moment tensor. Valence localized anionic states have spatial extensions slightly larger than the neutral system, whereas clear differences appear for diffuse states (or mixed valence dipole-bound states).

The basis set of atomic natural orbital, ANO-L type, contracted to C, N, O [4s3p1d] / H [2s1p] (hereafter, ANO-L 431/21) was chosen as a compromise between accuracy and computational cost. No symmetry requirements (C_1 symmetry) were employed in the computations. The geometries of the neutral, and the hydrogen dissociated $(U\text{-}H_1)^-$ and $(U\text{-}H_3)^-$ uracil anionic systems were optimized at the CASSCF level using an active space comprising the whole valence π system of the nucleobase, that is, 10 electrons distributed among 8 π molecular orbitals, namely CASSCF(10,8). Planar geometries were obtained for these anions. The ZPE was calculated in these minima at the same level of theory with the harmonic approximation. A larger active space was used within the CASSCF method for the computations of the *Potential Energy Curves* (PECs) along the N-H reactive coordinates, including four additional orbitals: the σ^* orbitals related to the N_1-H and N_3-H bonds plus two additional diffuse orbitals required to stabilize the active space along the PECs. In total, 11 active electrons and 12 active orbitals [hereafter, CASSCF(11,12)]. Eight states were averaged with equal weights within the CASSCF method for the PECs of the uracil anion. Several computational strategies were employed as required in order to determine the relevant PEC features to the DEA processes occurring at the N-H sites which are next described. i) The PECs between the neutral structure and the two $(U\text{-}H)^-$ anions were explored initially by means of the linear interpolation of internal coordinates (LIIC) procedure. This method allows the location of the crossing points between the PECs of the low-lying states of the anion which mediates the hydrogen loss phenomenon. The final hydrogen separation from the uracil molecular frame in the LIIC was set at 3.0 Å. ii) A constrained optimization with fixed N_3-H bond length was also performed at the CASSCF(11,12) level to analyze the role of the π_2^- anionic state in the DEA events. The geometry of the π_2^- state was optimized with fixed N_3-H values corresponding to the first point of

the LIIC curve (the ground-state equilibrium geometry of the uracil neutral molecule). All other degrees of freedom of the uracil molecule are allowed to relax. iii) CASSCF(11,12) minimum energy path (MEP) computations on the PEC of the π_2^- excited state of the uracil anion from the equilibrium geometry of the neutral system were carried out to describe the evolution of this state after electron attachment and the crossing with the dissociative σ^- excited state which ends in the fragmented nucleobase [$(U\text{-}H)^-$ + H]. MEPs were built as steepest descendent paths,[94, 95] in which each step implies the minimization of the energy on a hyperspherical cross section of the PEC centered on the initial geometry within a predefined radius of 0.09 au. Mass-weighted coordinates were used.

The CASPT2 method was employed at the geometries obtained in the LIIC, constrained optimizations, and MEPs to obtain the dynamic correlation for the ground state of the neutral uracil and the low-lying doublet states of the anionic system. In order to minimize weakly interacting intruder states, the imaginary level-shift technique with a parameter of 0.2 au was employed. The IPEA definition of the zeroth-order Hamiltonian in the CASPT2 method, with a value of 0.25 au was found previously to improve the results of EA in nucleobases and consequently was employed here.[57] All the calculations were performed with the CASSCF/CASPT2 method as implemented in the MOLCAS quantum-chemistry package of software.[69]

Results and Discussion During the electron attachment process in the uracil molecule, the system is initially ionized to one of the low-lying states of the anion producing a temporary anionic state with the geometry of the neutral molecule. This state must evolve towards a dissociative σ^- state localized in the N-H bonds in order to produce the DEA phenomenon, through a crossing region between the curves of the initial π^- and final σ^- states. The latter anionic state will drive the molecule to the separated hydrogen and $(U\text{-}H)^-$ fragments.

Table 8.1: Experimental data derived from electron transmission spectroscopy (ETS) and theoretical vertical electron affinities (eV) for the low-lying anionic states in uracil.

anion state	CASPT2//CASSCF ANO-L 431/21[a]	CASPT2//CASSCF ANO-L 4321/321[b]	CCSD(T)//CCSD aug-cc-pVDZ[b]	exp ETS[c]
π_1^-	-0.69	-0.61	-0.64	-0.22
$\mathrm{diff}^-_{N_1H}$	-0.84			
diff^-	-1.69			
$\mathrm{diff}^-_{N_3H}$	-1.74			
π_2^-	-2.24			-1.58

[a]This Work
[b]Reference [57]
[c]Reference [178]

Table 8.1 compiles the present computed CASPT2 *Vertical Electron Affinities* (VEAs) for the five low-lying states of the anion at the geometry of the neutral system, together with the related data measured in ETS experiments[178] and reference values from a previous work.[57] The results obtained here for the lowest-lying VEA are in agreement with the data computed at higher levels of theory. Systematic differences are found between all the theoretical VEAs and the vertical attachment energies measured in ETS experiments. Such discrepancies were carefully analyzed in the previous study on the accurate determination of the EAs in nucleobases.[57] The authors concluded that the ETS technique underestimates the relative energy of the anion states with respect to the neutral molecule by near 0.3-0.4 eV in nucleobases, which is in agreement with other studies.[57, 175, 178]

The present VEAs show that the lowest VB-anionic state has π character and three diffuse states are located between the π_1^- and π_2^- states. As shall be seen below, two of these diffuse states ($\text{diff}^-_{N_1H}$ and $\text{diff}^-_{N_3H}$) are connected to the dissociative anionic states in which the electron occupies the σ^* orbitals of the N_1-H and N_3-H bonds, respectively. However, at the equilibrium geometry of the neutral uracil, these solutions correspond to diffuse or mixed valence and dipole-bound states and the VEA cannot be considered reliable. The spatial extensions $<r^2>$ of these states are 177, 184, and 172 au^2, for $\text{diff}^-_{N_1H}$, diff^-, and $\text{diff}^-_{N_3H}$, respectively, which are clearly much larger than the results obtained for the valence localized π_1^- and π_2^- anionic states and neutral S_0 ground state at this geometry, with values in the range 106-128 au^2.

Table 8.2 compiles the results obtained by Märk and co-workers with the P2MP2 [179] and G2MP2 methods,[174, 180] data from Li *et al.* with the DFT and CBS-Q methods,[181] and the present computed CASPT2 values. The theoretical results establish the threshold for N_1-H and N_3-H DEA at 0.6-0.8 and 1.2-1.4 eV, respectively. Such energies correspond to the lowest band and the origin of the broad band in the DEA cross sections.[179] These thresholds, therefore, allow an estimate of the regions in the experimental spectrum related to each one of the N-H dissociation processes. While energies below 1.2-1.4 eV are sufficient to break the N_1-H bond, higher energies are needed for the

Table 8.2: Dissociation energies (in eV) corresponding to the N_1-H and N_3-H bonds computed at different levels of theory. ZPE energies are included at the same level of theory.

	P2MP2[a]	G2MP2[b]	DFT[c]	CBS-Q[c]	CASPT2[d]
N_1-H	0.8	0.8	0.71	0.84	0.61
N_3-H	1.4	1.4	1.25	1.36	1.21

[a]Reference [179]
[b]References [174, 180]
[c]Reference [181]
[d]This Work

N_3-H cleavage. The same conclusions were obtained from experiments with methylated uracil and thymine nucleobases at the N_1 and N_3 positions,[174] where the activation of the dissociative N_1-H path was exclusively achieved at energies below the calculated 1.4 eV threshold, whereas at energies above that value, dissociation at the N_3-H site became accessible.

The mechanism provided for the N_1-H DEA process can be seen in Figure 8.2. Four anionic states are represented in the figure: the lowest-lying π_1^- and π_2^- VB states and the dissociative σ^- states. In the π_1^- and π_2^- states the attached electron is placed in a valence π orbital, whereas the antibonding σ^* orbital of the N_1-H and N_3-H bonds is occupied in the σ-like type states of the anion. The π_1^- state related to the population of this antibonding orbital rapidly crosses with the $\sigma_{N_1H}^-$ yielding the $(\pi_1^-/\sigma_{N_1H}^-)_{CI}$ crossing. From there onwards the system follows the $\sigma_{N_1H}^-$ state towards the complete cleavage of the N_1-H bond. It can therefore be envisaged that surface crossings are the driving forces behind the dissociation process. The σ^- states have intrinsic dissociative character, although it is not clear from Figure 8.2. The reasons are related to the mixed valence and dipole-bound solutions found for these states in the Franck-Condon region which cause an underestimation of the energies. In Figure 8.2, the $\sigma_{N_1H}^-$ state shows initially spatial extensions, $<r^2>$, around 172 au^2 (dashed line), becoming localized at larger N_1-H distances with an average $<r^2>$ of 138 au^2. This latter value is close to the result found for the

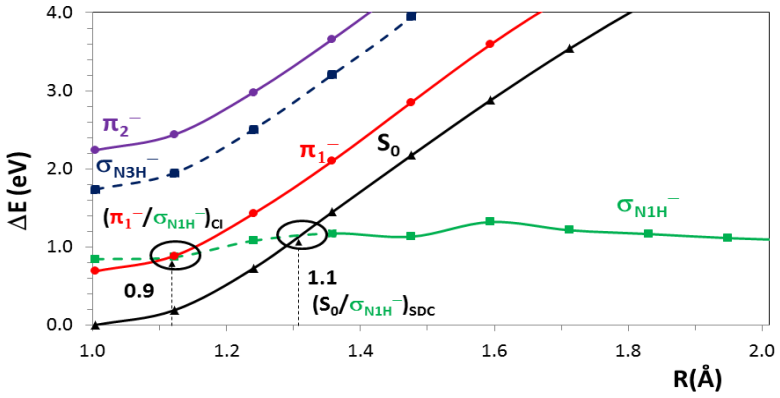

Figure 8.2: CASPT2 potential energy curves (PECs) for the ground state of the uracil neutral molecule (S_0) and the π_1^-, π_2^-, $\sigma_{N_1H}^-$, and $\sigma_{N_3H}^-$ states of the anionic species along the N_1-H reactive coordinate. Dashed lines indicate points with mixed valence and dipole-bound states (see text).

dissociated anion (133 au^2), which is a stable anion (positive EA), and the $<r^2>$ extensions computed for the π_1^- (128-134 au^2), π_2^- (124-125 au^2), and S_0 (95-111 au^2) states.

Figure 8.3 depicts the potential energy curves built for the N_3-H dissociation process. In this case, the crossing of the π_1^- state with the $\sigma_{N_3H}^-$ occurs at larger distances where the π_1^- state has already gone up in energy yielding the $(\pi_1^-/\sigma_{N_3H}^-)_{CI}$ crossing at 2.1 eV, in agreement with the experimental evidence that points at higher energies for the N_3-H dissociation as compared to the N_1-H bond. It can therefore be concluded that the π_1^- state also represents a possible way to populate the N_3-H dissociation channel. MEP computations are also available showing the possibility of a crossing from an initially populated π_2^-, making the precursor states for this bond cleavage twofold. As in the previous case, the $\sigma_{N_3H}^-$ is not well defined at close distances. The first points of the $\sigma_{N_3H}^-$ state have $<r^2>$ around 169 au^2 and next the electron becomes localized in the valence space, with an averaged $<r^2>$ value of 129 au^2, in line with the results obtained for the dissociated anion and the π_1^-, π_2^-,

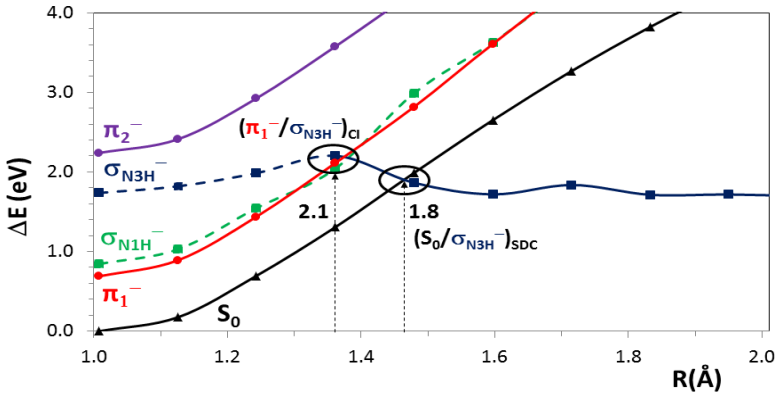

Figure 8.3: CASPT2 potential energy curves (PECs) for the ground state of the uracil neutral molecule (S_0) and the π_1^-, π_2^-, $\sigma_{N_1H}^-$, and $\sigma_{N_3H}^-$ states of the anionic species along the N_3-H reactive coordinate. Dashed lines indicate points with mixed valence and dipole-bound states (see text).

and S_0 states (126, 126-133, 124-125, and 105-109 au^2, respectively). Hence, descriptions on the starting region of the $\sigma_{N_1H}^-$ and $\sigma_{N_3H}^-$ states with the employed methodology are only approximate, whereas the points at larger N-H bond lengths and the π_1^- and π_2^- anion states are accurately determined.

Figure 8.4 displays the CASSCF and CASPT2 energies for the relevant neutral and anionic states of uracil in the DEA process along the π_2^- state MEP. The uracil molecule suffers out-of-plane distortions, reaching an equilibrium structure of the π_2^- state with ring-puckering at the end of the MEP. This point, with an energy of 1.59 eV, is obviously the lowest energy structure of the π_2^- anion state and therefore no possible contributions to the DEA process can take place from this state at lower energies. Regarding the mixing between the states, at the CASSCF level, the π_2^- state is still higher in energy than the σ^- states and no crossings are present along the path. However, dynamical electron correlation does not contribute with the same value to the π^- and σ^- states (strong differential correlation effects) and the scenario is completely different at the CASPT2 level. A CI point appears now between the π_2^- and

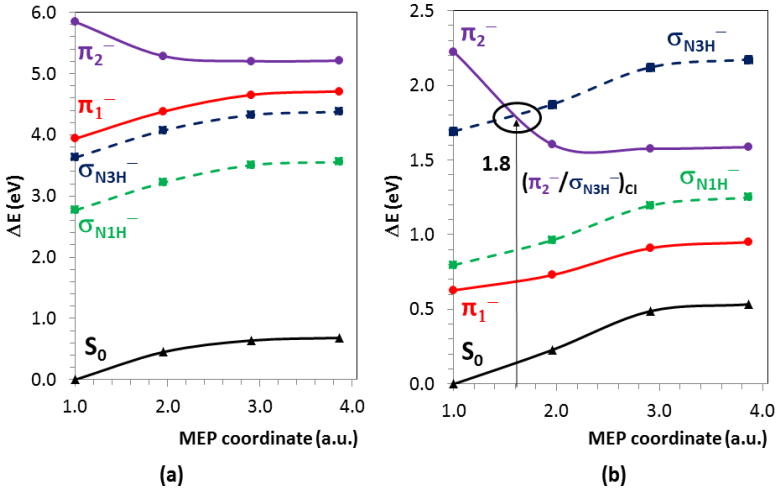

Figure 8.4: CASSCF (a) and CASPT2 (b) energies for the ground state of the uracil neutral molecule (S_0) and the π_1^-, π_2^-, $\sigma_{N_1H}^-$, and $\sigma_{N_3H}^-$ states of the anionic species along the minimum energy path (MEP) of the π_2^- state. Dashed lines indicate points with valence and dipole-bound mixing (see text).

$\sigma_{N_3H}^-$ states, $(\pi_2^-/\sigma_{N_3H}^-)_{CI}$, which is the structure responsible for funneling of the system toward the dissociation of the N_3-H bond (cf. Figure 4b). The equilibrium structure of the π_2^- state is only 0.2 eV below $(\pi_2^-/\sigma_{N_3H}^-)_{CI}$, although the results for the crossing should be taken with caution due to the difficulties of the method to properly describe the σ^- state. The N_3-H elongation will be driven by the $\sigma_{N_3H}^-$ dissociative state which is populated via the $(\pi_2^-/\sigma_{N_3H}^-)_{CI}$. Taking into account the VEA of the π_2^- anion state (2.24 eV), the energy for its equilibrium structure (1.59 eV), and the $(\pi_2^-/\sigma_{N_3H}^-)_{CI}$ crossing which activate the DEA mechanism (above 1.8 eV), the region of the broad band maximum in the cross sections of the DEA experiments (around 1.8-2 eV) can be mainly ascribed to the π_2^- state. The π_1^- state might also contribute to this band and the DEA process at the N_3-H site, although a lower-energy dissociative route is possible for such state, involving the N_1-H bond breaking. Therefore, the N_1-H DEA will be the main reactive path

driven by the π_1^- state, whereas the π_2^- state can only participate in the N_3-H DEA.

CIs between σ^- and π^- anionic states have been proposed above as responsible for the activation of the hydrogen loss phenomena. In addition, singlet-doublet crossings (SDCs) between the ground state of the neutral system and the σ^- states of the anion might also be relevant. In such SDC regions, there exists an energy resonance between the anion and the neutral system plus the electron at infinite distance from the molecular frame. Hence, conversions between both situations are energetically possible. In the N_1-H and N_3-H PECs, the approximate $(S_0/\sigma^-_{N_1H})_{SDC}$ and $(S_0/\sigma^-_{N_3H})_{SDC}$ crossings appear at around 1.1 eV (cf. Figure 8.2) and 1.8 eV (cf. Figure 8.3), respectively, which corresponds to the energy range of the experimental bands.[174, 179, 180]

Conclusions The CASSCF/CASPT2 method has been applied to determine the mechanisms for DEA at the N_1-H and N_3-H sites of uracil involving VB anions. Several low-lying states of the anionic species, with π^- and σ^- character, have been studied. According to the findings obtained and other theoretical results from the literature,[174, 179–181] the energy threshold for the dissociation of the N_1-H and N_3-H bonds is 0.6-0.8 and 1.2-1.4 eV, respectively, which agrees with the experimental observations. Regarding the mechanism for the DEA phenomena, once the low-energy electron becomes attached into an empty π^* orbital within the uracil monomer, crossing points between the PECs of the π^- and the σ^- states appear at energies around the maxima of the lowest peak and the broad band in the DEA cross sections. The corresponding conical intersections (CIs) can be interpreted as the points which activate the internal conversion to the dissociative σ^- states, therefore driving the system towards the $(U\text{-}H)^-$ ion plus a H atom. The π_1^- state is mainly involved in the cleavage of the N_1-H bond, since it crosses the $\sigma^-_{N_1H}$ state at 0.89 eV. Still, it can participate in the N_3-H DEA via a second CI, $(\pi_1^-/\sigma^-_{N_3H})_{CI}$, in the region of the broad band, around 2 eV. The π_2^- state is predicted, however, as the main state responsible for this shape having a crossing point with the $\sigma^-_{N_3H}$ state, $(\pi_2^-/\sigma^-_{N_3H})_{CI}$, in this region.

Dissociative electron attachment processes in uracil have then been studied by means of high level *ab initio* quantum chemical methods, capable to provide a meaningful description of the process under scrutiny. Such processes are triggered by the presence of low-energy electrons, coming from previous ionization processes given in the cellular system, such as those investigated for water in the previous chapter. The N-H dissociation can be considered as precursor of DNA/RNA-strand breaks, thus highlighting the importance of charge-transfer and charge-separation processes in biochemical systems, as those found in the charge transfer mechanism of the DNA chain or the charge separation occurring in the production of ordered water, respectively.

Chapter 9

Photochemical Studies on the DNA and RNA Nucleobases

> Almost all aspects of life are engineered at the molecular level, and without understanding molecules we can only have a very sketchy understanding of life itself.
>
> *Francis Crick*

The photophysical and photochemical properties of DNA/RNA nucleobases construct the main avenue of research of the QCEXVAL group over the last decade. During these years, the photochemical decay paths of all nucleobases have been unveiled and their different experimental spectroscopical features tackled with theoretical models yielding its respective molecular counterparts.

Much work has been done and yet plenty of controversy still surrounds the main intrinsic photochemical properties of the nucleobases, which provide unique information to ascertain the effect of light irradiation upon the cellular system as well as in current nano- and biotechnological applications. In order to provide a firm grip towards the understanding of the main photochemical decay routes present in DNA, a revisit of all previous work done in our group and others has been gathered rendering a more complete vision of the process and validating the previous efforts performed in the characterization of the main decay paths present in the nucleobases.

Ongoing work is being performed by using more correlated methods (CASPT2 //CASPT2) that would yield a more accurate picture of the process, yet initial estimates validate the previous methodologies employed (CASPT2// CASSCF) in the results so far attained. Therefore, we shall focus on the already characterized mechanisms that will be described in the current chapter. [182]

Computational Details Complete Active Space Self Consistent Field (CAS SCF) and Complete Active Space Second-Order Perturbation Theory (CASPT 2) methods have been used in this study as implemented in the MOLCAS-7 package of programs.[69] Two different types of basis sets were used: i) a segmented type of basis set, namely the 6-31G and 6-31G* basis sets, and ii) a set of generally contracted ones of the ANO-S O,C,N(10s,6p,3d)/H(7s3p) and ANO-L O,C,N(14s,9p,4d,3f)/H(8s4p3d) type, with the contraction O,C,N[3s2p 1d]/H[2s1p] for the former, and O,C,N[3s2p1d]/H[2s1p] and O,C,N[4s3p2d]/ H[2s1p] for the latter. The basis sets have been selected in order to provide a benchmark informing about the smallest basis set required for the correct characterization of the pyrimidines PEH.

Multiconfigurational wave functions have been built by adding the whole π occupied space of the pyrimidines and adding three of the π virtual orbitals making it a CASSCF(10,8) active space. Further calculations with three ex-

tra π virtual orbitals were performed to assess the convergence of the active space for such computations, at the CASSCF(10,11) level. Two roots have been averaged in the CASSCF wave function unless specified otherwise, with no symmetry restrictions as required to avoid any symmetry breaking problems and to allow any type of geometry distortion during the optimization procedures. On the other hand, CASPT2(10,8) and CASPT2(10,11) second-order perturbative corrections have been added on top of their corresponding CASSCF energies, maintaining all inactive electrons frozen during the perturbation step and making use of the zeroth order Hamiltonian as originally implemented.[65]

Geometry optimizations and MEPs have been obtained at both CASSCF and CASPT2 levels. CASSCF MEPs have been computed with analytical gradients and subsequent CASPT2 calculations have been performed on the converged geometries along the path in order to add the dynamical correlation, as usually performed nowadays in photochemical studies following the CASPT2//CASSCF protocol. CASPT2 optimizations and MEPs, on the other hand, have been computed making use of numerical gradients as implemented in the MOLCAS package.[69] An estimate of a given photochemically relevant conical intersection (CI) has been performed by determining the Crossing Point along the corresponding MEP.

Results and Discussion The main deactivation pathways of pyrimidines (thymine, uracil, and cytosine) are assembled together hinting towards a common decay route present in all pyrimidines.[7, 8] A schematic view of the process can be seen in Figure 9.1. Upon light absorption, the nucleobases possess an energy excess in their excited states, which drive the molecule through a puckering motion towards a crossing with the ground state in an ethene-like conical intersection, liberating subsequently the remaining extra energy through vibrational motions while reaching the bottom of the ground state potential energy surface being therefore photostable.[1, 2] This main decay route is outlined to explain the ultrafast decay component found experimentally for all pyrimidines.

As can be seen in Figure 9.1, direct population of the $\pi\pi^*$ bright state of the systems triggers an ultrafast decay route towards a ring-puckering CI with the ground state. This pathway can be directly accessed from the FC region as it is directly populated upon light absorption, yet other avenues of interest shall also be pursued to render a complete description of the mechanisms governing photoinduced phenomena in DNA/RNA. Whereas the direct decay of the $\pi\pi^*$ state justifies the ultrafast component, two other slower signals have been registered in the nucleobases which shall also be analyzed.

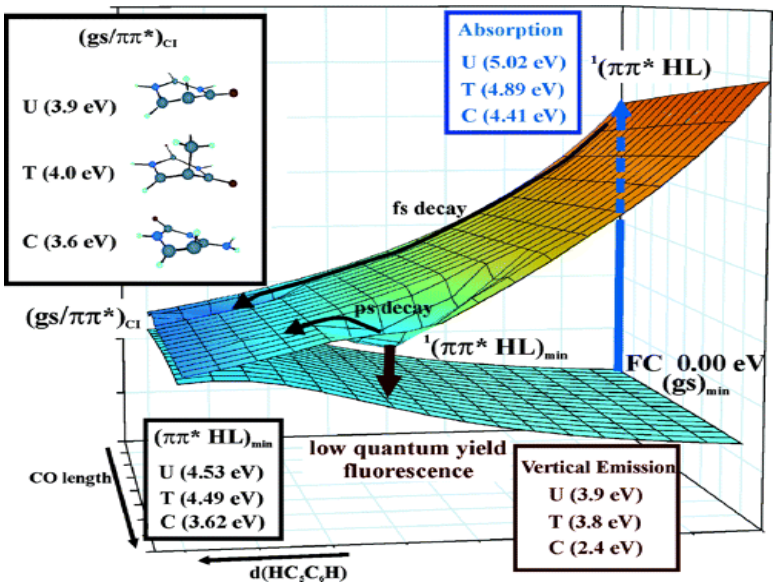

Figure 9.1: Schematic view of the photochemistry of the Pyrimidine nucleobases.

In both thymine and uracil, the $\pi\pi^*$ vertical state happens to be higher in energy compared to the $n\pi^*$, corresponding to the excitation from a lone pair orbital located in an oxygen. The transition from the ground state to the lower-lying $n\pi^*$ is characterized by a small oscillator strength, which justifies tackling the $\pi\pi^*$ bright state directly instead. However, the $n\pi^*$ state may still play an important role in the photochemical mechanisms of both thymine and uracil, as it appears in the way between the $\pi\pi^*$ state and the ground

state. It can therefore be envisaged that a $(\pi\pi^*/n\pi^*)_{CI}$ occurs on the way towards the previously reported ethene-like CI thus providing the following scenarios: i) to go through this conical intersection and towards the ethene-like crossing with the ground state, or ii) to follow the $n\pi^*$ state after the crossing yielding a different CI with the ground state. This alternative route might explain the picosecond timescale signal present in both thymine and uracil, leaving the explanation of the slower nanosecond fingerprint to the triplet state and the several intersystem crossing processes present in the systems that allow its population and highlight its relevance. In cytosine, the $\pi\pi^*$ state appears as the lowest-lying excited state thus validating the approach taken and emphasizing the population of the $\pi\pi^*$ as the driving force underlying the photochemical activity in cytosine. In contrast to uracil and thymine, the nucleobase at hand does not behave accordingly as it presents a small barrier on the path towards its ethene-like CI which hints at a lack of dynamical correlation to yield the proper description for the system. Ongoing work is being carried out to cover many different possibilities in order to ascertain its properties.

Focusing on thymine, Figure 9.2 shows the computed MEPs by using different basis sets and active spaces, increasing systematically the electron correlation. The calculations are performed using the full π-valence active space, and three extra virtual orbitals were subsequently added in order to check the results upon the enlargement of the active space. The π-valence active space comprises ten electrons in eight orbitals, namely 10in8. With three extra correlating orbitals added, the active space becomes 10in11. As can be seen in Figure 9.2, the MEP computed with the 6-31G basis set leads to a planar minimum in the $^1(\pi\pi^*)$ hypersurface, as previously reported by Perun *et al.*[183] Employing the same active space but increasing the flexibility of the one-electron basis set employed yields a very different result though (see the MEP at CASPT2//CASSCF/6-31G* level in Figure 9.2). In this case, the system decays in a barrierless manner towards a ring-puckering CI. The same result holds true after adding extra correlating orbitals to the active space, thus validating the previous result obtained with the 6-31G* basis set. The illustration highlights the importance of the electronic correlation in order to give

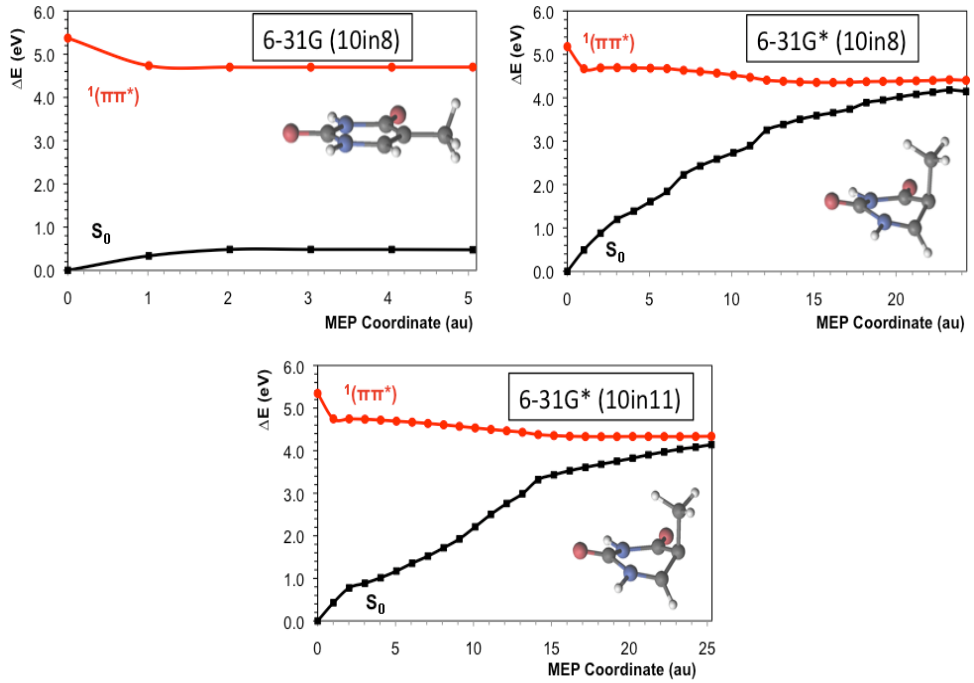

Figure 9.2: MEP computations of thymine at the CASPT2//CASSCF level of theory with the 6-31G, 6-31G*, and the 6-31G* basis sets and including three extra correlating orbitals (10,11) in the active space.

a proper description of the photochemical pathways actually accessible in a given molecular system. Moreover, the technique used to optimize the excited state also plays a key role. Optimizing $^1(\pi\pi^*)$ with the 6-31G* basis set yields a planar minimum, identical to the one found with the CASSCF/6-31G MEP, whereas the MEP at the CASSCF/6-31G* decays to the ring-puckering CI as shown in Figure 9.2. In other words, for thymine, determination of the different stationary points independently (geometry optimization for the ground state and the excited state, and CI search) at the CASSCF/6-31G* level, connecting those regions of the hypersurface subsequently by the LIIC procedure, will offer a view of the photochemistry through a minimum on the excited hypersurface. A similar situation occurs if the geometry optimizations are

carried out at the CASPT2 level.[184, 185] Nevertheless, CASPT2//CASSCF MEP computations employing a basis set with polarization functions on the C,N,O atoms will lead to a barrierless deactivation of the excited state towards the ground state. It is not therefore surprising the variety of results available in the literature depending on the computational strategy employed. In fact, the conclusions derived from those studies are strongly dependent on the strategy used. It is especially crucial in dynamics simulations, where the outcomes become biased with respect to the static strategy chosen. The findings emphasize the importance of the MEP technique to provide accurate estimates.

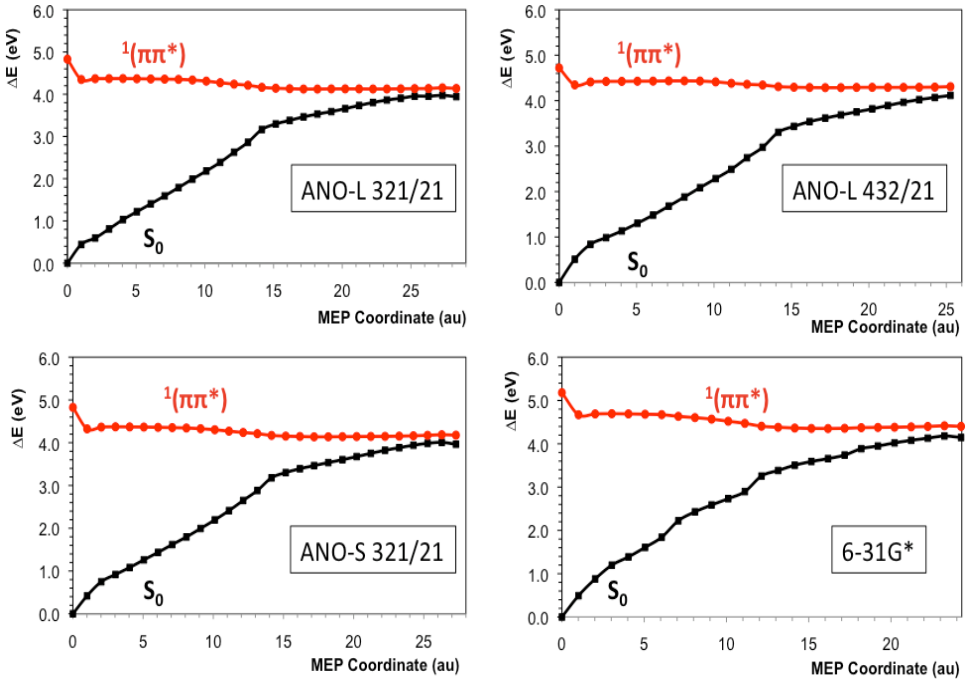

Figure 9.3: MEP computations of thymine at the CASPT2//CASSCF level of theory with the ANO-L 321/1, ANO-L 432/21, ANO-S 321/21, and 6-31G* basis sets.

A benchmark study on the influence of the basis set on the shape of the PEH has been performed yielding the results depicted in Figure 9.3. As can be seen, no qualitative changes are discernible by going from the 6-31G* to the ANO-L 432/21 basis sets, even though the energy values are overestimated on the former as compared to the results obtained by the ANO-type basis sets. Other properties highly dependent of the basis set, like the dipole moment, might change as well even though the resulting PEH might be qualitatively analogous. It can therefore be concluded that 6-31G* provides a proper description of the PEH and it is the basis set used to test the adequacy of the CASPT2//CASSCF protocol against the CASPT2//CASPT2, due to computational demands.

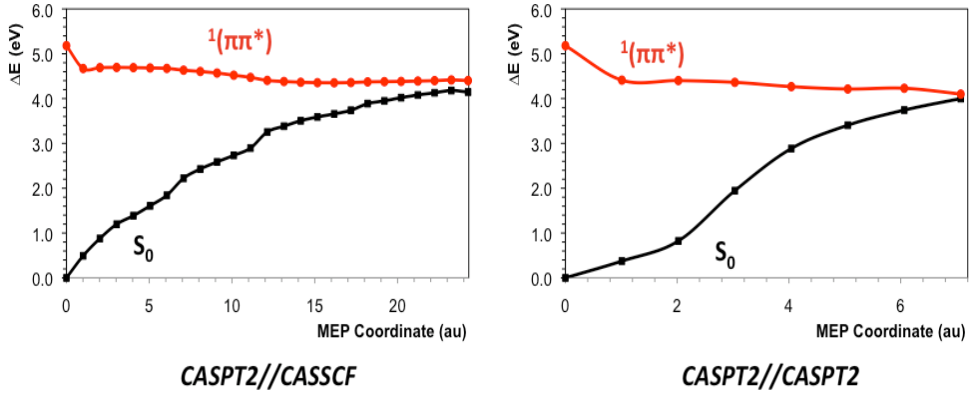

Figure 9.4: MEP computations of thymine at the CASPT2//CASSCF and CASPT2//CASPT2 levels of theory with the 6-31G* basis set.

Figure 9.4 depicts the MEP calculations of thymine at both the CASPT2//CASSCF and CASPT2//CASPT2 levels of theory. Both methods provide the same potential energy profile thus validating the use of the CASSCF method in thymine to map the PEH and subsequently applying the second-order correction on the energy to give correct energy-wise estimates. It can therefore be concluded that the dynamical electron correlation holds in this case a secondary role in determining the geometries along the PEH, as compared to the non-dynamical correlation provided by the multiconfigurational nature in the CASSCF method. This study is still ongoing and all avenues of inquiry

are being pursued to properly test the adequacy of the CASSCF method in mapping the PEH of the nucleobases, as well as to properly define the optimization algorithm that yields the better description of the photoinduced process occurring in the nucleobases upon light absorption.

The longer decay signals observed experimentally were tackled in the last decade by members of our group through the decays of the triplet states and their population mechanisms mediated by the $^1(\pi\pi^*)$ bright state.[186] Three different singlet-triplet crossing (STC) regions were characterized: i) one located just after the FC region with the $^{1,3}(n\pi^*)$ states with a sizable spin-orbit coupling (SOC) that favors the population of the triplet state, which subsequently decays together with the also populated $^1(n\pi^*)$ state towards a $^{1,3}(n\pi^*)$ minimum, ii) a minimum that can be surmounted via both the singlet and triplet manifolds and that holds an extremely large SOC populating the triplet state and subsequently reaching a minimum in the $^3(\pi\pi^*)$ hypersurface, and iii) an extended STC region near the ring-puckering CI between the $^1(\pi\pi^*)$ and the ground state, with a relatively high SOC between the $^1(\pi\pi^*)$ and $^3(\pi\pi^*)$ states, which leads to the $^3(\pi\pi^*)$ minimum. From there, a small barrier was obtained to reach the $^3(\pi\pi^*)$ STC with the ground state, recovering the planarity of the molecule.

The most probable decay channels in the naturally occurring vs. the modified nucleobases were also analyzed earlier in the group in order to find the particular decay patterns of the canonical nucleobases.[187] MEP computations on the $^1(\pi\pi^*)$ PEH of 1-fluorothymine yielded a barrier along the path thus hindering the direct non-adiabatic deactivation towards the ring-puckering CI with the ground state. As explained by the authors, these findings speak in favor of the choice of the biological nucleobases by natural selection based on their resistance to photochemical damage. On the other hand, methyl substitutes were found to have a similar photochemistry as that previously found for the canonical nucleobases.

Conclusions In general, many critical points on the PEHs of the lowest-lying excited state of thymine have been located over the years, arising several reaction pathways for non-radiative decay. The CASPT2// CASSCF computations provide the most efficient decay mechanisms present in the pyrimidines as a direct relaxation of the $^1(\pi\pi^*)$ bright state towards a ring-puckering CI with the ground state without barriers,[7] which is predicted by such approach as the main decay channel (see Figure 9.1). In addition, preliminary results obtained in the present work at the CASPT2//CASPT2 level also points to barrierless relaxation towards the ground-state PEH. Of course, accurate dynamics would be required for a direct comparison of the theoretical determined lifetimes with those derived experimentally. The direct and barrierless nature of this relaxation route can be related to very short lifetimes as those observed in the molecular beams in the sub-picosecond timescale [23] [32] [33] [34]. Longer time constants might be attributed to the presence of a flat region far from the main decay, as displayed in Figure 9.1. Moreover, other relaxation routes involving switches to the lowest-lying $^1(n\pi^*)$ state and barriers are also expected to increase the time spent in the excited state. Finally, several STC regions with their large associated SOCs imply the low-lying triplet states as main actors in one of the secondary decay routes,[186] which may provide a rationale for the experimental signals of longer time.

The photophysical and photochemical pathways of the DNA/RNA pyrimidines are here covered focusing mainly on pyrimidines through thymine. The study consists on a revisit of the previous work performed in the QCEXVAL group over the last decade,[182] which helps elucidate the photostability of nucleobases after light irradiation. As can be seen, all nucleobases access their respective bright state, which drives them to the main decay ultrafast channel towards the ground state through a ring puckering conical intersection. Several other deactivation routes have been characterized, specially for the larger purine derivatives, adenine and guanine. The different timescales found experimentally for the nucleobases are rationalized through the study of the different decay channels available and the difficulties found to access them, which influence their timescales making them slower.

Ongoing studies are being performed to further validate all results shown with a more correlated computational strategy (CASPT2//CASPT2), in order to grasp a definitive answer as to the photochemical mechanisms prevailing in DNA/RNA.

Part III

Conclusions

Chapter 10

Summary

> To succeed, jump as quickly at opportunities as you do at conclusions.
>
> *Benjamin Franklin*

Several conclusions can be derived from the results attained in the current Thesis. In general lines, an improved theoretical characterization of the spectroscopical properties of water has been achieved. Both bulk and ordered water models have been tackled, unveiling their minimal units and their intrinsic intermolecular interactions that make them so different spectroscopically. Those minimal units holding the basic interaction characterizing both bulk and ordered water have been enlarged to propose more realistic models that bear most of the features registered experimentally. Excited states derived from ultraviolet and ionizing radiation have been taken on in order to achieve a more complete understanding of the systems at hand, as well as to provide insights into mediated processes such as the dissociative electron attachment in nucleobases due to the loose electrons generated in conditions of oxidative

stress, like the ionization processes in the surrounding water present in the biological systems.

Specifically, two different types of water dimer, the conventional and the π-stacked, have been assessed in order to characterize the smallest building blocks present in bulk and ordered water, respectively.

The conventional water dimer is composed by a hydrogen bond tying both molecules together and relates fairly well with the experimental fingerprint found for gas-phase water. It is therefore clear that the hydrogen bond plays a minor role in electronic excitations and that both water molecules of the dimer produce signals similar to the one found in the monomer. The photodissociation process centered in each of the water molecules leads to dissociation profiles analogous to that registered for the monomer. However, a more saturated hydrogen-bonded network could drive the energy upwards to the values of bulk water, even though the nature of the excited state related to the transition is likely to be different to the one governing the gas-phase peak. An intermolecular type of excitation, as a state of charge-transfer nature, could provide useful information about the dynamic equilibrium held in bulk water in which protons are continuously transfered from one monomer to another in a picosecond timescale.

The π-stacked dimer is, on the other hand, constructed through a π-stacking interaction between both oxygen atoms present in the water molecules. This produces a delocalization of the orbitals in the molecule, which provides a significant stabilization of the excited state towards a minimum from where the system is expected to absorb. This delocalization drives the excited states to closer distances between the monomers and releases the steric effects by releasing the hindrance through O-H stretching processes, as it has been registered experimentally. It can therefore be envisaged that the effect of light irradiation upon π-stacked water will bring the O-O distance closer and hence increase the ordering given in the system, as it has also been reported in the experiments. The faint blue fluorescence observed experimentally is ratio-

nalized through an excimer-type fluorescence where an extra water molecule hydrates each of the π-stacked ones. Ordered water is expected to be surrounded by positive charge beyond the structured-water region in the form of protons. To that end, positive charges and external electric fields are used to provide estimates of the environmental effects generated over such molecular arrangements. Slight quantitative and no qualitative changes are found as compared to the gas-phase model, thus highlighting the prevalence of the intrinsic properties of the system over environmental perturbations.

Both dimers can be mainly differentiated by their localized and delocalized character. Whereas the localized character of the conventional dimer gives rise to monomer-like photodissociation mechanisms where the hydrogen bond plays a minor role, in the delocalized one, the photophysical and photochemical picture drastically changes. A whole new photochemical view is arisen, in which light drives the water monomers closer to each other and generates a structuring process, providing a fluorescence feature not observed in bulk water.

An hexameric model has been built upon the knowledge gathered from the π-stacked water dimer. It provides an extension to make a more realistic model that accounts for certain features characterized experimentally for ordered water. It is constructed in an ice-like manner through the fusion of two hexagonal rings thus forming $[H_{19}O_{10}]^{\cdot -}$. The $[H_{19}O_{10}]^{\cdot -}$ basic unit yields a plethora of spectroscopical fingerprints in the range from near infrared to near ultraviolet. The negative nature of the subunit and the loss of a proton in order to form it provide the charge separation feature missing in the π-stacked water dimer and thus reproduce more faithfully the experimental evidence. Most of the excitation processes are centered in the O-H-O subunit placed at the center of the system, and whereas its excitation energies do provide similar values as those expected for ordered water, the oscilator strength associated to such transitions is exceedingly small. A larger model, generated by stacking two $[H_{19}O_{10}]^{\cdot -}$ units has therefore been assessed in order to generate the essential π-stacking interaction seen to be the key factor in the spectroscopical fingerprint registered for ordered water. This yields the $H_{38}O_{20}$ system, which

is optimized obtaining a pronounced stabilization mainly attained for going from two negatively charged radicals by singlet pairing of the electrons. Several transitions have again been computed for this new system, giving rise to a signal with a relatively large oscilator strength associated to the transition that provides the absorption fingerprint that characterizes ordered water. As in the case of the π-stacked water dimer, environmental effects are deemed to be rather small and the essentials of the interaction will prevail yielding slight changes with respect to the computed values yet providing the same explanation for the absorption feature. Ongoing work is being done in order to elucidate the mechanisms responsible for the blue fluorescence feature, another of the distinctive properties of ordered water.

Photoionized processes have also been tackled as they provide yet another mechanism for charge separation that might be important for the process underlying the ordering of water, as well as a way to generate low-energy released electrons that can then produce dissociative electron attachment processes in the DNA/RNA. The properties under study are related to the ionization potentials of water clusters, which are chosen as the system under scrutiny as they provide a perfect way to track the evolution of the properties whilst increasing the size of the cluster. Different hydrogen-bonded connectivities are investigated through the use of four different water hexamers showing distinct hydrogen-bonding motifs. Structurally, the O-O distance between clusters decreases as its size and hydrogen-bond connectivity increases, yielding values similar to those expected for bulk water. Cationic structures show even smaller O-O distances as compared to their neutral counterparts, except in those cases where the cyclic-like structural motif is broken. The photoionization processes remove an electron from the system triggering a series of proton transfer processes, releasing a larger amount of energy as the number of such proton transfer processes increase within the cluster. The vertical and adiabatic properties computed throughout the study are then compared to the experimental values yielded in bulk water for these properties. It is unveiled that these properties converge to the bulk water value as the size of the cluster and the hydrogen-bonding connectivity present in them are increased,

being specially close for the hexamer prism that shows a hydrogen-bonding coordination of 3 as compared to the estimates of 3.2-3.4 given for bulk water.

Further efforts have been addressed into the DNA/RNA front by investigating new photochemically driven processes, like the aforementioned dissociative electron attachment in uracil. A deep revisit into the main photochemical pathways in the DNA/RNA nucleobases has been also undergone, and new avenues of inquiry are currently being pursued towards the validation of the previously characterized decay channels and also towards the elucidation of the combined effect of distinct intermolecular interactions between nucleobases in their relative position in the DNA/RNA chain.

Dissociative electron attachments (DEAs) are characterized as processes capable of generating DNA/RNA-strand breaks. The loose electrons present in the environment in a cellular system, arisen from previous ionization processes such as those previously described in water, attach themselves to antibonding orbitals in the nucleobases causing N-H bond cleavages that might serve as precursors of a series of harmful events. The DEA process has been studied in uracil, providing a molecular mechanism capable of rendering the essentials characterized experimentally. Highly correlated *ab initio* quantum chemical methods are deemed mandatory to yield a proper description for the system under study. In this sense, two main π-doublet states are protagonist of both N_1-H and N_3-H DEA processes. The N_1-H bond cleavage is therefore triggered by the population of the π_1^- state, which crosses with both the $\sigma^-_{N_1-H}$ and the ground state along the dissociation process. The low-lying energetically-wise crossings allow to assign the lowest experimental signals to the dissociation of the N_1-H bond. The N_3-H bond, on the other hand, is triggered by both π^- states, driving the system to subsequent crossings with the $\sigma^-_{N_3-H}$ and ground states on the way towards the bond cleavage. These crossings are placed higher in energy than those found for the N_1-H bond, and are in agreement with the higher energy bands registered experimentally. It is therefore predicted that one π_1^- state is responsible for the bond-breaking process of the N_1-H bond whereas two π-states, π_1^- and π_2^-, lead towards dissociation of the N_3-H one, placed at higher energies.

The photophysical and photochemical properties of the DNA/RNA nucleobases are revisited and the main photochemical decay paths present are elucidated. These processes are crucial to claim the assessment of the photostability related to the nucleobases, which is deemed of paramount importance as a protection mechanism to shield us from radiation. The results represent the main research line of the QCEXVAL group and give rise to a clear mechanistic view of the deactivation processes present in the DNA/RNA nucleobases. Natural pyrimidines share a common deactivation feature. The protagonist $\pi\pi^*$ bright state is initially populated and decays towards an ethene-like conical intersection with the ground state, reaching back to the FC region and thus remaining unchanged. This main deactivation channel provides photostability and it is assigned to the ultrafast component registered in experiments. The slower picosecond decay components in uracil and thymine are assigned to crossings with the excited state $n\pi^*$ and subsequent crossing with the ground state, thus preserving the photostability through another pathway. The nanosecond component is attributed to the triplet states and their possible population upon light irradiation in the $\pi\pi^*$ state.

Several photophysical and photochemical processes have been studied throughout this Thesis, emphasizing the importance of water and water mediated processes *in vivo*. A deeper understanding of the spectroscopical features of water has been acquired and a model for ordered water has been put in place. Processes derived from UV irradiation and ionization have been studied in water systems providing information that might tie its properties to the ones already characterized in the QCEXVAL group for the DNA/RNA nucleobases. Further advancement in both fronts prospects the possibility of combining the knowledge gained over the years on the photochemical pathways in DNA/RNA with the expertise attained here for different types of water thus providing a more realistic picture of the processes taking place in the cellular systems.

Chapter 11

Resumen

En el transcurso de los últimos años hemos sido testigos de un florecimiento sin precedentes de estudios fotoquímicos de biomoléculas, tanto desde el punto de vista experimental, como teórico. Sin duda alguna, ello se debe a un aumento de precisión en las técnicas desarrolladas, las cuales son capaces de revelarnos con mayor exactitud el comportamiento de los sistemas bioquímicos claves, posibilitándonos un conocimiento cada vez más exhaustivo de los procesos fotoquímicos a nivel molecular.

Cuando una molécula ha sido irradiada por la luz, dispone de una energía en exceso en un estado electrónico excitado que puede emplear de diferentes formas: de manera productiva, es decir, evolucionando hacia la formación de fotoproductos, emitiendo radiación, o bien de forma improductiva desde el punto de vista de la reactividad química, es decir, relajando su estado excitado y disipando su energía al exterior, con lo que el sistema regresa a su estado de partida y es considerado fotoestable.[1, 2] La fotorreactividad en los estados electrónicos excitados de las biomoléculas da lugar a procesos que, dependiendo de los sistemas, pueden ser beneficiosos o perjudiciales desde el punto de vista biológico, como en los cromóforos pertenecientes a las moléculas de ADN

y ARN, las bases púricas y pirimídicas de los ácidos nucleicos, en las que existe un daño fotobiológico que puede conllevar fenómenos de mutagénesis y carcinogénesis, por ejemplo, a través de la fototautomerización por transferencia de átomos de hidrógeno en los pares de bases,[49, 50] aunque sus monómeros básicos presentan una notable fotoestabiliad,[6, 7] o en procesos de fotoadición de substancias exógenas que conducen a modificaciones genéticas,[3, 53] y que constituyen la base de las técnicas clínicas fototerapéuticas. El interés se extiende además a los aminoácidos y proteínas, a su comportamiento fotoquímico y a su relación con el ADN y su función replicadora, y no sólo en especies naturales sino modificadas, así como en secuencias mutadas de aminoácidos empleadas como herramientas de detección y manipulación en biología molecular, necesarias todas ellas para profundizar en el conocimiento básico en la genómica y proteómica moderna.

En el año 2003, el grupo QCEXVAL (*Quantum Chemistry for the Excited State of Valencia*) del Instituto de Ciencia Molecular de la Universitat de València fue pionero en la introducción del uso de los métodos *ab initio* altamente correlacionados como el CASPT2 en el estudio de la fotoquímica no adiabática de las bases de los ácidos nucleicos,[8] en las cuales los efectos de correlación electrónica dinámica no habían sido tenidos en cuenta hasta la fecha. Desde entonces, el grupo QCEXVAL ha profundizado notablemente en el campo de la fotoquímica y fotobiología teóricas, terreno en el cual el potencial del cálculo químico-cuántico preciso y predictivo no había sido desarrollado en su justa medida. El presente proyecto de Tesis abarca el estudio sistemático del agua, así como una extensión al trabajo realizado en el grupo QCEXVAL sobre las nucleobases, haciendo uso de los métodos altamente correlacionados previamente mencionados.

Desde su caracterización inicial a finales del siglo XIX, el agua ha sido, sin lugar a dudas, la substancia más estudiada de la que se tiene constancia. La determinación de su composición a cargo de Henry Cavendish y Antoine Laurent Lavoisier,[16] marca un punto de partida al que siguen una infinidad de análisis que aún perduran a día de hoy. Ya durante el siglo XX entre los grandes investigadores que hicieron del agua su objeto de estudio destacan

Henry Eyring y especialmente Linus Pauling, Premio Nobel de Química del año 1954, cuya mayor contribución resultó en la descripción de los enlaces químicos; dichos enlaces, a pesar de que describen la unión de los átomos de hidrógeno con el oxígeno, no son capaces de describir con exactitud la interacción de distintas moléculas de agua entre sí mediante los llamados *enlaces de hidrógeno*.[32] La constante evolución en la materia se aprecia con claridad con la reciente redefinición por la IUPAC de los mencionados *enlaces de hidrógeno* a los que se le atribuyen un papel crucial tanto en la mayoría de los procesos biológicos en los que interviene el agua, como de prácticamente todas las anomalías que presenta respecto a otros líquidos.

Por otro lado, es de señalar el agua como parte central de la vida ya que a su estructura se le atribuyen una gran variedad de procesos tanto biológicos como industriales. Entre otros destacan la ionización y solvatación de ácidos, bases, sales y moléculas orgánicas, la generación de micelas, el funcionamiento de las membranas biológicas, el mecanismo de pliegue de las proteínas, y muchas otras propiedades cuyos mecanismos exactos aún se desconocen.

No obstante la larga corriente de estudios y la importancia fundamental del agua hoy en día aún hay muchos interrogantes. En ese sentido muchos son los científicos que intentan actualmente explicar diversos fenómenos en relación con su estructura, dado que es la que se supone responsable de las excéntricas propiedades, tanto físicas como químicas, que presenta el agua: alta capacidad calorífica, se expande cuando se congela, presenta un máximo de densidad a 4 °C y una alta constante dieléctrica. Estos trabajos se centran principalmente en el estudio de *clústeres* de agua,[31] tanto teórica como experimentalmente, permitiendo la obtención de estructuras que explican parcial o completamente algunas de las propiedades antes mencionadas. Sin embargo, la estructura del agua líquida aún no ha podido ser determinada con total exactitud, siendo generalmente aceptado considerar que es un tetraedro flexible como su forma en hielo hexagonal donde el agua se encuentra rodeada de cuatro enlaces de hidrógeno, dos dadores y dos aceptores.[33] Esta aceptación general es cuestionada por muchos y las críticas se han intensificado en los últimos años a raíz de distintos estudios que sugieren otras alternativas muy diferentes; entre

los análisis que han creado más repercusión cabe destacar entre ellos el trabajo publicado por Nilsson y colaboradores en la revista científica Science en 2004,[37] en el que se defiende una estructura lineal unida por tan sólo dos enlaces de hidrógeno, y la gran cantidad de trabajos que suscitó dicha estructura tanto a nivel teórico como experimental que conforman el punto de partida en el cual se basa el presente proyecto.

Hasta la fecha la inmensa mayoría de los estudios relacionados con el agua se han centrado principalmente en el agua en su estado fundamental y, por tanto, los estudios tanto fotofísicos como fotoquímicos basados en los estados electrónicos excitados son escasos, especialmente desde un punto de vista teórico. La interacción del agua con la radiación es un tema de suma importancia dado que estamos rodeados de luz y expuestos a ella la mayoría del tiempo, puesto que el ser humano está formado al menos en un 70% por agua, dependiendo de la edad. La introducción de la presencia explícita del agua en estos sistemas se ha abordado en esta Tesis a través del estudio de los *clústeres* de agua, implicando progresivamente un mayor número de moléculas y complejidad.

El espectro de distintas biomoléculas en disolución acuosa y sus posibles mecanismos de interacción con la luz, que principalmente consisten en transferencias tanto de carga como de hidrógeno, han sido extensamente estudiadas y diversos mecanismos han sido sugeridos. El papel de los estados electrónicos excitados sigue pasando por alto y sus mecanismos de acción detallados aún se desconocen.

En concreto, se pretende contribuir a la comprensión de la fotoquímica del agua mediante el cálculo de parámetros energéticos y propiedades moleculares en *clústeres* de distintos tamaños, así como también en diferentes conformaciones describiendo distintos tipos de interacción. Asimismo, se analiza mediante el uso de la metodología químico cuántica distintas estructuras para las diferentes fases del agua de forma que se alcance una mayor comprensión de sus propiedades, ya que éstas son necesarias para la continuación de los estu-

dios en biomoléculas, haciendo especial énfasis y ampliando el conocimiento previamente adquirido sobre los cromóforos del ADN/ARN.

Metodología Para obtener una mejor comprensión de los fenómenos que ocurren en los *clústeres* de agua cuando éstos son irradiados con luz UV-Visible es necesaria la observación del efecto de dicha energía sobre la estructura electrónica del sistema molecular correspondiente. Los problemas electrónicos pueden ser tratados mediante la química computacional.[61] En concreto, los métodos *ab initio* han demostrado ser altamente eficientes y precisos a la hora de describir la estructura electrónica de los sistemas moleculares. Para una buena representación de los estados electrónicos excitados se requiere la consideración de varias configuraciones electrónicas en su referencia, lo que sugiere el uso de métodos multiconfiguracionales. El método elegido para la descripción de la estructura básica de los estados excitados del sistema ha sido el CASSCF,[68] que permite describir la correlación electrónica de largo alcance (correlación no dinámica). Con objeto de realizar un tratamiento exhaustivo del resto de la correlación, denominada correlación dinámica (correlación electrónica a corto alcance), se ha empleado el método de multirreferencia perturbativa CASPT2.[64, 65] El uso de ambas metodologías, extensamente empleadas y contrastadas a lo largo de los años por el grupo QCEXVAL dentro del cual se ha realizado la presente Tesis Doctoral,[66] nos ha permitido profundizar en algunos aspectos relevantes de la fotofísica y fotoquímica del agua así como también de los componentes del ADN/ARN. Para proporcionar una descripción precisa de dichos procesos fotoquímicos, se han empleado técnicas avanzadas de caracterización de las rutas de decaimiento más probables a través del cómputo de caminos de mínima energía (MEPs)[94] y la localización de regiones o puntos de cruce más relevantes entre las superficies de energía potencial. El uso de metodologías de dinámica molecular no-adiabática nos ha permitido finalmente facilitar estimaciones de la evolución temporal de los sistemas tras ser irradiados por la luz.[100]

Los métodos previamente descritos han sido calibrados con métodos monoreferenciales contrastados para la descripción del estado electrónico fundamental,

como son la teoría de perturbaciones de Møller-Pesset (MP2) o el método de los *clústeres* acoplados (Coupled Cluster).[125] En los cálculos del presente trabajo de Tesis se han empleado los paquetes de programas químico-cuánticos GAUSSIAN y MOLCAS,[69, 108] de amplia difusión, los cuales incluyen la mayoría de las herramientas necesarias para los estudios realizados.

Resultados y Conclusiones La presente Tesis Doctoral provee una serie de conclusiones basadas en los estudios realizados durante la realización de la misma. En líneas generales, la caracterización teórica de las propiedades espectroscópicas asociadas al agua ha sido completada y ampliada. Diferentes modelos de agua han sido postulados para mejorar la cararcterización tanto del agua tipo *bulk*,[23–25] generalmente asignada a la parte interior del agua líquida en la que hay un equilibrio dinámico de los enlaces de hidrógeno, así como del agua denominada estructurada u ordenada, presente generalmente en la interfase entre sólidos y biomoléculas y la propia agua líquida.[42] Ambos tipos de agua poseen señales espectroscópicas diferenciables, caracterizadas claramente desde un punto de vista experimental, facilitando datos experimentales con los que comparar los valores teóricos aquí postulados para proveer de un mecanismo molecular que explique las diferentes señales medidas en ambas.

Los estados excitados del agua que han sido investigados son aquellos involucrados en procesos originados por la irradiación con luz UV y por la ionización, facilitando su comprensión e impulsando el conocimiento en áreas relacionadas como lo son el daño en el ADN/ARN, de vital importancia para la investigación biomédica. En contreto, la adición de electrones disociativa en nucleobases (*dissociative electron attachment*) debida a los electrones sueltos tras la ionización del agua y su posterior fijación en nucleobases, es precursora de roturas de las cadenas de ADN/ARN.

Específicamente, dos tipos de modelos basados en el dímero de agua, el convencional caracterizado experimentalmente en la década de los ochenta por Dycke y colaboradores,[102] y el denominado apilamiento π (π-*stacked*), han

sido estudiados con objeto de caracterizar los modelos más reducidos que contienen las interacciones intermoleculares intrínsecas responsables de las diferentes señales espectroscópicas adquiridas experimentalmente.

El dímero de agua convencional se caracteriza por estar formado por dos aguas unidas a través de un enlace de hidrógeno.[102] Sus características espectroscópicas han sido obtenidas asemejándose más a aquellas conocidas sobre el agua en fase gas a las propiedades conocidas del agua líquida. Es por ello que podemos considerar que el efecto del enlace de hidrógeno entre ambas aguas va a jugar un papel secundario en los estados electrónicos excitados que definen las señales espectroscópicas medidas experimentalmente, dando lugar a señales similares a las que se tendrían con moleculas de agua individuales.[118] Los procesos de fotodisociación presentes tras la irradiación con luz son análogos a aquellos observados para la molécula de agua,[27] en este caso centradas en cada una de las dos aguas que forman el dímero.[109] Una mayor presencia de enlaces de hidrógeno podría cambiar ligeramente las propiedades fotofísicas y fotoquímicas del sistema, pudiéndose acercar posiblemente a las señales experimentales registradas para el agua líquida de tipo *bulk*,[23–25] basadas posiblemente en excitaciones a estados de transferencia de carga, que podrían explicar el equilibrio dinámico presente en el agua *bulk*, pero no al rango de señales que describen el agua ordenada.

El dímero de tipo π ha sido construido a través del apilamiento de ambas moléculas de agua a través de sus átomos de oxígeno.[129] Dicha interacción produce una deslocalización de los orbitales en la molécula, que resulta en una gran estabilización de su estado electrónico excitado más bajo en energía. La deslocalización promueve el acercamiento de las aguas a través de su estado excitado y lo libera de efectos estéricos a través del alargamiento de los enlaces O-H (*stretching*), tal y como se muestra en el experimento.[42] El efecto de la luz irradiada sobre el dímero apilado π reduce la distacia O-O entre los monómeros de agua aumentando el grado de ordenación presente. La fluorescencia azul observada experimentalmente puede ser explicada a través de un modelo excimérico, en el que las aguas del dímero apilado π son solvatadas lateralmente por otros dos monómeros formando un tetrámero. El entorno

que rodea el dímero π apilado y las capas de agua ordenada está formado por protones,[150] que han sido simulados aplicando campos eléctricos externos y acercando cationes de litio al sistema, obteniendo el mismo resultado a nivel cualitativo que en el vacío, reforzando así las propiedades caracterizadas para el dímero apilado π como algo intrínseco del agua y no como un efecto debido a un determinado entorno.

Ambos dímeros se diferencian principalmente en la localización de sus orbitales. Mientras que en el dímero de agua convencional los orbitales se localizan en ambas aguas dando lugar a procesos fotofísicos y fotoquímicos análogos a los previamente vistos para el monómero de agua,[109] en el dímero apilado π los orbitales se deslocalizan cambiando por completo su comportamiento.[129] Dicho cambio se manifiesta, por ejemplo, en la aparición de una señal espectroscópica debida a la fluorescencia, algo que no ocurre en el dímero convencional.

Basado en los conocimientos adquiridos en el estudio de los dímeros de agua, se postula un modelo hexamérico del agua de mayor tamaño consistente con las propiedades anómalas determinadas experimentalmente para el agua ordenada. El modelo se basa en la fusión de dos anillos de agua obteniendo la especie $[H_{19}O_{10}]^{\cdot -}$. Se predice que dicha unidad contiene toda una serie de estados electrónicos excitados en un amplio rango que abarca el infrarrojo cercano y el ultravioleta que la convierten en una unidad altamente flexible para absorber luz. La naturaleza aniónica de dicha especie y la pérdida del protón que genera la carga negativa ayuda a explicar la separación de carga generada en las capas de agua ordenada, algo que no tenía en cuenta el dímero apilado π.[129] La mayoría de los procesos de excitación se centran en la subunidad O-H-O que se halla en el centro de la molécula, obteniéndose valores energéticos similares a los determinados experimentalmente para el agua ordenada pero con fuerza de oscilador excesivamente pequeña. Por ello, y teniendo en cuenta la relevancia del apilamiento π en el dímero de agua se ha estudiado el modelo formado por dos subunidades $[H_{19}O_{10}]^{\cdot -}$ apiladas, resultando en la molécula $H_{38}O_{20}$. Este nuevo sistema se ve altamente estabilizado al emparejar los electrones separados de los monómeros de $[H_{19}O_{10}]^{\cdot -}$. Se han calculado para este

sistema toda una serie de excitaciones, obteniéndose una transición con una gran fuerza de oscilador, coincidente energéticamente con el valor experimental obtenido para el agua ordenada.[42] Al igual que ocurre en el dímero apilado π, los efectos del entorno se consideran relativamente pequeños y despreciables a nivel cualitativo obteniendo un modelo para la absorción observada experimentalmente. La presencia de la señal fluorescente está siendo investigada en la actualidad, representando la otra señal característica del agua ordenada.

Los procesos de fotoionización han sido estudiados ya que proveen de otro mecanismo de separación de carga que puede ser de importancia a la hora de generar capas de agua ordenada, así como para generar electrones de baja energía que pueden producir roturas de cadena en el ADN/ARN. Las propiedades estudiadas corresponden a los potenciales de ionización de los *clústeres* de agua.[125] El conjunto de agregados de agua se ha escogido de tal forma que nos permitiera analizar la evolución de las propiedades calculadas al aumentar el tamaño y la conectividad a través de enlaces de hidrógeno.[31] Estructuralmente, la distancia O-O entre las aguas en los *clústeres* decrece a medida que se aumenta el tamaño del *clúster* así como también su conectividad. Las estructuras catiónicas muestran una tendencia decreciente similar a la observada para las estructuras neutras, excepto en los casos en los que la estructura cíclica se rompe. Los procesos de fotoionización substraen un electrón de los *clústeres* iniciando una serie de procesos de transferencia de protones que liberan una gran cantidad de energía, mucho mayor cuanto mayor es el número de transferencias protónicas. Las propiedades tanto verticales como adiabáticas son comparadas con valores experimentales y teóricos hallados en la bibliografía así como valores asignados a agua tipo *bulk*. Los potenciales de ionización convergen con el valor experimental obtenido para el agua tipo *bulk* a medida que se aumenta el tamaño del *clúster* y su conectividad. Es especialmente cercano el caso del hexámero prisma que está constituido por tres enlaces de hidrógeno, siendo el valor más similar a los 3.2-3.4 enlaces de hidrógeno estimado para la agua *bulk*.[35]

La línea de investigación basada en el estudio de los procesos fotoquímicos en el ADN/ARN ha sido continuada a través del análisis de los mencionados pro-

cesos de adición de electrones disociativa en el uracilo.[175] Posteriormente se ha procedido a revisar el conocimiento obtenido durante la última década en el seno del grupo QCEXVAL acerca de los caminos de decaimiento fotoquímicos que describen la fotoestabilidad de las nucleobases,[182] contrastándolos con métodos más precisos disponibles en la actualidad con tal de validar los mecanismos previamente elucidados.

La adición de electrones disociativa es un proceso considerado como precursor de roturas de las cadenas de ADN/ARN. Los electrones sueltos presentes en el entorno celular son generados en procesos previos de fotoionización como los anteriormente descritos para el agua.[125] Dichos electrones pueden añadirse a los orbitales antienlazantes de las nucleobases, causando la rotura de enlaces N-H, cosa que puede desencadenar en toda una serie de procesos nocivos. La adición de electrones disociativa ha sido estudiada en el caso concreto de uracilo, facilitando una visión molecular y mecanística de los eventos que ocurren en la nucleobase para dar lugar a la señal experimental previamente registrada. Los métodos *ab initio* altamente correlacionados permiten proveer una descripción suficientemente precisa de estos procesos y por tanto han sido los métodos escogidos para el estudio. La rotura de ambos enlaces N-H (N_1-H y N_3-H) presentes en uracilo ha sido investigada. La rotura del enlace N_1-H se asigna a la señal experimental energéticamente más baja y es asociada a un estado de tipo π^- que cruza con el estado $\sigma^-_{N_1H}$, el cual impulsa la rotura del enlace N_1-H. Un mecanismo similar se deduce para la rotura del enlace N_3-H, en cuyo caso hay dos estados π^- que pueden poblar el estado $\sigma^-_{N_3H}$ a través del cual se disocia el hidrógeno. Dichos cruces se encuentran a mayor energía del cruce obtenido para el mecanismo de disociación del enlace N_1-H, en concordancia con los datos experimentales. De esta manera, podemos predecir que un estado π^- es el responsable de la rotura del enlace N_1-H mientras que dos estados π^- se encargan de la disociación del enlace N_3-H, que se encuentra a mayor energía.[174, 179, 180]

Las propiedades fotofísicas y fotoquímicas del ADN/ARN y sus procesos asociados, cruciales para la determinación de la fotoestabilidad inherente de las nucleobases y los mecanismos de protección que nos protegen de la radiación UV, han sido reconsiderados, ampliándose con métodos más correlacionados que se encuentran a nuestro alcance gracias a la tecnología hoy en día disponible. Estos resultados representan la línea principal de investigación del grupo QCEXVAL y nos permiten diseñar mecanismos teóricos que explican la desactivación de las nucleobases tras ser irradiadas con luz. Las pirimidinas, las nucleobases en la que nos hemos centrado, poseen un mecanismo común de desactivación a través de su estado brillante $\pi\pi^*$,[7] el cual tras ser irradiado decae hasta una intersección cónica con el estado fundamental con una distorsión geométrica similar a la conocida para el eteno, en el que el doble enlace C=C se estira y gira dando lugar a una conformación de anillo plegado, que posteriormente decae a la geometría Franck-Condon en el estado electrónico fundamental recobrando la planaridad. Esta vía de desactivación es la principal conocida para las pirimidinas y se asocia a una señal experimental encontrada en la escala de los femtosegundos. En el trabajo de Tesis, ha sido refrendada con métodos más correlacionados de los usados inicialmente, en concreto con la estrategia computacional CASPT2//CASPT2, validando los mecanismos previamente establecidos con el protocolo CASPT2//CASSCF. La segunda señal experimental observada en el rango de los picosegundos se puede asociar al cruce del estado brillante $\pi\pi^*$ con el estado $n\pi^*$ y la subsiguiente evolución hacia el estado fundamental, mientras que la componente más lenta, y asociada a los nanosegundos, se asigna directamente a mecanismos de población de los estados tripletes y a su evolución hacia su cruce con el estado fundamental.[186]

Diversos procesos fotofísicos y fotoquímicos han sido estudiados a lo largo de la presente Tesis Doctoral, enfatizando en la importancia del agua y de los procesos mediados por agua. Se ha profundizado en el conocimiento acerca de las propiedades espectroscópicas del agua, tanto del agua *bulk* como agua estructurada. El efecto de la luz UV e ionizante han sido estudiado en sistemas acuosos y en nucleobases, aportando información esencial que nos permita entender con mayor precisión los procesos intrínsecos del agua, de las bases

nucleicas y del posible papel del agua en sistemas biológicos. El avance en ambos frentes, tanto sobre el agua como en el ADN/ARN, promete precisar a largo plazo un mecanismo más realista de los procesos que ocurren en el sistema celular.

Appendix A

Basis Set Superposition Error

Basis Set Superposition Error (BSSE) appears due to the truncation and therefore the finite character of the basis sets. As the atoms of interacting molecules or separate molecules approach one another, their basis functions overlap. Each monomer *borrows* functions from other nearby components, effectively increasing its basis set and artificially improving the calculation.

In order to correct this effect, several treatments have been proposed and the one used in this work is the one developed by Boys and Bernardi.[87] This technique is based on the calculation of both monomers separately with their own basis sets and then subtract its own energy with that obtained by computing the energy in the presence of the dimer's basis set. This allows the quantification of the extra-energy given in the dimer, and therefore allows the correction of the energy by simple subtraction of this energy to the total value obtained for the dimer.

For a water dimer, such as the π-stacked water dimer discussed in this work, the non-corrected E_b is given by

$$E_b(WW^*) = E_W + E_{W^*} - E_{W^*W} \tag{A.1}$$

with E_W, and E_{W^*} being the total energies of the ground (W), and excited state (W^*) of the water molecule, both at the equilibrium geometry of the ground-state water, and E_{W^*W} representing the total energy of the W^*W dimer at a given geometry. The corrected counterpoise binding energy ($CP - E_b$) comes from the expression:

$$CP - E_b(W^*W) = E_b(W^*W) - [CP - BSSE(W^*W)] \tag{A.2}$$

where

$$\begin{aligned} CP - BSSE(W^*W) = E_{W^*}(W, R = \infty) - E_{W^*}(W, R = W^*W) \\ + E_W(W^*, R = \infty) - E_W(W^*, R = W^*W) \end{aligned} \tag{A.3}$$

Here the notation $E_{W^*}(W, R = W^*W)$ indicates the energy of W^* computed in the ghost orbitals of W at the geometry of W^*W, whereas in $E_{W^*}(W, R = \infty)$ the ghost orbitals of W are at *infinite* distance of W^*. The ghost orbitals imply the presence of the dimer's basis set in the calculation of the energy for the monomer, as has been stated above.

Appendix B

List of Publications

Paper I

***Ab initio* determination of the ionization potentials of water clusters $(H_2O)_n$ (n=2-6)**
Javier Segarra-Martí, Manuela Merchán, and Daniel Roca-Sanjuán
The Journal of Chemical Physics, **136**, 244306 (2012).

Paper II

On the N_1-H and N_3-H Bond Dissociation in Uracil by Low Energy Electrons: A CASSCF/CASPT2 Study
Israel González-Ramírez, Javier Segarra-Martí, Luis Serrano-Andrés, Manuela Merchán, Mercedes Rubio, and Daniel Roca-Sanjuán
Journal of Chemical Theory and Computation, **8**, 2769 (2012).

Paper III

On the photophysics and photochemistry of the water dimer
Javier Segarra-Martí, Daniel Roca-Sanjuán, Manuela Merchán, and Roland Lindh
The Journal of Chemical Physics, **137**, 244309 (2012).

Paper IV

Towards the Understanding at the Molecular Level of the Structured-Water Absorption and Fluorescence Spectra: A Fingerprint of π-Stacked Water
Javier Segarra-Martí, Pedro B. Coto, Mercedes Rubio, Daniel Roca-Sanjuán, and Manuela Merchán
Molecular Physics, **111**, 1308 (2013).

Paper V

Excitation of nucleobases from a computational perspective I: reaction paths
Angelo Giussani, Javier Segarra-Martí, Daniel Roca-Sanjuán, and Manuela Merchán
Topics in Current Chemistry, Accepted.

Paper VI

On the Hexagonal Ice-like Model of Structured Water: Theoretical Analysis of the Low-lying Excited States
Javier Segarra-Martí, Daniel Roca-Sanjuán, and Manuela Merchán
Submitted.

List of Figures

List of Tables

List of Acronyms

A

AC Avoided Crossing

AIMD *Ab Initio* Molecular Dynamics

AIP$_e$ Adiabatic Ionization Potential

AIP$_0$ Adiabatic Ionization Potential with ZPE corrections

ANO Atomic Natural Orbital

B

BFGS Broyden-Fletcher-Goldfarb-Shanno update

BMU Basic Molecular Unit

BSSE Basis Set Superposition Error

C

CAS Complete Active Space

CASPT2 Complete Active Space Second-Order Perturbation Theory

CASSCF Complete Active Space Self Consistent Field

CC Coupled Cluster

CCSD Coupled Cluster Singles and Doubles

CCSDT Coupled Cluster Singles Doubles and Triples

CCSD(T) Coupled Cluster Singles Doubles and Triplets perturbationally added

CGTO Contracted Gaussian-Type Orbital

CI Configuration Interaction
Conical Intersection

CSF Configuration State Function

D

DB Dipole-Bound

DCI Doubly excited Configuration Interaction

DEA Dissociative Electron Attachment

DFT Density Functional Theory

DNA Deoxyribonucleic acid

DZ Double Zeta

E

EA Electron Affinity

EOM Equation of Motion

ETS Electron Transmission Spectroscopy

eV electron-Volt

F

f Oscillator strength

fs femtosecond

FC Franck-Condon

FCI Full Configuration Interaction

G

GTO Gaussian-Type Orbitals

H

H HOMO

HF Hartree Fock

HOMO Highest Occupied Molecular Orbital

I

IC Internal Conversion

I_h Hexagonal Ice

IP Ionization Potential

IRC Intrinsic Reaction Coordinate

ISC Intersystem Crossing

IVR Intramolecular Vibrational Relaxation

L

L LUMO

LIIC Linear Interpolation of Internal Coordinates

LS-CASPT2 Level-Shift Complete Active Space Second-Order Perturbation Theory

LUMO Lowest Unoccupied Molecular Orbital

M

MBPT Many Body Perturbation Theory

MCSCF Multiconfiguration Self Consistent Field

MD Molecular Dynamics

MECI Minimum Energy Conical Intersection

MECP Minimum Energy Crossing Point

MEP Minimum Energy Path

MM Molecular Mechanics

MO Molecular Orbital

MP2 Møller-Plesset second-order Perturbation Theory

MRCC Multireference Coupled Cluster

MRCI Multireference Configuration Interaction

MS-CASPT2 Multi-State Complete Active Space Second-Order Perturbation Theory

MO-LCAO Molecular Orbitals as Linear Combination of Atomic Orbitals

N

nm nanometers

NO Natural Orbital

O

P

PHF Projected Hartree Fock

PMP Projected Møller Plesset Perturbation Theory

PES Potential Energy Surface

PEH Potential Energy Hypersurface

PEC Potential Energy Curve

Q

QM Quantum Mechanics

QM/MM Quantum Mechanics/Molecular Mechanics

QCEXVAL Quantum Chemistry for the Excited State of Valencia

R

RNA Ribonucleic acid

ROHF Restricted Open Hartree Fock

RASSCF Restricted Active Space Self Consistent Field

S

SCF Self Consistent Field

SCI Singly excited Configuration Interaction

SDCI Singly and Doubly excited Configuration Interaction

SDTCI Singly, Doubly, and Triply excited Configuration Interaction

STO Slater-Type Orbitals

$\mathbf{S}_0$ Singlet Ground State

$\mathbf{S}_1$ First energetically-wise singlet excited state

$\mathbf{S}_2$ Second energetically-wise singlet excited state

$\mathbf{S}_3$ Third energetically-wise singlet excited state

$\mathbf{S}_4$ Fourth energetically-wise singlet excited state

SONO Singly Occupied Natural Orbital

SDC Singlet Doublet Crossing

STC Singlet Triplet Crossing

T

TZ Triple Zeta

TS Transition State

TDM Transition Dipole Moment

U

UHF Unrestricted Hartree Fock

UMP2 Unrestricted Møller Plesset second-order Perturbation Theory

V

VDZ Valence Double Zeta

VTZ Valence Triple Zeta

VIP Vertical Ionization Potential

VB Valence-Bound

VEA Vertical Electron Affinity

W

$\mathbf{W}_D$ Water donor

$\mathbf{W}_A$ Water acceptor

X

Y

Z

ZPE Zero Point Energy

Bibliography

[1] N. J. Turro. *Modern Molecular Photochemistry.* University Science Books, Sausalito, California, 1991.

[2] M. Klessinger and J. Michl. *Excited States and Photochemistry of Organic Molecules.* VCH Publishers Inc., New York, 1995.

[3] D. Roca-Sanjuán, G. Olaso-González, I. González-Ramírez, L. Serrano-Andrés, and M. Merchán. Molecular Basis of DNA Photodimerization: Intrinsic Production of Cyclobutane Cytosine Dimers. *Journal of the American Chemical Society*, 130:10768–10779, 2008.

[4] L. Serrano-Andrés, M. Merchán, and A. C. Borin. A Three-State Model for the Photophysics of Adenine. *Chemistry. A European Journal*, 12:6559–6571, 2006.

[5] L. Serrano-Andrés, M. Merchán, and A. C. Borin. Adenine and 2-Aminopurine: Paradigms of Modern Theoretical Photochemistry. *Proceedings of the National Academy of Sciences USA*, 103:8691–8696, 2006.

[6] L Serrano-Andrés, M Merchán, and A. C. Borin. A Three-State Model for the Photophysics of Guanine. *Journal of the American Chemical Society*, 130:2473–2484, 2008.

[7] M. Merchán, R. González-Luque, T. Climent, L. Serrano-Andrés, E. Rodríguez, M. Reguero, and D. Peláez. Unified Model for the Ultrafast Decay of Pyrimidine Nucleobases. *Journal of Physical Chemistry B*, 110(26471-26476), 2006.

[8] M. Merchán and L Serrano-Andrés. Ultrafast Internal Conversion of Excited Cytosine via the Lowest pipi* Electronic Singlet State. *Journal of the American Chemical Society*, 125:8108–8109, 2003.

[9] P. Ball. How to keep dry in water. *Nature*, 423:25–26, 2003.

[10] P. Ball. Water, water, everywhere? *Nature*, 427:19–20, 2004.

[11] P. Ball. Seeking the solution. *Nature*, 436:1084–1085, 2005.

[12] P. Ball. Water-an enduring mystery. *Nature*, 452:291–292, 2008.

[13] P. Ball. Water as an Active Constituent in Cell Biology. *Chemical Reviews*, 108:74–108, 2008.

[14] P. Ball. Water as a Biomolecule. *ChemPhysChem*, 9:2677–2685, 2008.

[15] P. Ball. *H2O: A biography of water*. Oring Publishing Co., 2000.

[16] A. L. Lavoisier. *Traité Élémentaire de Chimie*. 1789.

[17] Y. Maréchal. *The Hydrogen Bond and the Water Molecule*. Elsevier, Amsterdam, 2007.

[18] B. C. Garrett, D. A. Dixon, D. M. Camaioni, D. M. Chipman, M. A. Johnson, C. D. Jonah, G. A. Kimmel, J. H. Miller, T. N. Rescigno, P. J. Rossky, S. S. Xantheas, S. D. Colson, A. H. Laufer, D. Ray, P. F. Barbara, D. M. Bartels, K. H. Becker, K. H. Bowen, S. E. Bradforth, I. Carmichael, J. V. Coe, L. R. Corrales, J. P. Cowin, M. Dupuis, K. B. Eisenthal, J. A. Franz, M. S. Gutowski, K. D. Jordan, B. D. Kay, J. A. LaVerne, S. V. Lymar, T. E. Madey, C. W. McCurdy, D. Meisel, S. Mukamel, A. R. Nilsson, T. M. Orlando, N. G. Petrik, S. M. Pimblott, J. R. Rustad, G. K. Schenter, S. J. Singer, A. Tokmakoff, L.-S. Wang, , and T. S. Zwier. Role of Water in Electron-Initiated Processes and Radical Chemistry: Issues and Scientific Advances. *Chemical Reviews*, 105:355–390, 2005.

[19] O. Mishima and H. E. Stanley. The relationship between liquid, supercooled and glassy water. *Nature*, 396:329–335, 1998.

[20] P. Kumar, K. T. Wikfeldt, D. Schlesinger, L. G. M. Pettersson, and H. E. Stanley. The Boson peak in supercooled water. *Nature Scientific Reports*, 3:doi:10.1038/srep01980, 2013.

[21] G. H. Pollack. *The Fourth Phase of Water: Beyond Solid, Liquid, and Vapor*. Ebner and Sons, Seattle, WA, 2013.

[22] R. Mota, R. Parafita, Giuliani A., M.-J. Hubin-Franskin, J. M. C. Lourenço, G. Garcia, S. V. Hoffman, N. J. Mason, P. A. Ribeiro, M. Raposo, and P. Limao-Vieira. Water VUV electronic state spectroscopy by synchrotron radiation. *Chemical Physics Letters*, 416:152–159, 2005.

[23] V. Garbuio, M. Cascella, L. Reining, and O. Del Sole, R. Pulci. Ab Initio Calculation of Optical Spectra of Liquids: Many-Body Effects in the Electronic Excitations of Water. *Physical Review Letters*, 97:137402, 2006.

[24] V. Garbuio, M. Cascella, and O. Pulci. Excited state properties of liquid water. *Journal of Physics: Condensed Matter*, 21:033101, 2009.

[25] A. Hermann, W. G. Schmidt, and P. Schwerdtfeger. Resolving the Optical Spectrum of Water: Coordination and Electrostatic Effects. *Physical Review Letters*, 100:207403, 2008.

[26] K. Kobayashi. Optical Spectra and Electronic Structure of Ice. *Journal of Physical Chemistry*, 87:4317–4321, 1983.

[27] M. Rubio, L. Serrano-Andrés, and M. Merchán. Excited states of the water molecule: Analysis of the valence and rydberg character. *Journal of Chemical Physics*, 128:104305, 2008.

[28] K. Liu, J. D. Cruzan, and R. J. Saykally. Water Clusters. *Science*, 271:929–933, 1996.

[29] K. Liu, M. G. Brown, J. D. Cruzan, and R. J. Saykally. Vibration-Rotation Tunneling Spectra of the Water Pentamer: Structure and Dynamics. *Science*, 271:62–64, 1996.

[30] J. D. Cruzan, L. B. Braly, K. Liu, M. G. Brown, J. G. Loeser, and R. J. Saykally. Quantifying Hydrogen Bond Cooperativity in Water: VRT Spectroscopy of the Water Tetramer. *Science*, 271:59, 1996.

[31] R. Ludwig. Water: From Clusters to the Bulk. *Angewandte Chemie International Edition*, 40:1808–1827, 2001.

[32] G. R. Desiraju. A Bond by Any Other Name. *Angewandte Chemie International Edition*, 50:52–59, 2011.

[33] F. H. Stillinger. Water Revisited. *Science*, 209:451–457, 1980.

[34] T. D. Kühne and R. Z. Khaliullin. Electronic signature of the instantaneous asymmetry in the first coordination shell in liquid water. *Nature Communications*, 4:1450–1458, 2013.

[35] R. Kumar, J. R. Schmidt, and J. L. Skinner. Hydrogen bonding definitions and dynamics in liquid water. *Journal of Chemical Physics*, 126:204107, 2007.

[36] M. V. Fernández-Serra and E. Artacho. Electrons and hydrogen-bond connectivity in liquid water. *Physical Review Letters*, 96:016404, 2006.

[37] P. Wernet, D. Nordlung, U. Bergmann, M. Cavalleri, M. Odelius, H. Ogasawara, L. A. Näslund, T. K. Hirsch, L. Ojamäe, P. Glatzel, L. G. M. Pettersson, and A. R. Nilsson. The Structure of the First Coordination Shell in Liquid Water. *Science*, 304:995–999, 2004.

[38] C. J. T. de Grotthuss. Sur la décomposition de l'eau et des corps qu'elle tient en dissolution à l'aide de l'électricité galvanique. *Annales de Chimie (Paris)*, LVIII:54–74, 1806.

[39] A. Hassanali, M. K. Prakash, H. Eshet, and M. Parrinello. On the recombination of hydronium and hydroxide ions in water. *Proceedings of the National Academy of Sciences USA*, 108:20410–20415, 2011.

[40] A. Hassanali, F. Giberti, J. Cuny, T. D. Kühne, and M. Parrinello. Proton transfer through the water gossamer. *Proceedings of the National Academy of Sciences USA*, page doi: 10.1073/pnas.1306642110, 2013.

[41] H. Larzul, F. Gelebart, and A. Joahnin-Gilles. *Comptes Rendus*, 261:4701, 1965.

[42] B. H. Chai, J. M. Zheng, Q. Zhao, and G. H. Pollack. Spectroscopic Studies of Solutes in Aqueous Solution. *Journal of Physical Chemistry A*, 112:2242–2247, 2008.

[43] L. M. Hunter, D. Lewis, and W. H. Hamill. Low-energy electron reflection spectrometry for thin films of n-hexane, benzene, and ice at 77Â°K. *Journal of Chemical Physics*, 52:1733–1739, 1970.

[44] J. Carrasco, B. Santra, J. Klimes, and A. Michaelides. To Wet or Not to Wet? Dispersion Forces Tip the Balance for Water Ice on Metals. *Physical Review Letters*, 106:026101, 2011.

[45] J. Carrasco, A. Hodgson, and A. Michaelides. A molecular perspective of water at metal interfaces. *Nature Materials*, 11:667–674, 2012.

[46] C. C. Pradzynski, R. M. Forck, T. Zeuch, P. Slavícek, and U. Buck. A Fully Size-Resolved Perspective on the Crystallization of Water Clusters. *Science*, 337:1529–1532, 2012.

[47] J. D. Watson and F. H. C. Crick. A Structure for Deoxyribose Nucleic Acid. *Nature*, 171:737–738, 1953.

[48] M. Barbatti, A. J. A. Aquino, J. J. Szymczak, D. Nachtigallova, P. Hobza, and H. Lischka. On the relaxation mechanisms of UV-photoexcited DNA and RNA nucleobases. *Proceedings of the National Academy of Sciences USA*, 107:21453–21458, 2010.

[49] J. P. Gobbo, V. Saurí, D. Roca-Sanjuán, L. Serrano-Andrés, M. Merchán, and A. C. Borin. On the Deactivation Mechanisms of Adenine-Thymine Base Pair. *Journal of Physical Chemistry B*, 116:4089–4097, 2012.

[50] V. Saurí, J. P. Gobbo, J. J. Serrano-Pérez, M. Lundberg, P. B. Coto, L. Serrano-Andrés, A. C. Borin, R. Lindh, M. Merchán, and D. Roca-Sanjuán. Proton/Hydrogen Transfer Mechanisms in the Guanine-Cytosine Base Pair: Photostability and Tautomerism. *Journal of Chemical Theory and Computation*, 9:481–496, 2013.

[51] G. Olaso-González, D. Roca-Sanjuán, L. Serrano-Andrés, and M. Merchán. Toward the understanding of DNA fluorescence: The singlet excimer of cytosine. *Journal of Chemical Physics*, 125:231102, 2006.

[52] G. Olaso-González, M. Merchán, and L. Serrano-Andrés. The Role of Adenine Excimers in the Photophysics of Oligonucleotides. *Journal of the American Chemical Society*, 131:4368–4377, 2009.

[53] A. Giussani, L. Serrano-Andrés, M. Merchán, D. Roca-Sanjuán, and M. Garavelli. Photoinduced Formation Mechanism of the Thymine-Thymine (6-4) Adduct. *Journal of Physical Chemistry B*, 117:1999, 2013.

[54] I. González-Ramírez, D. Roca-Sanjuán, T. Climent, J. J. Serrano-Pérez, M. Merchán, and L. Serrano-Andrés. On the Photoproduction of DNA/RNA Cyclobutane Pyrimidine Dimers. *Theoretical Chemistry Accounts*, 128:705, 2011.

[55] D. Roca-Sanjuán, M. Merchán, and L. Serrano-Andrés. Modeling Hole Transfer in DNA: Low-lying Excited States of Oxidized Cytosine Homodimer and Cytosine-Adenine Heterodimer. *Chemical Physics*, 349:188–196, 2008.

[56] D. Roca-Sanjuán, M. Rubio, M. Merchán, and L. Serrano-Andrés. Ab initio determination of the ionization potentials of DNA and RNA nucleobases. *Journal of Chemical Physics*, 125:084302, 2006.

[57] D. Roca-Sanjuán, M. Merchán, L. Serrano-Andrés, and M. Rubio. Ab initio determination of the electronic affinities of DNA and RNA nucleobases. *Journal of Chemical Physics*, 129:095104, 2008.

[58] M. Rubio, D. Roca-Sanjuán, M. Merchán, and L. Serrano-Andrés. Determination of the Lowest-Energy Oxidation Site in Nucleotides: 2'-Deoxythymidine 5'-Monophosphate Anion. *Journal of Physical Chemistry B*, 110:10234–10235, 2006.

[59] M. Rubio, D. Roca-Sanjuán, L. Serrano-Andrés, and M. Merchán. Determination of the Electron-Detachment Energies of 2'-Deoxyguanosine 5'-Monophosphate Anion: Influence of the Conformation. *Journal of Physical Chemistry B*, 113:2451–2457, 2009.

[60] J. Simons. How Do Low-Energy (0.1-2 eV) Electrons Cause DNA-Strand Breaks? *Accounts of Chemical Research*, 39:772–779, 2006.

[61] A. Szabo and N. S. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory (Dover Books on Chemistry)*. Dover Publications, new edition, July 1996.

[62] T. J. Lee and P. R. Taylor. A diagnostic for determining the quality of single-reference electron correlation methods. *International Journal of Quantum Chemistry*, 23:199, 1989.

[63] B. O. Roos, Peter R. Taylor, and Per E. M. Siegbahn. A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach. *Chemical Physics*, 48(2):157–173, May 1980.

[64] K. Andersson, P. A. Malmqvist, B. O. Roos, A. J. Sadlej, and K. Wolinski. 2nd-Order Perturbation-Theory with a CASSCF Reference Function. *Journal of Physical Chemistry*, 94(14):5483–5488, JUL 12 1990.

[65] K. Andersson, P. A. Malmqvist, and B. O. Roos. Second-order perturbation theory with a complete active space self-consistent field reference function. *Journal of Chemical Physics*, 96:1218–1226, 1992.

[66] B. O. Roos, K. Andersson, M. P. Fülscher, P. A. Malmqvist, L. Serrano-Andrés, K. Pierloot, and M. Merchán. Multiconfigurational perturbation theory: Applications in electronic spectroscopy. In *Advances in Chemical Physics, Vol XCIII*, volume 93 of *Advances in Chemical Physics*, pages 219–331. John Wiley & Sons INC, 605 3RD AVE, NEW YORK, NY 10016, 1996.

[67] B. O. Roos. A new method for large-scale CI calculations. *Chemical Physics Letters*, 15(2):153–159, 1972.

[68] B. O. Roos. The Complete Active Space Self-Consistent Field Method and its Applications in Electronic Structure Calculations. In KP Lawley, editor, *Quantum Chemistry*, volume 67 of *Advances in Chemical Physics*, pages 399–445. 1987.

[69] F. Aquilante, L. De Vico, N. Ferré, G. Ghigo, P. A. Malmqvist, P. Neogrady, T. B. Pedersen, M. Pitonak, M. Reiher, B. O. Roos, L. Serrano-Andrés, M. Urban, V. Veryazov, and R. Lindh. MOLCAS 7: The Next Generation. *Journal of Computational Chemistry*, 31(1):224–247, 2010.

[70] J. Stalring, A. Bernhardsson, and R. Lindh. Analytical gradients of a state average MCSCF state and a state average diagnostic. *Molecular Physics*, 99(2):103–114, 2001.

[71] C. Angeli, R. Cimiraglia, S. Evangelisti, T. Leininger, and J.-P. Malrieu. Introduction of n-electron valence states for multireference perturbation theory. *The Journal of Chemical Physics*, 114(23):10252–10264, 2001.

[72] G. Ghigo, B. O. Roos, and P. A. Malmqvist. A modified definition of the zeroth-order Hamiltonian in multiconfigurational perturbation theory (CASPT2). *Chemical Physics Letters*, 396:142–149, 2004.

[73] B. O. Roos and K. Andersson. Multiconfigurational perturbation theory with level shift - the Cr2 potential revisited. *Chemical Physics Letters*, 245:215–223, 1995.

[74] N. Forsberg and P. A. Malmqvist. Multiconfiguration perturbation theory with imaginary level shift. *Chemical Physics Letters*, 274:196–204, 1997.

[75] J. Finley, P. A. Malmqvist, B. O. Roos, and L. Serrano-Andrés. The multi-state CASPT2 method. *Chemical Physics Letters*, 288:299–306, 1998.

[76] L. Serrano-Andrés, M. Merchán, and R. Lindh. Computation of conical intersections by using perturbation techniques. *The Journal of Chemical Physics*, 122:104107, 2005.

[77] W. J. Hehre, R. F. Stewart, and J. A. Pople. Self-consistent molecular-orbital methods. I. Use of gaussian expansions of slater-type atomic orbitals. *The Journal of Chemical Physics*, 51:2657–2665, 1969.

[78] W. J. Hehre, R Ditchfield, R. F. Stewart, and J. A. Pople. Self-consistent molecular orbital methods. IV. Use of gaussian expansions of slater-type orbitals. Extension to second-row molecules. *The Journal of Chemical Physics*, 52:2769–2774, 1970.

[79] D. H. Dunning. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *The Journal of Chemical Physics*, 90:1007–1023, 1989.

[80] R. A. Kendall, D. H. Dunning, and R. J. Harrison. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *The Journal of Chemical Physics*, 96:6796–6806, 1992.

[81] P. O. Widmark, P. A. Malmqvist, and B. O. Roos. Density matrix averaged atomic natural orbital (ANO) basis sets for correlated molecular wave functions. i. First row atoms. *Theoretica Chimica Acta*, 77:291, 1990.

[82] P. O. Widmark, B. J. Persson, and B. O. Roos. Density matrix averaged atomic natural orbital (ANO) basis sets for correlated molecular wave functions. ii. Second row atoms. *Theoretica Chimica Acta*, 79:419, 1991.

[83] R. Pou-Amérigo, M. Merchán, I. Nebot-Gil, P. O. Widmark, and B. O. Roos. Density matrix averaged atomic natural orbital (ANO) basis sets for correlated molecular wave functions III. First row transition metal atoms. *Theoretica Chimica Acta*, 92:149–181, 1995.

[84] K. Pierloot, B. Dumez, P. O. Widmark, and B. O. Roos. Density matrix averaged atomic narural orbital (ANO) basis sets for correlated molecular wave functions. IV. Medium size basis sets for the atoms H-Kr. *Theoretica Chimica Acta*, 90:87, 1995.

[85] F. Aquilante, L. Boman, J. Bostrom, H. Koch, R. Lindh, A. Sanchez-De Merás, and T. B. Pedersen. *Cholesky decomposition techniques in electronic structure*

theory, chapter 13, pages 301–343. Linear-Scaling Techniques in Computational Chemistry and Physics. Springer Science+Business Media, 2011.

[86] J. Bostrom, F. Aquilante, T. B. Pedersen, and R. Lindh. Ab Initio Density Fitting: Accuracy Assessment of Auxiliary Basis Sets from Cholesky Decompositions. *Journal of Chemical Theory and Computation*, 5:1545–1553, 2009.

[87] S. F. Boys and F. Bernardi. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Molecular Physics*, 19(4):553–566, 1970.

[88] E. B. Wilson, J. C. Decius, and P. C. Cross. *Molecular Vibrations: the Theory of infrared and Raman Vibrational Spectra*. Dover Publications, 1980.

[89] R. Fletcher. *Practical Methods of Optimization*. Wiley, New York, 1987.

[90] D. J. Wales. *Energy Landscapes*. Cambridge University Press, 2003.

[91] H. B. Schlegel. Exploring potential energy surfaces for chemical reactions: an overview of some practical methods. *Journal of Computational Chemistry*, 24:1514–1527, 2003.

[92] K. Fukui. The path of chemical reactions-the IRC approach. *Accounts of Chemical Research*, 14:363–368, 1981.

[93] M. J. Bearpark, M. A. Robb, and H. B. Schlegel. A Direct Method for the Location of the Lowest Energy Point on a Potential Surface Crossing. *Chemical Physics Letters*, 223:269–274, 1994.

[94] L. De Vico, M. Olivucci, and R. Lindh. New General Tools for Constrained Geometry Optimizations. *Journal of Chemical Theory and Computation*, 1:1029–1037, 2005.

[95] J. M. Anglada and J. M. Bofill. A Reduced-Restricted-Quasi-Newton-Raphson Method for Locating and Optimizing Energy Crossing Points Between Two Potential Energy Surfaces. *Journal of Computational Chemistry*, 18:992–1003, 1997.

[96] K. Müller and L. D. Brown. Localization of saddle points and minimum energy paths by a constrained simplex optimization procedure. *Theoretical Chemistry Accounts*, 53:75–93, 1979.

[97] W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *The Journal of Chemical Physics*, 76:637, 1982.

[98] E. Wigner. On the Quantum Correction For Thermodynamic Equilibrium. *Physical Review*, 40:749, 1932.

[99] M. Barbatti, G. Granucci, M. Persico, M. Ruckenbauer, M. Vazdar, M. Eckert-Maksic, and H. Lischka. The on-the-fly surface-hopping program system Newton-X: Application to ab initio simulation of the nonadiabatic photodynamics of benchmark systems. *Journal of Photochemistry and Photobiology A: Chemistry*, 190(2-3):228–240, 2007.

[100] I. Schapiro, M. N. Ryazantsev, L. M. Frutos, N. Ferré, R. Lindh, and M. Olivucci. The Ultrafast Photoisomerizations of Rhodopsin and Bathorhodopsin Are Modulated by Bond Length Alternation and HOOP Driven Electronic Effects. *Journal of the American Chemical Society*, 133:3354–3364, 2011.

[101] J. L. McHale. *Molecular Spectroscopy.* Prentice Hall, 1999.

[102] J. A. Odutola and T. R. Dyke. Partially deuterated water dimers: Microwave spectra and structure. *The Journal of Chemical Physics*, 72:5062–5070, 1980.

[103] V. Poterya, M. Fárnik, M. Oncak, and P. Slavícek. Water photodissociation in free ice nanoparticles at 243 nm and 193 nm. *Physical Chemistry Chemical Physics*, 10:4835–4842, 2008.

[104] N. Goldman, R. S. Fellers, C. Leforestier, and R. J. Saykally. Water Dimers in the Atmosphere: Equilibrium Constant for Water Dimerization from the VRT(ASP-W) Potential Surface. *Journal of Physical Chemistry A*, 105:515–519, 2001.

[105] N. Goldman, C. Leforestier, and R. J. Saykally. Water Dimers in the Atmosphere II: Results from the VRT(ASP-W)III Potential Surface. *Journal of Physical Chemistry A*, 108:787–794, 2004.

[106] Y. Scribano, N. Goldman, R. J. Saykally, and C. Leforestier. Water Dimers in the Atmosphere III: Equilibrium Constant from a Flexible Potential. *Journal of Physical Chemistry A*, 110:5411–5419, 2006.

[107] M. Y. Tretyakov, E. A. Serov, M. A. Koshelev, V. V. Parshin, and A. F. Krupnov. Water Dimer Rotationally resolved Millimeter-Wave Spectrum Observation at Room Temperature. *Physical Review Letters*, 110:093001, 2013.

[108] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox. Gaussian 09 Revision C.01.

[109] J. Segarra-Martí, D. Roca-Sanjuán, M. Merchán, and R. Lindh. On the photophysics and photochemistry of the water dimer. *Journal of Chemical Physics*, 137:244309, 2012.

[110] P. Cabral do Couto and D. M. Chipman. Insights into the ultraviolet spectrum of liquid water from model calculations. *Journal of Chemical Physics*, 132:244307, 2010.

[111] A. L. Sobolewski and W. Domcke. Hydrated hydronium: a cluster model of the solvated electron? *Physical Chemistry Chemical Physics*, 4:4–10, 2002.

[112] D. M. Chipman. Stretching of hydrogen-bonded OH in the lowest singlet excited electronic state of water dimer. *Journal of Chemical Physics*, 122:044111, 2005.

[113] G. Avila, G. J. Kroes, and M. C. van Hemert. The photodissociation of the water dimer in the A band: A twelve-dimensional quasiclassical study. *Journal of Chemical Physics*, 128:144313, 2008.

[114] B. Chmura, Z. Lan, M. F. Rode, and A. L. Sobolewski. Photochemistry of the water dimer: Time-dependent quantum wave-packet description of the dynamics at the S1-S0 conical intersection. *Journal of Chemical Physics*, 131:134307, 2009.

[115] L. Valenzano, M. C. van Hemert, and G. J. Kroes. Photodissociation of the water dimer: Three-dimensional quantum dynamics studies on diabatic potential-energy surfaces. *Journal of Chemical Physics*, 123:034303, 2005.

[116] J. N. Harvey, O. J. Jung, and R. B. Gerber. Ultraviolet spectroscopy of water clusters: Excited electronic states and absorption line shapes of (H2O)n, n=2-6. *Journal of Chemical Physics*, 109:8747, 1998.

[117] K. Dressler and O. Schnepp. Absorption spectra of solid methane, ammonia, and ice in the Vacuum Ultraviolet. *Journal of Chemical Physics*, 33:270, 1960.

[118] G. Herzberg. *Molecular Spectra and Molecular Structure III.* van Nostrand, New York, 1966.

[119] R. Schinke. *Photodissociation Dynamics.* Cambridge University Press, Cambridge, 1993.

[120] K. Yuan, R. N. Dixon, and X. Yang. Photochemistry of the Water Molecule: Adiabatic versus Nonadiabatic Dynamics. *Accounts of Chemical Research*, 44:369–378, 2011.

[121] I. G. Gurtubay and R. J. Needs. Dissociation energy of the water dimer from quantum monte carlo calculations. *Journal of Chemical Physics*, 127:124306, 2007.

[122] A. Shank, Y. Wang, A. Kaledin, B. J. Braams, and J. M. Bowman. Accurate ab initio and hybrid potential energy surfaces, intramolecular vibrational energies, and classical IR spectrum of the water dimer. *Journal of Chemical Physics*, 130:144314, 2009.

[123] B. E. Rocher-Casterline, L. C. Chng, A. K. Mollner, and H. Reisler. Determination of the bond dissociation energy (D0) of the water dimer, (H2O)2, by velocity map imaging. *Journal of Chemical Physics*, 134:211101, 2011.

[124] C. G. Elles, I. A. Shkrob, R. A. Crowell, and S. E. Bradforth. Excited state dynamics of liquid water: Insight from the dissociation reaction following two-photon excitation. *The Journal of Chemical Physics*, 126:164503, 2007.

[125] J. Segarra-Martí, M. Merchán, and D. Roca-Sanjuán. Ab initio determination of the ionization potentials of water clusters (h2o)n (n=2-6). *Journal of Chemical Physics*, 136:244306, 2012.

[126] A. L. Sobolewski and W. Domcke. Photochemistry of water: The (H2O)5 cluster. *Journal of Chemical Physics*, 122:184320, 2005.

[127] J. Eisinger, M. Guéron, R. G. Shulman, and T. Yamane. Excimer fluorescence of dinucleotides, polynucleotides, and DNA. *Proceedings of the National Academy of Sciences USA*, 55:1015–1020, 1966.

[128] C. E. Crespo-Hernández, B. Cohen, P. M. Hare, and B. Kohler. Ultrafast Excited-State Dynamics in Nucleic Acids. *Chemical Reviews*, 104:1977–2020, 2004.

[129] J. Segarra-Martí, P. B. Coto, M. Rubio, D. Roca-Sanjuán, and M. Merchán. Towards the Understanding at the Molecular Level of the Structured-Water Absorption and Fluorescence Spectra: A Fingerprint of pi-Stacked Water. *Molecular Physics*, 111:1308–1315, 2013.

[130] C. D. Cappa, J. D. Smith, B. M. Messer, R. C. Cohen, and R. J. Saykally. Effects of Cations on the Hydrogen Bond Network of Liquid Water: New Results from X-ray Absorption Spectroscopy of Liquid Microjets. *Journal of Physical Chemistry B*, 110:5301–5309, 2006.

[131] Y. Marcus. Effect of ions on the Structure of Water: Structure Making and Breaking. *Chemical Reviews*, 109:1346–1370, 2009.

[132] X. Chen, T. Yang, S. Kataoka, and P. S. Cremer. Specific Ion Effects on Interfacial Water Structure near Macromolecules. *Journal of the American Chemical Society*, 129:12272–12279, 2007.

[133] A. T. Thomas and A. H. Elcock. Molecular Dynamics Simulations of Hydrophobic Associations in Aqueous Salt Solutions Indicate a Connection between Water Hydrogen Bonding and the Hofmeister Effect. *Journal of the American Chemical Society*, 129:14887–14898, 2007.

[134] K. J. Tielrooij, N. Garcia-Araez, M. Bonn, and H. J. Bakker. Cooperativity in Ion Hydration. *Science*, 328:1006–1009, 2010.

[135] C. Caleman, J. S. Hub, P. J. van Maaren, and D. van der Spoel. Atomistic simulation of ion solvation in water explains surface preference of halides. *Proceedings of the National Academy of Sciences USA*, 108:6838–6842, 2011.

[136] F. D. Brunet and H. A. Joly. Electron Paramagnetic Resonance Spectroscopic Evidence for the Interaction of HAlOH with Water Molecules. *Journal of Physical Chemistry A*, 116:4267–4273, 2012.

[137] J. Kim, G. Kim, and P. S. Cremer. Investigations of Polyelectrolyte Adsorption at the Solid/Liquid Interface by Sum Frequency Spectroscopy: Evidence for Long-Range Macromolecular Alignment at Highly Charged Quartz/Water Interfaces. *Journal of the American Chemical Society*, 124:8751–8756, 2002.

[138] Z. D. Schultz, S. K. Shaw, and A. A. Gewirth. Potential Dependent Organization of Water at the Electrified Metal-Liquid Interface. *Journal of the American Chemical Society*, 127:15916–15922, 2005.

[139] J. Gumbart, F. Khalili-Araghi, M. Sotomayor, and B. Roux. Constant electric field simulations of the membrane potential illustrated with simple systems. *Biochimica Biophysica Acta*, 1818:294–302, 2012.

[140] R. Harada, Y. Sugita, and M. Feig. Protein Crowding Affects Hydration Structure and Dynamics. *Journal of the American Chemical Society*, 134:4842–4849, 2012.

[141] D. M. Chipman. Hemibonding between Hydroxyl Radical and Water. *Journal of Physical Chemistry A*, 115:1161–1171, 2011.

[142] P. Vassilev, M. J. Louwerse, and E. J. Baerends. Ab initio molecular dynamics simulation of the OH radical in liquid water. *Chemical Physics Letters*, 398:212–216, 2004.

[143] H. Yoo, R. Paranji, and G. H. Pollack. Impact of Hydrophilic Surfaces on Interfacial Water Dynamics Probed with NMR Spectroscopy. *Journal of Physical Chemistry Letters*, 2:532–536, 2011.

[144] J. M. Zheng, W. C. Chin, E. Khijniak, and G. H. Pollack. Surfaces and interfacial water: Evidence that hydrophilic surfaces have long-range impact. *Advances in Colloid Interface Science*, 127:19–27, 2006.

[145] C. S. Chen, W. J. Chung, I. C. Hsu, C. M. Wu, and W. C. Chin. Force field measurements within the exclusion-zone of water. *Journal of Biological Physics*, 38:113–120, 2012.

[146] V. Tychinsky. High Electric Susceptibility is the Signature of Structured Water in Water-Containing Objects. *WATER*, 3:95–99, 2011.

[147] N. F. Bunkin, P. S. Ignatiev, V. A. Kozlov, A. V. Shkirin, S. D. Zakharov, and A. A. Zinchenko. Study of the Phase States of Water Close to Nafion Interface. *WATER*, 4:129–154, 2013.

[148] M. Rohani and G. H. Pollack. Flow through Horizontal Tubes Submerged in Water in the Absence of a Pressure Gradient: Mechanistic Considerations. *Langmuir*, 29:6556–6561, 2013.

[149] G. A. Kimmel, J. Matthiesen, M. Baer, C. J. Mundy, N. G. Petrik, R. S. Smith, Z. Dohnálek, and B. D. Kay. No Confinement Needed: Observation of a Metastable Hydrophobic Wetting Two-Layer Ice on Graphene. *Journal of the American Chemical Society*, 131:12838–12844, 2009.

[150] H. Chai, H. Yoo, and G. H. Pollack. Effect of Radiant Energy on Near-Surface Water. *Journal of Physical Chemistry B*, 113:13953–13958, 2009.

[151] J. E. M. McGeoch and M. W. McGeoch. Entrapment of water by subunit c of ATP synthase. *Journal of the Royal Society Interface*, 5:311–318, 2008.

[152] A. Michaelides and K. Morgenstern. Ice nanoclusters at hydrophobic metal surfaces. *Nature Materials*, 6:597–601, 2007.

[153] E. R. Lippincott, R. R. Stromberg, W. H. Grant, and G. L. Cessac. Polywater. *Science*, 164:1482–1487, 1969.

[154] M. Mella and D. C. Clary. Zero temperature quantum properties of small protonated water clusters (H2O)nH+ (n=1-5). *Journal of Chemical Physics*, 119:10048, 2003.

[155] J. Kim and K. S. Kim. Structures, binding energies, and spectra of isoenergetic water hexamer clusters: Extensive ab initio studies. *Journal of Chemical Physics*, 109:5886, 1998.

[156] K. Liu, M. G. Brown, and R. J. Saykally. Terahertz Laser Vibration-Rotation Tunneling Spectroscopy and Dipole Moment of a Cage Form of the Water Hexamer. *Journal of Physical Chemistry A*, 101:8995, 1997.

[157] H. Tachikawa. Ionization Dynamics of the Small-Sized Water Clusters: A Direct Ab Initio Trajectory Study. *Journal of Physical Chemistry A*, 108:7853, 2004.

[158] Y. V. Novakovskaya. Adiabatic Ionization of Water Clusters: Nonempirical Dynamic Model. *Russian Journal of Physical Chemistry A*, 81:225, 2007.

[159] H. Tachikawa. Ionization dynamics of a water dimer: specific reaction selectivity. *Physical Chemistry Chemical Physics*, 13:11206, 2011.

[160] R. Barnett and U. Landman. Structure and Energetics of Ionized Water Clusters: (H2O)n+, n=2-5. *Journal of Physical Chemistry A*, 101:164, 1997.

[161] L. Belau, K. R. Wilson, S. R. Leone, and M. Ahmed. Vacuum Ultraviolet (VUV) Photoionization of Small Water Clusters. *Journal of Physical Chemistry A*, 99:10075, 2007.

[162] M. Svoboda, O.and Oncák and P. Slavícek. Simulations of light induced processes in water based on ab initio path integrals molecular dynamics. i. Photoabsorption. *Journal of Chemical Physics*, 135:154301, 2011.

[163] S. Barth, M. Oncak, V. Ullrich, M. Mucke, T. Lischke, and P. Slavícek. Valence Ionization of Water Clusters: From Isolated Molecules to Bulk. *Journal of Physical Chemistry A*, 113:13519, 2009.

[164] P. A. Pieniazek, J. Van de Vondele, P. Jungwirth, A. I. Krylov, and S. E. Bradforth. Electronic Structure of the Water Dimer Cation. *Journal of Physical Chemistry A*, 112:6159, 2008.

[165] I. B. Müller and L. S. Cederbaum. Ionization and double ionization of small water clusters. *Journal of Chemical Physics*, 125:204305, 2006.

[166] S. P. de Visser, L. J. de Koning, and N. M. M. Nibbering. Reactivity and Thermochemical Properties of the Water Dimer Radical Cation in the Gas Phase. *Journal of Physical Chemistry*, 99:15444, 1995.

[167] C. Ferreira, H. F. M. C. Martiniano, B. J. Costa Cabral, and V. Aquilanti. Electronic Excitation and Ionization of Hydrogen Peroxide-Water Clusters: Comparison With Water Clusters. *International Journal of Quantum Chemistry*, 111:1824, 2011.

[168] Y. V. Novakovskaya and N. F. Stepanov. Small Charged Water Clusters: Cations. *Journal of Physical Chemistry A*, 103:3285, 1999.

[169] A. Mozumder. Ionization and excitation yields in liquid water due to the primary irradiation: Relationship of radiolysis with far UV-photolysis. *Physical Chemistry Chemical Physics*, 4:1451, 2002.

[170] B. Winter, R. Weber, W. Widdra, M. Dittmar, M. Faubel, and I. V. Hertel. Full Valence Band Photoemission from Liquid Water Using EUV Synchrotron Radiation. *Journal of Physical Chemistry A*, 108:2625, 2004.

[171] I. Watanabe, J. B Flanagan, and P. Delahay. Vacuum ultraviolet photoelectron emission spectroscopy of water and aqueous solutions. *Journal of Chemical Physics*, 73:2057, 1980.

[172] H. Hayashi, N. Watanabe, Y. Udagawa, and C.-C Kao. The complete optical spectrum of liquid water measured by inelastic x-ray scattering. *Proceedings of the National Academy of Sciences USA*, 97:6264, 2000.

[173] J. Cadet and P. Vigny. *Bioorganic Photochemistry*, volume 1, chapter The Photochemistry of Nucleic Acids, pages 1–273. John Wiley and Sons, 1990.

[174] S. Ptasinska, S. Denifl, P. Scheier, E. Illenberger, and T. D. Märk. Bond- and Site-Selective Loss of H Atoms from Nucleobases by Very-Lov-Energy Electrons (¡ 3eV). *Angewandte Chemie International Edition*, 44:6941–6943, 2005.

[175] I. González-Ramírez, J. Segarra-Martí, L. Serrano-Andrés, M. Merchán, M. Rubio, and D. Roca-Sanjuán. On the N1H and N3H Bond Dissociation in Uracil by Low Energy Electrons: A CASSCF/CASPT2 Study. *Journal of Chemical Theory and Computation*, 8:2769–2776, 2012.

[176] M. Rubio, M. Merchán, E. Ortí, and B. O. Roos. Theoretical Study of the Electronic Spectra of the Biphenyl Cation and Anion. *Journal of Physical Chemistry*, 99:14980, 1995.

[177] R. Pou-Amérigo, L. Serrano-Andrés, M. Merchán, E. Ortí, and N. Forsberg. A Theoretical Determination of the Low-Lying Electronic States of the p-Benzosemiquinone Radical Anion. *Journal of the American Chemical Society*, 122:6067, 2000.

[178] K. Aflatooni, G. A. Gallup, and P. D. Burrow. Electron Attachment Energies of the DNA Bases. *Journal of Physical Chemistry A*, 102:6205, 1998.

[179] G. Hanel, B. Gstir, S. Denifl, P. Scheier, M. Probst, B. Farizon, M. Farizon, E. Illenberger, and T. D. Märk. Electron Attachment to Uracil: Effective Destruction at Subexcitation Energies. *Physical Review Letters*, 90:188104, 2003.

[180] S. Denifl, S. Ptasinska, G. Hanel, B. Gstir, M. Probst, P. Scheier, and T. D. Märk. Electron attachment to gas-phase uracil. *Journal of Chemical Physics*, 120:6557, 2004.

[181] X. Li, L. Sanche, and M. D. Sevilla. Low Energy Electro Interactions with Uracil: The Energetics Predicted by Theory. *Journal of Physical Chemistry B*, 108:5472, 2004.

[182] A. Giussani, J. Segarra-Martí, D. Roca-Sanjuán, and M. Merchán. Excitation of nucleobases from a computational perspective I: reaction paths. *Topics in Current Chemistry*, page Submitted, 2013.

[183] S. Perun, A. L. Sobolewski, and W. Domcke. Conical Intersections in Thymine. *Journal of Physical Chemistry A*, 110:13238, 2006.

[184] H. R. Hudock, B. G. Levine, A. L. Thompson, H. Satzger, D. Townsend, N. Gador, S. Ullrich, A. Stolow, and T. J. Martínez. Ab Initio Molecular Dynamics and Time-Resolved Photoelectron Spectroscopy of Electronically Excited Uracil and Thymine. *Journal of Physical Chemistry A*, 111:8500, 2007.

[185] S. Yamazaki and T. Taketsugu. Nonradiative Deactivation Mechanisms of Uracil, Thymine, and 5-Fluorouracil: A Comparative ab initio Study. *Journal of Physical Chemistry A*, 116:491, 2012.

[186] J. J. Serrano-Pérez, R. González-Luque, M Merchán, and L. Serrano-Andrés. On the Intrinsic Population of the Lowest Triplet State of Thymine. *Journal of Physical Chemistry B*, 111:11880, 2007.

[187] L. Serrano-Andrés and M. Merchán. Are the Five Natural DNA/RNA Base Monomers a Good Choice from Natural Selection? A Photochemical Perspective. *Journal of Photochemistry and Photobiology C: Photochemistry Reviews*, 10:21–32, 2009.