Tesis Doctoral

UNIVERSITAT DE VALÈNCIA

INSTITUT DE CIÈNCIA MOLECULAR

Radiation damage in the DNA/RNA nucleobases:

A theoretical study

Tesis Doctoral Presentada por:
Israel Alberto González-Ramírez

València 2015

Programa de Doctorado:
121 B Química Teórica y Computacional II

Directores de Tesis:
Dra. Manuela Merchán, Dr. Daniel Roca-Sanjuán, and Dr. Luis Serrano-Andrés

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 Israel Alberto González Ramírez con el título **"Radiation damage in the DNA/RNA nucleobases: A theoretical study"** corresponde a su Tesis Doctoral y ha sido realizada bajo su dirección, junto con **D. Luis Serrano Andrés (1966-2010)**, Profesor del Departamento de Química Física y del Instituto de Ciencia Molecular de la *Universitat de València*, autorizando mediante este escrito la presentación de la misma.

En València, a 14 de Octubre de 2015

Dra. Manuela Merchán

Dr. Daniel Roca Sanjuán

Dr. Luis Serrano Andrés (1966-2010)

Doctorando: Israel Alberto González Ramírez

Contents

Acknowledgements

Every person you encounter along the way is a master, a great teacher, presenting you with the gift of wisdom from their own journey with one purpose only … to enrich yours.

Thanks to my Family, for inspiring me to reach higher and move further. Their love, teachings, sacrifices, advices, endurance, perseverance, and support will continue to motivate me in my ongoing path.

Thanks to Daniel Roca, Patricio González and Rubén Costa, for being not only good friends but brothers to me, both in science and in life within my time in Spain. You guys have been my pillars!!!

Thanks to Teresa Climent for all the words of encouragement to move forward.

Thanks to the Theoretical Chemistry Unit Professors Ignacio Nebot and José (Pepe) Sánchez for their help and consultation. Additionally, I want to express my gratitude to Professor Enrique Ortí for his words of fortitude and support.

Thanks to Mª Jesús Leiva for her help with the paperwork and the documentation processes always with a kind smile, promptitude and professionalism.

Thanks to the QCEXVAL group members Daniel Roca, Juan José Serrano, Gloria Olaso, Teresa Climent, Vicenta Sauri, Pedro Braña, Vicente Pérez, Angelo Giussani, Javier Segarra, Toni Francés, Mercedes Rubio, Remedios González for the help, their assistance, and for maintaining the curiosity of science alive in our conversations.

Thanks to Manuela Merchán for her inspiring scientific vision, her aid, and advices. Having her as my advisor has being an instrumental aspect of growing up professionally and personally. Her thoughts and

words of wisdom have being an influence and have made an impact on both the realms of science and life.

Thanks to Luis Serrano-Andrés for his great scientific mentorship in photochemistry and in the utilization of the Molcas software. His outstanding work, ethic, and passion for the field will continue to inspirit me. Thank you for the invigorating and stimulating conversations about science, math, world cultures, travels, and history. You were a truly remarkable scientist and an even more extraordinary person. You're deeply missed.

Thanks to the MAE AECI (July 25th, 2003 AECI Resolution, 206th BOE of August 28th, 2003) Grant, the MEC BES-2005-819 FPI Grant, the Projects CTQ2004-01739, CTQ2007-61260, CTQ2010-14892, CTQ2014-58624-P, and CSD2007-0010 Consolider Ingenio in Molecular Nanoscience of the Spanish MEC-FEDER and GV2015-057 of the Generalitat Valenciana for the necessary funding to make this work possible.

List of Papers

The work presented in this Thesis has been carried out in the QCEXVAL group of the Instituto de Ciencia Molecular (Universitat de València) in collaboration with Profs. Manuela Merchán, Luis Serrano-Andrés, Remedios González-Luque, Mercedes Rubio, Teresa Climent, Daniel Roca-Sanjuán, Juan José Serrano-Pérez, Gloria Olaso-González, Angelo Giussani, and Javier Segarra-Martí. The publications which compile the results obtained are listed below in the order of appearance in the Thesis. My contribution in the papers is also briefly described.

Publications included in the Thesis

Paper I. Molecular basis of DNA photodimerization: Intrinsic production of cyclobutane cytosine dimers. D. Roca-Sanjuán, G. Olaso-González, I. González-Ramírez, L. Serrano-Andrés, and M. Merchán, J. Am. Chem. Soc., 130, 10766-10779 (2008).

Paper II. Theoretical insight into the intrinsic ultrafast formation of cyclobutane pyrimidine dimers in UV-irradiated DNA: Thymine versus cytosine. J. J. Serrano-Pérez, I. González-Ramírez, P. B. Coto, M. Merchán, and L. Serrano-Andrés, J. Phys. Chem. B, 112, 14096-14098 (2008).

Paper III. The role of pyrimidine nucleobase excimers in DNA Photophysics and photoreactivity. I. González-Ramírez, T. Climent, J. J. Serrano-Pérez, R. González-Luque, M. Merchán, and L. Serrano-Andrés, Pure Appl. Chem., 81, 1695-1705 (2009).

Paper IV. Singlet-triplet states interaction regions in DNA/RNA nucleobase hypersurfaces. R. González-Luque, T. Climent, I. González-Ramírez, M. Merchán, and L. Serrano-Andrés, J. Chem. Theor. Comput., 6, 2103-2114 (2010).

Paper V. Cyclobutane pyrimidine photodimerization of DNA/RNA nucleobases in the triplet state. T. Climent, I. González-Ramírez, R. González-Luque, M. Merchán, and L. Serrano-Andrés, J. Phys. Chem. Lett., 1, 2072-2076 (2010).

Paper VI. On the photoproduction of DNA/RNA cyclobutane pyrimidine dimers. I. González-Ramírez, D. Roca-Sanjuán, T. Climent, J. J. Serrano-Pérez, M. Merchán, and L. Serrano-Andrés, Theor. Chem. Acc., 128, 705-711 (2011).

Paper VII. On the N1-H and N3-H bond dissociation in uracil by low energy electrons: A CASSCF/CASPT2 study. I. González-Ramírez, J. Segarra-Martí, L. Serrano-Andrés, M. Merchán, M. Rubio, and D. Roca-Sanjuán, J. Chem. Theor. Comput., 8, 2769-2776 (2012).

Additional publication not included in this Thesis

Paper VIII. Photophysics of 1-aminonaphthalene: A theoretical and time-resolved experimental study. R. Montero, A. Longarte, A. P. Conde, C. Redondo, F. Castaño, I. González-Ramírez, A. Giussani, L. Serrano-Andrés, and M. Merchán, J. Phys. Chem. A, 113, 13509-13518 (2009).

Comments on my contribution to the Papers of the Thesis

I have been responsible for the computations and writing of Papers III, VI, and VII. In the other papers (I, II, IV, and V), I have performed some of the calculations, mainly those related to the pyrimidines thymine and uracil, and also I have contributed in the discussion, analyses, and corrections of the manuscripts and results.

1 Introduction

"The beauty of the living thing
is not the atoms that go into it,
but the way those atoms
are put together."

Carl Sagan (1934-1996)
Astronomer and Astrophysicist

Nowadays, the terms deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) are not uncommon and have become household names. Common concepts in today's public knowledge are how the traits or characteristics of bacteria, viruses, plants, and animals are combined and passed onto the next generation, or their involvement on how species adapt through a long evolutionary processes to better fit their environment. This was not always the case. In the beginning of the 19th century in Europe, the concept of the "transformation" of species through long evolutionary processes was not as "accepted" as it is today, considering that it contradicted the strongly established Christian world view. The radical and controversial idea that one species could transmute into another was originally hinted by Erasmus Darwin, (Figure 1.1) a renowned English physician, and Charles Darwin's grandfather, in his book Zoonomia.[1,2] The notion was finally, fully developed and documented by all the biological evidence gathered by his grandson, Charles Darwin (Figure 1.1) throughout his sea voyages. Fully aware of the contradicting nature of his conclusions, and due to fears of being ostracized by both his peers and society, it took Charles Darwin almost twenty years to finally make the decision of revealing his ideas and compiled evidence. Jointly with Alfred Wallace, who independently arrived to the same conclusions as Darwin did, they presented their paper in 1858, on how the varieties of the species arise through the process of

natural selection. The publication was imprinted in the Journal of the Proceedings of the Linnean Society, Britain's leading Natural History body.[2]

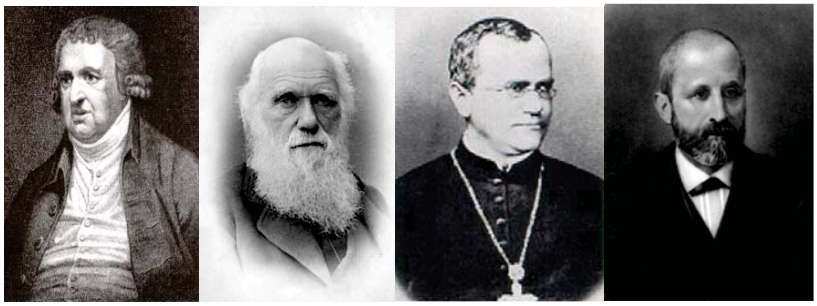


Figure 1.1 Portraits of (from left to right) Erasmus Darwin, Charles Darwin, Gregor Mendel, and Friedrich Miescher. Images from references 3, 4, 5, and 6, respectively.

Then, in 1865, seven years later, a Czech Augustinian Monk named Gregor Mendel (Figure 1.1) presented his paper, "Experiments on plant hybridization" to the Brunn Society for Natural Science. Mendel wanted to create hybrid pea plants and to observe the result. By observing the plants and keeping strict records, Mendel discovered and postulated the laws of heredity. He was the first to present the concepts of dominant and recessive traits. Furthermore, he established the use of statistical methods to predict hereditary information.[7]

Four years later, in 1869, a young Swiss physician by the name of Friedrich Miescher, (Figure 1.1) driven by the will of understanding the development of tissues, realized, that such questions could only be solved on the basis of chemistry. Knowingly, Miescher decided to pursue this line of research and moved to Tübingen, in Germany, after finishing his training as a physician. At the time, it was believed that proteins were the chemical structures responsible of heredity. Therefore, Miescher investigated the proteins contained in leucocyte cells. Nonetheless, in the midst of his research, he stumbled upon a substance with properties unrelated to those of proteins. The material obtained by Miescher, was the first DNA extraction ever performed. Miescher had collected the very first DNA sample. He ascribed the material to the cell's nucleus, thus naming it "nuclein". In 1874, he published the first DNA (Nuclein) description.[6,8] Miescher had discovered the ultimate material of life.

Unfortunately, his peers failed to pay attention to nuclein and the implications of his work since the vast majority of scientist were convinced that proteins were the carriers of genetic information. Yet, an important cornerstone was set through his findings.

Figure 1.2 Portraits of (from left to right) Phoebus Leven, Oswald Avery, Maurice Wilkins, and Rosalind Franklin. Images from references 9, 10, 11, and 11, respectively.

The experiments performed on nucleic acids in 1900-1919 by the Russian born chemist Phoebus Leven (Figure1.2) allowed establishing really crucial facts about these substances.[12] First, Leven characterized and remarked the existence of two main types of nucleic acids. His findings accounted for the distinction that some nucleic acids contained a ribose sugar (Figure 1.3), whereas others contained a different kind of sugar, namely deoxyribose (Figure 1.3). Thus, today the substances are named according to this chemical difference in their sugar composition. Ribonucleic acid is used for the one polymer containing the ribose sugar and deoxyribose acid is used for the polymer containing the oxygen lacking ribose.

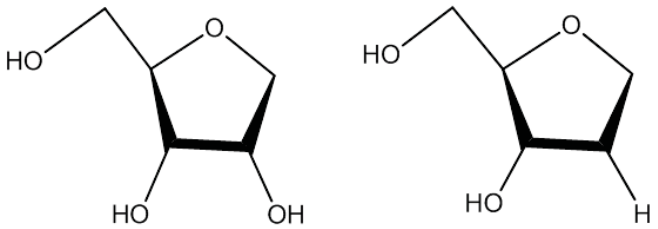


Figure 1.3 Structural representations of ribose and deoxyribose sugars respectively. Ribose sugar contains three hydroxyl (HO–) groups whereas deoxyribose contains only two.

Leven was also responsible for the identification and characterization of the building blocks contained in DNA[12] and RNA,[13] namely the nucleic bases, (Figure 1.4). He noticed that DNA contained the bases adenine, guanine, cytosine, and thymine, while RNA's composition contained the same nucleobases with the exception of thymine. Instead, RNA contained a uracil base. Levene was the first to propose that these nucleic acids were able to bind to form long chains. Sadly, he assumed that the amount of each of the four bases in DNA was the same, which led him to believe that DNA was a monotonous molecule incapable of carrying any useful information summing him up to the group of scientists who believed that the key to information transfer resided in proteins.

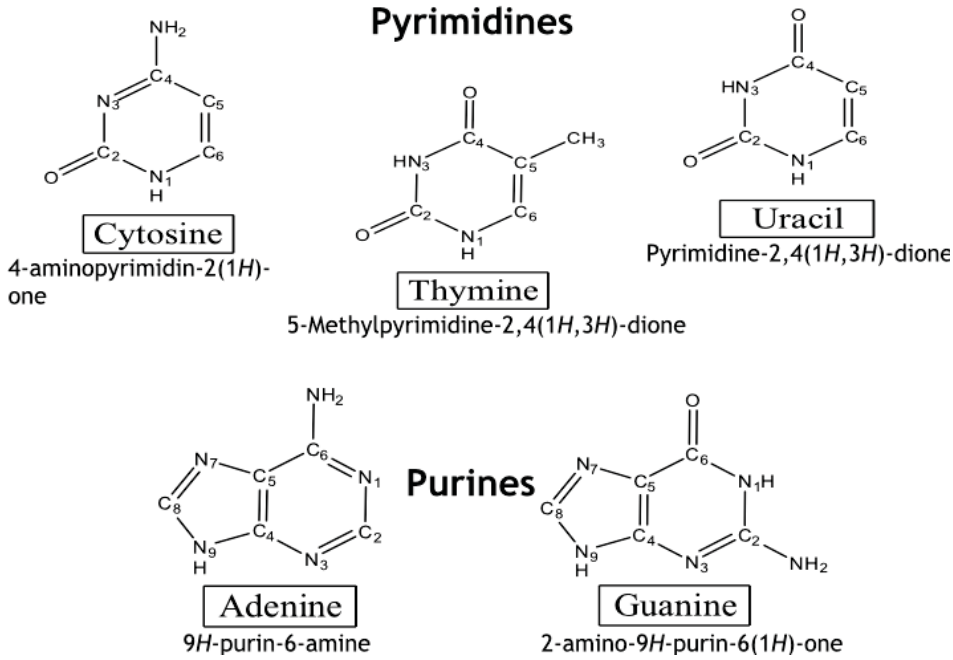


Figure 1.4 Representation of the nucleobases contained in DNA (cytosine, thymine, adenine, and guanine) and RNA (cytosine, uracil, adenine, and guanine) with their conventional name and, within parentheses, the nomenclature used by the International Union of Pure and Applied Chemistry (IUPAC).[14]

The true significance of the nucleic acids would remain misunderstood until a few decades later when in 1944 this paradigm would be changed once and for all by the work of Ostwald Avery (Figure 1.2), Colin McLeod, and Maclyn McCarty. They performed *in vitro* experiments on Pneumococcus bacteria. They introduced changes to the bacteria's nucleic acid material, transforming a Type II Pneumococcus into a Type III. Thus, demonstrating that DNA, and not proteins, is the real responsible of the hereditary information transmission in cells.[15] Their work triggered the ultimate race within the scientific community to unravel the mystery behind the structure of DNA. The structural array of the double helix of B-DNA was captured for the first time by the British crystallographer Rosalind Franklin (Figure 1.2) in 1952 while working at the Medical Research Council (MRC) at King's College. Maurice Wilkins (Figure 1.2), another crystallographer who was working at the MCR with DNA as well, showed Franlkin's X-Ray images to James Watson.[16] Watson and Crick started working on a plausible model based on the structural information obtained from the X-Ray image (Figure 1.5). This image was crucial for the development of the last piece of the puzzle, the deduction of DNA's structure, and the correct chemical connectivity that would allow the transfer of genetic material.

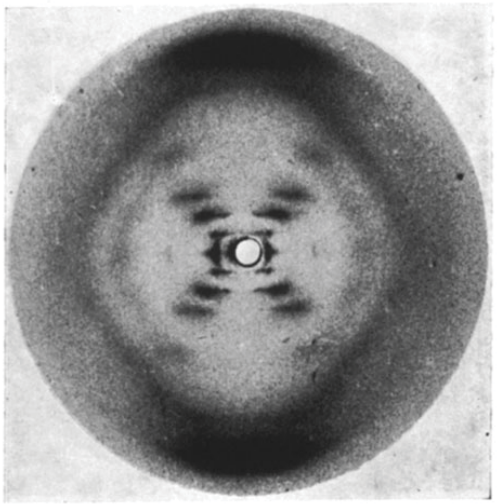


Figure 1.5 X-Ray Image of B-DNA obtained by Rosalind Franklin and published in Nature in 1953.[17]

After a struggling race of the scientific community to unravel the mystery behind the structure of the deoxyribonucleic acid (DNA), in 1953 James Watson and Francis Crick (Figure 1.6) astonished the world with their paper published in the journal Nature.[18] In that publication, they presented a model with the correct arrangement of the molecules that build up DNA. Their model was in complete agreement with the X-Ray images obtained and published in Nature by Rosalind Franklin and Maurice Wilkins in 1953.[17,19] Their assembly presented a double helix configuration (Figure 1.6) with a sugar-phosphate component on the outside of the double strand. That is, the deoxyribose sugars are bonded to the phosphate groups forming a long repetitive chain. This chain constitutes the outside skeleton of the DNA molecule. The inside of the chain is composed by the nucleic purine-pyrimidine base pairs which are connected together through hydrogen bonds (Figure 1.7).

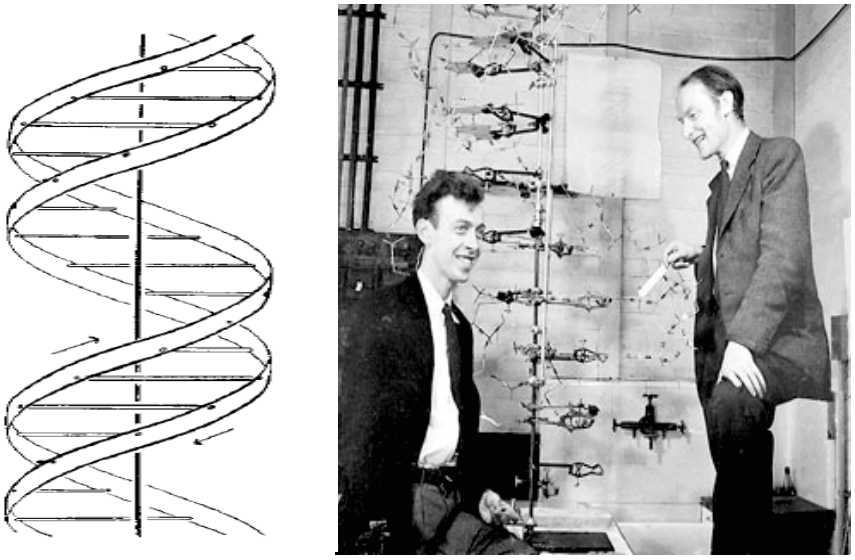


Figure 1.6 DNA model presented by James Watson and Francis Crick in Nature. From references 18 and 20, respectively.

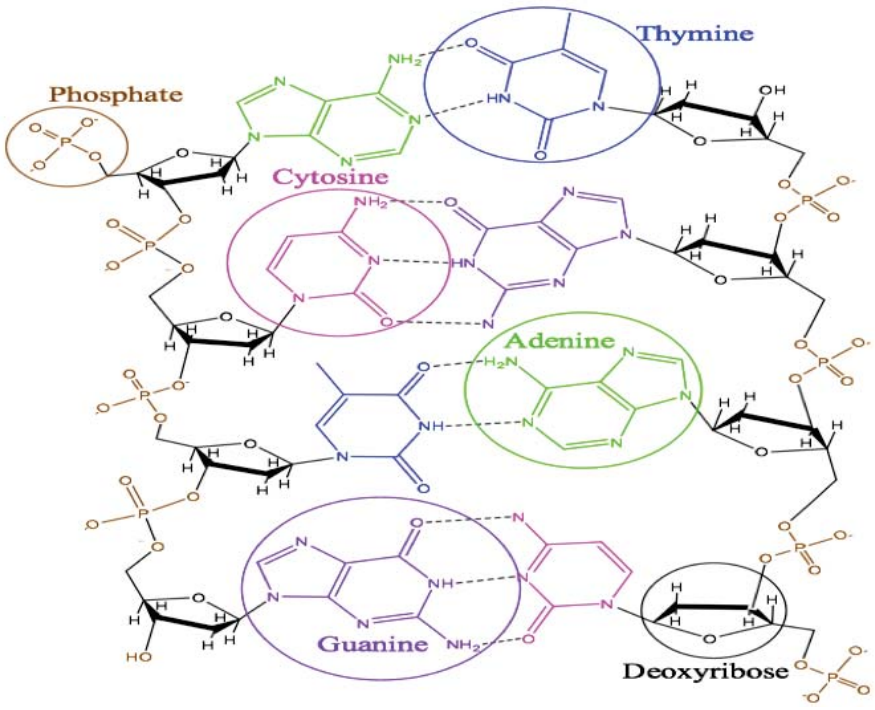


Figure 1.7 DNA chemical connectivity. The deoxyribose-phosphate components form the outside of the double strand whereas the inside of the chain is composed by the nucleic purine-pyrimidine base pairs connected together through hydrogen bonds.

The chemical structures of the base pairs concede them to be formed specifically by the connection of a certain purine with a precise pyrimidine through hydrogen bonds. The pyrimidines cytosine and thymine always combine distinctively with the purines guanine and adenine, respectively (Figure 1.7). It is this chemical connectivity, "the way that these atoms are put together", that allows for the structure to be a double strand helix and to have the specific structure that acquiesce the passing of the hereditary material from one generation to the next. On the other hand, RNA is responsible, among other things, for the transcription of information and the formation of proteins. It contains ribose sugars within its backbone (Figure 1.8) instead of the deoxyribose sugar found in DNA. Like DNA, RNA possesses the same purines (adenine and guanine) as well as the same pyrimidine (cytosine). In RNA, the pyrimidine thymine is replaced by a uracil.

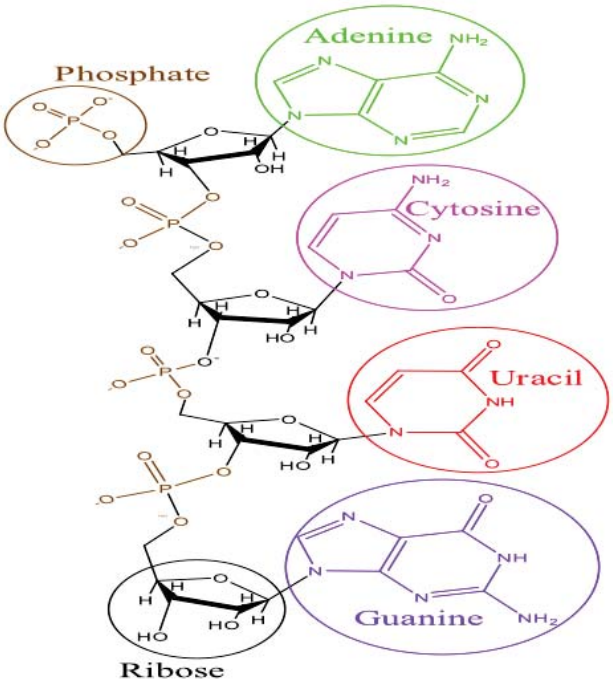


Figure 1.8 RNA chemical connectivity. RNA contains ribose sugars within its structure. RNA possesses the purines adenine and guanine as well as the pyrimidine cytosine. The thymine pyrimidine found in DNA is however replaced by uracil.

On their seminal paper published in the journal Nature in 1953, Watson and Crick suggested that in the case of RNA, it would be probably impossible to build a similar double helix model as they had for DNA due to the presence of an additional HO– group in the ribose sugar which would present a van der Waals repulsion.[18] This would be disproven in 1956 by the work of Alexander Rich and his co-worker David Davies (Figure 1.9) when they obtained the first RNA double helix by mixing a polyuridylic acid with a polyadennylic acid.[21] In this setting, it was eventually discovered that RNA adopted a conformation in the sugar where its C3 carbon formed an endo pucker, leaving a proper separation thus avoiding the HO– group repulsion. This characteristic in structure was similar to the one found in A-DNA.[22] Eventually other varieties of RNA as well as hybridized RNA/DNA were obtained in the following decades.[23,24] This new information on the chemical structure of DNA/RNA materialized the opportunity to study, in great detail, intrinsic

characteristics such as composition, energy, reactivity, as well as its interaction with light.

Figure 1.9 Alexander Rich and David Davies were responsible for obtaining the first double helix structure of RNA in 1956.[25,26]

1.1 DNA/RNA and light

The interaction of DNA/RNA with radiation is a very important subject of research given the fact that we are surrounded by light and exposed to it for the vast majority of the time. Some of the inquiries that come to mind are: What happens when nucleic acids are irradiated with light, particularly, with the ultraviolet (UV) or visible (Vis) range of energies? What happens to the bases when exposed to ionizing radiation energies? Unavoidably, understanding the innards of such inquiries became the starting points for this work.

It is well-known that the initial event during the interaction between the bases and the UV radiation is the population of excited electronic states. The UV energy is sufficient to populate the lower singlet excited states, preferably those with higher probability of absorption.[27] After the excitation, photophysical and photochemical processes will take place. In nucleobases, the measured lifetime related to the bright excited states is found to be in the femtosecond and picosecond scales, which prevents the system to react with other species due to the excess of energy.[28,29] In fact, the fluorescence quantum yields of the monomers of the nucleobases and its corresponding nucleotides and nuclesides are very low.[30] The experimental observations recorded at the end of the sixties indicate the existence of a type of mechanism that helps the bases to carry

out an ultrafast internal conversion (IC) of energy, which is dissipated to the environment in a non-radiative fashion, as the molecule returns to its ground state.[27,31] The use of novel spectroscopical techniques and accurate reaction path computational strategies to study these phenomena has latter confirmed the existence of non-radiative relaxation mechanisms in DNA/RNA isolated bases. Then, it is known now that the ultrafast IC process is due to the presence of energetically accessible regions that lie between the ground state and the excited states of lower energy. These regions are called conical intersections (CI) and are responsible for the highly efficient IC process observed.[28,29,32,33]

In nucleobase dimers and polymers, new photochemical properties arise, as compared to the isolated bases. Thus, while the absorption spectrum of DNA is similar to that of its constituent bases, its emission spectrum is qualitatively different. In the emission spectrum of the oligonucleotides with at least two stacked bases, there is a bathochromic (towards red) shift of fluorescence. This phenomenon was firstly described as excimer fluorescence by Eisinger and Shulman.[27] Furthermore, these excimers have been suggested by some authors to be precursors in the formation of photoproducts,[34,35] including bipyrimidines, which is one of the most common lesions in the genetic material.[28,29,34,36,37] The role of triplet electronic states (states with a longer lifetime), in the photochemistry of DNA and, in particular, the formation of cyclobutane pyrimidine dimers (CPDs) was also suggested by Cadet et al.[34] However, the detailed mechanisms of action were poorly understood.

Besides the effects of UV light in DNA/RNA, ionizing radiation can also produce damage to the double strand and/or its nucleic components. This can take place by either direct interaction with DNA/RNA or indirectly through the reaction of the nucleic acid with species formed in the vicinity of the genetic polymer in the ionization processes, the so-called reactive oxygen species (ROS). Damage caused by ROS and free radicals, in general, has been the subject of research from many biologists, physicists, and chemists during the last decades.[38] In this context, computational chemistry might provide useful insights on the mechanistic aspects of the oxidations and reductions actually occurring.

Summarizing, the interaction of electromagnetic radiation with living systems can trigger a whole new set of conditions within the cell from which we can find photochemical and ionization phenomena. New chemical species may be formed that can change the structure of the genetic material. The effects produced might be regarded as harmful if the evolutionary cell's mechanisms of coping with these events are rendered ineffective or unresponsive.

In the next sections, the main features of UV and ionizing radiation, as well as the damage caused in the DNA/RNA, shall be described in more detail.

1.2 DNA/RNA damage by UV light

In our planet, living organisms are protected and somewhat shielded against possible damage caused by the electromagnetic radiation coming from the Sun. This global safeguard is provided at the stratospheric level by the ozone layer. In recent years, this layer has being significantly lessen due to the increase of atmospheric contamination with pollutants such as chlorofluorocarbons, chlorocarbons, and organobromides,[39] thus, rendering our genetic material more vulnerable against this kind of radiation. [40,41]

The electromagnetic spectrum contains a variety of different energetic intensities ranging from the low energy of a Radio wave to the ionizing high energy of X-Rays and Gamma Rays (Figure 1.10). Within the electromagnetic spectrum, the UV region ranges from 100–400 nanometers (nm). The UV radiation spectrum is subdivided according to the wavelength in three segments designated as UV-A (320–400 nm), UV-B (295–320 nm), and UV-C (100–295 nm).[42,43] The absorption maximum of DNA is reached at 260 nm. Therefore, the most dangerous absorption of our genetic material would come from the UV-C region of the spectrum. Fortunately, the presence of the Earth's ozone layer is able to filter most of the shorter wavelengths radiation thus protecting us from exposure to the UV-C radiation range. On the other hand, UV-A and UV-B radiation do reach the Earth's surface. Although the damage to our genes is more serious by exposure to UV-C, many harmful changes can also occur by exposure to the lower frequencies of UV-A and UV-B.

Among large number of lesions caused to DNA by UV irradiation, the majority of cases falls into two specific kinds of reactions, namely dimerizations or adduct formation.[34,42,43,44,45] The reactive processes take place between adjacent pyrimidine bases mainly within the same chain of DNA/RNA.

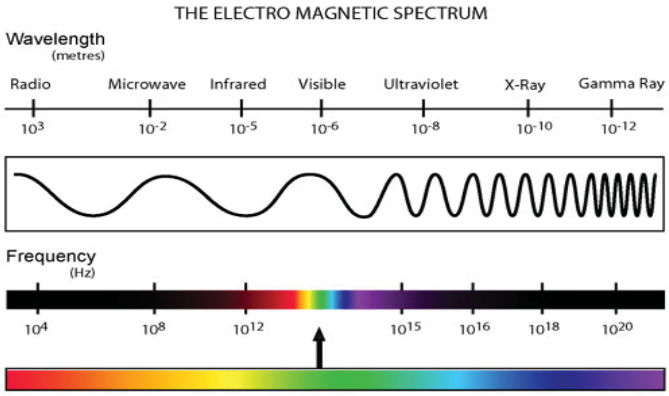


Figure 1.10 The electromagnetic spectrum of light. Wavelengths are given in meters and frequency in Hertz (Hz).[46]

In a dimerization, depending on their relative proximity, two consecutive bases obtain enough energy when they are UV irradiated to excite electrons from π bonding orbitals to π^* antibonding orbitals mainly localized in the C5=C6 and C5'=C6' double bonds (see labeling and numbering in Figure 1.4). The consequence is that the double bonds are broken to form C5–C5' and C6–C6' single bonds. This constitutes a [2+2] cycloaddition photoreaction (Figure 1.11) and the products are known as CPDs. Within the CPDs, two isomers are relevant. The first one is the *cis-syn*, where both bases pile up in a "sandwich" like type of arrangement and are orientated towards the same side of the newly formed carbon-carbon bonds (Figure 1.11). The second plausible orientation is the *trans-syn*. Here, the bases orientate towards opposite sides of the newly formed carbon-carbon bonds. In the double stranded DNA the *cis-syn* isomer is mainly formed.

The second type of lesion among nearby bases are adducts. The adduct is produced when the carbon at the position 6 of the base forms a bond with the carbon at the position 4 of the adjacent base. Due to their connectivity, these pyrimidines are designated as 6-4 [Pyrimidin-2'-one] and are often denoted as 6-4 photoproducts (Figure 1.11).

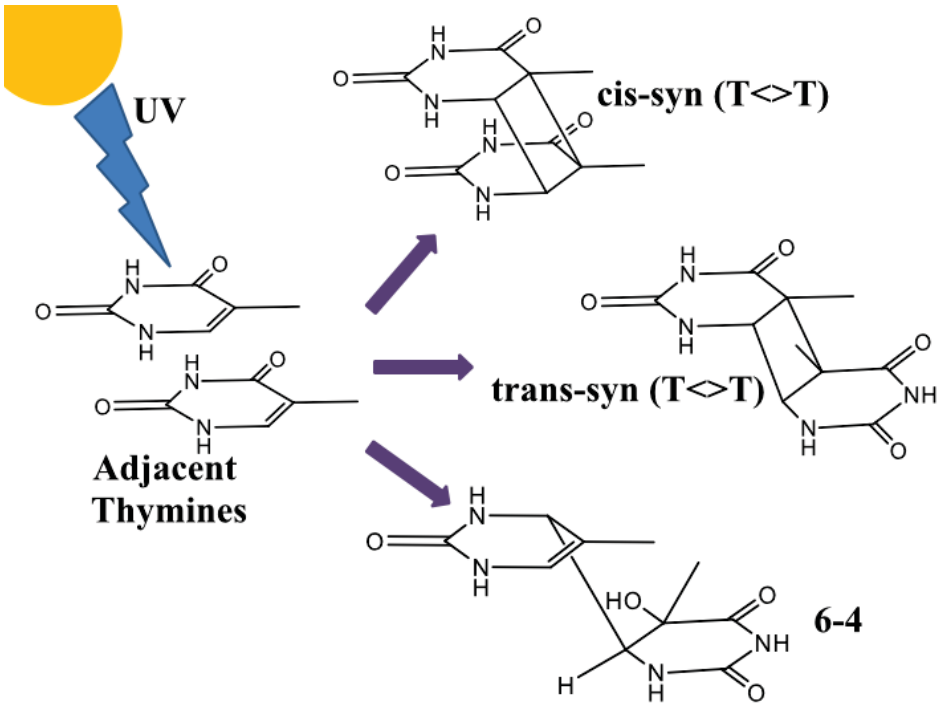


Figure 1.11 Structure of a CPD, the cyclobutane thymine (CBT), with its two diastereoisomers *cis-syn* and *trans-syn* adducts, and 6-4 [thymidine-2'-one].

CPD lesions occur more frequently than the photoadducts 6-4. Under UV-C radiation (254 nm) of a J/m^2 between 2 and 10 CPDs per million bases can be formed.[43] Two additional orders of magnitude of UV radiation of type B are required to produce the same result. The biological effects of the formation of these photoproducts might produce a halt in the nucleic acid replication or the introduction of a mutation that is present in the new strand when the replication process takes place.[47,48] Therefore, the study of the underlying mechanisms governing the CPDs photolesions is very important to better comprehend the DNA/RNA damage by UV light.

1.3 DNA/RNA damage by ionizing radiation

Interaction of the nucleic material with ionizing radiation can take place through the exposure to either natural or artificial sources. The main natural ionizing sources come from cosmic radiation and/or from

radioactive nuclei present on the planet. Conversely, artificial ionizing radiation sources come from man-made technology developed for applications in medical diagnostic and medical treatment (among others), such as X-rays and radiopharmaceuticals.

The interaction of ionizing radiation with the genetic material can cause damage to the latter either directly or indirectly. Direct ionizing radiation resulting from direct absorption of radiant energy by DNA may lead to the ionization of different substructures, such as nucleobases, sugars, nucleosides, nucleotides, oligonucleotides, etc.[42,49] The formation of excited and ionized species by this type of radiation mostly causes random damage to cellular components and induces a variety of lesions in the nucleic material (Figure 1.12). DNA samples irradiated at low temperatures indicate that the direct effects of ionizing radiation give rise mostly to localized cationic and anionic radicals on the nucleobases.[50,51,52]

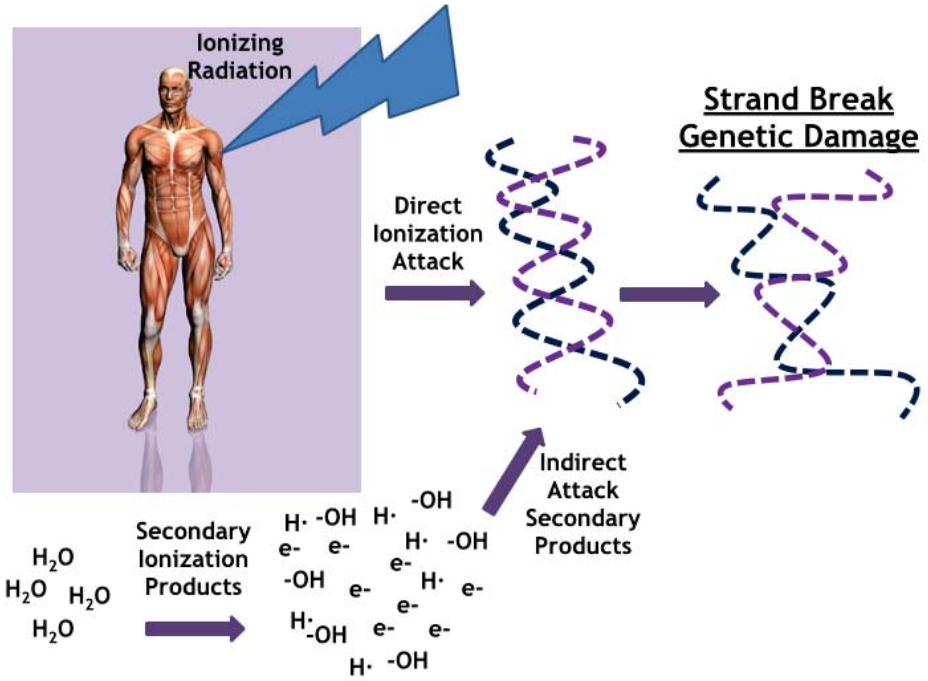


Figure 1.12 Schematic representation of the effects of ionizing radiation to the genetic material. Direct ionization attack may cause strand breaks and gene fragmentation. Ionization may also affect genes via the formation of radicals, free electrons, or ions in the surroundings of the nucleic acids, which in turn attack the genetic material causing strand breaks or gene fragmentation.

DNA may also react with species formed by the irradiation of molecules present in the vicinity of the nucleic acid, both of endogenous and exogenous origin.[42,49] In cells, the DNA is surrounded by an environment of many chemical species, such as water molecules, sugars, proteins, or products of aerobic metabolism (singlet oxygen, hydrogen peroxide, superoxide, etc.). Some of these molecules are themselves reactive and can damage DNA. In addition to that, ionizing radiation can become potential sources of reactive species. Because of the predominance of water in biological systems, the species produced by the radiolysis of water (e^-, $OH\cdot$, $H\cdot$, and H_3O^+) are the major sources of the indirect effect. These species interact with the DNA and facilitate a range of hydrogenation-dehydrogenation, hydroxylation-dehydroxylation, or fragmentation-dimerization processes. These reactions can trigger single strand breaks (SSBs) and double strand breaks (DSBs), which are some of the lesions of major biological impact.[53] Thus, while the SSB can be repaired using the other strand of the double helix as a template, the DSBs are very difficult to repair. Thus, a better understanding of the underlying mechanisms operating through these processes would provide the grounds for developing future treatments.

2 Quantum Chemistry: Methods

"It is no paradox that in our most theoretical moods we may be nearest to our most practical applications"

A.N. Whitehead (1861-1947)
Mathematician and Philosopher

The modeling of a molecular system using computers requires the use of mathematical algorithms designed to accurately describe the chemical nature of the sample under study and its possible interactions. The choice of a particular kind of methodology will depend on a number of factors, among them, we would highlight the size of the system (the number of atoms, electrons, etc.), the computational cost required to study it (the time it takes the implemented algorithm to produce results), and the desired accuracy. The goal of the research presented in this Thesis has been to model some of the intrinsic photoreactivity processes that take place in DNA and RNA after irradiating the system with UV light, ionizing it, or under conditions of reductive stress. To achieve this goal, we have designed the study of molecular systems with a maximum of two isolated nucleobases, using a high precision calculation level. Within the different kinds of quantum-chemical methods available to study such molecular interactions, *ab initio* (from the first principles) methods possess the accuracy required for a proper description of the phenomena. They arise from the application of quantum mechanics to chemical and molecular problems. Consequently, the theoretical methodology selected has relied mainly on the application of *ab initio* quantum mechanical (QM) methods. This chapter briefly reviews these methods, focusing mainly on the complete-active-space self-consistent field/complete-

active-space second-order perturbation theory (CASSCF/CASPT2) method, which is the one employed throughout.

2.1 The Schrödinger equation

Within the framework of quantum chemistry, a problem may be defined by means of the Schrödinger equation:

$$i\hbar\frac{d\psi}{dt} = \hat{\mathrm{H}}\psi \tag{2.1}$$

where $\hat{\mathrm{H}}$ represents the Hamiltonian, i.e., the sum of the kinetic and the potential energies, and ψ is the wave function, an object that contains all the relevant information of the given system formed by a certain number of nuclei (M) and electrons (N). Solving the Schrödinger equation gives the wave function as a function of the spatial coordinates and time $\psi(r,t)$. The square of such wave function is the probability of observing a particle at position *r* and time *t*.[54,55,56,57,58,59]

Many chemically interesting problems can be studied from the perspective of stationary states. In this case, the Schrödinger equation becomes:

$$\hat{\mathrm{H}}\Phi = E\Phi \tag{2.2}$$

where *E* represents the energy of the system or the resulting *eigenvalue*.

Usually, quantum chemical applications in spectroscopy and photochemistry employ a non-relativistic Hamiltonian and take into account the Born-Oppenheimer approximation. It is based on the fact that electrons and bare nuclei move at very different velocities because the latter are much heavier than the former. In this manner, the electrons can be described in the field of static or fixed possitive charges (nuclei). The resulting electronic Hamiltonian depends only on the position and spin of the electrons. Meanwhile, the relative location of the nuclei are considered as a parameter. Thus, the non-relativistic electronic Hamiltonian operator, $\hat{\mathrm{H}}_{elec}$, is defined through the following expression:

$$\hat{\mathrm{H}}_{elec} = -\sum_{i=1}^{N}\frac{1}{2}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}} \quad (2.3)$$

The first term of the equation corresponds to the kinetic energy of the electrons, the second one refers to the Coulomb attraction between electrons and nuclei, the third term represents the repulsion between electrons, and the fourth term provides the nuclear repulsion for a given geometry. Due to the Born-Oppenheimer approximation, the nuclei-nuclei ($Z_A Z_B / R_{AB}$) interaction can be regarded as static. This assumption renders the fourth term of the Hamiltonian as a constant. The physics behind the Hamiltonian hide what is known as the "electronic correlation". The Coulomb repulsion term given by the reciprocal of the distance between two electrons, increases in regions where electrons find themselves very close to one another, preventing the occupation of the same location in space of any two electrons. This means that the movement of any pair of electrons is not independent, but correlated. To say that two electrons are correlated is equivalent to stating that the probability of finding them in the same point in space is zero. The instantaneous position of the electron *i* forms the center of a region that prevents the electron *j* to occupy it. For this reason, it is said that each electron, described by the wave function, is surrounded by a "Coulomb hole".

Unfortunately, even for stationary systems, only two particle systems can be solved analytically. Numerical solutions of high accuracy (some essentially indistinguishable from the real one) can be obtained for small many-body problems by performing a great number of mathematical operations with aid of supercomputers. However, for systems of relatively larger molecular size, involving many or heavy atoms, and more interesting problems, such as chemical reactions, additional approximations are needed to solve the equation.

The simplest function used to describe a system is the many-electron "Slater determinant". The determinant is built by one-electron orthogonal wave functions that describe both an electron spatial distribution and its spin and are called "spin-orbitals". The electrons are fermions and hence must be described by antisymmetric wave functions, which is the case of determinants. The Slater determinant also fulfills the basic law of symmetry derived from the "indistinguishability principle"

because it allows the description of *N* electrons occupying *N* spin-orbitals without specifying which electron occupies each orbital. Because the non-relativistic Hamiltonian does not dependent on the electron spin, each spin-orbital can be expressed by the product of a spatial orbital and a spin function.

From a physical standpoint, the use of one-electron wave functions to describe the total electron density of the molecular system implies that we are in an independent electron model, uncorrelated, thus discarding the Coulomb hole. Nevertheless, the incorporation of the "antisymmetry condition" of the wave function via the Slater determinant causes electrons with the same spin function to be correlated, i.e., that the probability of finding two electrons with parallel spins at the same point in space is zero. This phenomenon is known as the "Fermi hole". We are, therefore, in an independent particle model where the behavior of certain electrons is not fully independent, as the Fermi hole simulates to some extent the Coulomb hole. As the movement of electrons with different spin function remains uncorrelated (there is a finite probability of finding two electrons with opposite spins at the same position in space), the wave function corresponding to a single Slater determinant is commonly known as an uncorrelated wave function.

2.2 *Ab initio* methods: Hatree-Fock and *post*-Hartree-Fock

Ab initio quantum methods are usually based on the Hartree-Fock (HF) approximation,[60,61] which constitutes, in most cases, the first step toward more specific procedures, playing a crucial role in Modern Quantum Chemistry. In fact, many of the quantum chemical methods can be considered as simplifications of the HF method or improvements of it. The HF method provides the mathematical tools necessary to determine the unknown spin-orbitals that provide the best Slater determinant within the framework of the variational principle. This principle states that the best wave function (one-determinant type) is that which fulfils the stationary condition, normally leading to the lowest-energy solution. The HF equation is nonlinear and must be solved iteratively through the self-consistent field (SCF) procedure. In practice, the HF equation is solved by introducing a finite set of spatial basis functions that can be expressed in different matrix equations: the Roothaan equations in the case of

restricted closed-shell determinants (restricted Hartree-Fock, RHF),[62] the Pople-Nesbet equations for unrestricted determinants (unrestricted Hartree-Fock, UHF),[63] and the Roothaan-Hartree-Fock equations for restricted open-shell determinants (restricted open-shell Hartree-Fock, ROHF).[64]

The HF approximation does not provide an accurate description of the molecular properties, mainly because it disregards the electronic correlation energy. The mathematical approximation introduced in the HF method has a simple physical interpretation: it is equivalent to consider that each electron interacts with the field produced by the nuclei and an average charge due to the other electrons. The actual interaction of the electrons with each other is then only considered in average. This makes that the motion of two electrons with opposite spin is uncorrelated, meaning that the probability of finding two such electrons in the same point of space is different from zero, which is clearly a physical nonsense. The correlation energy (E_c) is defined as the difference between the exact nonrelativistic energy of the system within the Born-Oppenheimer approximation (E_0) and the HF energy (E_{HF}) in the limit in which the set of functions employed meet the condition that it is a complete basis set. Since the use of a complete basis set is prohibitive or simply impossible, in practice, the correlation energy is usually computed as the difference between the energies calculated with a given method and the HF, both obtained with the same basis set.

The HF wave function, being a Slater determinant, is uncorrelated, which implies certain limitations on its applicability. Then, distinct techniques have been developed to further improve the description of electronic problems, giving rise to the so-called "post-Hartree-Fock" methods. Their quality is assessed in terms of the amount of electron correlation which is included (in addition to the basis set).

The HF method determines the best Slater determinant within a given basis set. It has been proven to be able to describe the ground state of many molecules in the vicinity of the equilibrium geometry. That is because most of such systems are well described with a single configuration. However, for all the other situations, the method yields poor results. A possible improvement of the HF, which increases the flexibility of the variational problem, involves the construction of a starting many-electron wave function containing more than one Slater

determinant. That multideterminantal wave function can be written generically as

$$\Psi = a_0\Phi_{HF} + \sum_{i=1} a_i\Phi_i, \qquad (2.4)$$

where the square of the coefficients a_i is the weight of each of the determinants in the linear expansion. Such determinants can be built by replacing the occupied molecular orbitals (MOs) by virtual orbitals which are available due to the fact that a number of orbitals equal to the dimension of the subspace spanned by the basis set employed is obtained in the resolution of the HF equations. The so-built new determinants are classified according to the number of occupied orbitals that have been replaced. Thus, the Slater determinants that have been obtained substituting *m* occupied orbitals are called *m*-excited determinants, i.e., singly (S), doubly (D), triply (T) excited determinant, and so on.

The principle behind the configuration interaction (CI) post-HF method[65] is to use the Slater determinants that can be obtained from the HF calculation to form a many-electron basis set in which the electronic wave function can be expanded linearly and carry out a variational determination of the coefficients of such expansion. If the electronic wave function is expressed as a linear combination of all the possible determinants that can be obtained from a HF calculation performed using a complete basis set, the CI strategy will give the exact solution of the electronic problem. Due to the impossibility of using a complete basis set, since it requires an infinite number of functions, this exact calculation cannot be carried out. Nevertheless, a CI calculation performed using all the orbitals obtained with a truncated one-electron basis set gives the exact result within the space spanned by the finite basis set. This type of calculation is called full configuration interaction (FCI).

The number of possible Slater determinants increases considerably with the dimension of the set of orbitals that can be used to construct them. Therefore, except for very small systems, the CI method must be used in a truncated way, in which the electronic wave function is expanded in the basis of only some types of *m*-excited determinants. For example, one common way to perform an approximate CI calculation is to consider only the singly- and doubly-excited determinants. This strategy

is named singly- and doubly-excited configuration interaction (SDCI) method. Although all forms of truncated CI methods usually improve results with respect to the simple HF calculation, they have a serious deficiency: they are not size consistent. The lack of this property leads to a non-physical description of systems of non-interacting molecules (the energy of such system is not equal to the sum of the isolated compounds that constitute them) and in general prevent the use of truncated CI methods like SDCI for the description of chemical process in which the relative energies of molecules of different size must be calculated.

A post-HF method that solves the size-consistency problems of truncated CI is the coupled cluster (CC).[66,67,68] This method reformulates the electronic Schrödinger equation as a nonlinear equation, enabling the computation of size-consistent high-precision approximations of the ground-state solutions for weakly correlated systems. It essentially takes the basic HF molecular orbital method and constructs multi-electron wave functions using the exponential cluster operator to account for electron correlation. The complexity of the equations and the cost of the computation increases sharply with the level of excitation operators taking into account. For many applications, the coupled cluster singles and doubles (CCSD), while relatively inexpensive, does not provide sufficient accuracy except for the smallest systems (approximately 2 to 4 electrons), and often an approximate treatment of triples is needed. The most well-known coupled cluster method that provides an estimate of triples by means of perturbative techniques is CCSD(T). It provides a highly accurate description of closed-shell molecules near the equilibrium geometry, although the method breaks down in more complicated situations where a multiconfigurational wave function is required to describe the electronic problem.

Apart from the CI and CC techniques developed to compute the electron correlation, a third procedure is also possible which makes use of the many-body perturbation theory in order to add corrections to the HF wave function and its energy. In this group of methods, the electronic energy and many-electron wave functions is written as a Taylor expansion in powers of a perturbation parameter. Then, different levels (first, second, etc.) of improvements or corrections arise depending on the terms of the Taylor expansion considered. Among these methods, the most common are those corresponding to the Møller-Plesset (MP) perturbation theory up to second and third order (MP2 and MP3, respectively).[69] The

MP2 method typically covers a large amount of the correlation energy and it is probably the cheapest method (computationally speaking) to treat electronic problems. Like many of the approaches on perturbation theory, MP works well when the perturbation is sufficiently small. Otherwise, the correlation energy can be overestimated (resulting energies below the exact energy, E_0). One of the great advantages of this method over the CI is that MP scale properly with the number of particles of the molecular system (correct N-dependence). In contrast, the disadvantage with respect to CI is the fact that whereas the latter methods result in a higher or upper approximation bound to the exact total energy of the system (due to the variational method employed), such guarantee does not exist in the perturbation theory methods.

So far, the methods described (CI, CC, and MP) are single-configuration methods, which employ as a reference a single Slater determinant. These methods are restricted to situations where a single reference wave function is adequate for the description of the chemical process. However, many other electronic structure problems, such as biradical systems, energy degeneracies, bond dissociation, heavy elements in which several degenerate states are present, or electronic excited states in general, cannot be properly described. These situations require a more flexible many-electron wave function. To achieve that, one possibility is to use an electronic wave function containing a relative small number of determinants in which not only the coefficients of the expansion are optimized but also the orbitals employed. That is the basic idea of the multiconfigurational self-consistent field (MCSCF) method,[58,59] in which the variational principle is used to determine both the expansion coefficients and the orbitals. The MCSCF method retrieve a part of the electron correlation called “static correlation” that is related to electronic configurations that have similar energies. In a subsequent step, the MCSCF wave function can be used as a reference in distinct calculation procedures which attempts to add the remaining electron correlation, the so-called “dynamic correlation”. As for the single-configurational methods, the CI, CC, or MP techniques can be used here, giving rise to the multi-reference configuration interaction (MRCI), multi-reference coupled cluster (MRCC), and multi-reference perturbation theory (MRPT) methods, respectively. The advantage of these methods is that they can generate very precise wave functions whether the reference includes sufficient MCSCF configurations. The disadvantages are the

high computational cost, mainly in the case of the MRCC, and, in the MRCI method, the fact that its truncations do not scale linearly with the number of particles.

2.3 The CASSCF/CASPT2 method

A variant of the MCSCF method that has become particularly popular because of its technical and conceptual simplicity is the CASSCF.[70,71,72] In this case, the selection of the configurations that will be included in the many-electron wave function is carried out by classifying the MOs in three sets: inactive, active, and secondary orbitals. The active MOs are typically chosen from the HF canonical spin-orbitals occupied and unoccupied with higher and lower energies, respectively, from a HF calculation. Nevertheless, we must bear in mind that the choice of the active space depends on the problem at hand. The inactive MOs are always doubly occupied, while the secondary MOs (also called virtual) have no electrons. Among the active MOs, a complete or full CI (CAS-CI) is performed. The CASSCF wave function is formed by a linear combination of all possible configurations that can be built by distributing the active electrons among the active orbitals that are consistent with a given spatial and spin symmetry, known as symmetry- and spin-adapted configuration state functions (CSFs). The correlation energy introduced by the CASSCF procedure usually corresponds to the non-dynamic correlation or static electronic correlation and includes near-degenerate CSFs. The goal of the CASSCF methods is not to retrieve a large part of the total correlation energy, but rather describe qualitatively all changes that occur in the correlation energy for a given process, which is achieved by a proper choice of the MOs to be correlated.

Once the CASSCF wave functions and energies are computed, the inclusion of dynamic electron correlation due to the short-range electron-electron interactions is needed in order to obtain accurate results and a quantitative description of the electronic problem. This can be achieved by using non-degenerate perturbation theory with the multiconfigurational zeroth-order wave function (CASSCF), which gives rise to the CASPT2 method.[73,74] The wave function is corrected up to first order, and the energy is corrected up to second order. The mathematical formulations are described as follow,

$$E = E_0 + \lambda E_{(1)} + \lambda E_{(2)} \quad (2.5)$$

$$E_{(1)} = \psi_0 V \psi_0 \text{ and } E_{(2)} = \psi_0 V \psi_{(1)} \quad (2.6)$$

$$\hat{\mathrm{H}}_0 - E_0 \psi_{(1)} = -\hat{\mathrm{H}} - E_0 \psi_0 \quad (2.7)$$

where λ is a parameter of the perturbation, V is the perturbation operator, ψ_0 and $\psi_{(1)}$ represent the zeroth-order and first-order wave functions, respectively.

In some practical applications, corrections over the zeroth-order Hamiltonian, $\hat{\mathrm{H}}_0$, must be applied due to the presence of strongly-interacting intruder states in the second-order calculations, which are normally related to large coefficients in the first-order expansion that leads to a low value of the reference weight.[55] The reference weight can be used as a simple and rapid criterion of the quality of the perturbation treatment, which should be similar for the different electronic states considered. Its value depends on the number of correlated electrons; thus, it decreases by enlarging the molecular system. Strongly-interacting intruder states are sometimes present in the treatment of excited states when the active space does not include all the π valence system, causing an unbalanced description of the system. This problem can be solved by enlarging the active space of the reference CASSCF wave function with the important orbitals.

Despite removing the presence of the strongly-interacting intruder states, the weight of the reference function might be still small due to the effects of accidental near-degeneracies, the so-called weakly-interacting intruder states. To minimize the presence of these weak intruder states, a common technique is to include an imaginary (IMAG) level shift in the $\hat{\mathrm{H}}_0$ that is subsequently back corrected in the second-order energy.[75,76,77]

Another common correction over the $\hat{\mathrm{H}}_0$ is the introduction of a shift parameter called ionization potential electron affinity (IPEA).[78] This is done to avoid the systematic overestimation of the correlation energy in open-shell systems with respect to close-shell cases. The use of an IPEA shift of 0.25 a.u. was proved to improve the description of radical cations and anions. For excited states, the improvement of this correction is, however, not well-established. Hence, the use of the conventional zeroth-order Hamiltonian is recommended for the characterization of the excited

states, since it is supported by more than one decade of spectroscopic and photochemical studies.

Finally, another issue to comment about the CASPT2 method is the non-orthogonality of the wave functions obtained for different electronic states. This means that the interaction between states is not properly considered at this level of theory. To account for the coupling of several electronic states, a multistate treatment can be carried out, in which an effective Hamiltonian matrix is built. The diagonal elements are the CASPT2 energies and the off-diagonal elements take into account the coupling up to second order of dynamic correlation energy. The orthogonal states can be obtained from the so-called multistate (MS)-CASPT2 solutions,[79] where the resulting asymmetric matrix is made symmetric, assuming that the off-diagonal terms are similar. This assumption must be verified in any particular case since if it is not valid, the MS may lead to non-physical results for both energies and wave functions.

The combination of the CASSCF and CASPT2 methods giving rise to the CASPT2//CASSCF protocol is probably one of the most useful and practical methodologies for studying spectroscopy and photochemistry.[56] In this protocol, geometry optimizations are carried out at the CASSCF level and then CASPT2 energies are computed at the optimized structures. The success of the CASPT2//CASSCF protocol comes from the fact that the computation of the energy gradients, which are very time and CPU-demanding, are carried out at the lower level of theory, which is accurate enough to determine the geometries. Only situations in which strong differential correlation effects are present cannot be treated with the CASPT2//CASSCF protocol and CASPT2//CASPT2 should be employed. In general, the most relevant advantages of the CASPT2//CASSCF methodology are, firstly, that it can be applied to study medium to relatively large systems, secondly, that it has no restrictions, and thus, it can be used to describe all type of states, degeneracies, etc., and, finally, that it has a general accuracy of around 0.2–0.3 eV.

2.4 Basis sets

All *ab initio* methods stated above make use of one-electron basis set functions in order to form the MOs through a linear combination of such functions, procedure related to the strategy known as molecular orbitals as linear combination of atomic orbitals (MO-LCAOs). These basis set functions are atom-centered and span most part of the theoretically complete Hilbert Space, which implies an infinite number of functions. Larger basis sets normally produce improved results due to the inherent better description and the higher flexibility provided, even though error cancellation could lead to significantly "good" results with relatively small basis set.

There are two basic types of one electron AO functions commonly used in electronic structure calculations, Slater type orbitals (Slater Type Orbitals, STOs)[80] and Gaussian type orbitals (Gaussian Type Orbitals, GTOs).[81] Whereas the STOs exhibit an exponential dependence with the distance between the nuclei and electrons and are more accurate as regards the representation of the MOs, GTOs exhibit an exponential dependence with the square of the distance between the nuclei and electrons. GTOs require, however, lower computational time for evaluating the two-electron integrals.

The smallest or minimal basis set that can be built employs only the minimum number of functions per atom in the description of its orbitals. For hydrogen, this means a single function s, while it takes two functions s (1s and 2s) and a set of functions p (2px, 2py and 2pz) for the next period in the periodic table. The next step to improve the basis set is doubling all functions, leading to the double-zeta basis set. In this case, two s functions (1s and 1s') are used for the hydrogen and four s functions (1s, 1s′, 2s, 2s′) and six p functions ($2p_x$, $2p_x'$, $2p_y$, $2p_y'$, $2p_z$, $2p_z'$) for the elements of the second row. A variant of such basis set is one that only doubles the number of valence orbitals, which are the most relevant in chemical problems, and leaves the description of the core orbitals to its simplest form. The valence double zeta (VDZ) basis set is then produced. The next step in increasing the number of functions leads to the triple zeta (TZ) or, if only the valence shell is improved, the valence triple zeta (VTZ). In some cases, such an increase of the number of functions for each atom is not enough to describe the phenomenon under study that

requires higher angular momentum functions. These functions are called polarization functions. The addition of a single set of p functions for hydrogen and d functions for heavier atoms to the DZ and VDZ basis sets gives rise to the double-zeta polarization (DZP) and valence double-zeta polarization (VDZP) basis sets, respectively. In certain situations in which it is intended to more accurately describe the outer part of the many-electron wave function, diffuse functions (basis functions with small exponents) are added. This type of functions is especially required in molecular systems where weakly-bound electrons are present, for example, in anions of highly external excited states.

In order to combine the advantages of the STOs and GTOs, another strategy can be used, which reduces the computational cost, as compared to the use of STOs, without significant loss of precision. In this strategy, the full set of basis functions, known as primitive GTOs (PGTOs), are grouped by forming fixed linear combinations (so-called contraction of the basis set) to yield a smaller set of contracted GTOs (CGTOs):

$$\phi_{CGTO} = \sum_i a_i \phi_{i,PGTO} \tag{2.8}$$

The contraction of a basis set reduces the number of variational parameters and thus increases the computational efficiency. An example of contraction is C,H(10s4p1d/2s1p)→C,H[3s2p1d/2s1p], where the notation ("primitive functions for elements of the first row"/"primitive functions for hydrogen")→["contracted functions for elements of the first row"/"contracted functions for hydrogen"] is employed.

There are two different schemes for contracting a set of PGTOs, the general and segmented contractions. In the first case, all primitive functions (of a given atom) of a particular angular momentum contribute to all the contracted functions with that angular momentum with different contraction coefficients. In the segmented scheme, each primitive function participates in one or a few contracted functions.

Distinct types of basis sets have been produced in the literature, which employ different approaches to determine the contraction

coefficients of equation 2.8. The most common are the Pople,[82] Dunning,[83,84] and atomic natural orbital (ANO) basis sets.[85]

The contraction coefficient in the Pople basis sets are obtained by performing ground-state non-correlated HF calculations. This type of basis functions is more widespread in their use by its presence in most quantum-chemical computational programs and extensive calibration available. They use segmented contraction and allow performing low computational cost calculations. An example of Pople basis set is 6-311G++(2df,2pd). It has a core formed by one CGTO as a linear combination of six PGTOs. The valence shell is described by three CGTOs (valence triple zeta), represented by three, one, and one PGTOs functions. The addition of diffuse functions is denoted by the symbol "+". The first "+" indicates a set of diffuse sp functions to the heavy atoms and the second ("++") corresponds to the extra addition of a diffuse s function in the hydrogen atoms. Finally, the polarization functions are indicated within parentheses. Two d and one f functions are added for the heavy atoms and two p and one d for the hydrogen atoms.

Unlike Pople-type basis sets, the most modern Dunning and ANO basis sets have the advantage that their contraction coefficients are optimized using quantum-chemical methods that take into account the correlation energy and have the ability to generate a sequence of basis sets that converges towards the basis set limit (correlation errors are reduced at each step of increased quality). Dunning basis sets consistent with the correlation (correlation consistent, cc), also using a segmented contraction scheme, were developed in order to recover the correlation energy of the valence electrons. They are designed in such way that those functions which contribute with similar amounts to the correlation energy are included simultaneously, regardless of the type of function. An example is the cc-pVDZ basis set that has C,H(9s4p1d/4s1p) PGTOs contracted to C,H[3s2p1d/2s1p]. These basis sets may grow in size by adding diffuse functions, in which case the prefix "aug-" is added to the acronym. Thus, the aug-cc-pVDZ basis set has a set of 1s, 1p, and 1d additional functions for the non-hydrogen atoms.

The ANO-type basis sets constitutes also powerful basis sets, whose contraction coefficients are optimized with correlated methods and using the general contraction scheme. The principle underlying this type of basis sets is the contraction of a large set of PGTOs in a relatively

small number of functions CGTOs using natural orbitals (NOs) obtained in correlated calculations on free atoms, typically at CISD or MCSCF. To increase the flexibility of the basis set, an average is done over NOs obtained from different situations in the atom: neutral, cation, anion, different excited states, with the addition of an electric field, etc. The PGTOs are combined in the ANO contraction scheme depending on the magnitude of the occupation numbers. Due to the diffuse and flexible character of the ANO-type basis set are particularly suitable for the treatment of excited and anionic states. Augmented correlation consistent (aug-cc) basis set are also appropriate in these situation, although the ANO basis sets usually lead to better results using a smaller number of functions. There are two main different sets of ANO-type basis set which differ in the number of primitive functions: ANO-L (large)[86,87,88] and ANO-S (short).[89] Furthermore, the ANO-RCC type,[90] have been taken into account relativistic effects in their construction. In the present Thesis, the notation used is referred to ANO basis sets; for example, ANO-L C,H[4s3p1d/2s1p] shall be denoted as ANO-L 431/21.

2.4.1 Basis set superposition error

The basis set superposition error (BSSE)[91] is a consequence of the fact that non-complete basis sets are used in the actual computations. In the calculation of dissociation energies, when the energy of a super-system (AB) is compared with the energies of the constituent fragments (A and B), the description of the fragments A and B in the super-system is better than that of the isolated fragments. This is due to the fact that a higher number of basis functions are considered in the computation of each fragment at the geometry of the super-system, since the basis functions of the adjacent fragment also contribute to the energy of each fragment.

One of the most common techniques employed to correct the BSSE is the counterpoise (CP) method proposed by Boys and Bernardi,[92] which was developed for describing dimerizations of atoms or rigid monomers. It is based on computing the extra stabilization energy that the fragments have in the dimer as compared to the isolated situation and then subtracting this energy to the total energy of the dimer.

For a super-system or dimer AB, the non-corrected binding energy E_b is calculated by means of the equation:

$$E_b(AB) = E_A + E_B - E_{AB}, \quad (2.9)$$

where, E_X (X=A,B) indicates the energy of the isolated X fragment at its equilibrium geometry and E_{AB} represents the total energy of the dimer AB at a particular geometry.

The corrected binding energy ($CP\text{-}E_b$) is obtained by subtracting the BSSE to the non-corrected E_b:

$$CP\text{-}E_b(AB) = E_b(AB) - [CP\text{-}BSSE(AB)], \quad (2.10)$$

where

$$CP\text{-}BSSE(AB) = E_A(B,R=\infty) - E_A(B,R=AB) + E_B(A,R=\infty) - E_B(B,R=AB). \quad (2.11)$$

Here, the geometry of the monomers is the same as that in the dimer. Then, E_A(B,R=AB) indicates that the energy of A is calculated with the ghost orbitals of B at the geometry of AB, whereas in E_A(B,R=∞) the ghost MOs of B are placed at an infinite of sufficiently large distance with respect to A.

3 Photophysical and Photochemical Processes

"For the rest of my life I will reflect on what light is."
Albert Einstein (1879-1955)
Physicist

When radiation interacts with matter, energy quanta are distributed among the different degrees of freedom that it possesses: translational, rotational, vibrational, electronic, or nuclear. In biomolecules, from the photochemical standpoint, we are interested especially in the vibrational and electronic states. Whereas the former can be populated by molecular collisions, the latter generally require the participation of radiation. Before absorbing energy from the electromagnetic radiation, the molecules are in their electronic ground state, with an energy threshold named zero-point vibrational energy. After irradiation, the molecule can deactive its excess of energy via several processes that can be either reactive or non-reactive, radiative or non-radiative, unimolecular or involving other species, etc. In this chapter we will focus firstly on the unimolecular processes and next, we will briefly describe a particular type of bimolecular process that involves the formation of a relatively stable structure in the excited state, the excimer.

3.1 Unimolecular processes

Figure 3.1 shows the basic phenomena that can take place in the molecule after absorption (Abs) of radiation.[93,94,95,96,97,98] Electronic states are represented by curves (actually hypersurfaces in many dimensions) which constitute the potential energy of the driving forces of the molecules determined by the different distributions of the charge density. Such energy functions of the coordinates are called potential

energy hypersurfaces (PEHs). The absorption of radiant energy from the ground state (S_0) is used to change the charge distribution of the molecule by exciting or relocating its electrons. Depending on the energy range of the radiation, the molecule reaches some energy levels or others that are known as electronic excited states (S_1, S_2, S_3, etc.). Not all the energy transitions are equally likely. Those states in which the transition is more probable are called bright states (the term dark is reserved for those with a negligible transition probability).

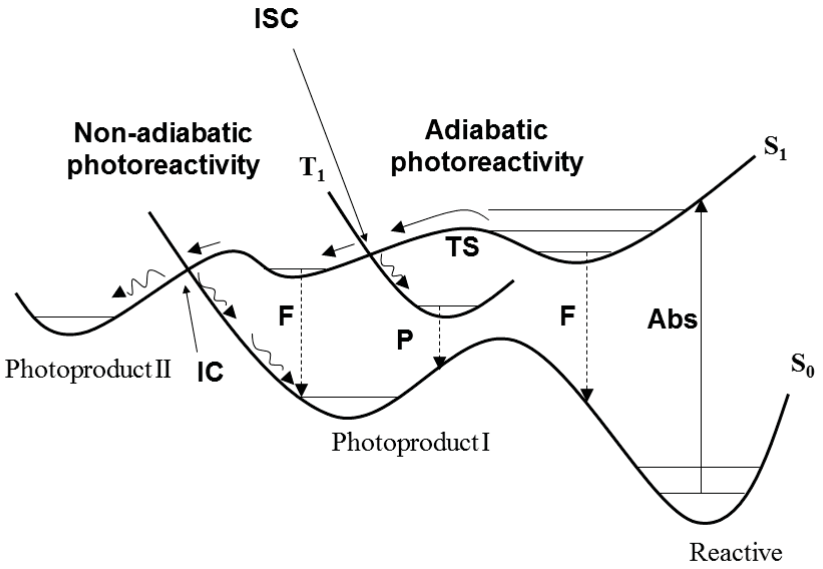


Figure 3.1 Scheme of the main photophysical and photochemical processes (see text).

The electronic states have a certain lifetime, i.e., the energy gained is distributed among the degrees of freedom and the molecule then drops out. As highlighted in Figure 3.1, the energy relaxation may occur through a great number of phenomena. We shall concentrate here on the lower electronic states, in particular, S_0 (ground state), S_1 (lowest excited singlet), and T_1 (lowest triplet). Higher states may or may not be populated, but the de-excitation paths are so fast that the main photochemical process occurs in practice from the lower energy excited states (Kasha rule).[99] The energy absorption is the fastest process of the

entire mechanism, in the femtosecond time scale (10^{-15} fs), since it involves only electron density redistribution and no nuclear reorganization. After energy absorption to S_1 (more likely than to T_1 because in this case they have different spin multiplicity) the majority of the molecules will be in an excited vibrational state. Hence, the subsequent process implies the dissipation of the excess energy through internal vibrations in the molecule, by the intramolecular vibrational relaxation process. The excess energy may be also transferred to solvent molecules in another process called vibrational energy transfer. This relaxation step is fast (10^{-14}–10^{-11}s), although the rate depends also on the molecular size. In this process, the molecule relaxes by adjusting the positions of its nuclei towards a more favorable geometry on S_1 like a stable minimum as shown in Figure 3.1. It is stated that the molecule, on average, goes through a minimum energy path from the initial position (sometimes called absorption or Franck-Condon region) to the S_1 stable minimum. Actually, the molecules oscillate around that route and other singular points of interest can be reached if the energy barriers present along the path are not high. Such routes are known as photoadiabatic, where the term applies to any adiabatic process which involves a single PEH and a path which can lead the molecule to different regions of the hypersurface, and ultimately relax giving rise to various photoreaction products.

All those molecules that reach a minimum region are trapped in a structural arrangement that relax through radiative emission to the lower energy state. The emission can occur between states of the same or different spin multiplicity, producing fluorescence (F) or phosphorescence (P), respectively. The corresponding decay times (called average radiative lifetimes, τ) are higher than those of other relaxation processes, ranging 10^{-9}–10^{-6} s (ns-μs) for fluorescence and 10^{-3}–10^{-2} s (ms-min) for phosphorescence. The ratio between the emitted and absorbed radiation is known as quantum yield (ϕ_i). Thus, systems with an intense emission have a quantum yield close to unity. Any reduction of this magnitude implies that part of the energy has found other ways to relax. Then, we say that the molecule undergoes a deactivation (quenching).

The most common deactivation mechanism is the non-radiative decay (radiationless decay) that the molecule undergoes when it finds a crossing (funnel) between two PEHs. In regions where two or more

electronic states are very close in energy, the molecule has a high probability to rapidly transfer the excess energy to the lower state. This transition is called IC or intersystem crossing (ISC) depending on whether the transfer is between states of the same or different spin multiplicity, respectively. The times associated with the IC and ISC processes are fast, 10^{-14}–10^{-11} s. Photochemical reactions that take place through these transfers are called non-adiabatic reactions. After the crossing, a certain percentage of the molecules returns to their initial state, and the rest may produce new photoproducts. The percentage will depend on the topology of the PESs at the crossing and also on dynamical aspects.

To study the processes displayed in Figure 3.1 from a theoretical standpoint, and to determine the main decay channels, we need to know the topology of the PEHs, i.e., the position of the singular points, minima, transition states, and crossings. The most favorable paths connecting these points and the nearby regions are also relevant information. In the present work, a time-independent approach is employed, which allow determining the driving forces of the photochemical phenomena. The computational strategies used to determine the main photochemical routes are described in the chapter 4.

3.2 Bimolecular processes: Excimers

The term "excimer" is defined as a dimer (*D*) of two identical molecules in which the monomers (*M*) are bound in the excited state, whereas they are unbound in the ground state in the absence of external constraints.[100,101,102] The excimer can be seen as a homodimer in which one of the monomers is electronically excited and then the excitation is delocalized over the two molecular systems:

$$^1M + {}^1M \xrightarrow{h\nu} {}^1M + {}^1M^* \Leftrightarrow ({}^1M + {}^1M)^* \equiv {}^1D^* \qquad (3.1)$$

Regarding the property of spin, these structures can be singlet or triplet. There are three manners in which a triplet excimer $^3D^*$ can be produced. First, it may be formed by collision between a triplet excited-state molecule ($^3M^*$) and a non-excited (^1M) system:

$$^1M + {}^3M^* \Leftrightarrow {}^3D^* \qquad (3.2)$$

The other two routes are via an ISC process between singlet and triplet states excimer (Scheme 3.3), or via an ISC occurring in one of the monomers followed by the π-stacking interaction with the other monomer in its ground state (Scheme 3.4):

$$^1D^* \xrightarrow{\text{ISC}} {}^3D^* \tag{3.3}$$

$$^1M + {}^1M^* \xrightarrow{\text{ISC}} {}^1M + {}^3M^* \Leftrightarrow {}^3D^* \tag{3.4}$$

Experimental studies have shown that the formation of excimers is common in aromatic systems and occurs in different media.[100] Due to the formation of excimers, the emission properties of the aromatic molecules in solution or condensed phases change with respect to those of the isolated systems. Thus, the fluorescence band appears at lower energies, whereas the absorption spectrum remains similar. The reason is that the excited state has lower energy in the dimer than in the separated monomers. Thus, the energy difference between the excited and ground states is lower in the dimer. In contrast, the absorption spectrum remains the same in the gas phase and in solution or condensed phases. This indicates that the excimer species is not present in the ground state.

Excimers taking place in aromatic molecules may not only affect their spectroscopic properties, but also they may play an important role in the photochemistry of these systems. Thus, the excimer-type interaction can become strong enough in some cases to give rise to a stable photodimer M_2:

$$(^1M + {}^1M)^* \equiv {}^1D^* \Leftrightarrow M_2 \tag{3.5}$$

This occurs, for example, in the case of anthracene solutions exposed to sunlight, which photodimerization is known since the nineteenth century. As shown in Scheme 3.5, the photodimers can subsequently be reversed to the original pair of individual molecules by using radiation of sufficient energy (photolysis reverse reaction). The presence of excimers or similar structures in biological systems, called here “bioexcimers” is much more common than expected. In DNA, its presence has been suggested in several experiments.[31,32,103,104]

4 Computational Strategies in Photochemistry

"Victorious warriors win first and then go to war, while defeated warriors go to war first and then seek to win"
Sun Tzu (544-496 B.C.E.)
Chinese General and Philosopher

In computational photochemistry, from a static standpoint, we are interested in describing the most probable evolution of the molecule after light absorption. Solving the Schrödinger equation for different electronic states at fixed values of the atomic coordinates of the molecular system under study will result in a set of data that relate parametrically the energy of the states and the nuclear coordinates. The representation of the obtained energies *vs.* the degrees of freedom gives rise to the so-called PEHs, as previously introduced in chapter 3. The topography and the transition probabilities between the PEHs are two quantum-chemical entities that are useful to interpret and predict spectroscopic and photochemical properties of molecular systems by using theoretical methods. Hence, its determination is the objective of many studies carried out in the field of Applied Theoretical Chemistry. In particular, the interest in dealing with quantum chemical PEHs, from a static standpoint, is mainly to locate certain singular points, to determine the relative energy between these points, and to obtain information about the transition probability among the electronic states.

4.1 Singular points of potential energy hypersurfaces

The first types of singular points that are relevant in computational photochemistry are the minima and first-order saddle points. They are stationary points on the PEHs, that is, geometries in which the gradients are zero. Minima are related to reaction reactants and products, which

possess certain stability due to the fact that the energy is trapped at these points. In this case, the matrix of the second energy derivatives (Hessian) has all the eigenvalues possitive. On the other hand, if this Hessian has *n* negative eigenvalues, it is a saddle point (SP) of *n* order, i.e., a point that is a maximum in *n* nuclear displacements and a minimum for the rest of distortions. The first-order SP is chemically related to the transition state (TS) of adiabatic reactions. The TS connects different regions of the PEHs, such as minima.

More complex singular points imply crossings between PEHs. These situations are hardly relevant in the ground-state chemistry, but play a crucial role in non-adiabatic photochemistry. The crossing points that occur between states of the same spin symmetry are called conical intersections (CIs).[96,105,106,107,108] They represent the most favorable situation for an efficient IC process since the probability of transfer is greater when the gap is lower.

In a system with $F = 3M–6$ degrees of freedom, a CI corresponds to a crossing between two PEHs along a hyperline of $F–2$ dimension. In the topological description of the CIs, the total space of coordinates (dimension *F*) is divided into: i) the intersection subspace of $F–2$ size, in which the two states have the same energy, and ii) the branching subspace of dimension 2, in which the energy degeneracy is lifted. These two dimensions are defined by the gradient difference vector, x_1, and the non-adiabatic coupling vector, x_2:

$$x_1 = \frac{\partial(E_1 - E_2)}{\partial Q} \tag{4.1}$$

$$x_2 = \left\langle \Psi_1 \left| \frac{\partial \Psi_2}{\partial Q} \right. \right\rangle \tag{4.2}$$

where *Q* represents the nuclear coordinates of the system.

The x_1 and x_2 vectors are both linear combinations of internal coordinates. While x_1 points to the direction with the more pronounced slopes of the upper and lower PEHs, the x_2 direction indicates the nuclear displacements that maximize the mixing of the adiabatic functions at the CI.

From a static standpoint, it is common to search the CI structure of lowest energy. However, the chemically interesting geometry is the one that is accessible along the main relaxation paths of the molecule. Since the upper or lower hypersurfaces at the CI region do not present zero gradients, as would be true for a stationary point, special algorithms are required to optimize the structure of the CI. Most of them are based on following the gradient projection on the $F-2$ intersection subspace orthogonal to x_1 and x_2 until it becomes zero.

Other crossings between PEHs involve electronic states with the same spatial symmetry and different spin multiplicity. The most common situation corresponds to crossings between singlet and triplet states (singlet-triplet crossings, STCs).[96] They can be related to the ISC phenomena, which are much less efficient than IC processes because the population transfer between states with different spin multiplicity is forbidden by the spin selection rules. Nevertheless, the spin and angular momenta might couple and thus increase the transition probability, as described in the next section.

Similarly to the STC points, we can also find situations in which the states that cross are a singlet and a doublet state (singlet doublet crossings, SDCs). Here, one should bear in mind that this type of singular point does not correspond exactly to an ISC because it implies a change in the total number of electrons of the molecular system. Nevertheless, the SDC may play an important role in the ionization processes of the molecule and it is worth taking it into account. Thus, the SDC represents the situation in which the process of adding (or removing) an electron does not require or release any energy due to the fact that the states are degenerate at this point. That is, the neutral molecule, having the particular geometry of this point, has the same stability as the corresponding monoanion (or monocation).

4.2 Energy transitions and transition probabilities

Once the singular points of the PEHs have been determined by using appropriate geometry optimization algorithms, the information that is subsequently required to study photoinduced phenomena is the relative

energy between the excited and ground states at the singular points and the transition probabilities between the electronic states.[96] Figure 4.1 shows a scheme with the theoretical quantities that provide relevant energetic information regarding the absorption and emission processes. The vertical absorption (E_{VA}) is the energy difference between the minimum of the ground state and the excited state at the equilibrium geometry of the former, the so-called Franck-Condon geometry. The electronic adiabatic transition (T_e) is the energy difference between the excited state and the ground state at their optimized equilibrium geometries. The zero-point vibrational energy (ZPVE) of the initial and final states has to be included to get the vibrational band origin (T_0) which can be directly compared with the experimental value. Finally, the vertical emission energy (E_{VE}) is the energy difference between the excited state and the ground state at the relaxed geometry of the former. The E_{VE} may be compared to the band maximum of the emission spectrum.

In the particular case of the ionization processes, transitions from the ground state of a neutral molecule to an ionic state corresponding to a positive or negative charge can be described by the ionization potentials (IPs) and electron affinities (EAs), respectively. We will only consider in the present Thesis the process of adding an extra electron. Then, the EA of a neutral molecule is defined as the energy involved in the process of adding an electron to a neutral system, such that a positive value refers to the situation in which the anion is more stable than neutral. Similarly to the absorption and emission energies, vertical and adiabatic quantities are also defined for the EAs. The vertical energy difference (VEA) corresponds to the electron transition from the ground state of the neutral system to a particular anionic state of the anion at the equilibrium geometry of the neutral system, while the electronic adiabatic energy difference (AEA_e) denotes the energy gap between the minima of both neutral and anionic states and reflects the geometrical relaxation occurring in the anionic state. The addition of the ZPVE correction leads to the AEA_0 value.

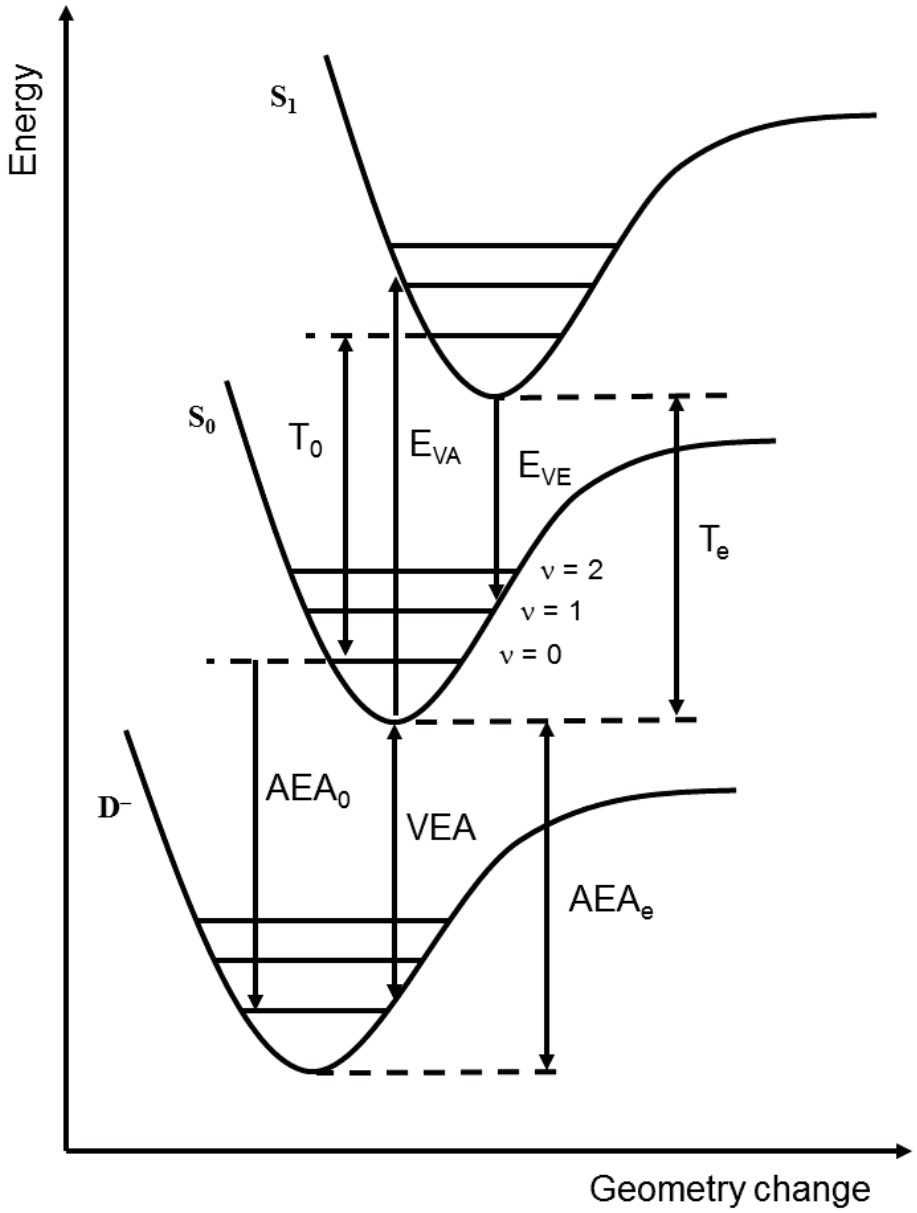


Figure 4.1 Diagram of the absorption, emission, and electron-attachment processes. Definitions of the theoretical quantities related to such processes are graphically shown through the electronic and vibrational potential energy levels (see text).

Positive VEAs indicate that the neutral molecule acts as an electron attractor. Then, the addition of the extra electron is energetically favored, and therefore, the anion can be created spontaneously. In this case, there are also positive values of AEA and the system is stabilized, that is, no electron undergoes spontaneous loss or autodetachment. Conversely, negative values represent anionic temporary states, known as transient or negative ionic states resonances that exist in short time periods and are prone to show autodetachment processes.

The probability of energy transfer between two electronic states is related to the strength of the interaction between the time-dependent electromagnetic field and the multipolar charge distribution of the molecular system, via the transition dipole moment (TDM). The transition probability between two states is proportional to the square of the TDM. Theoretically, this probability is known as the electronic oscillator strength, *f*. The TDM and *f* provide information on the polarization of the transition caused by radiation and the intensity of the bands observed in the experimental spectra through the Fermi Golden Rule.[109] The oscillator strength of vertical electronic transitions assuming the Franck-Condon principle can be estimated by

$$f = \frac{2}{3}\left|M_{if}\right|^2 \Delta E_{fi}\,, \tag{4.3}$$

where ΔE_{fi} corresponds to the energy difference between the final and initial states and M_{if} is the TDM between these states, defined as

$$M_{if} = \left\langle \Psi_i \middle| \hat{\mu} \middle| \Psi_f \right\rangle, \tag{4.4}$$

being $\hat{\mu}$ the electric dipole moment operator.[110]

The intensity of the spectral band corresponding to the transition between the spectroscopic states *i* and *f* is proportional to the oscillator strength between these states and thus the square of the matrix element of the transition (Fermi's Golden Rule). There are some selection rules that allow determining which transitions are dipole allowed or dipole forbidden. They are based on the symmetry and spin multiplicities of the states *i* and *f*. In the first case, only those M_{if} integrals belonging to the

totally symmetric irreducible representation will differ from zero. In the case of the spin, the electric dipole moment operator does not operate on the spin functions and, therefore, the integrals will be zero if the spin functions of the two states *i* and *f* are different. Thus, only transitions between states of the same spin multiplicity are allowed in principle. Nevertheless, spin-forbidden transitions can still be observed experimentally in certain conditions. For instance, singlet-triplet absorptions are recorded in electronic spectra, although with much less intensity than the singlet-singlet absorptions. In addition, phosphorescence emission and ISC are common processes occurring in certain molecules. This is due to the fact that the selection rules are based on a number of approximations, either in the resolution of the Schrödinger equation or in the treatment of radiation-matter interaction. For example, the spin-orbit coupling (SOC)[111] is not taken into account in the Hamiltonian of the non-relativistic Schrödinger equation used in the definition of the selection rules. The SOC corresponds to the interaction between the spin and orbital angular momenta. It allows the mixing of states of different spin multiplicity and therefore the interaction between them. The length of the SOC vector related to two states *i* and *f* is given by

$$SOC_{if} = \sqrt{\sum_{u} \left| \left\langle \mathrm{T}_{i,u} \left| \hat{\mathrm{H}}_{\mathrm{SO}} \right| \mathrm{S}_{f} \right\rangle \right|} \quad / \, u = x, y, z \,, \tag{4.5}$$

where $\hat{\mathrm{H}}_{\mathrm{SO}}$ is the spin-orbit Hamiltonian.

4.3 Photochemical reaction paths

So far, we have described the singular points (minimum, SP, CI, STC, and SDC) and the energy transitions among electronic states at these points, providing a static model of chemical reactivity. However, it must be noted at this point that an accurate description of the photophysics and photochemistry of a molecule requires verifying the connectivity and accessibility of the singular points after the irradiation event.[112] Here, we will briefly describe a few computational strategies that may help to achieve this target.

One of the simplest computational strategies to map the PEHs is to perform a linear interpolation of internal (LIIC) coordinates between the two points of interest and to determine the energy profile along this route. This strategy might be quite good when the structural transformation is simple, such as a bond breaking. However, the energy of the photochemical path might be largely overestimated in more complex geometrical changes. Constraint geometry optimizations may be useful to better describe the energy profiles when there is only one or two internal coordinates that change. In this case, a series of geometry optimizations is performed in which the reaction internal coordinate is fixed in each optimization and the rest of coordinates are allowed to relax. However, care must be taken here because the obtained points might not correspond to a connected path.

A highly-accurate computational strategy for determining the main photochemical reaction paths implies computing the intrinsic reaction paths (IRCs) or the minimum energy paths (MEPs).[56,113,114] While the former allows determining the lowest-energy route from a TS point, the latter might be used to describe the main decay channel from a non-stationary structure such as the point on the PEH of the bright excited state at the Franck-Condon geometry. The MEP computation follows the negative direction of the gradient of the energy. The algorithm[115] used in the studies in this Thesis implies determining the path of maximum decrease (steepest descendent path) and is based on a modification of the projected constraint optimization algorithm of Anglada and Bofill[116] and follows the approach of Müller-Brown.[117] In this computational strategy, each step requires the minimization of the PEH within the cross section of the hypersphere centered on the initial geometry and characterized by a predefined radius. The optimized structure is taken as the center of the new hypersphere radius and the same procedure is iterated until it reaches an energy minimum of the PEH. Mass-weighted coordinates are used, and therefore, the coordinate of the MEP corresponds to the so-called intrinsic reaction coordinate (IRC), measured in a.u., that is, bohr·amu$^{1/2}$.

4.4 Quantum chemistry programs

Several quantum-chemistry packages of software have been developed during the last decades, which include efficient algorithms to perform calculations with the methods and computational strategies

described in chapter 2 and 4, respectively. The program used in the present Thesis is MOLCAS,[118,119,120] which is a modular software mainly designed to study the chemistry of the excited electronic states. CASSCF geometry optimizations, CASSCF transition dipole moments (TDM), and CASPT2 energy calculations were computed in this Thesis with MOLCAS. The ANO-type basis sets included in the basis library of MOLCAS were used. From the CASSCF transition dipole moments and the CASPT2 vertical transition energies, oscillator strengths (*f*) were calculated. SOCs were computed within the atomic mean-field integral (AMFI) framework[121] and the states interaction CAS method (CASSI)[122] implemented in the MOLCAS package.

5 Objectives

"A goal is a dream with a deadline"
Napoleon Hill (1883-1970)
Author

The present work is aimed, in general, at determining the molecular basis of DNA damage induced by UV radiation and free electrons with low energies. In detail, the main objectives of the studies performed and reported in the articles presented in this Thesis are the following:

- To understand the photochemistry of cytosine oligomers arranged in a π-stacked orientation once they are irradiated with UV photons. Moreover, to determine the mechanisms of photo-induced production of cyclobutane cytosine (CBC) dimers via the singlet and triplet excited electronic states.

- To establish a comparison of the photo-dimerization mechanisms that take place in thymine and cytosine in the singlet manifold. Furthermore, to provide a rationale on the higher yield of CPD photoproduction observed experimentally for thymine as compared to cytosine.

- To study the excimer formation in the lowest-lying excited states of all the pyrimidine homodimers. Likewise, to show that by means of the π-stacking interaction, the formation of excimers is an intrinsic property of all pyrimidine nucleobases.

- To determine the triplet population mechanisms in isolated DNA/RNA nucleobases. Specially, to locate the regions of STC in

DNA/RNA nucleobases that are accessible along the main decay pathway of the bright state and that have significant SOC values.

- To study the photo-induced production of CPD photodimers on the triplet manifold in cytosine, thymine, and uracil homodimers.

- To make a comparison of the photoproduction mechanisms of DNA/RNA cyclobutane pyrimidine dimers in cytosine, thymine, uracil, and 5-methylcytosine via the singlet excited state.

- To study the dissociative electron-attachment (DEA) process in uracil caused by low energy electrons (0–3 eV). Expressly, to interpret the cross sections recorded experimentally for the dissociations taking place at the N_1–H and N_3–H sites of uracil.

6 Results

“However beautiful the strategy,
you should always look at the results”
Winston Churchill (1874-1965)
U.K. Prime Minister

In this chapter, we compile the seven publications that compile the results obtained in the studies carried out in the present Doctoral Thesis.

6.1 Paper I

Molecular basis of DNA photodimerization: Intrinsic production of cyclobutane cytosine dimers.

D. Roca-Sanjuán, G. Olaso-González, I. González-Ramírez, L. Serrano-Andrés and M. Merchán

J. Am. Chem. Soc., 130, 10766-10779 (2008).

JACS
ARTICLES
Published on Web 07/16/2008

Molecular Basis of DNA Photodimerization: Intrinsic Production of Cyclobutane Cytosine Dimers

Daniel Roca-Sanjuán, Gloria Olaso-González, Israel González-Ramírez, Luis Serrano-Andrés, and Manuela Merchán*

Instituto de Ciencia Molecular, Universitat de València, Apartado 22085, ES-46071 Valencia, Spain

Received April 25, 2008; E-mail: Manuela.Merchan@uv.es

Abstract: Based on CASPT2 results, the present contribution establishes for the first time that cytosine photodimer formation (C<>C) is mediated along the triplet and singlet manifold by a singlet−triplet crossing, $(T_1/S_0)_X$, and by a conical intersection, $(S_1/S_0)_{CI}$, respectively. The former can be accessed in a barrierless way from a great variety of photochemical avenues and exhibits a covalent single bond between the ethene $C_6-C_{6'}$ carbon atoms of each monomer. The efficiency of the stepwise triplet mechanism, however, would be modulated by the effectiveness of the intersystem crossing mechanism. The results provide the grounds for the understanding of the potential photogenotoxicity of endogenous and exogenous compounds via triplet−triplet sensitization, with a lower bound for cytosine oligonucleotides predicted to be 2.70 eV, and give support to the traditional view of the primary role of triplet excited states in the photochemistry of DNA, a well-known source of photoproducts in solution under triplet photosensitization conditions. The function played by singlet excimers (excited dimers) to explain both the red-shifted fluorescence and photoreaction is highlighted. A rationale on the pronounced wavelength dependence of the observed fluorescence is offered. Geometrical arrangements at the time of light irradiation close to, but energetically above, $(S_1/S_0)_{CI}$ are suggested as *reactive orientations* that become prone to produce C<>C directly, with no energy barrier. Because of the outstanding intrinsic ability of cytosine to form stable relaxed excimers, the system located near the bound relaxed excimer has to accumulate enough vibrational energy to surmount a small barrier of 0.2 eV to reach $(S_1/S_0)_{CI}$, making the overall process to proceed at a slower relative rate as compared to other compounds such as thymine, which is not susceptible of forming so stable excimers.

Introduction

Cyclobutane pyrimidine dimers (Pyr<>Pyr) formed by adjacent pyrimidine bases can be considered the most frequent lesion induced in ultraviolet (UV)-irradiated *cellular* DNA, occurring with a yield of formation 1 order of magnitude larger than pyrimidine 6-4 pyrimidone photoproducts.[1,2] Living organisms are able to reverse the photodamage by using DNA photolyase enzymatic repair, which catalyzes the cleavage of the $C_5-C_{5'}$ and $C_6-C_{6'}$ bonds of the formed Pyr<>Pyr, restoring the pyrimidines to their native state. The different DNA repair enzymes usually involve electron transfer from a catalytic cofactor to the dimer.[3] The Pyr<>Pyr photoproducts are lesions normally associated with various lethal biological responses happening at the cellular level since they inhibit DNA replication and transcription.[1,4,5] The different mechanisms proposed for Pyr<>Pyr production[1] involve singlet and triplet states of the monomers in solution and singlet states of vertical stacked nucleobases in the solid state, respectively. The efficiency of the photodimerization markedly depends on the experimental conditions (solvent, aggregation state, pH, degree of hydratation), the sequence of nucleotides, and the type (A-, B-like) of DNA conformation. It is also worth mentioning the vast amount of literature where the presence of Pyr<>Pyr photoproducts was detected.[1,6–10] It is, however, surprising that despite the importance of the matter just two recent high-level ab initio studies are available on this issue.[11,12] Those communications have independently suggested that the [2 + 2] cycloaddition photoreaction[13] for thymine (T) dimerization occurs via a barrierless concerted nonadiabatic mechanism on a singlet excited-state through a S_1/S_0 conical intersection (CI), which is the funnel for ultrafast nonradiative decay leading to T<>T.[11,12]

(1) Cadet, J.; Vigny, P. In *Bioorganic Photochemistry*; Morrison, H., Ed.; John Wiley & Sons: New York, 1990; Vol. 1, pp 1−272.
(2) Douki, T.; Cadet, J. *Biochemistry* **2001**, *40*, 2495–2501.
(3) Heelis, P. F.; Hartman, R. F.; Rose, S. D. *Chem. Soc. Rev.* **1995**, 289–297.
(4) Danilov, V. I.; Slyusarchuk, O. N.; Alderfer, J. L.; Stewart, J. J. P.; Callis, P. R. *Photochem. Photobiol.* **1994**, *59*, 125–129.
(5) Kraemer, K. H. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 11–14.
(6) Crespo-Hernández, C. E.; Cohen, B.; Hare, P. M.; Kohler, B. *Chem. Rev.* **2004**, *104*, 1977–2019.
(7) Crespo-Hernández, C. E.; Cohen, B.; Kohler, B. *Nature* **2005**, *436*, 1141–1144.
(8) Marguet, S.; Markovitsi, D. *J. Am. Chem. Soc.* **2005**, *127*, 5780–5781.
(9) Schreier, W. J.; Schrader, T. E.; Soller, F. O.; Gilch, P.; Crespo-Hernández, C. E.; Swaminathan, V. N.; Carell, T.; Zinth, W.; Kohler, B. *Science* **2007**, *315*, 625–629.
(10) Holman, M. R.; Ito, T.; Rokita, S. E. *J. Am. Chem. Soc.* **2007**, *129*, 6–7.
(11) Boggio-Pasqua, M.; Groenhof, G.; Schäfer, L. V.; Grubmüller, H.; Robb, M. A. *J. Am. Chem. Soc.* **2007**, *129*, 10996–10997.
(12) Blancafort, L.; Migani, A. *J. Am. Chem. Soc.* **2007**, *129*, 14540–14541.
(13) Klessinger, M.; Michl, J. *Excited States and Photochemistry of Organic Molecules*; VCH Publishers: New York, 1995.

10.1021/ja803068n CCC: $40.75

The major photoproduct induced in DNA by UV radiation is T<>T, but TT sites are not mutational hotspots.[2] In contrast, cytosine-cytosine (CC) sequences are sites of relatively frequent CC to TT tandem mutations, although the corresponding photoproducts (C<>C) are produced with relatively lower yields.[2] As a first step toward elucidating the distinct behavior of CC with respect to TT sites, we focus in this contribution on the characterization on theoretical grounds of the intrinsic mechanisms responsible for the production of cyclobutane cytosine (CBC) along both the triplet and singlet manifolds.

The nature of electronic excited states in base multimers depends on conformation and base sequence.[6–10,14–16] The electronic coupling between closely spaced bases is responsible for the distinct spectroscopic features such as the well-known hypochromism of the lowest-energy absorption band and the decrease of the ionization potential.[6,15] The coupling is not large enough, however, to affect significantly the absorption spectra of biopolymers. Thus, the DNA absorption spectrum closely resembles the sum of the spectra of its building blocks. For this reason, it is generally assumed that, when DNA is illuminated by UV radiation, light is initially absorbed by the nucleic acid bases leading mainly to excited states of the same multiplicity (singlet) as the respective ground state. The low-lying excited states of DNA bases that can be accessed by UV absorption lie near 5 eV.[6,15] Thanks to the great efforts of several groups during the last decades, together with the outstanding development of novel spectroscopic techniques seen more recently, it is well established by now the extremely short lifetimes, in the subpicosecond regime, of singlet excited states of nucleotides, nucleosides, and isolated purine and pyrimidine bases, suggesting the presence of fully operative ultrafast internal conversion (IC) channels.[6,9,14,17] Accordingly, the fluorescence quantum yields measured for these systems are small.[18] In the meantime, as a nice example of constructive interplay between quantum chemistry and experimental outcome, computational evidence has been able to successfully identify specific radiationless decay mechanisms in isolated DNA bases,[19–43] involving energetically accessible regions of so-called conical intersections (CIs) between the lowest excited and the ground state, which are essentially responsible for efficient internal conversion (IC).[13,44,45] These favorable IC processes represent a self-protection mechanism preventing the occurrence of chemical reactions induced by UV light. Photostability is therefore the primordial photophysical characteristic of the building blocks of DNA (as well as of their Watson−Crick canonical pairs[46]), a concomitant DNA property, crucial to understanding life on earth as we know it, that has probably evolved as the most optimal biochemical response of genetic material to sunlight exposure.

The fluorescence spectra of multimers are qualitatively different from that of the constituent nucleotides. The most striking photophysical attribute of base polynucleotides and DNA is the appearance of long-lived emissive states not found in base monomers. The red-shifted emission seen in base multimers was first termed *excimer fluorescence* by Eisinger et al.,[16] and it is observed for both the single- and double-stranded forms of polynucleotides. Eight years ago, Plessow et al.,[47] employing a novel picosecond laser approach, reported time- and wavelength-resolved fluorescence of different oligonucleotides and were able to make readily apparent the longer-decay components of the emission, that are assumed to have arisen from excimer formation. In this respect, we have recently reported that formation of bound cytosine excimers can be regarded as an intrinsic property of the cytosine dimer.[48] In particular, the computed vertical emission fully supports the excimer origin of the red-shifted fluorescence observed in cytosine-containing oligonucleotides.[16,47] A recent femtosecond excited-state absorption (fs-ESA) study of Kohler and co-workers[7] has shown that excimers are formed in high yields in a variety of synthetic DNA oligonucleotides and concludes that excited-state dynamics of A•T DNA is controlled by base stacking. On the other hand, Kwok et al.[14] have subsequently reported the first femtosecond combined time- and wavelength-resolved study on the ultraviolet-excited adenosine and a single-stranded adenine-containing oligonucleotide in aqueous solution,

(14) Kwok, W.-M.; Ma, C.; Phillips, D. L. *J. Am. Chem. Soc.* **2006**, *128*, 11894–11905.
(15) Eisinger, J.; Shulman, R. G. *Science* **1968**, *161*, 1311–1319.
(16) Eisinger, J.; Guéron, M.; Shulman, R. G.; Yamane, T. *Proc. Natl. Acad. Sci. U.S.A.* **1966**, *55*, 1015–1020.
(17) Canuel, C.; Mons, M.; Pluzzi, F.; Tardivel, B.; Dimicoli, I.; Elhanine, M. *J. Chem. Phys.* **2005**, *122*, 074316.
(18) Callis, P. R. *Annu. Rev. Phys. Chem.* **1983**, *34*, 329–357.
(19) Merchán, M.; Serrano-Andrés, L. *J. Am. Chem. Soc.* **2003**, *125*, 8108–8109.
(20) Merchán, M.; Serrano-Andrés, L.; Robb, M. A.; Blancafort, L. *J. Am. Chem. Soc.* **2005**, *127*, 1820–1825.
(21) Merchán, M.; González-Luque, R.; Climent, T.; Serrano-Andrés, L.; Rodríguez, E.; Reguero, M.; Peláez, D. *J. Phys. Chem. B* **2006**, *110*, 26471–26476.
(22) Serrano-Andrés, L.; Merchán, M.; Borin, A. C. *Chem. Eur. J.* **2006**, *12*, 6559–6571.
(23) Serrano-Andrés, L.; Merchán, M.; Borin, A. C. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, 8691–8696.
(24) Climent, T.; González-Luque, R.; Merchán, M.; Serrano-Andrés, L. *Chem. Phys. Lett.* **2007**, *441*, 327–331.
(25) Matsika, S. *J. Phys. Chem. A* **2004**, *108*, 7584–7590.
(26) Kistler, K. A.; Matsika, S. *J. Phys. Chem. A* **2007**, *111*, 2650–2661.
(27) Marian, C. M. *J. Chem. Phys.* **2005**, *122*, 104314.
(28) Tomic, K.; Tatchen, J.; Marian, C. M. *J. Phys. Chem. A* **2005**, *109*, 8410–8418.
(29) Marian, C. M. *J. Phys. Chem. A* **2007**, *111*, 1545–1553.
(30) Perun, S.; Sobolewski, A. L.; Domcke, W. *J. Am. Chem. Soc.* **2005**, *127*, 6257–6265.
(31) Perun, S.; Sobolewski, A. L.; Domcke, W. *J. Phys. Chem. A* **2006**, *110*, 13238–13244.
(32) Ismail, N.; Blancafort, M.; Olivucci, M.; Kohler, B.; Robb, M. A. *J. Am. Chem. Soc.* **2002**, *124*, 6818–6819.
(33) Blancafort, L.; Cohen, B.; Hare, P. M.; Kohler, B.; Robb, M. A. *J. Phys. Chem. A* **2005**, *109*, 4431–4436.
(34) Blancafort, L. *J. Am. Chem. Soc.* **2006**, *128*, 210–219.
(35) Chen, H.; Li, S. H. *J. Phys. Chem. A* **2005**, *109*, 8443–8446.
(36) Chen, H.; Li, S. H. *J. Chem. Phys.* **2006**, *124*, 154315.
(37) Nielsen, S. B.; Sølling, T. I. *ChemPhysChem* **2005**, *6*, 1276–1281.
(38) Gustavsson, T.; Banyasz, A.; Lazzarotto, E.; Markovitsi, D.; Scalamani, G.; Frisch, M. J.; Barone, V.; Improta, R. *J. Am. Chem. Soc.* **2006**, *128*, 607–619.
(39) Zgierski, M. Z.; Patchkovskii, S.; Lim, E. C. *J. Chem. Phys.* **2005**, *123*, 081101.
(40) Zgierski, M. Z.; Patchkovskii, S.; Fujiwara, T.; Lim, E. C. *J. Phys. Chem. A* **2005**, *109*, 9384–9387.
(41) Zgierski, M. Z.; Patchkovskii, S.; Lim, E. C. *Can. J. Chem.* **2007**, *85*, 124–134.
(42) Serrano-Pérez, J. J.; González-Luque, R.; Merchán, M.; Serrano-Andrés, L. *J. Phys. Chem. B.* **2007**, *111*, 11880–11883.
(43) Serrano-Andrés, L.; Merchán, M.; Borin, A. C. *J. Am. Chem. Soc.* **2008**, *130*, 2473–2484.
(44) Olivucci, M., Ed. *Computational Photochemistry*; Elsevier: Amsterdam, 2005.
(45) Domcke, W., Yarkony, D. R., Köppel, H., Eds. *Conical Intersections*; World Scientific: Singapore, 2004.
(46) Sobolewski, A. L.; Domcke, W.; Hättig, C. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 17903–17906.
(47) Plessow, R.; Brockhinke, A.; Eimer, W.; Kohse-Höinghaus, K. *J. Phys. Chem. B* **2000**, *104*, 3695–3704.
(48) Olaso-González, G.; Roca-Sanjuán, D.; Serrano-Andrés, L.; Merchán, M. *J. Chem. Phys.* **2006**, *125*, 231102.

providing clear evidence for the involvement of excimers in the excited relaxation pathways of adenine nucleotides.

The singlet excimer has been suggested by some authors to be a precursor to photodimerization.[1,4] In this sense, thymine dimerization has recently been determined by Schreier et al.[9] to be an ultrafast reaction along the singlet manifold, although no thymine excimers could be recorded earlier.[7] Therefore, the excited-state dimerization reaction occurs in competition with internal conversion processes to the electronic ground state. The role of triplet states in DNA chemistry, in particular on the formation of Pyr<>Pyr,[49,50] has been highlighted since it was first suggested by Cadet and co-workers.[1] Despite the fact that triplet formation has a low quantum yield, the longer-lived triplet states are crucial in the photochemistry and photophysics of DNA components, since they induce cyclobutane dimers at the bipyrimidine sites under triplet photosensitization conditions.[1,50] Another way for triplet state Pyr-formation is also possible. As it has been documented in detail for cytosine,[20] uracil,[24] and thymine,[42] the lowest triplet state can be populated along the ultrafast internal conversion by an intersystem crossing (ISC) mechanism. Apparently in contrast, recent time-resolved studies of thymine dimer formation by Marguet and Markovitsi[8] show that direct excitation of $(dT)_{20}$ leads to cyclobutane thymine dimers (T<>T) in less that 200 ns with a remarkably absence of any triplet absorption from the transient spectra of the oligonucleotide. It is clear that the origin and mechanisms of both excimer and photodimer formation at the molecular level are controversial and poorly understood.

In this scenario, a model built on the basis of predictive quantum chemical computations seems timely in order to understand the underlying basics at the molecular level of these relevant photoreactions. In particular, theory can shed light on whether excimers can be considered precursors of cyclobutane pyrimidine dimers. For this purpose, a well-established and sound quantum chemical ab initio method, namely the complete-active-space self-consistent-field second-order perturbation theory (CASPT2)[51–55] as implemented in the MOLCAS 6.0 software,[56] in conjunction with extended one-electron basis sets, was used.

Methods and Computational Details

The basis set of atomic natural orbital (ANO) type with the primitive set C,N,O(10s6p3d)/H(7s3p), the ANO-S set,[57] contracted to C,N,O[3s2p1d]/H[2s1p] was used throughout. Geometry optimizations, minimum reaction paths (MEPs), and determination of hypersurface crossings were carried out initially at the CASSCF level. The active space comprises the same π system employed earlier for ground-state cytosine.[19] It represents in the dimer a total of 16 electrons distributed among 14 molecular orbitals (MOs), that is, all the π system except for the deep π orbital localized on the NH_2 fragment of each cytosine, which was treated as inactive. Dynamic electron correlation was subsequently taken into account perturbatively at the second-order level through the CASPT2 method.[51] In order to mimic the actual interaction of pyrimidines in DNA, geometry optimizations were initially performed within the constraints of the C_s symmetry, thus allowing for an effective and natural interaction of two cytosine molecules in the biologically relevant cis−syn stereoisomer. Seven active MOs in each of the irreducible representation a′ and a″ (16 active electrons) were used in C_s symmetry. At the optimized geometries, the energies were computed at the CASPT2 level with no symmetry restrictions (C_1 symmetry), since wave function symmetry breaking is a prerequisite to describe correctly the asymptotic limit for the lowest electronic singlet and triplet transition of the two moieties. It is said that the calculation breaks symmetry when the computed electronic wave function has lower symmetry than that implied by the nuclear coordinates (see, e.g., ref 58). For the computations in C_1 symmetry, two additional π MOs were also kept inactive, since the occupation number of the corresponding natural orbitals when they were treated as active was practically 2.0. They correspond to the plus and minus linear combinations of the deeper all-in-phase π MOs of the two cytosine molecules. A CASSCF wave function of 12 active π electrons and 12 active π MOs was therefore employed, hereafter denoted as CASSCF(12,12). On the other hand, the corresponding results with the second-order corrections included shall be labeled as CASPT2(12,12). For the triplet and singlet states the state-average CASSCF(12,12) reference wave function involved two and four roots, respectively. In order to minimize weakly interacting intruder states, the imaginary level-shift technique with a parameter 0.2 au, has been employed.[59] MEPs have been built as steepest descendent paths in a procedure[60] which is based on a modification of the projected constrained optimization (PCO) algorithm of Anglada and Bofill[61] and follows the Müller−Brown approach.[62] Each step requires the minimization of the potential energy hypersurfaces on a hyperspherical cross section of the explored hypersurface centered on the initial geometry and characterized by a predefined radius. The optimized structure is taken as the center of a new hypersphere of the same radius, and the procedure is iterated until the bottom of the energy surface is reached. Mass-weighted coordinates are used, therefore the MEP coordinate corresponds to the so-called intrinsic reaction coordinate (IRC), measured in au, that is, bohr·amu$^{1/2}$. The singlet−triplet crossing $(T_1/S_0)_X$ and the conical intersection $(S_1/S_0)_{CI}$ were computed by using the restricted Lagrange multipliers technique as included in the MOLCAS-6.0 package[56] in which the lowest-energy point was obtained under the restriction of degeneracy between the two considered states.[60] The CASSCF structure computed for $(T_1/S_0)_X$ represents also a crossing at the CASPT2 level. It does not hold true for $(S_1/S_0)_{CI}$. Thus, the reported CASPT2 $(S_1/S_0)_{CI}$ conical intersection was obtained by exploring a grid of points along the distortions that involve the smallest values for the CASSCF gradients in the region of the crossing. Additional details on the computational strategy developed will be given in the next section as required.

As shown previously,[48] the inclusion of the basis set superposition error (BSSE) is crucial to accurately describe the binding energies. Here the effect was taken into account by using the

(49) Word, P. D.; Redmond, R. W. *J. Am. Chem. Soc.* **1996**, *118*, 4256–4263.
(50) Bosca, F.; Lhiaubet-Vallet, V.; Cuquerella, M. C.; Castell, J. V.; Miranda, M. A. *J. Am. Chem. Soc.* **2006**, *128*, 6318–6319.
(51) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. *J. Chem. Phys.* **1992**, *96*, 1218–1226.
(52) Serrano-Andrés, L.; Merchán, M.; Nebot-Gil, I.; Lindh, R.; Roos, B. O. *J. Chem. Phys.* **1993**, *98*, 3151–3162.
(53) Roos, B. O.; Andersson, K.; Fülscher, M. P.; Malmqvist, P.-Å.; Serrano-Andrés, L.; Pierloot, K.; Merchán, M. *Adv. Chem. Phys.* **1996**, *93*, 219–331.
(54) Merchán, M.; Serrano-Andrés, L. In *Computational Photochemistry*; Olivucci, M., Ed.; Elsevier: Amsterdam, 2005.
(55) Serrano-Andrés, L.; Merchán, M. In *Encyclopedia of Computational Chemistry*, Schleyer, P. v. R., Schreiner, P. R., Schaefer, H. F., III, Jorgensen, W. L., Thiel, W., Glen, R. C., Eds.; Wiley: Chichester, 2004.
(56) Andersson, K.; et al. *MOLCAS*, version 6.4; Department of Theoretical Chemistry, Chemical Centre, University of Lund: Lund, Sweden, 2006.
(57) Pierloot, K.; Dumez, B.; Widmark, P.-O.; Roos, B. O. *Theor. Chim. Acta* **1995**, *90*, 87–114.
(58) Merchán, M.; Pou-AméRigo, R.; Roos, B. O. *Chem. Phys. Lett.* **1996**, *252*, 405–414.
(59) Forsberg, N.; Malmqvist, P.-Å. *Chem. Phys. Lett.* **1997**, *274*, 196–204.
(60) De Vico, L.; Olivucci, M.; Lindh, R. *J. Chem. Theory Comput.* **2005**, *1*, 1029–1037.
(61) Anglada, J. M.; Bofill, J. M. *J. Comput. Chem.* **1997**, *18*, 992–1003.
(62) Müller, K.; Brown, L. D. *Theor. Chim. Acta* **1979**, *53*, 75–93.

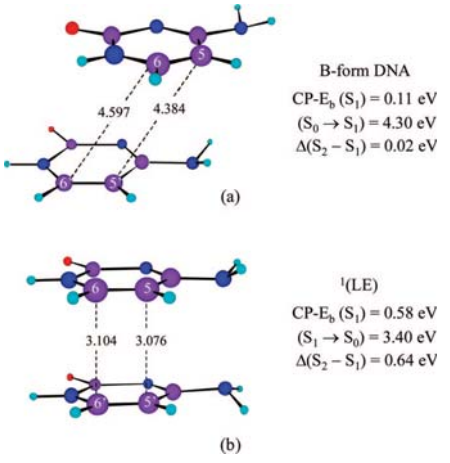


Figure 1. Structures for two cytosine molecules at the ground-state B-form DNA (a) and at the S_1 locally excited state ^{1}(LE)-type structure in a face-to-face orientation (b) (see text). The computed CASPT2 binding energies for S_1 corrected for the BSSE through the counterpoise correction, CP-E_b, the vertical absorption transition energy ($S_0 \rightarrow S_1$), the vertical fluorescence ($S_1 \rightarrow S_0$), together with the splitting for the two lowest excited states, $\Delta(S_2 - S_1)$, are also included. Interatomic distances in Å.

counterpoise correction (CP).[63] The binding energy and the corrected counterpoise binding energy will be denoted E_b and CP-E_b, respectively (see Supporting Information).

All the computations have been carried out by using the MOLCAS 6.0 quantum-chemical software.[56]

Results and Discussion

Singlet and Triplet Cytosine Excimers. Figure 1 shows two cytosine molecules in the ground-state B-type DNA conformation[64] and the structure for the locally excited-state relative minimum on S_1, hereafter ^{1}(LE), computed at the CASPT2 level with respect to the intermolecular distance in the face-to-face π-stacked arrangement of the two moieties, maintaining the monomers at the optimized ground-state CASSCF(8,7) geometry.[48]

As can be readily seen from the molecular drawings, the ground-state stacking in the B-form is somewhat different from the idealized "sandwich" geometry required for producing a fully stabilized "excimer".[65] In the ground-state B-form DNA, the interatomic distances $R(C_5-C_{5'})$ and $R(C_6-C_{6'})$ are about 1.5 and 1.3 Å larger than those in ^{1}(LE), whereas the dihedral angle $\angle C_5-C_6-C_{6'}-C_{5'}$ varies from 0° in ^{1}(LE) to 38.6° in B-DNA. *The lowest singlet excited-state in the parallel excimer ^{1}(LE) is 0.47 eV more stable than in the B-form.* Accordingly, the binding energy (CP-E_b) for the former, ^{1}(LE), increases considerably, about half an eV, with respect to the B-form. In the B-form the vertical S_1 state is weakly bound, just by 0.11 eV, whereas the S_2-S_1 splitting is small, reflecting a weak coupling between the states. As expected, the vertical transitions, 4.30 eV (S_1) and 4.32 eV (S_2), are slightly red-shifted as compared to the lowest vertical singlet−singlet transition of the cytosine monomer. At the same level of theory, the ground-state CP-E_b in the B-form is computed to be 0.04 eV, whereas it is found −0.43 eV at the ^{1}(LE)-geometry. In fact, the potential energy curve with respect to the intermolecular distance of two cytosine molecules in the face-to-face orientation has been shown for the ground-state to be repulsive.[48] The vertical emission, calculated at the ^{1}(LE) structure, 3.40 eV, is consistent with the maximum of the red-shifted fluorescence observed for dinucleotides, polynucleotides, and DNA (3.2−3.4 eV).[16,47] On the other hand, the pronounced S_2-S_1 gap for ^{1}(LE), 0.64 eV, points out to an efficient coupling between the two states. Although at the ^{1}(LE) geometry the S_2 becomes unbound by 0.06 eV, the potential energy curve for the S_2 state has a minimum at longer interatomic distances.[48]

Taking into account the inherent flexibility of DNA and related oligonucleotides, competitive ^{1}(LE)-type orientations might be present at the time of UV-irradiation. Because of the concomitant increased stability in the lowest singlet excited-state at those parallel arrangements, geometries around the ^{1}(LE)-type structure can be considered the best candidates as precursors of photodimers. It seems that the ideal twist angle between successive base pairs makes the geometry of B-DNA (and A-DNA) nonreactive. According to recent experimental evidence,[9] the static Pyr-Pyr conformations and not conformational motions after photoexcitation determines the formation of Pyr<>Pyr photoproducts. Within the model proposed by Schreier et al.,[9] the relatively smaller degree of flexibility of A-DNA compared to B-DNA to achieve the right orientations that become prone to react can be related to the greater resistance of A-DNA to Pyr<>Pyr formation. As shown by these authors, dimerization occurs only for thymine residues that are already in a reactive arrangement at the instant of excitation, because the rate of photoproduct formation by favorably oriented thymine pairs is much faster that the rate of orientation change. A similar situation can therefore be assumed in cytosine oligomers. From the results compiled so far, the ^{1}(LE)-type cytosine excimer is revealed as a reactive intermediate, a possible source of the CBC photoproduct, and consequently the ^{1}(LE) excimer has been taken as the starting point for the study of the dimerization reaction occurring along the singlet manifold (see below). It is also worth pointing out that singlet cytosine excimers may be involved in the photoinduced formation of the pyrimidine 6-4 pyrimidone photoadducts (Pyr(6-4)Pyr), whose precursor is an oxetane adduct, and other minor photohydrated products.[1] For cytosine in particular, and in contrast to cellular DNA, in the isolated system the formations of C<>C and C(6-4)C become competitive.[2] For a complete elucidation of the relative importance of those mechanisms, modeling of the environment and explicitly considering the Watson−Crick pairing should be most probably required since the yield formation depends on the type of nucleotide, the isolated or cellular nature of the DNA biopolymer, and the range of the UV radiation.[8,9,66,67] Future research shall be addressed in those directions.

A parallel study has been performed for the two lowest triplet states. The numerical results are listed in Table 1, where for the sake of completeness the related findings for the two lowest singlet excited states are also included. The potential energy curves at the CASPT2(12,12)+BSSE level are depicted in Figure 2. The corresponding BSSE-uncorrected results can be

(63) Boys, S. F.; Bernardi, F. *Mol. Phys.* **2002**, *100*, 65–73.
(64) Lu, X.-J.; Olson, W. K. *Nucleic Acids Res.* **2003**, *31*, 5108–5121.
(65) Birks, J. B. *Rep. Prog. Phys.* **1975**, *38*, 903–974.
(66) Douki, T.; Court, M.; Sauvaigo, S.; Odin, F.; Cadet, J. *J. Biol. Chem.* **2000**, *275*, 11678–11685.
(67) Mouret, S.; Baudouin, C.; Charveron, M.; Favier, A.; Cadet, J.; Douki, T. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 13765–13770.

Table 1. Binding Energy (E_b), Basis Set Superposition Error (BSSE) Obtained through the Counterpoise Method (CP-BSSE), and Corrected Binding Energy CP-E_b, Computed at the CASPT2(12,12) Level for the Lowest Triplet and Singlet Excited States of the Cytosine Excimer[a]

state	$R(C_5-C_{5'})$	E_b[b]	CP-BSSE	CP-E_b
At the $(T_1)_{min}$ Structure:				
T_1	3.029	1.15	0.98	0.17
S_0	3.029	0.41	0.97	−0.56
vertical emission[c]: 2.93 eV (=3.66−0.17−0.56)				
0−0 transition[c]: 3.49 eV (=3.66−0.17)				
At the $(T_1)_{min}$+BSSE structure: ^{3}(LE)				
T_1	3.304	1.06	0.84	0.22
S_0	3.304	0.61	0.82	−0.21
vertical emission[c]: 3.23 eV (=3.66 - 0.22 - 0.21)				
0−0 transition[c]: 3.44 eV (=3.66−0.22)				
At the $(S_1)_{min}$ Structure:[d]				
S_1	2.954	1.51	0.97	0.54
S_0	2.954	0.29	0.97	−0.68
vertical emission[c] S_0 S_1: 3.19 eV[d] (=4.41−0.54−0.68)				
0−0 transition[c]: 3.87 eV (=4.41−0.54)				
At the $(S_1)_{min}$+BSSE Structure: ^{1}(LE)				
S_1	3.076[d]	1.47	0.89	0.58[d]
S_0	3.076	0.47	0.90	−0.43
vertical emission[c]: 3.40 eV[d] (=4.41−0.58−0.43)				
0−0 transition[c]: 3.83 eV (=4.41−0.58)				

[a] Distances in Å and energies in eV. The triplet and singlet locally excited states ^{3}(LE) and ^{1}(LE) discussed in the text are identified. [b] Binding energy with respect to the corresponding states of the monomers. [c] Vertical excitations in the monomer: $S_0 \rightarrow T_1$ (3.66 eV) and $S_0 \rightarrow S_1$ (4.41 eV). [d] Taken from Olaso-González et al.[48]

found in the Supporting Information. The accurate theoretical treatment of these singlet and triplet excimers becomes particularly challenging since it requires inclusion of electron dynamic correlation, flexible enough one-electron basis sets, wave functions with no symmetry constraints in order to achieve the correct asymptotic limit, and BSSE corrections. The influence of the latter factor is apparent by inspection and comparison of the two sets of results displayed in Figure 2 and in the Supporting Information.

At the highest level of theory employed, the binding energy for the lowest triplet state (T_1) computed at the CASPT2(12,12)+BSSE level is 0.22 eV, with a predicted vertical emission (phosphorescence) of 3.23 eV and a 0−0 triplet−singlet transition of 3.44 eV. Consequently, it is concluded that the triplet excimer is bound, although the binding energy is reduced about 60% with respect to the lowest singlet excimer. Interestingly, as can be noted from Figure 2, the S_1 and T_2 states are involved in a singlet−triplet crossing around the intermolecular distance 3.0−3.4 Å, precisely the distances expected for the ground-state biopolymer.[6,15] Thus, T_2 could be populated through an ISC mechanism, becoming then deactivated toward T_1 via a CI facilitated with the breathing movement of the own DNA. Apart from the near-degeneracy found, the magnitude of the spin−orbit coupling (SOC) is also relevant to asses the efficiency of the ISC process, which would be strongly affected by the actual environment of the biopolymer. The possibility of excimer formation arises from the Watson−Crick structure in which hydrogen-bonded pairs A-T and G-C are situated inside a double helix, the backbone formed by two sugar-phosphate chains. One turn of the helix involves 10 base pairs and is 34 Å high. Thus, the interplanar distance between neighboring base pairs is about 3.4 Å, a value which is often found in excimer-type organic crystals.[65,68] The structure for the locally excited-state relative minimum on T_1 computed at the CASPT2(12,12)+BSSE level with respect to the intermolecular distance in the face-to-face arrangement (maintaining the monomers at the optimized ground-state geometry) shall be accordingly denoted as ^{3}(LE). As compiled in Table 1, the ^{3}(LE) structure has a $R(C_5-C_{5'})$ distance of 3.304 Å. Similarly, as in the case of the singlet manifold, the ^{3}(LE)-type cytosine excimer has been taken as the starting structure to study the intrinsic reactivity of the system along the triplet manifold.

To summarize, from the results compiled so far, the ^{1}(LE)- and the ^{3}(LE)-type cytosine excimers are revealed as promising, local, two-by-two reactive intermediates easily accessible in DNA single- and double strands that may be the source of the CBC dimer photoproducts, and consequently they have been taken as the starting point for the study of the photoinduced dimerization reaction occurring along the singlet and triplet manifold, respectively. Of course, the LE-type structures might not be the unique arrangements susceptible of initiating photoreactions in such complex hypersurfaces, but a priori they are clear candidates and bear the requirements to produce C<>C, and we shall therefore focus next on how they evolve. In fact, as shall be discussed below, the T_1-MEP and S_1-MEP computations from the ^{3}(LE) and ^{1}(LE) structures lead to stationary points on the respective hypersurfaces close to regions where an efficient radiationless decay toward the ground state can take place.

Photodimerization of Cytosine along the Triplet Manifold. A detailed analysis of the CASSCF wave functions for the LE states, that is, the T_1 and S_1 states at the LE geometries, shows that both the triplet ^{3}(LE) and the singlet ^{1}(LE) excited states can be described as belonging to the irreducible representation ^{3,1}A″, as appropriate, of the C_s symmetry group. The CASPT2 energies for the states ^{3}(LE) and ^{1}(LE) computed within the geometrical constraints of C_s and with no restrictions (C_1 symmetry) are actually the same. In other words, no symmetry breaking of the respective wave functions is made apparent at the geometries of the ^{3}(LE) and ^{1}(LE) states as it happens markedly when going to longer intermolecular distances of the two cytosine moieties. In fact, only working in C_1 symmetry the degeneracy between the two lowest triplet (T_2, T_1) and singlet (S_2, S_1) states can be achieved at the asymptotic limit (cf. Figure 2). Inspired by this finding, CASSCF geometry optimizations and MEP computations for the lowest triplet and singlet states were carried out in practice for the ^{3,1}A″ states (C_s symmetry). In order to make sure that a balanced treatment is given in all the explored regions of the hypersurfaces considered, the CASPT2 results reported here correspond always to computations carried out in C_1 symmetry. Furthermore, consistency with the results presented for the excimers in the previous section requires also inclusion of the BSSE correction. Therefore, BSSE has been introduced in the calculation of the relative stability between the different structures discussed. This is an important remark to bear in mind, because as the structural difference between two given stationary points becomes larger, the larger the difference between the computed relative stability with or without the BSSE correction would be. The procedure is certainly somewhat cumbersome, but it has the advantage that the results obtained for the states are consistent and directly comparable among them at the distinct geometries, from CBC to infinity separation of the two cytosine molecules.

(68) Klöpffer, W. In *Organic Molecular Photophysics*; Birks, J. B., Ed.; Interscience: London, 1973; pp 357−402.

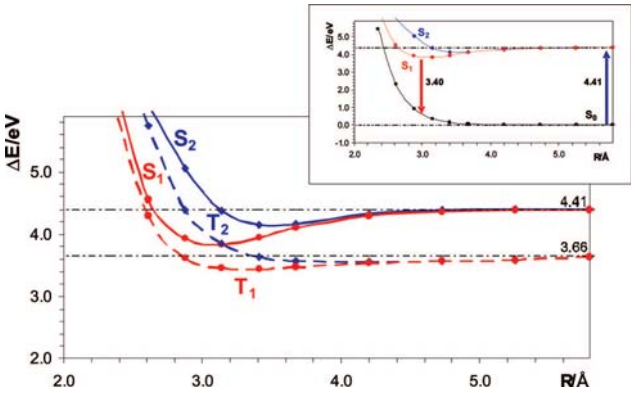


Figure 2. CASPT2+BSSE potential energy curves built with respect to the intermolecular distance $R(C_5-C_{5'})$ of two face-to-face π-stacked cytosine molecules involving the ground and the lowest two triplet and two singlet excited states. The inset, obtained at the same level of theory,[48] illustrates the emission event related to the observed red-shifted fluorescence in cytosine oligonucleotides.[16,47]

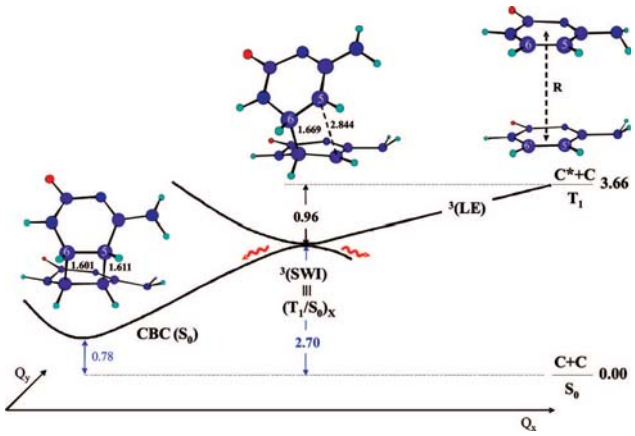


Figure 3. Proposed scheme, based on actual CASPT2 results, for the decay path of the lowest triplet excited state (T_1) of the cytosine dimer through the triplet locally excited state ^{3}(LE) and the stepwise intermediate ^{3}(SWI) leading to ground-state cyclobutane cytosine (CBC) via an intersystem crossing mechanism (ISC). The main intermolecular geometric parameters (in Å) are included. The remaining numerical values (in eV) correspond to relative energies, as indicated, with respect to two ground-state cystosine molecules separated by $R \cong 10.5$ Å. At the ^{3}(SWI)-optimized structure a singlet−triplet crossing, $(T_1/S_0)_X$, takes place. The Q_x coordinate is mainly related to the average intermolecular distance of $R(C_5-C_{5'})$ and $R(C_6-C_{6'})$, whereas Q_y is associated to the remaining degrees of freedom.

The overall proposed photodimerization decay path taking place along the triplet manifold is shown schematically in Figure 3. On the right-hand side of Figure 3, C + C denote two ground-state cytosine molecules separated by about 10.5 Å, both at the equilibrium geometry of ground-state cytosine. In C* + C one monomer (C*) is electronically excited. Transition from the ground to the lowest excited-state at this intermolecular distance thus corresponds to the vertical transitions $S_0 \rightarrow T_1$, 3.66 eV, in agreement with earlier results for isolated cytosine.[20] The CASPT2 results for the MEP computation along the T_1 state starting from the ^{3}(LE) state are depicted in Figure 4, where the evolution of the S_0 and T_2 states is also included. Remarkably, the T_1-MEP leads directly in a barrierless fashion to a structure coincident to that obtained by optimizing the geometry, denoted as a triplet stepwise intermediate, ^{3}(SWI), which has a clear degeneracy with the ground state. In other words, the triplet state is coincident with a triplet−singlet crossing, $(T_1/S_0)_X$, a region of the hypersurface where decay to the ground state becomes particularly favored.

The stationary point ^{3}(SWI) is characterized by the formation of a single covalent bond between the $C_6-C_{6'}$ atoms computed to be 1.669 Å, whereas the $C_5-C_{5'}$ interatomic distance stays elongated, about 2.8 Å (see Figure 3). Thus, ^{3}(SWI) cannot be really considered an excimer but an intermediate toward the formation of CBC.

Starting from an ample range of parallel arrangements around the ^{3}(LE) structure, geometry optimization of the ^{3}A″ state leads unambiguously to the ^{3}(SWI) structure, pointing out to the presence of many related reactive orientations. Of course, if the symmetry restrictions are released and initially the two cytosine molecules are considerably far apart, at about an intermolecular distance of 7 Å or larger, a local relaxed triplet in one of the

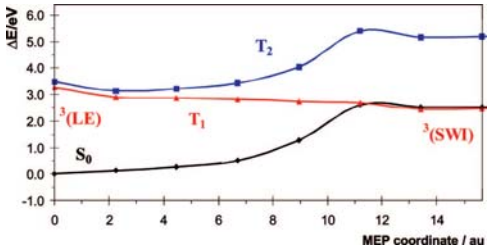


Figure 4. Low-lying triplet excited states of the cytosine dimer computed at the CASPT2//CASSCF level along the minimum energy path (MEP) of the T_1 state from the geometry of the triplet locally excited state ^{3}(LE). The T_1-MEP ends at the stepwise intermediate ^{3}(SWI) (see Figure 3), and it is isoenergetic with the ground state (S_0).

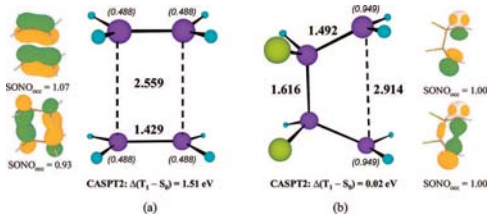


Figure 5. CASSCF(4,4) optimized geometry for the lowest triplet state of the ethene dimer (a) and the fluorethene dimer (b). The occupation number of the corresponding CASSCF singly occupied natural orbital ($SONO_{occ}$) is also included, as well as the CASPT2(4,4) energy difference computed between the lowest triplet and ground state. Main spin population derived from Mulliken analysis is given within parentheses. Distances in Å.

monomers is obtained, whereas the other maintains the ground-state equilibrium structure. As documented in detailed elsewhere,[20] the lowest triplet state of the monomer has $\pi\pi^*$ character and a band origin of 3.04 eV, which represents an stabilization of about 0.6 eV with respect to the vertical singlet−triplet transition.

In order to get further insight into the source and nature of the species ^{3}(SWI), comparable results for the lowest triplet state of ethene and fluorethene dimers (ED and FED hereafter) are helpful and illustrative. The T_1 state for ED and FED has been optimized at the CASSCF(4,4) level employing as active space the π valence molecular orbitals (MOs) (four electrons active) and the same ANO-type basis set as for cytosine. As can be readily seen from Figure 5, the equilibrium geometry for the lowest T_1 state of the FED molecule resembles to that of ^{3}(SWI), whereas for ED the structure of lowest T_1 state (HOMO−LUMO type) is rectangular. Furthermore, the single bond formed in T_1 of FED involves the intermolecular substituted carbon atoms, and the spin density is localized on the remaining carbon atoms. The T_1-FED equilibrium structure corresponds clearly to a biradical, and not surprisingly, it is degenerate with the ground state, both at the CASSCF and CASPT2 levels. On the contrary, for ED the T_1 state is about 1.5 eV above S_0, having the spin density of the two unpaired electrons equally shared by the four carbon atoms. Interestingly, according to the nature of the states, the optimized geometries for T_1-ED and T_1-FED are actual minima. A similar situation can be expected for ^{3}(SWI) in the cytosine dimer.

The singlet−triplet degeneracy occurring at the equilibrium structure ^{3}(SWI) can be therefore understood on the basis of the biradical character of the triplet/singlet states, having the unpaired electrons located on the C_5 and $C_{5'}$ atoms, yielding also a rationalization for the relatively longer $C_5-C_{5'}$ bond distance with respect to that calculated for $C_6-C_{6'}$. Taking into consideration that the C_6 atom (and $C_{6'}$) is attached to a nitrogen atom, whereas C_5 (and $C_{5'}$) is linked to a carbon atom, the C_6 and $C_{6'}$ centers are relatively more electron deficient than the C_5 and $C_{5'}$ atoms, since the electron affinity of nitrogen is higher than that of carbon. The single bond-making process between C_6 and $C_{6'}$ is therefore favored. As it occurs in FED, the substituted carbon atoms compensate for the lack of electron density by getting closer to the other monomer, causing in their quest for restoring the missed electron density the birth of a new bond.

The ^{3}(SWI) state has a CP-E_b of 0.96 eV and lies 2.70 eV above two ground-state cytosine molecules, which represents a stabilization of 0.3 eV with respect to the adiabatic triplet state of isolated cytosine. The 2.70 eV energy can be considered a lower bound for the triplet energy of cytosine in DNA. It can be envisaged that exogenous photosensitizers would populate the relaxed triplet state of the monomer, which may interact attractively with ground-state cytosine, and the so-formed dimer would evolve toward ^{3}(SWI), precursor of CBC. The possibility of endogenous photosensitization by triplet energy transfer in DNA between different nucleobases has also to be kept in mind. In either case, the required energy to access the ^{3}(SWI) state can be related to the threshold observed experimentally in a given compound to become a potential DNA photodamager via C<>C formation. The computed result for the ^{3}(SWI) state of cytosine, 2.70 eV, is consistent with the triplet energy of thymine in DNA deduced experimentally, 2.80 eV.[50] Triplet−triplet energy transfer is very important and common in chemical reactions.[13] It is utilized to specifically populate the triplet state of the reactant. This process is referred to as photosensitization, and the donor is called a triplet sensitizer. It is important to recall that, for efficient energy transfer to take place, the donor must absorb in the region of interest, undergo efficient ISC, and have a triplet-state energy higher than that of the acceptor, that is in this case, the ^{3}(SWI) species. The present results offer a nice rationale to the known fact that pyrimidine dimers are formed in solution under triplet photosensitization conditions.[1]

The computed intermediate ^{3}(SWI) thus represents a channel for photodimer formation from the triplet state of π-stacked cytosine in DNA and provides the basic understanding of potential phototogenotoxicity via triplet−triplet sensitization. It is known that UVA radiation (3.10−3.87 eV) preferentially induces the production of cyclobutane dimers at TT sites without any detectable formation of Pyr(6-4)Pyr photoproducts. Direct singlet population cannot take place at such low irradiation energies, and the process should then proceed through a triplet-energy transfer photosentization mechanism.[67] As preliminary calculations support, a similar pathway for the formation of the ^{3}(SWI) intermediate can be expected to take place for dimers including thymine. At higher irradiation energies, UVB or UVC, the singlet excited states of the pyrimidine molecule are accessible, and the mechanism of intrinsic population of the lowest triplet state of the monomer becomes operative. Since the efficiency of triplet-state formation has been determined to be larger in isolated thymine than in cytosine,[20,42] it is possible to rationalize the preference of the TT sites, and to a lesser extent CT, TC, and CC, to generate the cyclobutane photoproduct.

The presence of the $(T_1/S_0)_X$ crossing clearly favors the ISC to the ground state, but the actual efficiency of the decay process

along the triplet manifold will also rely on the enhancement of the SOC, estimated to be just a few cm^{-1} at the in-vacuo ^{3}(SWI) structure. In this respect, the solvent (or, in general, the environment) is expected to play a crucial role in the ISC process.[69] Under favorable ISC conditions, the decay would most probably take place on a subpicosecond range, which is considerably less than the 200 ns employed in the time-resolved study of thymine dimer formation.[8] We support therefore the suggestion made by Marguet and Markovitsi[8] in relation to the possibility that the ultrafast reactivity of the triplet state to yield cyclobutane dimers occurs with quasi-unit efficiency.

Photodimerization of Cytosine along the Singlet Manifold. Considering the seminal work reported by Bernardi, Olivucci, and Robb in 1990 on predicting forbidden and allowed excited-state [2 + 2] cycloadditon reactions of two ethene molecules, decay toward ground-state CBC from the singlet excimer should take place via a CI.[70,71] The ethene−ethene photochemical cycloaddition reaction path has been documented to occur in a single step via a rhomboid ($C_{2\,h}$ symmetry) conical intersection,[72] and a similar situation is expected to occur here as it has already been documented for the thymine dimer.[11,12] Imposing C_s symmetry in the study represents therefore a clear approximation in this case, since in principle the rhomboid arrangement cannot be reached. The cut along the singlet reaction path within C_s symmetry would lead to a region of the S_1 surface used to be referred as "pericyclic minimum" because it was thought that, in general, the touching between the states was likely weakly avoided everywhere, even without symmetry constraints.[13] Thanks to the extensive work performed by these authors and many others,[44,45] the exact location of the bottom of the respective funnel has been determined in considerable detail for many different processes and has been demonstrated to be of fundamental significance in many photoreactions. In this respect, it should be emphasized that exploring the excited-state potential energy surfaces by using geometries from the ground-state path might lead to meaningless results, and consequently, the model should not be employed. Currently, the importance of the crucial role played by CIs in modern photochemistry is well established theoretically and supported experimentally.

If one is able to connect the different relevant stationary points to CBC within C_s symmetry constraints, the geometry around the touching region would be then a good guess for the computation of the actual CI (C_1 symmetry). The computational strategy has been found quite advantageous, especially taking into consideration that the present study is addressed to get further insight into the intrinsic possible photodimerization paths along the singlet manifold occurring in oligonucleotides and DNA. The C_s symmetry constraints somehow simulate for two cytosine monomers the restrictions that those would have in the actual biopolymer. In order to be fully aware of the complexity of the study, it is worth mentioning that unconstrained geometry optimization of two stacked cytosine molecules leads for the lowest singlet states to structures with orientations unlikely to occur in the single and double strands of DNA. Those unconstrained structures are solely appropriate to be compared to the gas-phase data, a target that is out of the scope of the current research and shall not be further discussed. Past work has overcome this problem by keeping fixed the $C_5-C_{5'}$ and $C_6-C_{6'}$ bonds and the corresponding dihedral angle to those data known for a B-DNA form.[11,12,73,74] For the reasons indicated above, we have first preferred mapping the full photoreaction within C_s symmetry, which nicely gives the basic clues of the nonradiative decay path taking place along the S_1 hypersurface toward the ground state.

On the basis of CASPT2 results, the overall photoinduced production of CBC along the singlet manifold is schematically displayed in Figure 6. On the right-hand side of the drawing, the transition $S_0 \rightarrow S_1$ at 4.41 eV corresponds to that of the isolated cytosine molecule. The related oscillator strength for the singlet−singlet absorption is 0.09. The findings are consistent with earlier results for the system at the same level of theory.[19,21,48] Using a similar notation as in Figure 3, C + C represent two ground-state cytosine molecules separated at about 10.5 Å, both at the equilibrium geometry of ground-state cytosine, whereas in C* + C one monomer (C*) is in the lowest electronically excited (singlet, in Figure 6) state. As shown in Figure 7, MEP computations along the singlet excited state starting from the ^{1}(LE) excimer leads in a barrierless form to a relaxed excimer, hereafter labeled like ^{1}(C*C). In the MEP ^{1}(LE) - ^{1}(C*C) the energy of the ground-state is not significantly affected (cf. Figure 7). Recall at this point that CASSCF MEP geometry optimizations were carried out within C_s symmetry for the $1\,^1A''$ state, and punctual CASPT2 calculations at the converged structures were obtained from averaging the lowest four roots with no symmetry constraints. As it is clearly verified from Figure 8, the relaxed excimer ^{1}(C*C) connects with the lowest singlet excited-state of the photodimer CBC, and it is apparent that from ground-state CBC the S_1-MEP evolves in a steepest descendent fashion toward the relaxed structure ^{1}(C*C). For the sake of completeness, results for the S_2 and S_3 states are also included in Figures 7 and 8. For CBC, the lowest vertical transition, computed to be at 4.57 eV with oscillator strength 0.070, is expected to contribute to the photoreversibility process observed in pyrimidines.[10] The energy of the ground state (S_0) along the S_1-MEP of Figure 8 increases progressively and, after reaching a maximum, decreases. Around the proximity of the MEP coordinate 4.5 au where the maximum for S_0 is accomplished, the S_1 and S_0 states are placed relatively close, being the corresponding geometry a good candidate as initial trial to search for a crossing between the respective singlet hypersurfaces. In this manner, the stationary point for the actual conical intersection $(S_1/S_0)_{CI}$ displayed in Figure 6 was obtained in C_1 symmetry. The funnel $(S_1/S_0)_{CI}$ provides the 2-fold channel for ultrafast internal conversion toward the ground-state of the system, that is, to CBC and to the separated monomers, although just the former has been emphasized in Figure 6.

The relaxed excimer ^{1}(C*C) has a binding energy (CP-E_b) of 1.10 eV, and it is stabilized with respect to ^{1}(LE) by about ~0.5 eV. In order to reach the conical intersection $(S_1/S_0)_{CI}$ from ^{1}(C*C), the system has to surmount a barrier height of 0.2 eV. Because of the pronounced CP-E_b and the presence of a barrier, the relaxed species ^{1}(C*C) can be expected to be fluorescent. The predicted fluorescence features from ^{1}(LE) and ^{1}(C*C), 3.40 and 2.76 eV, respectively, may help to rationalize

(69) Hare, P. M.; Crespo-Hernández, C. E.; Kohler, B. *J. Phys. Chem. B* **2006**, *110*, 18641–18650.
(70) Bernardi, F.; De, S.; Olivucci, M.; Robb, M. A. *J. Am. Chem. Soc.* **1990**, *112*, 1737–1744.
(71) Bernardi, F.; Olivucci, M.; Robb, M. A. *Acc. Chem. Res.* **1990**, *23*, 405–412.
(72) Celani, P.; Robb, M. A.; Garavelli, M.; Bernardi, F.; Olivucci, M. *Chem. Phys. Lett.* **1995**, *243*, 1–8.
(73) Durbeej, B.; Eriksson, L. A. *Photochem. Photobiol.* **2003**, *78*, 159–167.
(74) Zhang, R. B.; Eriksson, L. A. *J. Chem. Phys. B* **2006**, *110*, 7556–7562.

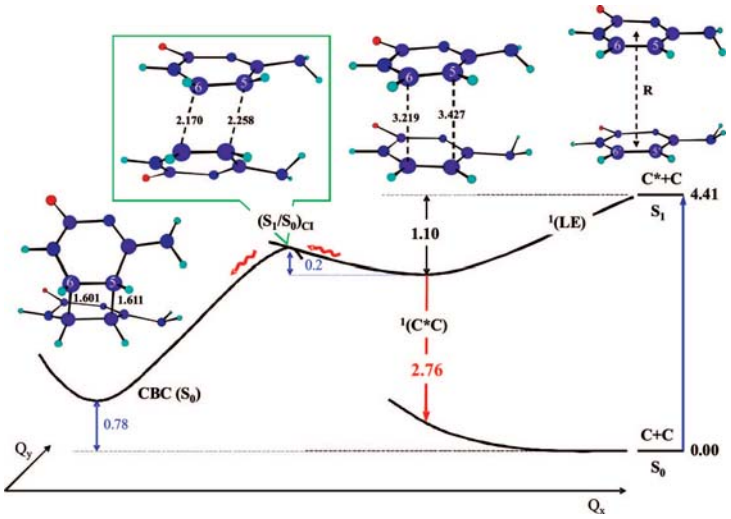


Figure 6. CASPT2 results for the computed decay path of the lowest singlet excited state (S_1) of the cytosine dimer involving the relaxed excimer ^{1}(C*C) through the conical intersection $(S_1/S_0)_{CI}$ leading to ground-state cyclobutane cytosine (CBC). The photoreaction competes with the fluorescence from ^{1}(C*C). The main intermolecular geometric parameters (in Å) are included. The remaining numerical values (in eV) correspond to relative energies, as indicated, with respect to two ground-state cystosine molecules separated $R \cong 10.5$ Å. The Q_x coordinate is mainly related to the average intermolecular distance of $R(C_5-C_{5'})$ and $R(C_6-C_{6'})$, whereas Q_y is associated with the remaining degrees of freedom.

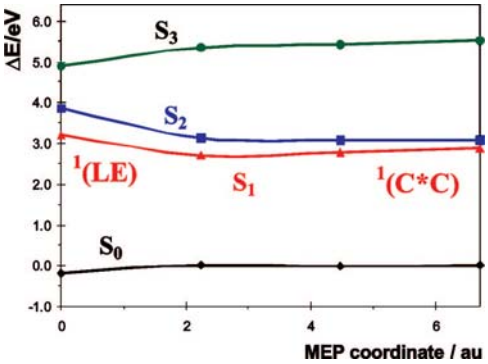


Figure 7. Low-lying singlet excited states of the cytosine dimer computed at the CASPT2//CASSCF level along the minimum energy path (MEP) of the S_1 state from the geometry of the singlet locally excited state ^{1}(LE). The S_1-MEP ends at the relaxed excimer ^{1}(C*C) (see Figure 6).

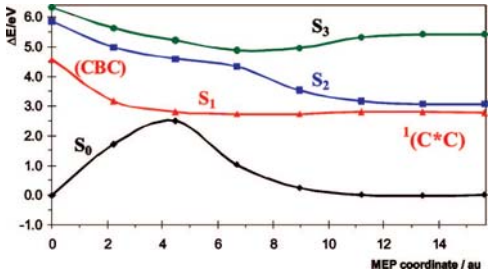


Figure 8. Low-lying singlet excited states of the cytosine dimer computed at the CASPT2//CASSCF level along the minimum energy path (MEP) of the S_1 state from the cyclobutane cytosine (CBC) dimer at its ground-state equilibrium geometry. The S_1-MEP ends at the relaxed excimer ^{1}(C*C) (see Figure 6).

the dramatic wavelength dependence seen by using fluorescence polarization measurements of the pure excimer fluorescence observed for CpC in neutral ethylene glycol/water glasses at low temperatures.[75,76] Thus, when the excitation wavelength (λ_e) used is 250 nm (4.96 eV), the recorded fluorescence wavelength (λ_f) becomes 310 nm (4.00 eV), whereas at $\lambda_e = 300$ nm (4.13 eV) the observed emission is $\lambda_f = 460$ nm (2.70 eV). Even if the results computed in vacuo are not strictly comparable to those data derived from glycol/water glasses, it is tempting to anticipate qualitatively that when λ_e is shorter (case A) or longer (case B) than the absorption maximum (λ_{max}) of the monomer, computed at 4.41 eV, the fluorescence observed for CpC could be related to the ^{1}(LE)-type excimer or to a relaxed ^{1}(C*C)-type excimer, respectively. The predicted fluorescence from the relaxed excimer ^{1}(C*C), 2.76 eV, is indeed consistent with case B (2.70 eV).[75,76] The observed variation in the polarization character could then be attributed basically to the different nature of the systems responsible of the fluorescent features in the unrelaxed (case A) and relaxed (case B) excimers.

The presence of a conical intersection as the main actor in the photoinduced production of bipyrimidine lesions along the singlet manifold is intrinsic to any ethene-like dimer system such as the cytosine dimer. For this reason, examination of the results for the crossing $(S_1/S_0)_{CI}$ in the ethene dimer itself, employing the same tools and computational strategies as described above, is illustrative. Figure 9 shows the $(S_1/S_0)_{CI}$

(75) Callis, P. R. *Chem. Phys. Lett.* **1973**, *19*, 551–555.
(76) Wilson, R. W.; Callis, P. R. *J. Phys. Chem.* **1976**, *80*, 2280–2288.

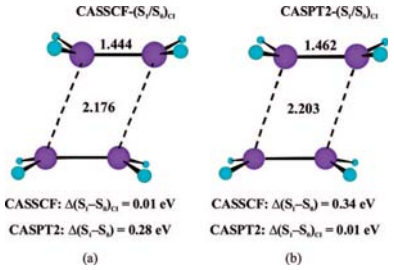


Figure 9. Conical intersection $(S_1/S_0)_{CI}$ of the ethene dimer and energy difference between the implied states at the different levels of theory (a) at the CASSCF(4,4) optimized geometry, and (b) at the CASPT2(4,4) optimized geometry. Distances in Å.

optimized geometry computed at the CASSCF(4,4) and CASPT2(4,4) levels of theory by using analytical and numerical gradients, respectively. The $(S_1/S_0)_{CI}$ structures are of rhomboid type and agree with previous results based on large-scale MRCI and CASPT2 computations.[77,78] As can be readily seen in Figure 9, dynamic electron correlation accounted for at the CASPT2 level tends to increase the intra- and interbond distances. On the other hand, at the optimized geometry for the CASSCF crossing, the energy difference between the two states becomes 0.28 eV at the CASPT2 level, whereas at the conical intersection computed at the CASPT2 level, the CASSCF energy difference is 0.34 eV. A similar situation happens for the cytosine dimer, where the CASSCF and CASPT2 conical intersections are placed in different regions of the hypersurface. The differential electron correlation between S_0 and S_1 at the region of crossing is relevant both qualitatively and quantitatively. While $(S_1/S_0)_{CI}$ is below the relaxed excimer ^{1}(C*C) at the CASSCF level, the former is separated by a barrier height of 0.2 eV from the latter at the CASPT2 level. In other words, the pathway is predicted barrierless by using the CASSCF approach. In contrast the more advanced CASPT2 method offers a completely different picture (cf. Figure 6). The ethenic bond lengths (C_5-C_6) and ($C_{5'}$-$C_{6'}$) at the $(S_1/S_0)_{CI}$ geometry, 1.471 Å and 1.422 Å, respectively, are elongated with respect to the relaxed ^{1}(C*C) excimer, 1.399 Å, and shorter than those in CBC, 1.539 Å. In the meantime, a progressive decrease of the intramonomer bond distances $C_5-C_{5'}$ and $C_6-C_{6'}$ takes place from ^{1}(C*C) to CBC passing through $(S_1/S_0)_{CI}$. These structural changes reflect the conversion process involving the transformation of two carbon−carbon double bonds into single bonds, together with the simultaneous formation of two new single carbon−carbon bonds.

While the photoinduced [2 + 2] cycloaddition of two stacked cytosine molecules proceeds through a stepwise mechanism in the triplet manifold, the photoreaction occurs via a concerted mechanism on the lowest singlet excited state. The former is mediated by a singlet−triplet crossing and the latter takes place through a singlet−singlet conical intersection, which are the funnels for the ultrafast nonradiative decay leading to CBC. The singlet−triplet crossing can be accessed barrierless, but the efficiency of the process relies on the effectiveness of the ISC mechanism. On the other hand, a small barrier (0.2 eV) has to be overcome along the studied pathway in the singlet manifold to reach the conical intersection, but there might be many different orientations in the vicinity of the crossing $(S_1/S_0)_{CI}$ prone to react directly with no barrier. Figure 6 just shows that from the most plausible reactive species, a relaxed excimer characterized by a large binding energy, a small barrier to reach the funnel does exist. Figure 10 displays a scheme of CBC formation through both singlet and triplet manifold. Within the present context, a key question rises: Why is thymine more reactive than cytosine? Because of the methyl group in thymine, the corresponding thymine excimers are expected to be less stable than those of cytosine. In this respect, thymine can be considered more reactive toward the formation of photoinduced dimers than cytosine because the lack of competitive stable thymine excimers. The present picture is supported by independent experimental research. On the one hand, excimer contributions to the total fluorescence yield were found less prominent in $(dT)_{15}$ than in $(dC)_{15}$.[47] On the other hand, the transient absorption signals at 570 nm from $(dT)_{18}$ and its 5′-mononucleotide TMP agree within experimental uncertainty, that is, formation of thymine excimers was not observed.[7] Both facts point out to the absence of thymine excimers. Consequently, the decay of the thymine-base monomer via internal conversion becomes the main deactivation route, whereas for cytosine oligomers it competes with the formation of stable excimers bearing well characterized photophysical properties.[16,47] In addition, as we have shown here, cytosine excimers can also be considered as precursors of mutagenic photoproducts. Thus, when a system is able to form stable singlet excimers placed at energies below the crossing with the ground state, as it is the case of cytosine, formation of photoproducts along the singlet mechanism should proceed at a slower relative rate for these orientations, because of the presence of a barrier to be overcome, making the full process less efficient (see Figure 10). Furthermore, taking into consideration that a very recent study based on femtosecond time-resolved infrared spectroscopy shows that thymine dimers are fully formed around 1 ps after UV excitation,[9] it is tempting to propose that the intrinsic ultrafast photoreactivity of the thymine dimer is partially due to the absence of stable thymine excimers, since, in practice, basically all the relative orientations of the two monomers close to the region where the corresponding conical intersection[11,12] takes place become potentially *reactive*. In accordance with the fact that no stable excimers are formed, the initial *reactive orientations* (shadowed regions in Figure 10) are expected to reach the funnel in a barrierless mode because they are at higher energies than the conical intersection $(S_1/S_0)_{CI}$. Nevertheless, the accessibility of the area around the funnel would be ultimately dictated by the inherent flexibility of the biopolymer under consideration.

Conclusions

The present contribution emphasizes the importance of excimers to understand not only the distinct photophysics of oligonucleotides, as well as DNA itself, but also to rationalize the intrinsic and distinct photoinduced reactivity of cytosine and thymine toward the formation of cyclobutane dimers, recognized as one of the most common processes leading to DNA damage under UV irradiation. Taking as starting structures the locally excited ^{3}(LE) and ^{1}(LE) excimers, which have a face-to-face arrangement at the ground-state geometry of the cytosine monomer, it is shown on the basis of high-level ab initio computations that a stepwise intermediate, ^{3}(SWI), is produced with no barrier in the triplet manifold, whereas the barrierless

(77) Serrano-Andrés, L.; Merchán, M.; Lindh, R. *J. Chem. Phys.* **2005**, *122*, 104107.
(78) Dallos, M.; Lischka, H.; Shepard, R.; Yarkony, D. R.; Szalay, P. G. *J. Chem. Phys.* **2004**, *120*, 7330–7339.

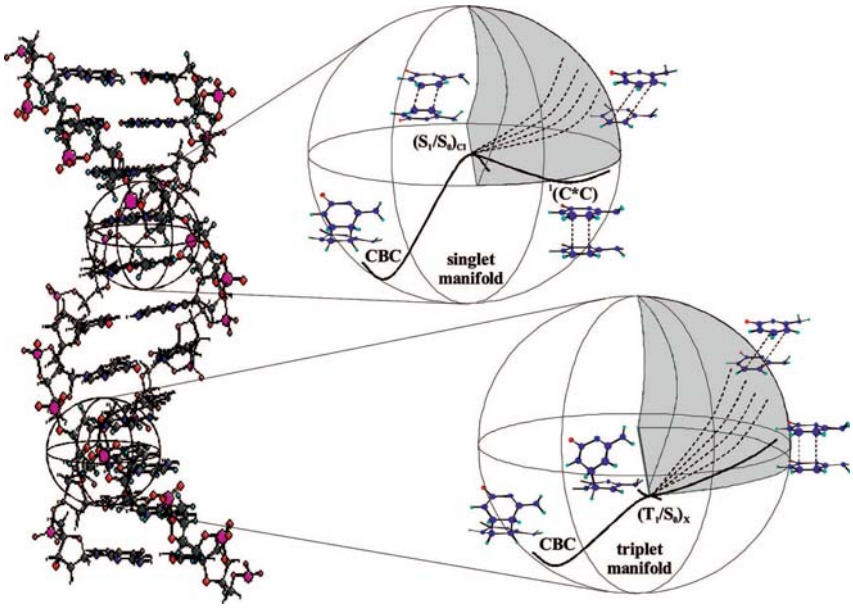


Figure 10. Scheme of the photodimerization process of the π-stacked cytosine dimer along the singlet and the triplet manifold. The shadowed volumes in the sphere represent regions of the DNA strand with reactive orientations, in which the decay path lies above the singlet−triplet crossing $(T_1/S_0)_X$ and the conical intersection $(S_1/S_0)_{CI}$.

pathway along the lowest singlet hypersurface yields the relaxed excimer ^{1}(C*C). The structure ^{3}(SWI) has a covalent single bond between the carbon atoms $C_6-C_{6'}$, with the $C_5-C_{5'}$ bond length kept long. The required energy to reach ^{3}(SWI) from two isolated ground-state cytosine monomers, 2.70 eV, is related to the threshold observed experimentally for a given photosensitizer to become a potential DNA photodamager, and it is in agreement with the triplet energy of thymine in DNA deduced experimentally, 2.80 eV.[50] At the ^{3}(SWI) structure a singlet−triplet crossing $(T_1/S_0)_X$ takes place, which mediates the nonradiative deactivation toward the ground state by an intersystem crossing mechanism (ISC). Under favorable conditions for spin−orbit coupling (SOC) and according to the barrierless profile computed, the decay is predicted to take place on a subpicosecond range. Because the ^{3}(SWI) intermediate is obtained from many different initial structures (see Figure 10), the findings can also be applied in situations where no excimers are expected to be present. Thus, the results offer a nice rationale to the known fact that pyrimidine dimers are formed in solution under triplet photosensitization conditions[1] and the lack of any triplet absorption in the transient spectra reported for $(dT)_{20}$ due to the high efficiency of the process compared to the time resolution used to monitor the thymine dimer formation.[8]

The *shearing-type* conical intersection involving the lowest singlet and the ground state becomes the cornerstone to understand the formation of photoproducts along the singlet manifold. In order to reach the funnel, the system placed at ^{1}(C*C) has to overcome a barrier of 0.2 eV. In competition to the photoreaction, fluorescence from the ^{1}(C*C) becomes also possible. The fluorescent features from the locally excited ^{1}(LE) and the relaxed ^{1}(C*C) excimers help to rationalize the pronounced wavelength dependence observed in solution by using fluorescence polarization techniques.[75,76] The presence of a barrier does not imply however that the overall process of formation of cyclobutane cytosine (CBC) is hindered. Instead, analysis of the current theoretical and experimental information at hand suggests as *reactive orientations* those that at the time of light irradiation are close but energetically above the shearing-type conical intersection $(S_1/S_0)_{CI}$, which is concomitant to Pyr<>Pyr formation (see Figure 10). In this sense, the lack of stable excimers represents a larger potential of the system to achieve reactive arrangements, which is translated into a relative higher rate. Such a situation is accomplished by thymine.[7,47] On the contrary, as it occurs in cytosine, stable excimers may decrease the effectiveness of photoproduct formation in the singlet manifold, since the orientations of the monomers around the region of the relaxed excimer might not be so reactive and an excess of vibrational energy is required to surmount the corresponding barriers and, as a result, the CBC formation becomes globally less effective (see Figure 10). In summary, stable singlet excimers slow down the efficiency of dimer formation, as for cytosine, whereas the absence of excimers, such as thymine, may indirectly favor the photoreaction. The scheme serves to rationalize the reason why a reaction which proceeds via triplet state in solution may have a singlet-state precursor when the biochromophores are held together, as is the case in frozen solutions or in a biopolymer.[15] It may also be useful to design health care phototherapeutic nucleobase-based drugs addressed to either enhance or decrease pyrimidine dimer formation by using, for instance, appropriate nucleobases derivatives. Not surprisingly, bioexcimers are ubiquitously used by Nature. They play a fundamental role in photobiology by modulating the

charge redistribution in the lowest excited state, something which is crucial to understand electron transfer in photosynthesis[79] and, as shown here, to rationalize the distinct intrinsic reactivity of DNA nucleobases toward intrastrand pyrimidine dimerization.

Acknowledgment. The research reported has been supported by the MEC-FEDER projects CTQ2007-61260 and CSD2007-0010 Consolider-Ingenio in Molecular Nanoscience. D.R.S., G.O.G., and I.G.R. gratefully acknowledge Ph.D. fellowships from the Spanish MEC.

Supporting Information Available: Further computational details and wave function descriptions, complete ref 56, Cartesian coordinates for the optimized structures, and plots on BSSE-uncorrected curves (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA803068N

(79) Olaso-González, G.; Merchán, M.; Serrano-Andrés, L. *J. Phys. Chem. B* **2006**, *110*, 24734–24739.

Molecular Basis of DNA Photodimerization: Intrinsic Production of Cyclobutane Cytosine Dimers

Daniel Roca-Sanjuán, Gloria Olaso-González, Israel González-Ramírez
Luis Serrano-Andrés, and Manuela Merchán*
Instituto de Ciencia Molecular
Universitat de València
Apartado 22085, ES-46071 Valencia, Spain

Supporting Information

* To whom correspondence should be addressed. E-mail: Manuela.Merchan@uv.es

Additional details

a. Basis Set Superposition Error.

The inclusion of the basis set superposition error (BSSE) is crucial to accurately describe binding energies.[1] Here the effect was taken into account by using the counterpoise correction (CP).[2]

The binding energy (E_b) was obtained as follows:

$$E_b\,(CC^*) = E_C + E_{C^*} - E_{C^*C} \qquad (1)$$

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

$$\text{CP-}E_b\,(C^*C) = E_b\,(C^*C) - [\text{CP-BSSE}(C^*C)] \qquad (2)$$

where

$$\text{CP-BSSE}\,(C^*C) = E_{C^*}\,(C, R=\infty) - E_{C^*}\,(C, R=C^*C) + E_C\,(C^*, R=\infty) - E_C\,(C^*, R=C^*C). \qquad (3)$$

In [CP-BSSE (C*C)] the geometry of the monomers is kept to that of the dimer. Thus, the notation E_{C^*} (C, R=C*C) indicates the energy of C* computed in the ghost orbitals of C at the geometry of C*C, whereas in E_{C^*} (C, R= ∞) the ghost MOs of C are at infinity distance of C*. In this manner the influence of the variation of geometry is accounted for in the BSSE treatment. The findings discussed correspond to CASPT2 results with inclusion of BSSE.

b. Excimer wave function details

In terms of the natural orbitals of the CASSCF(12,12) wave function computed without symmetry restrictions, the ^{1}(LE) state is described mainly by the singlet excited configuration HOMO-like (π) $\rightarrow$ LUMO-like (π^*) with a weight of 69%. In classic literature excimers are usually described in terms of the momomers wave function Ψ_{MM^*} and Ψ_{M^*M} (exciton states) mixed with ion-par or charge-transfer (CT) states. Thus, the exciton states are stabilized by the interaction with CT states ($\Psi_{M^+M^-}$, $\Psi_{M^-M^+}$ in the valence bond sense), so the excimer can be described by a wave function of the form:

$$\Psi_{excimer} = c_1 (\Psi_{M^*M} \pm \Psi_{MM^*}) + c_2 (\Psi_{M^+M^-} \pm \Psi_{M^-M^+}) \qquad (4)$$

The magnitude of the interaction will depend on the contribution of one or other terms at the different levels of theory. The interaction at the CASSCF level is considerably weak, since the relative treatment of the exciton and CT states is unbalanced. When dynamic correlation effects are taken into account at the CASPT2 level the interaction is notably enhanced because of the better description of the CT states, which clearly favor the interaction. In this sense the ^{1}(LE) and ^{1}(C*C) states can be considered largely stabilized by the CT contributions.

c. Complete references

Present Ref. 3 corresponds to complete Ref. 56 of the paper.

References

[1] Olaso-González, G.; Roca-Sanjuán, D.; Serrano-Andrés, L.; Merchán, M. *J. Chem. Phys.* **2006**, *125*, 231102.

[2] Boys, S. F.; Bernardi, F. *Mol. Phys.* **2002**, *100*, 65-73.

[3] Andersson, K.; Barysz, M.; Bernhardsson, A.; Blomberg, M. R. A.; Carissan, Y.; Cooper, D. L.; Cossi, M.; Fülscher, M. P.; Gagliardi, L.; de Graaf, C.; Hess, B.; Hagberg, G.; Karlström, G.; Lindh, R.; Malmqvist, P.-Å.; Nakajima, T.; Neogrády, P.; Olsen, J.; Raab, J.; Roos, B. O.; Ryde, U.; Schimmelpfennig, B.; Schütz, M.; Seijo, L.; Serrano-Andrés, L.; Siegbahn, P. E. M.; Stålring, J.; Thorsteinsson, T.; Veryazov, V.; Widmark, P.-O. *MOLCAS*, version 6.4; Department of Theoretical Chemistry, Chemical Centre, University of Lund: Lund, Sweden, **2006**.

Table S1. Cartesian coordinates x, y, z (in Å) of the stationary points optimized for the cytosine-cytosine dimer. CASPT2(12,12) total energy (Et) is also included.

S_0 Cytosine CASSCF(8e,7MOs)/ANO-S N,C,O[3s2p1d]/H[2s1p]
C_1 symmetry
Two cytosine molecules at about 10.5 Å, Et: -787.712497 au

N	–1.516732	–0.309951	5.302381
N	0.544315	0.867374	5.295448
N	2.529617	–0.269525	5.216023
C	–0.845807	0.906780	5.297600
C	1.154242	–0.267140	5.279228
C	0.474159	–1.556649	5.274839
C	–0.874191	–1.516454	5.288717
O	–1.464901	1.934123	5.303139
H	-2.514863	-0.261380	5.305898
H	2.945682	0.613548	5.443847
H	2.986034	-1.055067	5.633929
H	1.019989	-2.484251	5.249586
H	-1.498588	-2.396429	5.285647

S_0 CBC CASSCF(16e,14MOs)/ANO-S N,C,O[3s2p1d]/H[2s1p]
C_s symmetry ($1^1A'$)
symmetry distinct nuclear coordinates
Et: -787.719392 au

N	-1.174345	-1.029715	1.490393
N	0.129599	0.934492	1.915796
N	2.406920	0.893383	1.803909
C	-1.130331	0.320690	1.789649
C	1.177985	0.332582	1.520261
C	1.209686	-0.991614	0.805464
C	-0.139671	-1.732355	0.800324
O	-2.129488	0.935668	2.015623
H	-2.110021	-1.362973	1.372321
H	2.362412	1.864277	2.051934
H	3.160813	0.664517	1.193323
H	2.027307	-1.585923	1.203890
H	-0.067738	-2.746890	1.188110

Table S1 (continuation).

^{1}(C*C) CASSCF(16e,14MOs)/ANO-S N,C,O[3s2p1d]/H[2s1p]
C_s symmetry (1^{1}A")
symmetry distinct nuclear coordinates
Et: -787.622748 au

N	-1.161413	-1.007774	1.555709
N	0.053495	1.018591	1.666698
N	2.348887	1.009608	1.900174
C	-1.118763	0.388400	1.593804
C	1.175093	0.285123	1.735271
C	1.187043	-1.117378	1.713599
C	-0.044779	-1.771420	1.609303
O	-2.213961	0.952745	1.536585
H	-2.075177	-1.414661	1.544275
H	2.241896	1.964584	1.615856
H	3.164230	0.582935	1.509505
H	2.101285	-1.680877	1.795042
H	-0.167479	-2.840455	1.636827

^{3}(SWI) CASSCF(16e,14MOs)/ANO-S N,C,O[3s2p1d]/H[2s1p]
C_s symmetry (1^{3}A")
symmetry distinct nuclear coordinates
Et: -787.638110 au

N	-1.171893	-0.988398	1.414216
N	0.054851	0.922932	2.181740
N	2.340724	0.827457	2.429911
C	-1.175869	0.305164	1.894268
C	1.147527	0.258783	1.999938
C	1.196629	-1.053376	1.421808
C	-0.038664	-1.643466	0.834312
O	-2.204127	0.879234	2.108651
H	-2.085204	-1.320023	1.185379
H	2.238904	1.802557	2.644810
H	3.130652	0.647126	1.842686
H	2.137971	-1.568375	1.326759
H	-0.081143	-2.700196	1.099079

Table S1 (continuation).

^{1}(LE) CASSCF(16e,14MOs)/ANO-S N,C,O[3s2p1d]/H[2s1p]
C_s symmetry ($1^{1}A''$)
symmetry distinct nuclear coordinates
Et: -787.604485 au

N	-1.516732	-0.309951	1.565526
N	0.544315	0.867374	1.558593
N	2.529617	-0.269525	1.479168
C	-0.845807	0.906780	1.560745
C	1.154242	-0.267140	1.542373
C	0.474159	-1.556649	1.537984
C	-0.874191	-1.516454	1.551862
O	-1.464901	1.934123	1.566284
H	-2.514863	-0.261380	1.569043
H	2.945682	0.613548	1.706992
H	2.986034	-1.055067	1.897074
H	1.019989	-2.484251	1.512731
H	-1.498588	-2.396429	1.548792

^{3}(LE) CASSCF(16e,14MOs)/ANO-S N,C,O[3s2p1d]/H[2s1p]
C_s symmetry ($1^{3}A''$)
symmetry distinct nuclear coordinates
Et: -787.612157 au

N	-1.516732	-0.309951	1.679591
N	0.544315	0.867374	1.672658
N	2.529617	-0.269525	1.593233
C	-0.845807	0.906780	1.674810
C	1.154242	-0.267140	1.656438
C	0.474159	-1.556649	1.652049
C	-0.874191	-1.516454	1.665927
O	-1.464901	1.934123	1.680349
H	-2.514863	-0.261380	1.683108
H	2.945682	0.613548	1.821057
H	2.986034	-1.055067	2.011139
H	1.019989	-2.484251	1.626796
H	-1.498588	-2.396429	1.662857

Table S1 (continuation).

$(S_1/S_0)_{CI}$ CASPT2(12e,12MOs)/ANO-S N,C,O[3s2p1d]/H[2s1p]
C_1 symmetry
Et: -787.615430 au

N	-1.010357	-0.938526	1.374109
N	0.216111	1.061564	1.885637
N	2.501652	1.008861	1.985595
N	-1.316163	-1.042429	-1.639090
N	0.021909	0.918940	-1.883374
N	2.299864	0.824910	-1.661319
C	-0.999626	0.410853	1.764951
C	1.299595	0.385881	1.679900
C	1.363299	-0.950570	1.149350
C	0.075606	-1.630282	0.940933
C	-1.242906	0.307538	-1.911943
C	1.054044	0.252130	-1.523427
C	1.011069	-1.144932	-1.072649
C	-0.239786	-1.809361	-1.199015
O	-2.046761	0.934420	2.014116
O	-2.222202	0.937754	-2.184338
H	-1.916728	-1.357502	1.357855
H	2.409386	2.002508	2.081092
H	3.281640	0.739545	1.424221
H	2.216916	-1.563062	1.393375
H	0.028109	-2.703211	1.042184
H	-2.239244	-1.424656	-1.630706
H	2.265633	1.818443	-1.794221
H	2.988863	0.531607	-1.003051
H	1.917737	-1.714694	-1.163296
H	-0.271408	-2.863456	-1.426889

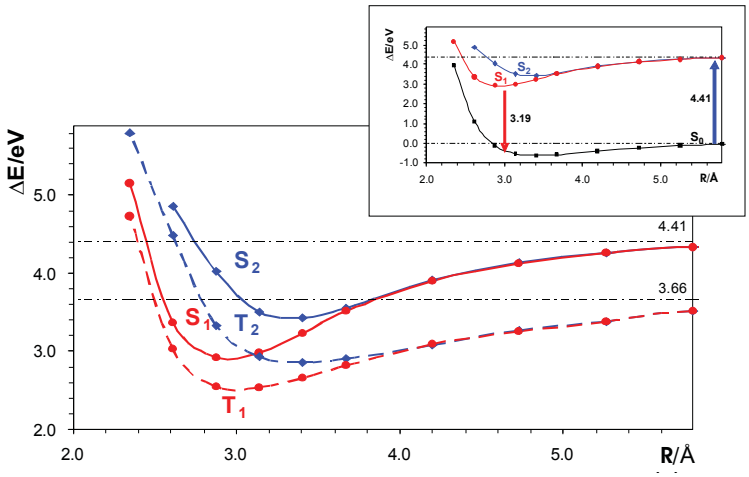


Figure S1. *CASPT2 potential energy curves built with respect to the intermolecular distance $R(C_5\text{-}C_5)$ of two face-to-face π-stacked cytosine molecules involving the ground and the lowest two triplet and two singlet excited states (BSSE uncorrected). The inset includes the singlet states obtained at the same level of theory.*[1]

6.2 Paper II

Theoretical insight into the intrinsic ultrafast formation of cyclobutane pyrimidine dimers in UV-irradiated DNA: Thymine versus cytosine.

J. J. Serrano-Pérez, I. González-Ramírez, P. B. Coto, M. Merchán and L. Serrano-Andrés

J. Phys. Chem. B, 112, 14096-14098 (2008).

THE JOURNAL OF PHYSICAL CHEMISTRY B LETTERS

2008, *112,* 14096–14098
Published on Web 10/18/2008

Theoretical Insight into the Intrinsic Ultrafast Formation of Cyclobutane Pyrimidine Dimers in UV-Irradiated DNA: Thymine versus Cytosine

Juan José Serrano-Pérez, Israel González-Ramírez, Pedro B. Coto, Manuela Merchán, and Luis Serrano-Andrés*

Instituto de Ciencia Molecular, Universitat de València, Apartado 22085, ES-46071 Valencia, Spain

Received: July 30, 2008

The higher formation yields measured in the ultrafast photoinduced formation of cyclobutane thymine dimers (T<>T) with respect to those of cytosine (C<>C) are explained, on the basis of ab initio CASPT2 results, by the existence in thymine of more reactive orientations and a less efficient photoreversibility, whereas in cytosine the funnel toward the photolesion becomes competitive with that mediating the internal conversion of the excited-cytosine monomer.

Among the possible photoreactions that pyrimidine bases of nucleic acids may undergo on ultraviolet (UV) irradiation, cyclobutane thymine dimers (T<>T) formed by intrastrand adjacent thymine bases constitute one of the major photoinduced lesions, particularly in cellular DNA.[1] Despite the fact that there are repair mechanisms for photodamaged sections of the DNA sequence, the UV irradiation of cells can result in mutation or death. In contrast to thymine−thymine (TT) sites, which are not actual mutational hot spots, cytosine−cytosine (CC) sequences are sources of relatively frequent CC-to-TT tandem mutations, although the corresponding photoproducts (C<>C) are produced with relatively lower yields.[1] Femtosecond spectroscopy has proved that thymine dimerization is an ultrafast photoreaction in which T<>T dimers are fully formed ∼1 ps after UV illumination, pointing to an excited-state reaction that is approximately barrierless *for bases that are properly oriented at the instant of light absorption*.[2] From a theoretical standpoint, relevant aspects of the [2 + 2] cycloaddition photoreaction forming the respective cyclobutane pyrimidine dimers have been analyzed for both thymine[3,4] and cytosine.[5] The concerted nonadiabatic photoreaction is mediated by a conical intersection (CI) involving the lowest singlet excited and the ground state, hereafter $(S_1/S_0)_{CI}$, which is related to the expected funnel for ultrafast nonradiative decay leading to T<>T and C<>C. There is, however, an elusive question still open. *Why is the photoinduced formation of T<>T globally more efficient than that producing C<>C?* Since the efficiency of the photodimerization markedly depends on the experimental conditions, the sequence of nucleotides, and the type (A-, B-like) of DNA conformation, the full response to this question is truly challenging. In order to get further insight into this complex issue, in the present contribution, we focus on whether the distinct photochemical behavior of TT and CC sites can be understood on the basis of the intrinsic molecular characteristics of the systems. The present research anticipates that *the relative stability of the formed excimers with respect to the placement of* $(S_1/S_0)_{CI}$ *is the main effect responsible at the molecular level for the different efficiency observed in the production of T<>T versus C<>C.*

The results discussed next were obtained by using the CASPT2 method with the active space of 12 π active electrons/12 π active orbitals, including the basis set superposition error (BSSE) through the counterpoise (CP) correction, CASPT2(12,12)+BSSE results. The ANO-S basis set with the contraction scheme C,N,O[3s2p1d]/H[2s1p] was employed throughout. Geometry optimizations were carried out for the ground state of the T<>T dimer, for a delocalized excimer $^1(TT)_{exc}$, and the crossing $(S_1/S_0)_{CI}$ (see the Supporting Information for details). In addition, the lowest-lying excited states were computed at the geometrical arrangements of the B-form DNA, $(TT)_B$. All of the calculations were performed using the MOLCAS-6.0 package.[6−8] Figure 1 compiles the main findings for TT. For proper comparison, results on CC at the same level are also included.[5]

The CP-corrected binding energy (CP-E_b) for the $^1(TT)_B$ state is computed to be 0.29 eV, about 3 times larger than that obtained for $^1(CC)_B$.[5] At the TT ground-state B-form DNA, the transition to the lowest excited singlet state (4.60 eV) becomes, as expected, slightly red-shifted as compared with the lowest singlet−singlet transition of the monomer (4.89 eV).[9] On the other hand, the S_2-S_1 gap is 0.1 eV larger for TT than that for CC, reflecting a more efficient coupling between the two states in the former. If the TT system is in the B-form DNA at the time of irradiation, the pathway from $^1(TT)_B$ toward the funnel $(S_1/S_0)_{CI}$ (path I in Figure 1) can be related to the actual decay path taking place in the biopolymer, which is predicted to be barrierless on the basis of the energy calculations derived from the linearly interpolated structures between those two geometries.[4] This also holds true for B-like arrangements energetically close to the B-form. Since DNA has a highly flexible backbone, motions such as the rise of stacking, torsional oscillation, and helix bending will continuously bring a given bipyrimidine pair into a favorable geometry for dimerization. Clearly, π-stacking facilitates formation of excimer states. In particular, the most favorable structure for producing a fully stabilized excimer corresponds to the idealized sandwich geometry.[5] The relaxed

* Author to whom correspondence should be addressed. E-mail: Luis.Serrano@uv.es.

10.1021/jp806794x CCC: $40.75

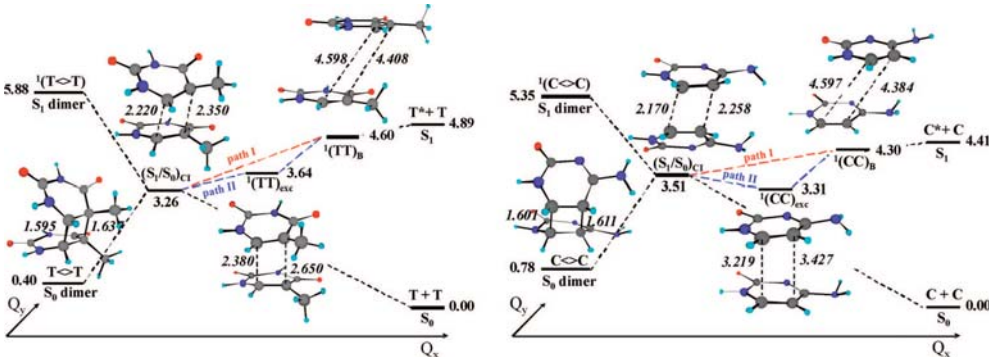


Figure 1. Relative energies (in eV) computed at the CASPT2(12,12)+BSSE level, with respect to two noninteracting ground-state thymine monomers, for the lowest excited state (S_1) at the ground-state B-form DNA, $^1(TT)_B$, at the relaxed geometry of the delocalized excimer, $^1(TT)_{exc}$, and at the ground-state structure of the T<>T dimer (left). The conical intersection $(S_1/S_0)_{CI}$, the ground state of the dimer, and the lowest transition of the monomer are also included. The main intermolecular geometric parameters (C−C distances in Å) are given in italics. The corresponding scheme for cytosine is on the right. The Q_x coordinate is mainly related to the average intermolecular distance, whereas Q_y is associated to the remaining degrees of freedom.

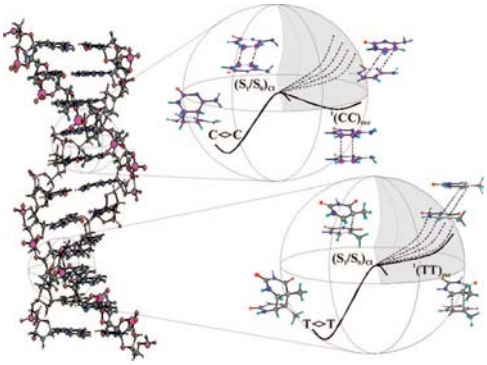


Figure 2. Scheme of the photodimerization process for π-stacked cytosine (top) and thymine (bottom) dimers along the singlet manifold. The shadowed volumes in the sphere represent regions of the DNA strand with reactive orientations, in which the decay path lies above the conical intersection $(S_1/S_0)_{CI}$.

delocalized excimer $^1(TT)_{exc}$ bears the largest overlap between the monomers; consequently, the computed CP-E_b, 1.25 eV, considerably increases with respect to $^1(TT)_B$. Independently from how the $^1(TT)_{exc}$ state is achieved, it can be directly deactivated through the funnel in a barrierless fashion (path II in Figure 1). The situation for CC is somewhat different. In order to reach the CI, the $^1(CC)_{exc}$ state has to surmount a barrier of 0.2 eV. The presence of a barrier does not imply, however, that the overall process of C<>C formation is forbidden. It simply predicts that the existence of stable excimer-like states below $(S_1/S_0)_{CI}$ decreases the effectiveness of photoproduct yield in the singlet manifold, since an excess of vibrational energy is required to overcome the barrier. As a result, production of C<>C becomes as a whole comparatively less effective than that of T<>T (see Figure 2).

The lack of stable excimers represents an intrinsic potential of the TT system to lead to photoproducts with a higher formation yield. For TT, no structure could be indeed determined whose S_1 was placed below the CI. This finding is supported by experimental evidence. Excimers were not observed in $(dT)_{18}$, that is, the single-stranded DNA 18-mer containing consecutive thymidine (dT) residues.[10] The strongest fluorescence for the T-containing oligonucleotide can just be attributed to monomer fluorescence. Accordingly, we conclude that in TT any possible structural arrangement becomes in principle prone to be a *reactive orientation* at the time of light irradiation, which is *defined as those energetically above the shearing-type CI*. In contrast, for CC sites, a certain percentage of arrangements shall not be so reactive due to the existence of the $^1(CC)$ excimers. On the basis of the red-shifted emission (the so-called excimer fluorescence) seen in the base multimers 15-mer, it was concluded that C and adenine (A), whose intrastrand excimer states were found in high yields whether stacked with itself or with T,[10] have a similar tendency to form excimers.[11] In principle, a similar kinetic model can then be used for C- and A-containing oligonucleotides; namely, every excitation in a base stack decays to an excimer, while every excitation of an unstacked base decays by internal conversion to the ground state of the monomer.[10] CC sites have, however, a striking uniqueness: the $(S_1/S_0)_{CI}$ leading to the C<>C formation at 3.51 eV is in the same energy range as that of the monomer, 3.6 eV.[9] It means that both decays may be competitive, making the production of C<>C less effective. For TT, the $(S_1/S_0)_{CI}$ and the monomer funnels are energetically far apart: 3.26 (cf. Figure 1) and 4.0 eV,[9] respectively, and no competition is established. In addition, the S_1 state is anticipated to play a more relevant role in the direct photoreversibility of CC sites than in TT sequences, since the associated oscillator strength for the transition of the former is somewhat larger (0.070 vs 0.024), with the ^{1}(C<>C) and ^{1}(T<>T) states located vertically at 4.57 and 5.48 eV, respectively. In this manner, the $^1(CC)_{exc}$ state may be repopulated within the middle UV range.

In summary, three distinct features have been deduced which account at a molecular level for the higher formation yield of T<>T with respect to C<>C observed experimentally, that provide an expansion of our understanding of this type of DNA photolesions: (i) TT has more reactive orientations than CC; (ii) photoreversibility by direct absorption to S_1 is expected to be less efficient for TT; (iii) in CC, the funnel toward C<>C production becomes competitive with the funnel that mediates the internal conversion of the excited-cytosine monomer,

whereas, for TT, the decay of the excited monomer becomes relevant only for unstacked thymine bases. By following these general guidelines, a number of designed derivatives can be envisaged with potential use in different areas of interest, from health care phototherapeutic treatments to industrial technological oriented purposes.

Acknowledgment. The research reported has been supported by the MEC-FEDER projects CTQ2007-61260 and CSD2007-0010 Consolider-Ingenio, and Juan de la Cierva programme (PBC).

Supporting Information Available: Computational details and Cartesian coordinates (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Douki, T.; Cadet, J. *Biochemistry* **2001**, *40*, 2495.
(2) Schreier, W. J.; Schrader, T. E.; Soller, F. O.; Gilch, P.; Crespo-Hernández, C. E.; Swaminathan, V. N.; Carell, T.; Zinth, W.; Kohler, B. *Science* **2007**, *315*, 625.
(3) Boggio-Pasqua, M.; Groenhof, G.; Schäfer, L. V.; Grubmüller, H.; Robb, M. A. *J. Am. Chem. Soc.* **2007**, *129*, 10996.
(4) Blancafort, L.; Migani, A. *J. Am. Chem. Soc.* **2007**, *129*, 14540.
(5) Roca-Sanjuán, D.; Olaso-González, G.; González-Ramírez, I.; Serrano-Andrés, L.; Merchán, M. *J. Am. Chem. Soc.* **2008**, *130*, 10768.
(6) Andersson, K.; Barysz, M.; Bernhardsson, A.; Blomberg, M. R. A.; Carissan, Y.; Cooper, D. L.; Cossi, M.; Fülscher, M. P.; Gagliardi, L.; de Graaf, C.; Hess, B.; Hagberg, G.; Karlström, G.; Lindh, R.; Malmqvist, P.-Å.; Nakajima, T.; Neogrády, P.; Olsen, J.; Raab, J.; Roos, B. O.; Ryde, U.; Schimmelpfennig, B.; Schütz, M.; Seijo, L.; Serrano-Andrés, L.; Siegbahn, P. E. M.; Stålring, J.; Thorsteinsson, T.; Veryazov, V.; Widmark, P.-O. *MOLCAS, version 6.4*; Department of Theoretical Chemistry, Chemical Centre, University of Lund: Lund, Sweden, 2006.
(7) Karlström, G.; Lindh, R.; Malmqvist, P.-Å.; Roos, B. O.; Ryde, U.; Veryazov, V.; Widmark, P.-O.; Cossi, M.; Schimmelpfennig, B.; Neogrady, P.; Seijo, L. *Comput. Mater. Sci.* **2003**, *28*, 222.
(8) Veryazov, V.; Widmark, P. O.; Serrano-Andrés, L.; Lindh, R.; Roos, B. O. *Int. J. Quantum Chem.* **2004**, *100*, 626.
(9) Merchán, M.; González-Luque, R.; Climent, T.; Serrano-Andrés, L.; Rodríguez, E.; Reguero, M.; Peláez, D. *J. Phys. Chem. B* **2006**, *110*, 26471.
(10) Crespo-Hernández, C. E.; Cohen, B.; Kohler, B. *Nature* **2005**, *436*, 1141.
(11) Plessow, R.; Brockhinke, A.; Eimer, W.; Kohse-Höinghaus, K. *J. Phys. Chem. B* **2000**, *104*, 3695.

JP806794X

Theoretical Insight into the Intrinsic Ultrafast Formation of Cyclobutane Pyrimidine Dimers in UV-Irradiated DNA: Thymine versus Cytosine

Juan José Serrano-Pérez, Israel González-Ramírez, Pedro B. Coto, Manuela Merchán, Luis Serrano-Andrés*

Instituto de Ciencia Molecular, Universitat de València, Apartado 22085, ES-46071 Valencia, Spain

Luis.Serrano@uv.es

Supporting Information

* Autor to whom correspondence should be addressed: Luis.Serrano@uv.es

Computational details

a. CASSCF and CASPT2 calculations.

Geometry optimizations of the ground-state T<>T dimer and the delocalized excimer $^1(TT)_{exc}$ have been studied by using the complete-active-space self-consistent-field (CASSCF) method comprising as active space a total of 16 π electrons distributed among 14 π molecular orbitals (MOs). The totally symmetric π MO of each thymine was kept inactive. In order to mimic the actual interaction of pyrimidines in DNA, geometry optimization has been carried out within C_s symmetry constraints, allowing so for an effective and natural interaction of two thymine molecules in the biologically relevant cis-syn stereoisomer. At the optimized geometries, the energies were computed with no symmetry restrictions (C_1 symmetry), since wave function symmetry breaking is a prerequisite to describe correctly the asymptotic limit for the lowest electronic singlet transition of the two moieties. It is said that the calculation breaks symmetry when the computed electronic wave function has lower symmetry than that implied by the nuclear coordinates. For the computations in C_1 symmetry, two additional π MOs were also kept inactive, since the occupation number of the corresponding natural orbitals when they were treated as active was practically 2.0. A CASSCF wave function of 12 active π electrons and 12 active π MOs was therefore employed, hereafter denoted as CASSCF(12,12). In this contribution conical intersection crossings were obtained as minimum energy crossing points (MECPs), obtained by using the restricted Lagrange multipliers technique as included in the MOLCAS-6.0 package[1] in which the lowest-energy point was obtained under the restriction of degeneracy between the two considered states.[2] In addition, an standard arrangement has been employed for the B-form DNA, $(TT)_B$.[3]

Using the C_1 state-average CASSCF(12,12) wave functions for three roots, dynamic electron correlation has been subsequently taken into account perturbatively at the second-order level through the CASPT2 method,[4,5,6] labeled as CASPT2(12,12). In order to minimize weakly interacting intruder states, the imaginary level-shift technique, with IMAG=0.2 au, has been employed.[7] The reported CASPT2 $(S_1/S_0)_{CI}$ conical intersection

was obtained by exploring a grid of points along the distortions that involve the smallest values for the CASSCF(12,12) gradients in the region of the crossing.[8] The obtained structure is consistent with previously obtained CASSCF results[9] but having larger intermonomer distances, as it occurs in ethene dimer,[8,10] an effect that is directly related to inclusion here of dynamic correlation.

The basis set of Atomic Natural Orbital (ANO) type with the contraction scheme C,N,O[3s2p1d]/H[2s1p] was used throughout.[11] Basis set superposition error (BSSE) was taken into account by using a modified counterpoise (CP) approach based on localized molecular orbitals, specifically designed for correlated approaches (see next sextion).[12] The results discussed shall be generically denoted as CASPT2(12,12)+BSSE and the corresponding CP-corrected binding energy as CP-E_b.

All the computations have been carried out by using the MOLCAS 6.0 quantum-chemical software.[1]

b. Basis Set Superposition Error.

The inclusion of the basis set superposition error (BSSE) is crucial to accurately describe binding energies.[13] Here the effect was taken into account by using the counterpoise correction (CP).[14]

The binding energy (E_b) was obtained as follows:

$$E_b\,(TT^*) = E_T + E_{T^*} - E_{T^*T} \qquad (1)$$

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

$$\text{CP-}E_b\,(T^*T) = E_b\,(T^*T) - [\text{CP-BSSE}(T^*T)] \qquad (2)$$

where

$$\text{CP-BSSE (T*T)} = E_{T^*}(T, R=\infty) - E_{T^*}(T, R=T^*T) + E_T(T^*, R=\infty) - E_T(T^*, R=T^*T). \quad (3)$$

In [CP-BSSE (T*T)] the geometry of the monomers is kept to that of the dimer. Thus, the notation E_{T^*} (T, R=T*T) indicates the energy of T* computed in the ghost orbitals of T at the geometry of T*T, whereas in E_{T^*} (T, R= ∞) the ghost MOs of T are at infinity distance of T*. In this manner the influence of the variation of geometry is accounted for in the BSSE treatment. The findings discussed correspond to CASPT2 results with inclusion of BSSE.

c. Geometries

Table S1 compiles the optimized geometries and their computed CASPT2 energies.

References

[1] Andersson, K., Barysz, M., Bernhardsson, A., Blomberg, M. R. A., Carissan, Y., Cooper, D. L., Cossi, M., Fülscher, M. P., Gagliardi, L., de Graaf, C., Hess, B., Hagberg, G., Karlström, G., Lindh, R., Malmqvist, P.-Å. , Nakajima, T., Neogrády, P., Olsen, J., Raab, J., Roos, B. O., Ryde, U., Schimmelpfennig, B., Schütz, M., Seijo, L., Serrano-Andrés, L., Siegbahn, P. E. M., Stålring, J., Thorsteinsson, T., Veryazov, V., Widmark, P. -O. *MOLCAS, version 6.0*; Department of Theoretical Chemistry, Chemical Centre, University of Lund, P.O.B. 124, S-221 00 Lund: Sweden, 2006.

[2] De Vico, L.; Olivucci, M.; Lindh, R. *J. Chem. Theory Comp.* **2005**, *1*, 1029-1037.

[3] Lu, X.-J.; Olson, W. K. *Nucleic Acids Res.*, **2003**, *31*, 5108-5121.

[4] Andersson, K., Malmqvist, P. -Å.; Roos, B. O. *J. Chem. Phys.* **1992,** *96*, 1218-1226.

[5] Roos, B. O., Andersson, K., Fülscher, M. P., Malmqvist, P.-Å., Serrano-Andrés, L., Pierloot, K., Merchán, M. *Adv. Chem. Phys.* **1996**, *93*, 219-331.

[6] Merchán, M.; Serrano-Andrés, L. In *Computational Photochemistry,* ed. Olivucci, M., Elsevier, Amsterdam, 2005.

[7] Forsberg, N.; Malmqvist, P.-Å. *Chem. Phys. Lett.* **1997**, *274*, 196-204.

[8] Roca-Sanjuán, D.; Olaso-González, G.; González-Ramírez, I.; Serrano-Andrés, L. *J. Am. Chem. Soc.* **2008**, in press.

[9] Boggio-Pasqua, M.; Groenhof, G.; Schäfer, L. V.; Grubmüller, H.; Robb, M. A. *J. Am. Chem. Soc.* **2007**, *129*, 10996-10997.

[10] Serrano-Andrés, L.; Merchán,M.; Lindh, R. *J. Chem. Phys.* **2005**, *122*, 104107.

[11] Pierloot, K., Dumez, B., Widmark, P.-O.; Roos, B. O. (1995) *Theor. Chim. Acta* **1995**, *90*, 87-114.

[12] Van Duijneveldt, F. B., Van Duijneveldt-Van de Rijdt, J. G. C. M.; Van Lenthe, J. H. *Chem. Rev.* **2004**, ***94***, 1873-1885.

[13] Olaso-González, G.; Roca-Sanjuán, D.; Serrano-Andrés, L.; Merchán, M. *J. Chem. Phys.* **2006**, *125*, 231102.

[14] Boys, S. F.; Bernardi, F. *Mol. Phys.* **2002**, *100*, 65-73.

Table S1. Cartesian coordinates x, y, z (in Å) of the stationary points optimized for the cytosine-cytosine dimer. CASPT2(12,12) total energy (Et) is also included.

S_0 Thymine CASSCF(8e,7MOs)/ANO-S N,C,O[3s2p1d]/H[2s1p]
C_1 symmetry
Two thymine molecules at about 22 Å, Et: -905.842746 au

N	-0.732356	1.031478	1.499145
N	-1.118134	-1.240959	1.499349
C	-1.663092	0.015776	1.499908
C	0.247451	-1.477747	1.500391
C	1.154751	-0.484077	1.499508
C	2.641891	-0.702205	1.500204
C	0.648709	0.902545	1.499442
O	1.352132	1.876155	1.500610
O	-2.843758	0.219643	1.500798
H	0.529155	-2.517058	1.501808
H	-1.101588	1.961794	1.499721
H	-1.768089	-1.997177	1.498473
H	2.871350	-1.765157	1.494060
H	3.095037	-0.250930	2.380649
H	3.096774	-0.240693	0.626048

S_0 T<>T CASSCF(16e,14MOs)/ANO-S N,C,O[3s2p1d]/H[2s1p]
C_s symmetry ($1^1A'$)
symmetry distinct nuclear coordinates
Et: -905.871904 au

N	-0.663032	1.034929	1.687450
N	-1.060511	-1.250889	1.493713
C	-1.557050	-0.016821	1.808581
C	0.165186	-1.500364	0.797693
C	1.199282	-0.351630	0.818750
C	2.523897	-0.707456	1.493337
C	0.689844	0.949306	1.423407
O	1.409690	1.882208	1.629656
O	-2.676235	0.161052	2.191298
H	0.589352	-2.426619	1.176232
H	-0.991511	1.910444	2.049657
H	-1.757302	-1.967895	1.502049
H	2.913191	-1.650289	1.121174
H	2.366217	-0.807822	2.566969
H	3.259216	0.073234	1.327812

Table S1 (continuation).

$^1(TT)_{exc}$ CASSCF(16e,14MOs)/ANO-S N,C,O[3s2p1d]/H[2s1p]
C_s symmetry ($1^1A''$)
Symmetry distinct nuclear coordinates. Et: -905.763511 au

N	-0.690402	1.005415	1.678946
N	-1.038640	-1.279440	1.503768
C	-1.589768	-0.014225	1.726182
C	0.255607	-1.509223	1.190208
C	1.213985	-0.445747	1.324791
C	2.653161	-0.755307	1.632312
C	0.714226	0.872395	1.634770
O	1.386410	1.849858	1.854621
O	-2.762095	0.109610	1.924902
H	0.573793	-2.537725	1.277865
H	-1.050281	1.922076	1.863143
H	-1.701012	-2.029394	1.541463
H	2.959491	-1.696647	1.181048
H	2.817442	-0.833152	2.711560
H	3.298194	0.033500	1.258784

$(S_1/S_0)_{CI}$ CASPT2(12e,12MOs)/ANO-S N,C,O[3s2p1d]/H[2s1p]
C_1 symmetry
Et: -787.615430 au

N	-1.586509	-0.762859	1.035520
N	0.601105	1.836724	0.935071
N	-1.574258	-1.085655	-1.268716
N	-0.474168	1.805148	-1.129034
C	-2.303789	-0.761344	-0.136622
C	-0.461589	2.268135	0.160892
C	-0.204711	-1.201709	-1.278956
C	0.443879	0.886483	-1.662139
C	0.448362	-1.590449	-0.009889
C	1.471430	0.376338	-0.788221
C	1.422267	-2.747310	0.033472
C	2.722856	-0.162422	-1.417427
C	-0.328674	-1.359972	1.209969
C	1.611573	0.972227	0.556133
O	0.040338	-1.639826	2.320221
O	2.538846	0.742898	1.281807
O	-3.466261	-0.490408	-0.202444
O	-1.303787	2.995501	0.600121
H	0.208146	-1.604691	-2.185655
H	0.656099	0.994170	-2.716163
H	-2.125853	-0.604634	1.865804
H	0.638885	2.201934	1.866729
H	-2.058402	-0.923889	-2.127567
H	-1.285539	2.066501	-1.651431
H	1.953028	-2.851814	-0.907793
H	3.333424	0.685195	-1.736233
H	0.902850	-3.688475	0.233672
H	2.498767	-0.762334	-2.297095
H	2.147820	-2.596394	0.829002
H	3.292271	-0.746922	-0.708417

6.3 Paper III

The role of pyrimidine nucleobase excimers in DNA Photophysics and photoreactivity.

I.González-Ramírez, T. Climent, J. J. Serrano-Pérez, R. González-Luque, M. Merchán and L. Serrano-Andrés

Pure Appl. Chem., 81, 1695-1705 (2009).

Pure Appl. Chem., Vol. 81, No. 9, pp. 1695–1705, 2009.
doi:10.1351/PAC-CON-08-08-31

The role of pyrimidine nucleobase excimers in DNA photophysics and photoreactivity*

Israel González-Ramírez, Teresa Climent, Juan José Serrano-Pérez, Remedios González-Luque, Manuela Merchán, and Luis Serrano-Andrés‡

Instituto de Ciencia Molecular, Universitat de València, Apartado de Correos 22085, ES-46071 Valencia, Spain

Abstract: Quantum chemical studies using the accurate CASPT2//CASSCF procedure show that π-stacked interactions in biochromophores such as pyrimidine (Pyr) DNA/RNA nucleobases pairs yield excimer-like situations which behave as precursors of processes like charge transfer (CT) or photoreactivity and are the source of the emissive properties in DNA. Examples are the CT between adjacent DNA nucleobases in a strand of oligonucleotides and the photodimerization taking place in cytosine (C) pairs leading to cyclobutanecytosine (CBC) mutants. These processes take place through nonadiabatic photochemical mechanisms whose evolution is determined by the presence and accessibility of conical intersections (CIs) and other surface crossings between different electronic states.

Keywords: bioexcimers; charge transfer; DNA lesions; DNA photochemistry; quantum chemistry.

INTRODUCTION

It is by now clearly established that π-stacked interactions between piled DNA and RNA nucleobases play a basic role in the stability, dynamics, and reactivity of the genetic material [1]. We have recently reported [2] that such types of conformations are favorable sources for the formation of excimers or exciplexes, that is, molecules which are homo- or heterodimers, respectively, associated in an excited electronic state and dissociative in the ground state [3]. The existence of excimers/exciplexes (excimer will be used as a general term hereafter) is common in different fields, including biochemistry and photobiology [1,3–5]. In particular, we have recently studied the binding energies in different states of the neutral and ionic cytosine (C) homodimer, and the existence of a bound excited state was proved by means of quantum chemical calculations in various relative conformations of the homodimer [6]. Binding and interaction change dramatically with the relative orientation of the two stacked C molecules. For instance, at the B-DNA-like type of orientations the binding energy for the low-lying excited singlet state, S_1, is just 0.11 eV, a value that increases to 0.58 eV in the so-called ^{1}(LE) conformation, a sandwich-like structure obtained by symmetrically approaching the molecules along the intermonomer distance, R, and to 1.10 eV in the conformation we coined $(C^*C)_{exc}$, the most stable structure for the excimer obtained by geometry optimization [6]. The stabilization is favored in the two latter orientations because of the maximum overlap obtained between the π-clouds in the homodimer. As

*Paper based on a presentation at the XXII[nd] IUPAC Symposium on Photochemistry, 28 July–1 August 2008, Gothenburg, Sweden. Other presentations are published in this issue, pp. 1615–1705.

‡Corresponding author: E-mail: Luis.Serrano@uv.es

will be shown below, the presence of so low-lying conformation in the excited state will have important consequences in the photoinduced reactivity of the nucleobases.

Before going deeper into the consequences of the excimer formation in DNA environments, and how they behave as precursors in the formation of DNA-base photoproducts [6,7], participating in the charge-transfer (CT) processes in DNA, we shall firstly establish their existence as an intrinsic property of all nucleobases, something that will be illustrated here for the homodimers of C, uracil (U), and thymine (T). Later, the importance of excimers in the rationalization of DNA emission properties and dynamics, in the formation of cyclobutane pyrimidine (Pyr) dimers, often leading to mutations, and in the CT process taking place along the DNA strands shall be discussed.

METHODOLOGY

Our methodology involves the multiconfigurational CASSCF and CASPT2 quantum chemical methods and highly accurate ANO-S-type one-electron basis sets contracted to C,N,O [3s2p1d]/H [2s1p], a strategy which has proved its accuracy [8–11]. An active space of 12 electrons in 12 orbitals was employed in the description of the CASPT2 potential energy curves (PECs) for the ground and low-lying singlet excited states of the different homodimers along the intermolecular distance between the two π-stacked systems using a C_s ground-state geometrical arrangement and C_1 symmetry to obtain the wavefunction. The inclusion of the basis set superposition error (BSSE) correction, estimated by the counterpoise correction [2], was required to get the proper description of the system. Geometry optimizations, minimum energy paths (MEPs), and determination of hypersurface crossings in the dimers were carried out initially at the CASSCF level using 16 electrons distributed among 14 molecular orbitals (MOs), that is, all the π-system except for the all-in-phase deep π-orbital localized on each monomer, which was treated as inactive. No symmetry constraints were imposed on the wavefunction. Additional computational details can be found elsewhere [2,6,7,12]. All calculations employed the MOLCAS 6.0 program quantum chemistry package [13,14].

RESULTS AND DISCUSSION

Excimer formation in π-stacked Pyr nucleobases

Our initial goal is to prove that all three Pyr nucleobases—C, T, and U—yield excimers when stacked through their π-structure. Figures 1 and 2 display the CASPT2 PECs for the ground and low-lying and triplet singlet excited states of the U homodimer along the intermolecular distance between the two π-stacked systems. A face-to-face sandwich-type arrangement has been selected because, taking into account the maximum overlap between the π-clouds, it represents the most favorable conformation to yield an excimer, which, on the other hand, is also present in other structures, for instance, the B-DNA form. As noted in Fig. 1, the lack of BSSE causes an overestimation of the states binding energies, a situation that is corrected in Fig. 2. Whereas the ground state is dissociative, the lowest singlet excited state becomes clearly bound (0.48 eV) at the corrected BSSE-CASPT2(12,12) level, with a S_1 minimum placed at $R(C_5–C_{5'})$ = 2.910 Å. In the PECs, we have chosen to monitor the intermolecular distance $R(C_5–C_{5'})$, which is particularly relevant in the formation of cyclobutane photodimers. In the figures, energies are referred to two ground-state separated U molecules. In the asymptotic limit, S_1 and S_2 become degenerate. They are related to the equivalent situations U + U* and U* + U, where U and U* represent the ground-state U and its lowest singlet excited state, respectively.

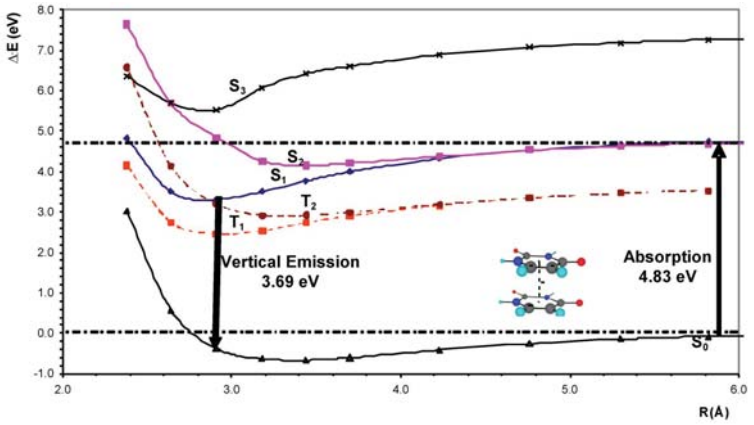


Fig. 1 CASPT2(12,12)/ANO-S C,N,O[3s2p1d]/H[2s1p] PECs built with respect to the intermolecular distance considering the center of mass of two face-to-face π-stacked U molecules.

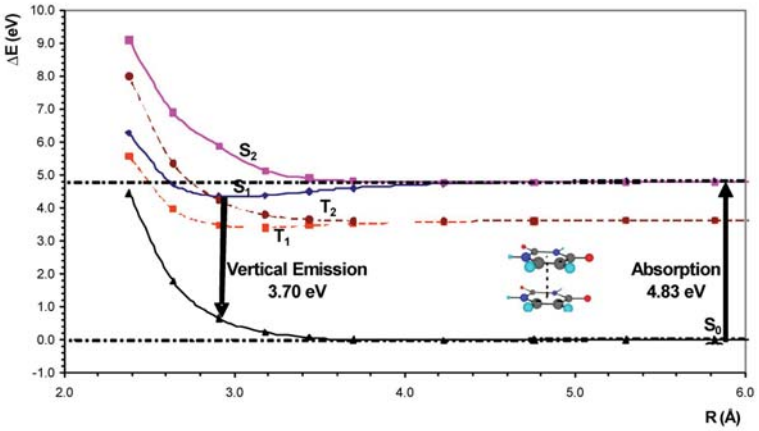


Fig. 2 Corrected BSSE-CASPT2(12,12)/ANO-S C,N,O[3s2p1d]/H[2s1p] PECs built with respect to the intermolecular distance considering the center of mass of two face-to-face π-stacked U molecules.

Similar behavior has been found for all three Pyr nucleobases. The binding energies (BSSE corrected) are compiled in Table 1.

The homodimers of the three natural Pyr nucleobases display similar behavior with respect to their binding properties, although the formation of excimers in T is hampered by the steric problems consequence of the presence of the methyl group. In particular (see Table 1) for the comparable profile as C and U, the binding energy for the T homodimer lowers to 0.36 eV and the corresponding mini-

mum takes place at a larger distance, 3.703 Å. In any case, the important conclusion obtained is that all three homodimers give rise to excimers, and, as mentioned above, the interaction occurs at many other relative orientations of the two nucleobases. Considering, therefore, the inherent flexibility of the DNA strand, there will be a large number of relative orientations in which the formation of excimers will be highly favored, even increasing the binding energy, as it was found for C (1.10 eV) upon optimization of the geometry for the S_1 state of the dimer [6].

Table 1 BSSE-corrected and -uncorrected binding energies (E_b), intermonomer distance (R_{min}), and VE for the face-to-face π-stacked homodimers of C, T, and U.

Homodimer	E_b/eV	BSSE-E_b/eV	R_{min}/Å	VE/eV
Cytosine	1.55	0.58	3.076	3.40
Thymine	1.05	0.36	3.703	4.38
Uracil	1.51	0.48	2.910	3.70

Red-shifted emission in DNA: Excimer radiative deactivation

Eisinger et al. [15] determined that in chains of oligonucleotides a source of emission was found at energies red-shifted from the weak features of the monomers, which were reported around 4.0 eV in water [16]. More detailed studies followed. For instance, the fluorescence maximum observed in aqueous solution for the dimer $d(C)_2$ and the 15-mer $d(C)_{15}$ (λ_{max} = 385 nm; 3.22 eV) [17] is considerably red-shifted as compared to that of the monomer (λ_{max} = 313 nm; 3.96 eV) [18]. Nowadays, the available experimental data for the fluorescence band maxima in dinucleotides, polynucleotides, and DNA range 3.2–3.4 eV [17].

It is, therefore, clear that π-stacked nucleobases give rise to a new source of fluorescence related to the association of the monomers. Our description of excimers in the homodimers of the Pyr nucleobases at different orientations of the moieties and the CASPT2 computed vertical emissions (VEs) supports the excimer origin of the red-shifted fluorescence observed in Pyr oligonucleotides [15,17], whereas preliminary studies indicate that excimers are also present for purine nucleobase dimers. Our best estimates have differences near 0.2 eV with respect to the recorded emission maxima, however, it is worth recalling that the computed vertical transition does not have experimental counterpart, and for a truly correct comparison with experiment, vibrational resolution of the band should be computed in order to determine the band maximum.

The evidence of the excimer origin of the emissive properties of DNA highlights the importance of π-stacking and excimer formation in the modulation of DNA relaxation dynamics, which will combine emission with nonradiative decays through accessible CIs connecting the ground and low-lying singlet excited state—such as those found for the monomers at 3.6 (C), 3.9 (T), 4.0 (U), 4.0 (adenine, A), and 4.3 eV (guanine, G) [7,11,19–21]—favoring hole and electron transfer along the strand, and formation of lesions as covalently bound Pyr dimers [1,6].

Charge transport along the DNA strand: Micro-hopping mechanism

Nucleobase cations and anions can be formed via direct ionization of the DNA strand as primary radical product, by means of the exposure of living matter to ionizing radiation and, also indirectly, as secondary radical products of the nucleobase interaction with reactive species present in DNA and its vicinity, that is, endogenous compounds that have been originated from irradiation of other DNA components or water: sugar and phosphate radicals, $^{\bullet}OH$, the hydrated electron, $H^{\bullet}$, and H_3O^+ [22,23]. The energy required for the formation of charged nucleobases, related to the corresponding ionization potential (IP) and electron affinity (EA), has been evaluated both experimentally and theoretically. Once

the charged DNA bases are formed, the produced hole or electron can migrate up to long distances through the π-stacked structure of the double helix [24,25]. This phenomenon is known in general as charge transport or transfer (CT), and it is usually denoted hole transfer (HT) and excess electron transfer (EET) when it takes place under oxidative and reductive stress, respectively. Because of its biological implications and its technological and medical applications, CT in DNA has been during the last decades an area of increasing interest [26–29].

Whereas it is currently established by using different spectroscopic, biochemical, electrochemical, and photophysical studies that DNA provides a favorable medium for CT [30–33], the detailed mechanism on how it takes place is not yet known. The available models have been particularly focused on HT and, among the different hypotheses suggested, three main mechanisms have been put forward: (1) coherent single-step tunneling process between the initial and final charged nucleobases [34–36] (2) random-walk multistep hopping process between the initial and final sites mediated by G [37–39] and, in special cases, by A [40], (3) and polaron-like hopping process [41–44]. We have recently applied the highest available levels of theory to produce accurate theoretical (in vacuo) values for IPs and EAs [45,46], and our goal here is to summarize our preliminary efforts to establish a comprehensive model to rationalize at molecular level the CT process. Focusing on the HT type of mechanism, and in order to get further insight into the mechanism of CT, we have analyzed at the CASPT2 level the lowest doublet states of different oxidized nucleobase homo- and heterodimers in different orientations, from the highly overlapped π-stacked face-to-face arrangement (see Figs. 1 and 2) to the B-DNA form (see Figs. 3 and 4) and their evolution upon different displacements [47–49]. As an illustration we show here the case of the cationic A–C heterodimer ($C_{3'}A_{5'}^{+}$) in the B-DNA form [12] along the rise coordinate, that is, the displacement along the DNA helix axis [48].

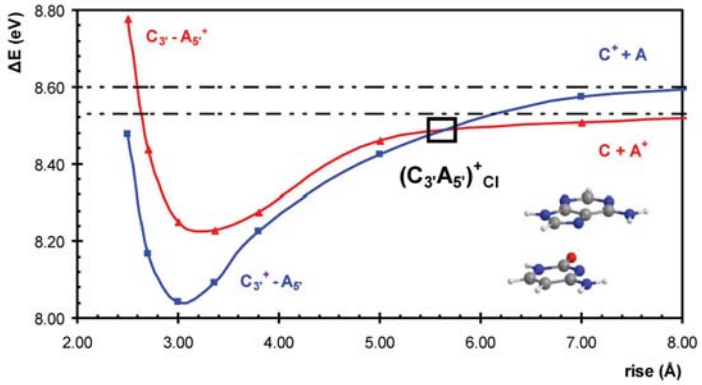


Fig. 3 BSSE-corrected CASPT2/ANO-S C,N,O [321]/H [21] PECs, built with respect to the intermonomer distance (rise) for the two low-lying doublet states of the oxidized C–A heterodimer $(C_{3'}A_{5'})^{+}$. The CI between the ground and lowest excited state of the dimer is represented by $(C_{3'}A_{5'})^{+}_{CI}$ and marked with a black square.

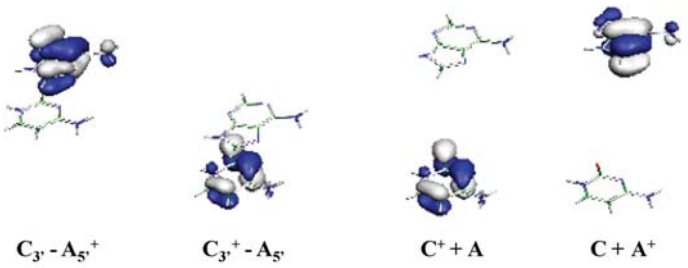


Fig. 4 Singly occupied natural orbitals (SONOs) of the ground and lowest excited state of the cationic B-DNA form of the A–C heterodimer $(C_{3'}A_{5'})^+$ at a geometry close to the average distance in DNA (3.4 Å) (left) and at large distances (right).

By inspection of the PECs for the $C_{3'}A_{5'}^+$ heterodimer, it is evident that an oxidized C is strongly stabilized by the presence of A. As it can be seen in Fig. 3, in the heterodimer analyzed here, $(C_{3'}A_{5'}^+)$, the cationic C and A are bound by the adjacent neutral nucleobase with decreases in their IPs of 0.56 eV at *rise* = 3.00 Å and 0.30 eV at *rise* = 3.36 Å, respectively. Since the Pyr nucleobase is more stabilized than the purine one and such stabilization is higher than the difference between the IPs of the isolated nucleic bases, the two low-lying states of the system, $C_{3'}^+A_{5'}$ and $C_{3'}A_{5'}^+$, which are characterized by the location of the hole in the C or A, respectively, cross-over in a point that takes place at intermolecular distances around 5.5–6.0 Å and corresponds to the conical intersection (CI) that mediates the HT between both nucleobases, $(C_{3'}A_{5'})^+_{CI}$. This result implies that the order of the IPs of C and A is inverted when these nucleic bases appear in adjacent positions of a DNA chain, and therefore the electron is removed from the Pyr nucleobase when the A–C heterodimer is oxidized. The remaining hole of the oxidation can be filled again by the promotion of one electron from the A partner by means of a thermally activated increase of the intermolecular distance between both nucleobases through the $(C_{3'}A_{5'})^+_{CI}$. This behavior is illustrated by the composition of the singly occupied MOs displayed in Fig. 4. Similar conclusions can be obtained by studying other nucleobase dimers upon other displacements, isolated or within the oligomer strand and under different environmental conditions, something which is currently under analysis by using QM/MM procedures. In all cases, the cornerstone of the mechanism is the presence of PEC CI crossings enabling ultrafast hole (or electron) transfer to take place. It is also true that the degree of transfer between the two states will be modulated by the strength of the electronic coupling between them, which is, together with the energy degeneration, the main condition for efficient charge or energy transfer. Comparison between the strength of the energy coupling between the different nucleobase dimers will be the goal of future research.

As a conclusion, the formation of stable excimer-like structures, revealed to be an intrinsic property of stacked nucleobases, also provides the framework to assemble the process of charge migration within the same basic unique scheme or unified theory. In this context, we propose a *micro-hopping* mechanism for CT along a single strand of nucleobase molecules formulated as a sequence of steps. Starting from an arrangement of π-stacked monomers at typical intermolecular distances of 3.1–3.4 Å, and considering the inherent flexibility of the DNA helix, in each step the two adjacent nucleobases may change their relative orientations by, for instance, increasing and decreasing their rise or twisting distance, by means of a thermally activated process, leading the system to reach accessible CIs between the lowest doublet states of the system and to transfer the excess electron (or hole charge) between the two molecules in an ultrafast manner. Therefore, charge migrates from an initial charged nucleobase to

the adjacent moiety in one step and between the initial and final charged DNA-sites in a sequence of such steps, thus making possible the CT process along the strand. That the process is more or less favorable will be surely related to the intrinsic properties of the nucleotide. It is therefore understandable that G sites, displaying the highest IPs, behave as hole charge trapping locations where the CT process is hindered, because they need a strongly distorted interacting conformation in order to cross with another state and transfer the hole to the neighbor moiety. Further research has to be performed in order to fully validate this hypothesis, extending the study to all combination of nucleobases and environments and considering the role of the Watson–Crick nucleobase pairs.

Photoreactivity in DNA: Formation of mutated Pyr dimers

One of the main motivations for studying the excited states of nucleic acids relies on the observation that UV illumination causes lesions and mutations due to photochemical modifications, the most common involving cycloaddition reactions of Pyrs T and C. Although the production of T–T cyclobutane dimers is most frequent, those involving C lead to mutation. T dimerization has recently been determined [50] to be an ultrafast reaction along the singlet manifold, whereas time-resolved studies of T dimer formation [51] show that direct excitation of $(dT)_{20}$ leads to cyclobutane T dimers (T <> T) in less that 200 ns with a remarkably absence of any triplet absorption from the transient spectra of the oligonucleotide. These evidences are in contrast to previous suggestions about the role of triplet states in DNA chemistry, in particular on the formation of Pyr complexes, Pyr <> Pyr [52], and are also at odds with the well-known dimerization at the bipyrimidine sites under triplet photosensitization conditions [53]. Effective population of the triplet manifold is feasible, and it has been documented in detail for C [54], U [55], and T [56] along the ultrafast internal conversion channel. It is therefore not surprising that in addition to the presence of singlet excimer states which will behave as precursors of the photodimerization process, we have found corresponding triplet excimer states. For instance, in the C homodimer, the binding energy for the lowest triplet state computed at the CASPT2 level (plus BSSE) is 0.22 eV, with a predicted VE (phosphorescence) of 3.23 eV and a 0–0 triplet–singlet transition of 3.44 eV. Consequently, it is concluded that the triplet excimer is also bound, although the binding energy is reduced about 60 % with respect to singlet excimer. The possibility of excimer formation arises from the Watson–Crick structure in which hydrogen-bonded pairs A–T and G–C are situated inside a double helix, the backbone formed by two sugar-phosphate chains. One turn of the helix involves 10 base pairs and is 34 Å high. Thus, the interplanar distance between neighboring base pairs is about 3.4 Å, a value that is often found in excimer-type organic crystals [3].

Because of the increased stability of the lowest excited state, geometries around the face-to-face sandwich-type structure can be considered as the best candidates as precursors of photodimers. It seems that the ideal twist angle between successive base pairs makes the geometry of the B-DNA (and A-DNA) nonreactive. According to recent experimental evidence, the static Pyr-Pyr conformations and not conformational motions after photoexcitation determines the formation of (Pyr <> Pyr) photoproducts. Within the model proposed by Schreier et al. [50], the relatively smaller degree of flexibility of A-DNA compared to B-DNA to achieve the right orientations that become prone to react has been related to the greater resistance of A-DNA to (Pyr <> Pyr) formation. As shown by these authors, dimerization occurs only for T residues that are already in a reactive arrangement at the instant of excitation because the rate of formation by favorably oriented T pairs is much faster that the rate of orientation change. A similar situation can therefore be assumed in C oligomers. We have taken the face-to-face arrangements, both for the singlet and triplet excimers, as the starting point for the study of the dimerization reaction occurring along the singlet and triplet manifold, respectively, in order to clarify at the molecular level the controversial and poorly understood mechanism of photodimer formation.

Figure 5 displays a scheme, based on the CASPT2 results, on the photodimerization of two C molecules taking place along both the singlet and the triplet manifold. The photoreactions taking place belong to the class of the [2+2] photocycloadditions. Starting from the triplet, MEP computations from

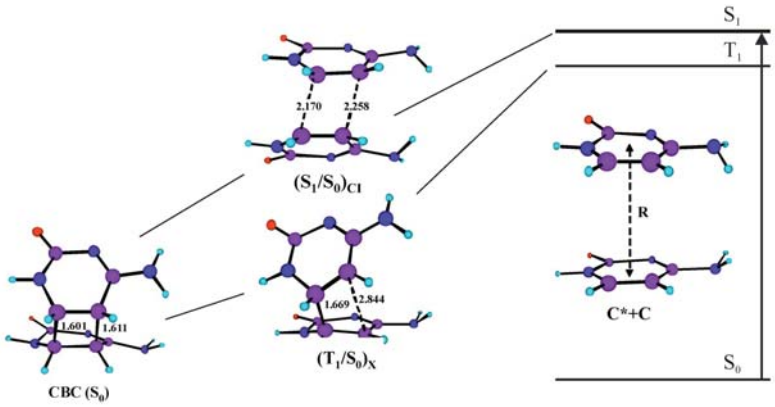


Fig. 5 Scheme based on CASPT2 calculations for the photoinduced formation of a CBC from a π-stacked C excimer (C* + C) and along the singlet and triplet manifold, involving a CI $(S_1/S_0)_{CI}$ and a singlet–triplet intermediate $(T_1/S_0)_X$, respectively.

the face-to-face arrangement lowest-energy minima of T_1 lead directly in a barrierless fashion to a triplet step-wise intermediate, which is characterized by the formation of a single covalent intermonomer bond computed to be 1.669 Å, whereas the other intermonomer CC distance remains elongated, about 2.8 Å. Remarkably, at this optimal structure the ground state of the system becomes degenerate. In other words, the triplet state is coincident with a triplet–singlet crossing, $(T_1/S_0)_X$, a region of the hypersurface where decay to the ground state becomes particularly favored. Such singlet–triplet degeneracy can be understood on the basis of the biradical character of the triplet–singlet states, having the unpaired electrons located on the bound C atoms. The intermediate triplet state lies 2.70 eV above two ground-state C molecules, an energy difference that can be considered a lower bound for the triplet energy of C in DNA.

It can be envisaged that exogenous photosensitizers could populate the relaxed triplet state of the monomer, which will subsequently evolve toward the intermediate and then toward the formation of the mutated dimers. Thus, the required energy can be related to the threshold observed experimentally for a given compound to become a potential DNA photodamager via (C <> C) or (T <> T) formation. The computed result for the singlet–triplet crossing structure of C, 2.70 eV, is fully consistent with the triplet energy of T in DNA deduced experimentally, 2.80 eV [53]. The intermediate represents, thus, a channel for photodimer formation from the triplet state of π-stacked C (and presumably also for T) in DNA and provides the basic understanding of potential photogenotoxicity via triplet–triplet sensitization. The efficiency of the overall process along the triplet manifold will be affected by the magnitude of the spin-orbit coupling (SOC), which is directly related to the efficiency of the intersystem crossing (ISC) process, and that would strongly rely on the actual environment of the biopolymer.

Regarding the mechanism of C photodimer formation along the singlet manifold, and in line with the similar excited-state [2+2] cycloadditon reactions of two ethene molecules [57], a CI is the funnel of decay toward ground-state cyclobutane C dimer (CBC) from the singlet excimer. Computations from different unrelaxed excimer lead to such CI structure. The process in the triplet manifold is also expected to follow a steepest descent path as it occurs along the singlet hypersurface. Thus, the current view supports the hypothesis that the dimerization photoreaction of two C molecules occurs barrierless, both on the singlet and triplet hypersurfaces. It would depend on the experimental conditions whether

the singlet or triplet mechanism becomes activated, fully operative, or even competitive with each other. The different mechanisms proposed in the literature involve singlet and triplet states of the monomers and vertical stacking to account for dimerization in solution and solid state, respectively. All those are here supported on the basis of CASPT2 results [6,58,59]. The efficiency of the photodimerization would markedly depend on the experimental conditions (solvent, aggregation conditions, pH, degree of hydratation), the sequence of nucleotides, the type (A-, B- C-like) of DNA conformation [50,52]. Also important for the relative efficiency of the dimer formation is the presence of relaxed excimer structures at energies lower than the CI, as it occurs in C in contrast to T, a feature that lowers the formation yield of CBC. In fact, if dimer formation occurs with reasonable yields between monomeric solute molecules in solution, the dimer must have a triplet precursor, because singlet lifetimes simply are not long enough to permit excited bimolecular reactions to occur [60]. Nevertheless, as Eisinger and Shulman have emphasized [60], the same reaction which proceeds via triplet state in solution may have a singlet-state precursor when the biochromophores are held together, as is the case in frozen solutions or in a biopolymer. Theory predicts that the photoinduced reactions both on the singlet and triplet hypersurfaces are essentially barrierless, and singlet and triplet excimers play an active role in the photophysics outcome and in the photochemical properties of C-containing biopolymers. The present results also offer a nice rationale to the known fact that Pyr dimers are formed under triplet photosensitization conditions [52].

ACKNOWLEDGMENTS

Financial support is acknowledged from projects CTQ2007-61260 and CSD2007-0010 Consolider-Ingenio in Molecular Nanoscience of the Spanish MEC/FEDER. I.G.R. acknowledges a Ph.D. grant from the Spanish MEC (FPI).

REFERENCES

1. C. E. Crespo-Hernández, B. Cohen, P. M. Hare, B. Kohler. *Chem. Rev.* **104**, 1977 (2004).
2. G. Olaso-González, D. Roca-Sanjuán, L. Serrano-Andrés, M. Merchán. *J. Chem. Phys.* **125**, 231102 (2006).
3. J. B. Birks. *Rep. Prog. Phys.* **38**, 903 (1975).
4. W. Klöpffer. "Intramolecular excimers", in *Organic Molecular Photophysics*, J. B. Birks (Ed.), p. 357, Wiley-Interscience, London (1973).
5. G. Olaso-González, M. Merchán, L. Serrano-Andrés. *J. Phys. Chem. B* **110**, 24734 (2006).
6. D. Roca-Sanjuán, G. Olaso-González, I. González-Ramírez, L. Serrano-Andrés, M. Merchán. *J. Am. Chem. Soc*. **130**, 10768 (2008).
7. L. Serrano-Andrés, M. Merchán. "Photostability and photoreactivity in biomolecules: Quantum chemistry of nucleic acid base monomers and dimers", in *Radiation Induced Molecular Phenomena*, J. Leszczynski, M. Shukla (Eds.), Springer, Berlin (2008).
8. K. Andersson, P.-Å. Malmqvist, B. O. Roos. *J. Chem. Phys.* **96**, 1218 (1992).
9. L. Serrano-Andrés, M. Merchán, I. Nebot-Gil, R. Lindh, B. O. Roos. *J. Chem. Phys*. **98**, 3151 (1993).
10. B. O. Roos, K. Andersson, M. P. Fülscher, P.-Å. Malmqvist, L. Serrano-Andrés, K. Pierloot, M. Merchán. *Adv. Chem. Phys.* **93**, 219 (1996).
11. L. Serrano-Andrés, M. Merchán, A. C. Borin. *Proc. Natl. Acad. Sci. USA* **103**, 8691 (2006).
12. D. Roca-Sanjuán, L. Serrano-Andrés, M. Merchán. *Chem. Phys.* **349**, 188 (2008).
13. G. Karlström, R. Lindh, P.-Å. Malmqvist, B. O. Roos, U. Ryde, V. Veryazov, P.-O. Widmark, M. Cossi, B. Schimmelpfennig, P. Neogrady, L. Seijo. *Comp. Mater. Sci.* **28**, 222 (2003).
14. V. Veryazov, P.-O. Widmark, L. Serrano-Andrés, R. Lindh, B. O. Roos. *Int. J. Quantum Chem.* **100**, 626 (2004).
15. J. Eisinger, M. Guéron, R. G. Shulman, T. Yamane. *Proc. Natl. Acad. Sci. USA* **55**, 1015 (1966).

16. P. R. Callis. *Annu. Rev. Phys. Chem.* **34**, 329 (1983).
17. R. Plessow, A. Brockhinke, W. Eimer, K. Kohse-Höinghaus. *J. Phys. Chem. B* **104**, 3695 (2000).
18. P. R. Callis. *Annu. Rev. Phys. Chem.* **34**, 329 (1983).
19. M. Merchán, R. González-Luque, T. Climent, L. Serrano-Andrés, E. Rodríguez, M. Reguero, D. Peláez. *J. Phys. Chem. B* **110**, 26471 (2006).
20. L. Serrano-Andrés, M. Merchán, A. C. Borin. *Chem.—Eur. J.* **12**, 6559 (2006).
21. L. Serrano-Andrés, M. Merchán, A. C. Borin. *J. Am. Chem. Soc.* **130**, 2473 (2008).
22. M. Aida, M. Kaneko, M. Dupuis. In *Computational Molecular Biology*, J. Leszczynski (Ed.), Elsevier, Amsterdam (1999).
23. S. D. Wetmore, L. A. Eriksson, R. J. Boyd. In *Theoretical Biochemistry: Processes and Properties of Biological Systems*, L. A. Eriksson (Ed.), Elsevier, Amsterdam (2001).
24. M. E. Núñez, D. B. Hall, J. K. Barton. *Chem. Biol*. **6**, 85 (1999).
25. D. Ly, L. Sanii, G. B. Schuster. *J. Am. Chem. Soc*. **121**, 9400 (1999).
26. S. Delaney, J. K. Barton. *J. Org. Chem*. **68**, 6475 (2003).
27. E. M. Boon, J. K. Barton. *Curr. Opin. Struct. Biol*. **12**, 320 (2002).
28. C. R. Treadway, M. G. Hill, J. K. Barton. *Chem. Phys*. **281**, 409 (2002).
29. E. B. Starikov. *Int. J. Quant. Chem*. **77**, 859 (2000).
30. M. E. Núñez, J. K. Barton. *Curr. Opin. Chem. Biol*. **4**, 199 (2000).
31. G. B. Schuster. *Acc. Chem. Res*. **33**, 253 (2000).
32. B. Giese. *Annu. Rev. Biochem*. **71**, 51 (2002).
33. F. D. Lewis, R. L. Letsinger, M. R. Wasielewski. *Acc. Chem. Res*. **34**, 159 (2001).
34. R. A. Marcus, N. Sutin. *Biochim. Biophys. Acta* **811**, 265 (1985).
35. M. Ratner. *Nature* **397**, 480 (1999).
36. J. Jortner, M. Bixon, T. Langenbacher, M. E. Michel-Beyerle. *Proc. Natl. Acad. Sci. USA* **95**, 12759 (1998).
37. B. Giese. *Acc. Chem. Res*. **33**, 631 (2000).
38. Y. A. Berlin, A. L. Burin, M. A. Ratner. *J. Am. Chem. Soc*. **123**, 260 (2001).
39. J. Joseph, G. B. Schuster. *J. Am. Chem. Soc*. **128**, 6070 (2006).
40. B. Giese, J. Amaudrut, A.-K. Köhler, M. Spormann, S. Wessely. *Nature* **412**, 318 (2001).
41. P. T. Henderson, D. Jones, G. Hampikian, Y. Kan, G. B. Schuster. *Proc. Natl. Acad. Sci. USA* **96**, 8353 (1999).
42. E. M. Conwell, S. V. Rakhmanova. *Proc. Natl. Acad. Sci. USA* **97**, 4556 (2000).
43. E. M. Conwell, S. M. Bloch. *J. Phys. Chem. B* **110**, 5801 (2006).
44. E. M. Conwell, S. M. Bloch, P. M. McLaughlin, D. M. Basko. *J. Am. Chem. Soc*. **129**, 9175 (2007).
45. D. Roca-Sanjuán, M. Rubio, M. Merchán, L. Serrano-Andrés. *J. Chem. Phys.* **125**, 084302 (2006).
46. D. Roca-Sanjuán, M. Merchán, L. Serrano-Andrés, M. Rubio. *J. Chem. Phys*. **129**, 095104 (2008).
47. R. Chandrasekaran, S. Arnott. In *Landolt-Börnstein Numerical Data and Functional Relationships in Science and Technology, Group VII/1b, Nucleic Acids*, W. Saenger (Ed.), p. 31, Springer-Verlag, Berlin (1989).
48. S. Arnott. In *Oxford Handbook of Nucleic Acid Structure*, S. Neidle (Ed.), p. 1, Oxford University Press, Oxford (1999).
49. X.-J. Lu, W. K. Olson. *Nucleic Acids Res.* **31**, 5108 (2003).
50. W. J. Schreier, T. E. Schrader, F. O. Koller, P. Gilch, C. E. Crespo-Hernández, V. N. Swaminathan, T. Carell, W. Zinth, B. Kohler. *Science* **315**, 625 (2007).
51. S. Marguet, D. Markovitsi. *J. Am. Chem. Soc.* **127**, 5780 (2005).
52. J. Cadet, P. Vigny. "The photochemistry of nucleic acids", in *Bioorganic Photochemistry*, H. Morrison (Ed.), John Wiley, New York (1990).

53. F. Bosca, V. Lhiaubet-Vallet, M. C. Cuquerella, J. V. Castell, M. A. Miranda. *J. Am. Chem. Soc.* **128**, 6318 (2006).
54. M. Merchán, L. Serrano-Andrés, M. A. Robb, L. Blancafort. *J. Am. Chem. Soc.* **127**, 1820 (2005).
55. T. Climent, R. González-Luque, M. Merchán, L. Serrano-Andrés. *Chem. Phys. Lett.* **441**, 327 (2007).
56. J. J. Serrano-Pérez, R. González-Luque, M. Merchán, L. Serrano-Andrés. *J. Phys. Chem. B* **111**, 11880 (2007).
57. F. Bernardi, M. Olivucci, M. A. Robb. *Acc. Chem. Res.* **23**, 405 (1990).
58. M. Boggio-Pasqua, G. Groenhof, L. V. Schäfer, H. Grubmüller, M. A. Robb. *J. Am. Chem. Soc.* **129**, 10996 (2007).
59. L. Blancafort, A. Migani. *J. Am. Chem. Soc.* **129**, 14540 (2007).
60. J. Eisinger, R. G. Shulman. *Science* **161**, 1311 (1968).

6.4 Paper IV

Singlet-triplet states interaction regions in DNA/RNA nucleobase hypersurfaces.

R. González-Luque, T. Climent, I. González-Ramírez, M. Merchán and L. Serrano-Andrés

J. Chem. Theor. Comput., 6, 2103-2114 (2010).

JCTC Journal of Chemical Theory and Computation

Singlet−Triplet States Interaction Regions in DNA/RNA Nucleobase Hypersurfaces

Remedios González-Luque, Teresa Climent, Israel González-Ramírez, Manuela Merchán, and Luis Serrano-Andrés*

Instituto de Ciencia Molecular, Universitat de València, Apartado 22085, ES-46071 Valencia, Spain

Received March 26, 2010

Abstract: The present study provides new insight into the intrinsic mechanisms for the population of the triplet manifold in DNA nucleobases by determining, at the multiconfigurational CASSCF/CASPT2 level, the singlet−triplet states crossing regions and the main decay paths for their lowest singlet and triplet states after near-UV irradiation. The studied singlet−triplet interacting regions are accessible along the minimum energy path of the initially populated singlet bright $^1\pi\pi^*$ state. In particular, all five natural DNA/RNA nucleobases have, at the end of the main minimum energy path and near a conical intersection of the ground and $^1\pi\pi^*$ states, a low-energy, easily accessible, singlet−triplet crossing region directly connecting the lowest singlet and triplet $\pi\pi^*$ excited states. Adenine, thymine, and uracil display additional higher-energy crossing regions related to the presence of low-lying singlet and a triplet $n\pi^*$ state. These funnels are absent in guanine and cytosine, which have the bright $^1\pi\pi^*$ state lower in energy and less accessible $n\pi^*$ states. Knowledge of the location and accessibility of these regions, in which the singlet−triplet interaction is related to large spin−orbit coupling elements, may help to understand experimental evidence such as the wavelength dependence measured for the triplet formation quantum yield in nucleobases and the prevalence of adenine and thymine components in the phosphorescence spectra of DNA.

1. Introduction

Phosphorescence spectra of DNA at low temperatures have been established as consisting of two basic components which originate mainly from thymine and, to a lesser extent, from adenine.[1−3] Although triplet state formation and phosphorescence data of individual nucleobases and different derivatives in several media and conditions have been reported and reviewed,[4−7] including recent studies employing external photosensitizers,[8−10] the specifics of the intrinsic population mechanism of the triplet manifold in each of the nucleobases has not been understood so far. The different fates of their triplet states, explaining, for instance, the prevalence of two of the bases in the phosphorescence spectra of DNA, the absence of triplet guanine signals, or the triplet state involvement in the fast relaxation processes of nucleobases, in particular for thymine,[11] have still to be elucidated. Triplet states of molecular systems are frequent intermediates in important photoinduced reactions. Both their usual biradical character and relatively long lifetimes make them reactive species prone to interacting with other compounds.[12] Triplet states of DNA/RNA purine and pyrimidine nucleobases are not an exception, and they have been determined to participate in UV-promoted photoreactions as the formation of phototherapeutic nucleobase-pharmacon adducts[13] or the photodimerization of pyrimidine nucleobases, considered to be the most frequent genetic lesion taking place after UV-light irradiation.[7,14−16] Since most of the recent attention has been focused on the rapid dynamics of the initially populated singlet states of DNA/RNA nucleobases,[17−22] their intersystem crossing (ISC) mechanisms and triplet states' decay processes are only now starting to be analyzed.[23−26] The present study aims to present a unified scheme, based on quantum chemical grounds, for the description of the main

* Corresponding author fax: (+34) 96-3544427, e-mail: Luis.Serrano@uv.es.

10.1021/ct100164m
Published on Web 05/28/2010

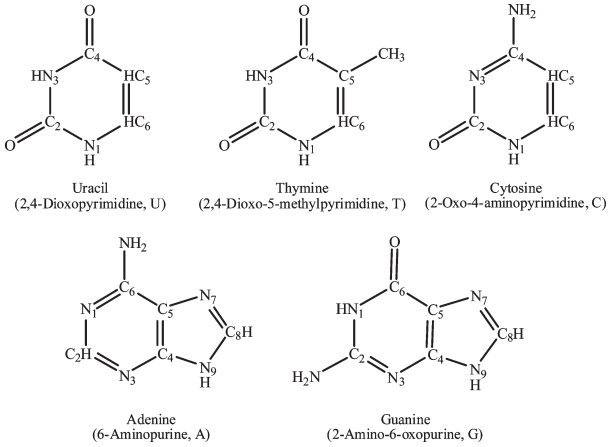


Figure 1. Structure, labeling, common name, IUPAC name, and acronym used for the five natural DNA/RNA nucleobases.

decay pathways for the singlet and triplet states of the five natural DNA/RNA nucleobases, thymine (T), uracil (U), cytosine (C), adenine (A), and guanine (G) (see Figure 1), locating the singlet−triplet crossing regions and computing the related spin−orbit coupling terms in order to provide insight into the intrinsic mechanisms of triplet state population in these molecules and to help rationalize the observed experimental data.

The triplet state population may proceed via endogenous or exogenous photosensitization from other triplet species or by efficient intersystem crossing (ISC) from the initially excited singlet state. There is an essential consensus that efficient radiationless transitions among states of the same multiplicity leading to internal conversion (IC) take place in the close vicinity of conical intersection (CI) regions and that the probability for the decay and the IC rates relate to the size of the nonadiabatic coupling matrix elements between the states.[27−29] The situation is more complex for the computation of ISC rates. In this case, the efficiency of the interaction between states of different multiplicities, for instance, singlet−triplet, seems to be reasonably well described by the Fermi Golden Rule, which relates the strength of the interaction to the extent of the vibronic spin−orbit coupling (SOC) factors and the Franck−Condon (FC) weighted density of states.[27] Recent studies of Marian and co-workers[26,30,31] have proved that the efficiency of an ISC process relies on a subtle balance of effects, including an enlarged density of vibrational states and a proper overlap of vibrational wave functions which, in turn, enhance the vibronic SOC effects. The decrease of the energy gap between singlet and triplet states, and in particular the presence of singlet−triplet degeneracies, crossing regions, especially when related to the existence of low-energy out of plane vibrational modes, is a good indication of a high density of states, and it is therefore conceivable that singlet−triplet crossings play an important role for increased ISC population transfer rates.[12] This relevance is well-known in the emerging field of multistate reactivity,[27,32,33] in which the presence of singlet−triplet crossings and the occurrence of corresponding ISC processes in the vicinity of the ground-state transition state regions become crucial for the enhancement of the reaction rates.[34] As they compete with generally faster internal conversion processes, intersystem crossings or spin crossovers can also be expected to be more efficient in energy trapping regions, for instance, near singlet states minima or sloped singlet−singlet conical intersections.[35,36] Full reaction dynamics calculations including in a balanced and accurate way nonadiabatic and spin−orbit coupling effects for polyatomic systems like those considered here have not been performed yet. Until those studies are available, calculations of ISC rates in which the vibronic spin−orbit and overlap coupling effects are considered give the best information about the efficiency of the ISC process.[26,30] Our goal in the present research is to determine the presence and accessibility of the singlet−triplet degeneracy regions in natural nucleobases along the main singlet decay pathways and provide hints of their relevance for ISC by computing also electronic SOC terms.

The strategy employed here starts by obtaining the minimum energy paths (MEPs) leading from the primary step of the photochemical process after UV light absorption in DNA nucleobases, being basically the population of the spectroscopically bright singlet excited state, here always the so-called $^1(\pi\pi^*\ L_a)$ state, toward the singlet−triplet degeneracy regions and finally the lowest-energy and reactive triplet excited state $^3(\pi\pi^*\ L_a)$, and calculating electronic SOC terms between relevant states.

Recent quantum-chemical *ab initio* CASPT2 studies have provided a unified model for the rapid internal conversion (IC) of the singlet excited DNA/RNA nucleobases manifold[18,20,29,37−46] that allowed a proper rationalization of the experimental findings.[17,47] The observed ultrafast decay component in all natural nucleobases, both in the gas phase and in solution, can be interpreted in terms of the barrierless character of the minimum energy path (MEP) associated with the lowest singlet state of the $\pi\pi^*$ type, $^1(\pi\pi^*\ L_a)$, toward a conical intersection (CI) with the ground state, $(gs/\pi\pi^*)_{CI}$. Secondary decay paths involving the lowest $^1n\pi^*$ state and

even a higher $^1\pi\pi^*$ state have been also identified.[20,23,29,38−41] Within the context of the photochemical reaction path approach[48] and the current theoretical paradigm for nonadiabatic photochemistry,[28,29] it is possible to analyze how the lowest triplet state can be reached efficiently by finding the singlet−triplet crossing (STC) regions more easily accessed from the FC MEP on the $^1(\pi\pi^*$ $L_a)$ state, which represents the major deactivation path responsible of the rapid IC process detected in the molecule. Further studies combining the calculation of ISC rates and wave packet evolution will have to determine how efficient actually are our proposed channels. The obtained results suggest that enhancements in the population yield of the lowest triplet state of the natural DNA/RNA nucleobases can be related to the presence in three of them, T, U, and A, of more ISC channels along the singlet state MEP, in particular those related with low-lying singlet and triplet $n\pi^*$ states that act as intermediate population switchers, unlike what occurs in C and G. The obtained scheme may help to understand how the intrinsic population of the lowest triplet state can take place in vacuo for all the nucleobases, why T and A triplet states seem to prevail on the DNA phosphorescence spectrum and can be expected to have a larger quantum yield of formation (ϕ_{ISC}) than the other nucleobases, and what the molecular basis is for the detected wavelength dependence of ϕ_{ISC}.[7] Since the calculations have been performed in vacuo, without the explicit consideration of solvent effects, the answer provided here can be regarded as a characteristic molecular property of the nucleobases, which might be expected to be somewhat disturbed by the specific environment in solution, in a solid, in vitro, or in vivo. The presence of the same ultrafast decays has been, however, identified in strands of oligonucleotides in solution,[49] probably related with the channels of the monomers in relatively unstacked nucleobases.[50]

II. Methods and Computational Details

The present calculations include CASSCF geometry optimizations, MEPs, CIs, and STC searches, followed by multiconfigurational perturbation theory, CASPT2, calculations at the optimized geometries. SOC terms and transition dipole moments (TDM) have also been computed. Radiative lifetimes have been estimated by using the Strickler−Berg relationship,[51] as explained elsewhere,[52] although their applicability is restricted to cases where radiative deactivation predominates. Their magnitude, otherwise, is only indicative of the prospective emissive characteristics of the state related with the TDM values. For the sake of consistency with previous calculations on the singlet states of the systems, the same one-electron basis sets and active spaces were employed. For the pyrimidine T, U, and C and purine A and G nucleobases, basis sets of the ANO-S type contracted to C,N,O[3s2p1d]/H[2s1p] and 6-31G(d,p) were used, respectively. The final results can be described as CASPT2-(14,10) for T, U, and C, involving an active space of 14 electrons distributed in 10 orbitals, with all valence $\pi\pi^*$ and lone-pair orbitals, and CASPT2(14,12) for A and G, which include all $\pi\pi^*$ orbitals except those related to the deepest canonical orbital plus two lone-pair orbitals. Other active spaces were employed in the optimization procedures, following a strategy which was proved successful previously. More detailed technical aspects of the calculations can be found in our previous papers[23−25,37−39] and in the Supporting Information (SI). All the reported calculations used the quantum-chemical methods implemented in the MOLCAS 7 package.[53,54]

Table 1. Computed Properties for the Low-Lying Singlet and Triplet Excited States of Adenine

	vertical transition (eV)		band origin (T_e, eV)		
State	CASSCF	CASPT2[a]	CASSCF	CASPT2	τ_{rad}[b]
$^1(n\pi^*)$	5.95	4.96 (0.004)	4.88	4.52	334 ns
$^1(L_b\ \pi\pi^*)$	5.56	5.16 (0.004)	4.92	4.83	251 ns
$^1(L_a\ \pi\pi^*)$[c]	7.03	5.35 (0.175)			
$^3(L_a\ \pi\pi^*)$	3.77	4.00	3.52	3.36[d]	359 ms
$^3(n\pi^*)$	5.38	4.91	4.84	4.41	
$^3(\pi\pi^*)$	5.07	4.95			

[a] Oscillator strengths within parentheses. [b] Computed using the Strickler−Berg approximation. See SI. [c] Geometry optimization leads directly to a conical intersection with the ground state, (gs/$\pi\pi^*$)CI, at 4.0 eV. See refs 29 and 38. [d] Phosphorescence band origin and maximum in solution/glasses: 3.43 and 3.05 eV, respectively. See refs 58 and 59.

III. Results and Discussion

The research effort in our group has been focused in recent years on the main singlet decay channels involving DNA/RNA nucleobases as well as several derivatives.[18,20,37−39] In addition, studies were reported on the lowest triplet population mechanisms of the pyrimidine nucleobases thymine,[24,26] uracil,[25,26] and cytosine.[23] Other recent theoretical studies on the vertical and adiabatic energies of the nucleobases' triplet states have also been reported.[55,56] In the present paper, we outline a unified scheme describing prospective population paths of the triplet manifold in all five natural DNA/RNA nucleobases T, U, C, A, and G, in order to obtain an overall model able to explain the common and the distinct behavior of the systems. Fully new results on the triplet states of the purine nucleobases A and G shall be presented, whereas our previous studies on T and U and new complementary calculations on C will be used and commented upon. The following subsections describe the results for each of the nucleobases. The most relevant conclusions are summarized in the last section.

A. Population of the Triplet Manifold in Adenine. Table 1 compiles vertical transitions, band origins, oscillator strengths, and radiative lifetimes computed for the transitions to the singlet and triplet states of adenine at the CASSCF and CASPT2 levels of theory. Unless indicated, CASPT2 results will be used in the discussion. Both at the FC region and adiabatically, the lowest-lying singlet excited state is of the $n\pi^*$ ($n_N\pi^*$) type, whereas the one carrying the largest intensity for the related transition, and therefore getting initially most of the population at low energies almost up to 6.0 eV, is the $^1(\pi\pi^*)$ HOMO (H) → LUMO (L) (hereafter L_a) singlet excited state at 5.35 eV. The ultrafast nonradiative decay undergone by adenine in the femtosecond range[17,47] can be rationalized by the barrierless character of the path on this state leading from the FC region toward a CI seam with the ground state, $(gs/^1\pi\pi^*)_{CI}$,[29,38,41,43,57] and it is shown also here in Figure 2. Unlike simple geometry optimizations,

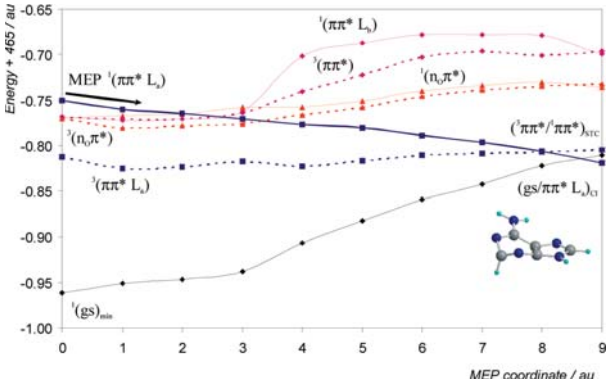


Figure 2. Evolution of the ground and lowest singlet and triplet excited states for adenine from the FC geometry along the $^1(\pi\pi^*\ L_a)$ MEP.

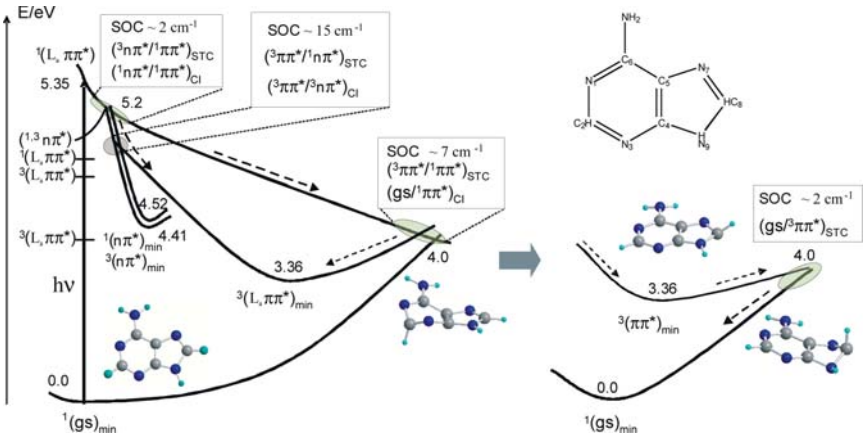


Figure 3. Scheme, based on CASPT2 results, of the photochemistry of adenine focused on the population of the lowest-energy triplet state. Unless otherwise stated, $^1\pi\pi^*$ represents the $^1L_a\ \pi\pi^*$ state.

the use of the MEP technique guarantees the absence of energy barriers along the lowest-energy path. The structure of the CI at the end of the MEP can be characterized as methanamine-like, involving combined stretching and twisting of the $C_2{=}N_3$ bond (analogous to a ethene-like CI).[37−39,60] The presence of an accessible CI explains also the low fluorescence quantum yield ($\sim\phi_F = 10^{-4}$) detected for adenine with a band origin near 4.4 eV in water.[17] This weak emission can be related to the presence of a more polar 7H isomer in solution.[29,38] A nonfluorescent $^1(n\pi^*)$ minimum is found at 4.52 eV (see Table 1) with a minor contribution to the emissive properties. Similar vertical and adiabatic energy values have been found at other levels of theory.[21,40,41,43−45,61]

Triplet $\pi\pi^*$-type states typically lie much lower in energy (here, the lowest one is placed near 1.3 eV) than their singlet counterparts, unlike for $n\pi^*$-type states, in which a small exchange integral term leads the triplet to be just slightly below the corresponding singlet state. In adenine, for instance, the lowest-energy $^3(n\pi^*)$ state lies, both vertically and adiabatically, less than 0.1 eV below its singlet analogous state. The consequences for the triplet photophysics of the system are important. Direct singlet $^1(\pi\pi^*)$−triplet $^3(\pi\pi^*)$ energy transfer seems unlikely in the FC region, where the molecule is almost planar, because of both the large energy gap and low electronic SOC terms (<0.1 cm^{-1}). The presence of two almost degenerate singlet and triplet $n\pi^*$-type states at the ground-state geometry can be, however, of high relevance. Along the main decay pathway on S_1, $^1(\pi\pi^*\ L_a)$, the state becomes degenerate with different triplet states. As it can be seen in Figures 2 and 3, along the $^1(\pi\pi^*\ L_a)$ state MEP, two singlet−triplet crossings are described: one at 5.2 eV with the $^3(n\pi^*)$ triplet state, $(^3n\pi^*/^1\pi\pi^*)_{STC}$, and another at 4.0 eV, further along the relaxation path and near the methanamine-like CI with the ground state. The latter crossing involves directly the lowest $^3(\pi\pi^*)$ T_1 triplet state, $(^3\pi\pi^*/^1\pi\pi^*)_{STC}$, and it has a structure displaying the same type of envelope puckered geometry[39] with a stretched and twisted double $C_2{=}N_3$ bond, as at the $(gs/^1\pi\pi^*)$ CI.[29,38] At

these two STC regions, the computed electronic SOC terms are 2 (^{3}n$\pi^*/^1\pi\pi^*$) and 7 cm^{-1} ($^3\pi\pi^*/^1\pi\pi^*$). These values can be considered in agreement with the qualitative El-Sayed rules, which pointed to large SOC terms for states of different natures and small otherwise.[62] El-Sayed rules were developed for molecules near the FC region, where most of the (organic) molecules considered were planar, and their identity, $\pi\pi^*$, nπ^*, etc., could be qualitatively described as such. Far from the FC region, in particular, close to a strongly distorted and puckered geometry like the ^{2}E CI, the same rules are not so easy to apply. For instance, the $\pi\pi^*$ state at this region, due to the out-of-plane distortion, has a close diradical character with two electrons in orbitals that are almost perpendicular to each other, the same as the nπ^* state in the FC region. This effect is particularly true for the low-energy singlet−triplet crossing region, which will be shown to be common in all nucleobases. The presence of the STCs combined with large electronic SOC terms are necessary, but not sufficient, conditions to guarantee efficient ISC processes, but they are good indications of relevant regions in which the population transfer toward the triplet states may take place, provided that the wave packet remains there for a long enough time for the ISC process to take place. The high-energy (∼5.2 eV) $^1\pi\pi^*-^3$nπ^* STC area, not far from the FC absorption region, fulfills those conditions. On the other hand, recent reaction dynamics calculations suggest[45] that the region of the (gs/$^1\pi\pi^*$)$_{CI}$ (reached in femtoseconds), where also the STC takes place, represents an area in which the system stays trapped for some time (due to the structure of the CI) until the population switch toward the ground state takes place, which could also explain the slower picosecond channel observed in nucleobases.[17] Figure 3 includes a scheme describing the population of T_1 based on our CASPT2 calculations.

From each one of the STC regions, we have computed corresponding MEPs along the populated triplet states, ^{3}(nπ^*) and $^3(\pi\pi^*)$, for the suggested high- and low-energy ISC channels, respectively (they can be found in the SI). Soon, along the MEP on ^{3}(nπ^*), a crossing with the lowest-lying $^3(\pi\pi^*)$ state takes place. The corresponding CI, (^{3}n$\pi^*/^3\pi\pi^*$)$_{CI}$, represents another funnel for efficient energy transfer within the triplet manifold. Additionally, as the singlet ^{1}(nπ^*) state lies very close to the triplet counterpart and their PEHs run almost parallel, an STC (^{1}n$\pi^*/^3\pi\pi^*$) also occurs at that region. Considering that the computed SOC term in this case rises to 15 cm^{-1}, the corresponding ISC process toward the $^3(\pi\pi^*)$ state should be considered very favorable. A subsequent MEP from the (^{3}n$\pi^*/^3\pi\pi^*$)$_{CI}$ along the ($^3\pi\pi^*$) PEH leads to the lowest triplet state minimum (see SI). Regarding the STC described at 4.0 eV, the MEP computed from the ($^3\pi\pi^*/^1\pi\pi^*$)$_{STC}$ along the ($^3\pi\pi^*$) state leads directly to the minimum of the triplet state (see SI). The involvement of a dark singlet nπ^* state on adenine relaxation dynamics was previously suggested by other authors to explain slow decay features.[17,63,64]

After the lowest triplet state is populated by any of the previous ISC processes, the system is finally expected to evolve toward the triplet state minimum, $^3(\pi\pi^*)_{min}$ (see Figure 3), which is characterized by a structure with almost planar rings but with the terminal hydrogen C_8H lifted near 40° and with an increased bond length C_2N_3 of 1.389 Å (compared to 1.311 Å in the ground state), in agreement with previous estimations.[55] The reactivity that could be attributed to this triplet state originates from its biradical character on C_2 and N_3. The minimum is placed at 3.36 eV adiabatically (see Table 1) from the ground state optimized minimum, a value consistent with the measured phosphorescence band origin in solution at 3.43,[58] and other theoretical results.[43,55] We have also located the singlet−triplet crossing connecting the $^3(\pi\pi^*)$ and the ground state and mapped the MEP leading from such an STC toward $^3(\pi\pi^*)_{min}$ (see SI). The crossing is placed near 4.0 eV from the ground state minimum, which means that there is a barrier of near 0.6 eV (14.0 kcal/mol) to reach (gs/$^3\pi\pi^*$)$_{STC}$ from $^3(\pi\pi^*)_{min}$. The distortion of the five-membered ring is larger at the STC point, and the computed electronic SOC is somewhat low, ∼2 cm^{-1}, suggesting for the triplet state a long lifetime and a slow relaxation, becoming therefore prone to reacting or transfering its energy by photosensitization mechanisms.[8−10]

In summary, we have identified in adenine (see Figure 3) three possible intrinsic ISC channels toward the lowest triplet state which can be easily accessed from the main barrierless MEP for singlet decay dynamics, two of them mediated by nπ^* states. In all three cases, the magnitude of the computed SOC terms between the relevant states is high enough to suggest an efficient population of the triplet manifold in adenine upon UV irradiation. This type of $^1\pi\pi^*/^3\pi\pi^*$ ISC mechanism via intermediate nπ^* states can be suggested here as favorable, even far from the FC region, as it has been recently reported also for other biological chromophores such as isoalloxazine[65] and psoralen.[66] Both mechanisms described here can in any case contribute to the overall population of the lowest triplet state. In principle, in different environments, such as in polar solvents, it is expected that the nπ^*-type excited state will become destabilized with respect to $\pi\pi^*$-type excited states.[67] Despite those effects, both singlet and triplet nπ^*-type states are estimated to lie in the solvent below the $^1(\pi\pi^*$ $L_a)$ state at the FC geometry,[68] guaranteeing the existence of the STC crossing upon decay along the $^1(\pi\pi^*$ $L_a)$ state. Intersystem crossing quantum yields have been measured by means of nanosecond laser photolysis in adenine to be 0.23×10^{-2} higher than in guanine.[7] Likewise, phosphorescence quantum yields of 4.5×10^{-2} for adenine in frozen solutions at 77 K have been reported, slightly higher than for guanine, 3.6×10^{-2} and 2×10^{-2},[69,70] and lower than thymine.[7] For the purine nucleobases, the ISC yield has been measured to be lower in the nucleotide.[7] Also in adenine,[71] although less clearly documented as in pyrimidine nucleobases, a wavelength dependence of the intersystem crossing quantum yield in nucleobases has been reported, as it can be expected by the contribution of the three (at excitation energies higher than 5.0 eV) or just the lowest-energy (at energies close to 4.0 eV) ISC mechanisms. This point requires further experimental confirmation.

B. Population of the Triplet Manifold in Guanine. The same strategy as for adenine has been followed in the calculations of guanine. Table 2 lists the main spectroscopic properties of the lowest-lying singlet and triplet states of the

Table 2. Computed Properties for the Low-Lying Singlet and Triplet Excited States of Guanine

	vertical transition (eV)		band origin (T_e, eV)		
state	CASSCF	CASPT2[a]	CASSCF	CASPT2	τ_{rad}[b]
$^1(L_a\ \pi\pi^*)$[c]	6.36	4.93 (0.158)			
$^1(n_O\pi^*)$	5.70	5.54 (0.002)	4.04	4.56	6800 ns
$^1(L_b\ \pi\pi^*)$	7.04	5.72 (0.145)	6.07	5.69	5 ns
$^3(L_a\ \pi\pi^*)$	3.97	4.11	3.13	3.15	3562 ms
$^3(\pi\pi^*)$	5.08	4.76			
$^3(\pi\pi^*)$	5.41	5.14			
$^3(n_O\pi^*)$	5.82	5.30	4.64	4.17	

[a] Oscillator strengths within parentheses. [b] Computed using the Strickler−Berg approximation. See SI. [c] Geometry optimization leads directly to a conical intersection with the ground state, (gs/$\pi\pi^*$)CI, at 4.3 eV. See ref.[39]

molecule. As compared to adenine, a couple of important aspects of the electronic structure of guanine have to be highlighted. First of all is the low energy displayed by the $^1(\pi\pi^*\ L_a)$ HOMO → LUMO state, placed at 4.93 eV at the FC region as the lowest-energy feature. The value of the related oscillator strength, 0.158, indicates that this is the bright singlet state basically populated in the low-energy absorption spectrum, and that the relevant photophysics of the system will take place along the MEP on such a state. The second aspect is related to the high-energy of the low-lying nπ^* states, which are placed near 0.6 (singlet) and 0.4 (triplet) eV above the $^1(\pi\pi^*\ L_a)$ state (even higher in solution). As is clear from Table 2, and also from Figure 4, the gap between the initially populated singlet state and the nπ^* states is much larger than in adenine. At the FC region, it is therefore expected that an ISC process relating the $^1(\pi\pi^*\ L_a)$ and ^{3}nπ^* states is less favorable than for adenine.

Figure 4 displays the MEP from the FC structure and along the $^1(\pi\pi^*\ L_a)$ state. At the beginning of the MEP, the singlet state only crosses with the second triplet $^3(\pi\pi^*)$ state. The computed electronic SOC terms are small (<0.1 cm^{-1}), and only strongly coupled vibronic terms would enhance in this region the ISC rate. Near point 9 of the MEP, the singlet state crosses with the lowest triplet state, as it occurred in adenine. The STC region, placed adiabatically at 4.3 eV, is not far from the CI between the singlet and the ground state. The corresponding SOC terms are much larger here, 8 cm^{-1}, and therefore a more efficient ISC process leading directly to the population of the lowest $^3(\pi\pi^*)$ state can be therefore expected, or at least proposed. As compared with adenine, however, the overall population of the triplet manifold cannot be expected to be favorable. Even when the $^3(n_O\pi^*)$ excited state minimum lies lower in energy than the $(gs/\pi\pi^*)_{CI}$, and therefore a crossing with the $^1(\pi\pi^*)$ state takes place at some other region, the key point is that such a crossing cannot be easily accessed from the photochemically relevant MEP, that is, the main decay path for singlet deactivation. As a matter of fact, we have computed the STC crossing structure $(^3n_O\pi^*/^1\pi\pi^*)_{STC}$, which lies almost degenerate with the computed $(^1n_O\pi^*/^1\pi\pi^*)_{CI}$ (see ref 39), at 4.6 eV, but far from the main MEP region, because it represents the stretching and twisting of the C_6N_1 bond. Even when such a structure, in which the SOC is large enough, 8 cm^{-1}, can be accessed with excess energy, it cannot be considered as favorable as those reached via the MEP-related channels.

For the sake of completeness, we have connected the mentioned STC points with the minimum of the lowest $^3(\pi\pi^*)$ state by computing the corresponding MEPs: (i) from the computed $(^3\pi\pi^*/^1\pi\pi^*)_{STC}$ and $(^3n_O\pi^*/^1\pi\pi^*)_{STC}$ structures along the $^3(\pi\pi^*)$ and $^3(n_O\pi^*)$ states, leading to their respective minima, (ii) from the computed $(^3n_O\pi^*/^3\pi\pi^*)_{CI}$ to the $^3(\pi\pi^*)$ minimum, and (iii) from the singlet−triplet $(^{1gs}/^3\pi\pi^*)_{STC}$ toward the final $^3(\pi\pi^*)$ minimum. All them are possible paths leading to the population of the lowest triplet state, although we emphasize that, unlike adenine, only the lowest-lying 4.3 eV ISC mechanism related to the $(^3\pi\pi^*/^1\pi\pi^*)$ STC should be initially considered efficient, because it is the only one taking place in the proximity of the main $^1(\pi\pi^*\ L_a)$ MEP (see Figure 5 for a scheme of the triplet photophysics in guanine). Finally, the $^3(\pi\pi^*)$ minimum has been connected through a corresponding MEP with the STC with the ground state, $(gs/^3\pi\pi^*)_{STC}$. Although the SOC terms at this point are higher than in adenine, the barrier from the minimum, placed at 3.15 eV, is too large (0.85 eV) to expect an efficient decay to the ground state. All computed MEPs can be found in the SI. At the $^3(\pi\pi^*)$ minimum, the molecule displays a slightly puckered envelope structure on the six-membered ring,[39] with the C_2N_3 bond having a biradical character and enlarged up to 1.438 Å, as compared to 1.286 Å at the FC ground-state geometry.

It has to be finally mentioned that guanine is the only natural nucleobase in which no phosphorescence data or triplet state formation has been reported for the parent compound, although intersystem crossing[7] and phosphorescence quantum yields of 0.042 and 0.095 have been reported for the nucleoside and nucleotide in ethanol,[6] 5 to 7 times larger than the fluorescence quantum yields. It has to be remembered also that the natural keto form of 9H-guanine is not the most stable in the gas phase and that other close tautomers can contribute to the measurements for the isolated system,[39,72] not in an oligomer sample.

C. Population of the Triplet Manifold in Pyrimidine Nucleobases: Thymine, Uracil, and Cytosine. For the sake of brevity, we will discuss the triplet manifold population of the pyrimidine nucleobases together within the same framework. The computational strategies followed have been those described above for adenine and guanine. As uracil has a state structure and triplet photophysics very similar to that of thymine, we will refer to our previous results[25] and concentrate on the latter. Thymine has, at the FC region, a low-lying $^1(n_O\pi^*)$ state (basically related to the O_4 atom), placed 0.2 eV below the spectroscopic $^1(\pi\pi^*\ L_a)$ HOMO → LUMO state, this one lying at 4.89 eV with a related oscillator strength of 0.167 (see Table 3). The photophysical mechanisms proposed for the population of the lowest triplet state will be very similar to those already explained for adenine, as Table 3 and Figures 6 and 7 can confirm. Once more, the key point is that three different STC regions can be easily accessed through the main decay path of the energy, as it is the FC $^1(\pi\pi^*\ L_a)$ MEP, being prospective channels for ISC toward the lowest-lying triplet state.

Soon after the beginning of the MEP (see Figure 6), the $^1(\pi\pi^*\ L_a)$ state crosses with both singlet and triplet nπ^* states. Apart from a possible IC toward the singlet $^1(n_O\pi^*)$

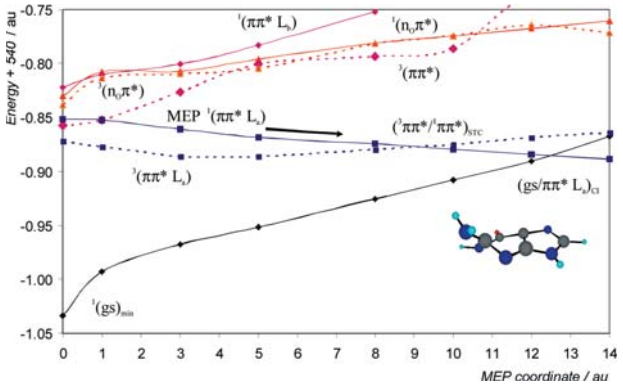


Figure 4. Evolution of the ground and lowest singlet and triplet excited states for guanine from the FC geometry along the $^1(\pi\pi^*\ L_a)$ MEP.

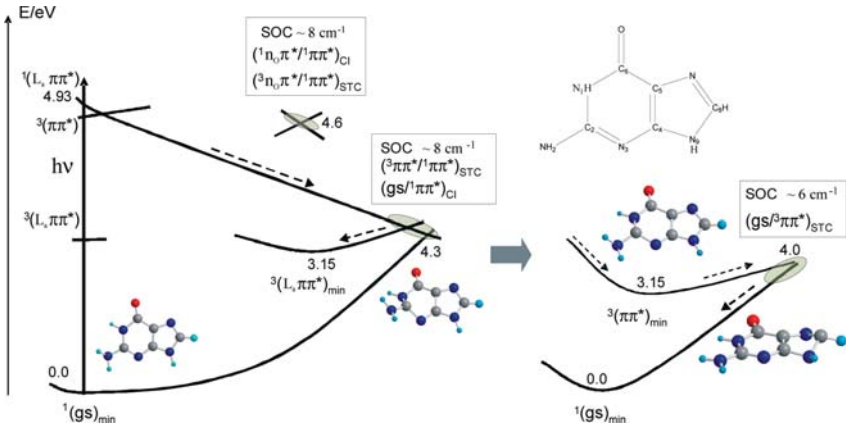


Figure 5. Scheme, based on CASPT2 results, of the photochemistry of guanine focused on the population of the lowest-energy triplet state. Unless otherwise stated, $^1\pi\pi^*$ represents the $^1L_a\ \pi\pi^*$ state.

Table 3. Computed Properties for the Low-Lying Singlet and Triplet Excited States of Thymine[a]

	vertical transition (eV)		band origin (T_e, eV)		
state	CASSCF	CASPT2[b]	CASSCF	CASPT2	τ_{rad}[c]
$^1(n_O\pi^*)$[d]	5.41	4.77 (0.004)	4.23	4.05	2501 ns
$^1(\pi\pi^*\ L_a)$	6.52	4.89 (0.167)	6.07	4.49	9 ns
$^1(\pi\pi^*)$	7.36	5.94 (0.114)			
$^3(\pi\pi^*\ L_a)$	3.95	3.59	2.99	2.87	17 ms
$^3(n_O\pi^*)$[d]	5.21	4.75	3.84	3.93	
$^3(\pi\pi^*)$	5.86	5.14			

[a] See also ref 24. [b] Oscillator strengths within parentheses. [c] Computed using the Strickler−Berg approximation. See SI. [d] Involving basically O_4 (in ortho position with methyl group).

state through a corresponding CI, this region may be responsible for the first ISC process taking place in thymine at high energies (4.8 eV), in which the $^3(n_O\pi^*)$ state could be populated from the initially activated singlet $\pi\pi^*$ state. The SOC terms, computed as 8 cm^{-1}, point to the efficiency of the process. Another MEP computed from this crossing and along the $^3(n_O\pi^*)$ PEH leads the system toward the minimum of this state, in whose neighborhood we have found the conical intersection with the lowest triplet state, $(^3n_O\pi^*/^3\pi\pi^*)_{CI}$, near 3.9 eV. As the singlet and triplet $n\pi^*$ PEHs are always very close along the MEP, near the CI we have also found the $(^1n_O\pi^*/^3\pi\pi^*)$ STC. In case some population reaches the $^1(n_O\pi^*)$ state via the higher-energy crossing with $^1(\pi\pi^*\ L_a)$—and a decay path through this dark intermediate has been recently reported[68]—the energy switch toward the lowest triplet state would be extremely favorable, because the computed SOC term increases in the $(^1n_O\pi^*/^3\pi\pi^*)_{STC}$ region to 61 cm^{-1}. It is possible to confirm our suggestions about the effectiveness of this type of mechanism thanks to the recent study by Etinski et al.,[26] which has established the efficiency of the $(^1n_O\pi^*/^3\pi\pi^*)_{STC}$ ISC channel by computing vibrational FC factors and ISC rates. Either by triplet−triplet IC or by singlet−triplet ISC, the final population process of the lowest $^3(\pi\pi^*)$ state should be considered to be extremely favorable. As in the other nucleobases, a low-energy STC region lies close to the end of the FC $^1(\pi\pi^*$

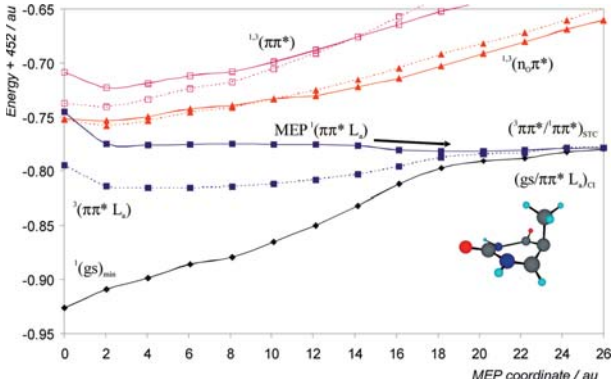


Figure 6. Evolution of the ground and lowest singlet excited states for thymine from the FC geometry along the $^1(\pi\pi^*\ L_a)$ MEP.

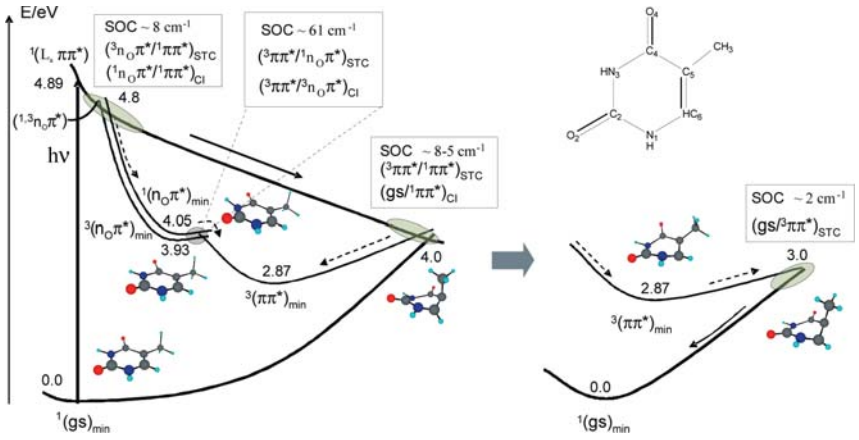


Figure 7. Scheme, based on CASPT2 results, of the photochemistry of thymine focused on the population of the lowest-energy triplet state. Unless otherwise stated, $^1\pi\pi^*$ represents the $^1L_a\ \pi\pi^*$ state.

L_a) MEP, at 4.0 eV. As observed in Figure 6, this area of the PEH is flat and extended close to the end of the MEP. The SOC values computed at different points along the path range from 5 to 8 cm^{-1}. The efficiency of the process would be also high if, as in adenine, the wave packet decaying through the singlet manifold is delayed in the region of the singlet−singlet CI. The present model allows for an understanding of the reported wavelength dependence on the ISC quantum yield in nucleobases, surely caused by the location of the two STC interacting regions and their accessibility upon the initial excitation conditions. In the case of thymine, the value increases from 3.9×10^{-3} at 280 nm (4.43 eV), where only the lowest-energy channel can be reached, to 5.2×10^{-2} at 240 nm (5.17 eV),[7,73] where both described channels are accessible.

As for the purine nucleobases, MEPs connecting the different critical points have been computed (see SI). The lowest triplet state may be populated by any of the previous ISC processes. At the state minimum, the molecule displays a distorted structure with a ring deformation including the dihedral angle $C_2N_1C_6C_5$ as 44° and an increased bond length C_5C_6 of 1.494 Å with certain biradical character. The minimum is placed at 2.87 eV adiabatically from the ground state optimized minimum, a value somewhat lower than the 3.2 eV estimated for the location of the triplet state for the thymine mononucleotide in aqueous solution at room temperature[9] and consistent with previous theoretical determinations at around 2.8−3.0 eV.[74] As a final aspect of the evolution along the triplet manifold in thymine, we have located the singlet−triplet crossing connecting the $^3(\pi\pi^*)$ and the ground state and mapped the MEP leading from such an STC toward $^3(\pi\pi^*)_{min}$ (see SI). The crossing is placed near 3.0 eV from the ground state minimum, which means that there is a barrier of 0.13 eV (3.0 kcal/mol) to reach $(gs/^3\pi\pi^*)_{STC}$ from $^3(\pi\pi^*)_{min}$, and the molecule recovers there the planarity. Although the computed electronic SOC is somewhat low, ~2 cm^{-1}, a barrier which is smaller than that for purines may explain the shorter triplet lifetimes

Table 4. Computed Properties for the Low-Lying Singlet and Triplet Excited States of Cytosine

	vertical transition (eV)		band origin (T_e, eV)		
state	CASSCF	CASPT2[a]	CASSCF	CASPT2	τ_{rad}[b]
$^1(L_a\ \pi\pi^*)$[c]	5.22	4.41 (0.069)	4.14	3.62	30 ns
$^1(n_O\pi^*)$	5.23	4.95 (0.001)	3.68	3.72	1200 ns
$^1(n_N\pi^*)$[d]	5.59	5.06 (0.003)			
$^1(L_b\ \pi\pi^*)$	6.17	5.89 (0.106)			
$^3(L_a\ \pi\pi^*)$	3.64	3.53	2.85	2.98	437 ms
$^3(\pi\pi^*)$	4.87	4.45			
$^3(n_O\pi^*)$	5.13	4.63	3.49	3.66	
$^3(n_N\pi^*)$	5.31	4.94			

[a] Oscillator strengths within parentheses. [b] Computed using the Strickler−Berg approximation. See SI. [c] The MEP to the minimum and the CI, $(gs/\pi\pi^*)_{CI}$, at 3.6 eV, are competitive. See ref 37. [d] Geometry optimization leads directly to a CI with the ground state, $(gs/n_N\pi^*)$CI. See ref 76.

measured for pyrimidine (∼0.6 s) than for purine (∼2.0 s) nucleobases in ethanol glasses.[6] Similar conclusions can be derived for uracil, which has a state structure and properties very similar to those of thymine.[25,26]

Regarding cytosine, the values in Table 4 help to understand (and predict to some extent) the behavior of its triplet photophysics. As in guanine, cytosine has a lowest-lying singlet $^1(\pi\pi^*\ L_a)$ state, whose initial interaction with the $n\pi^*$ states, placed higher in energy, will not be strong either at the FC region or along the $^1(\pi\pi^*\ L_a)$ decay pathway (see Figure 8). The singlet relaxation in C is somewhat more complex than in the other nucleobases. The presence of a low-lying planar minimum for the $^1(\pi\pi^*\ L_a)$ state at 3.62 eV, nearly isoenergetic with the ethene-like $(gs/\pi\pi^*)_{CI}$, generates several competitive decay paths, as has been analyzed before.[23,37,75] The possibilities for displaying different ISC processes are therefore larger, but always at low, not at high, energies like, for instance, in thymine, uracil, or adenine. In particular, we show in Figure 8 a linear interpolation in internal coordinates (LIIC) path from the FC region toward the ethene-like CI with the ground state. The barrier along the $^1(\pi\pi^*\ L_a)$ state, computed 2.5 kcal mol^{-1} as a higher bound, is very small, and in practice the path can be considered barrierless. As in the other nucleobases, an STC between the lowest $\pi\pi^*$ states takes place close to the CI, at 3.6 eV, yielding a SOC term value of 6 cm^{-1}. In a previous study,[23] we analyzed ISC processes taking place at other low-energy regions, obtaining also large SOC values and expectedly favorable situations for the lowest triplet population.

As a result of the excited state structure in C, obtained at the CASPT2 level, the photophysical scheme for the population of the lowest triplet state of the molecule can be summarized in Figure 9. Unlike in the other two pyrimidine nucleobases, where three basic channels for the possible triplet manifold population were found, one at high energies (close to FC and $n\pi^*$ mediated) and another at low energies (caused by the common ethene-like CI type of decay present in all nucleobases), in C, only low-energy channels seem to be accessible. This feature could probably help to explain the absence of cytosine (guanine too) components in DNA phosphorescence at low temperatures,[1−3] and also the generally lower phosphorescence quantum yields obtained for cytosine and its derivatives as compared to other nucleobases.[6] The same trends are obtained for ISC yields from flash photolysis experiments in nucleotides, although not in nucleobases.[7] Higher yields of $n\pi^*$ formation have been suggested for cytosine than thymine,[68] but theoretical evidence indicates that the higher-lying $n\pi^*$ states of cytosine will be less accessible from the main relaxation pathways than in thymine due to the large potential energy barriers found in the former.[76]

IV. Summary and Conclusions

Calculation of PEHs for the low-lying singlet and triplet states of natural DNA/RNA nucleobases adenine, guanine, thymine, uracil, and cytosine at the *ab initio* multiconfigurational CASPT2//CASSCF quantum-chemical level have been carried out in order to help to establish general mechanisms for the population of the triplet manifold of the molecules. The proposed framework is an attempt to rationalize the reported triplet states properties of DNA components, in particular the measurement of larger quantum yields of phosphorescence than of fluorescence in the individual systems,[4,7] the observed wavelength dependence of the triplet state formation,[7,73] or the prevalence of adenine and thymine components in the phosphorescence signals of DNA at low temperatures.[1−3] It can be considered that an efficient ISC channel is easily accessible from the regions close to the main decay pathway of the initially populated singlet state. We have analyzed the accessibility of the ISC channels for the population of the lowest triplet state along such a pathway, a strategy that requires computation of minimum energy paths on the different states and determination of singlet−triplet crossings and conical intersections. This is, however, only a necessary but not sufficient condition to establish the efficiency of an ISC process. Computation of vibronic contributions to the ISC rates and reaction dynamic calculations establishing the temporal evolution of the system are encouraged in a close future in order to unambiguously determine if the proposed accessible singlet−triplet crossing regions fulfill all the requirements: close singlet−triplet energies, a high density of vibronic states, large vibronic contributions to the spin−orbit coupling terms, and regions where the population gets trapped for long enough of a time to allow the ISC process to take place in competition with the internal conversion decay, for instance, close to the FC region, to a singlet state minimum, or near a sloped conical intersection. Recent ISC rate calculations on thymine and uracil[26] confirm the main role of some of our proposed ISC mechanisms in these systems.

Our results indicate that three STC regions can be easily accessed from the singlet main decay pathway in adenine, thymine, and uracil, two of them located at high energies and mediated by the presence of lowest-lying singlet and triplet $n\pi^*$ states, and a third one at low energies close to the end of the main MEP on the $^1(\pi\pi^*)$ singlet excited state and the ethene-like (pyrimidines) or methanamine-like (purines) conical intersection of this state with the ground state. These three regions are proposed as prospective ISC channels. At least those related to the $^1n\pi^*-^3\pi\pi^*$ STC seem to be confirmed as such by recent calculations on ISC rates on pyrimidine nucleobases.[26] Additionally, the wavelength

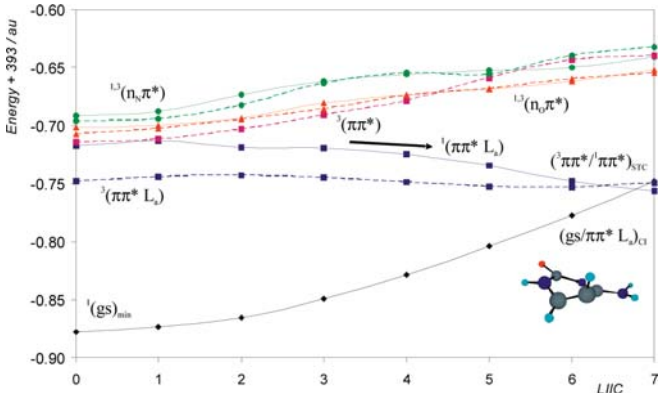


Figure 8. Evolution of the ground and lowest singlet and triplet excited states of cytosine from the FC geometry to the (gs/$\pi\pi^*$ La)$_{CI}$ along an LIIC path competitive with the $^1(\pi\pi^*$ La) MEP.

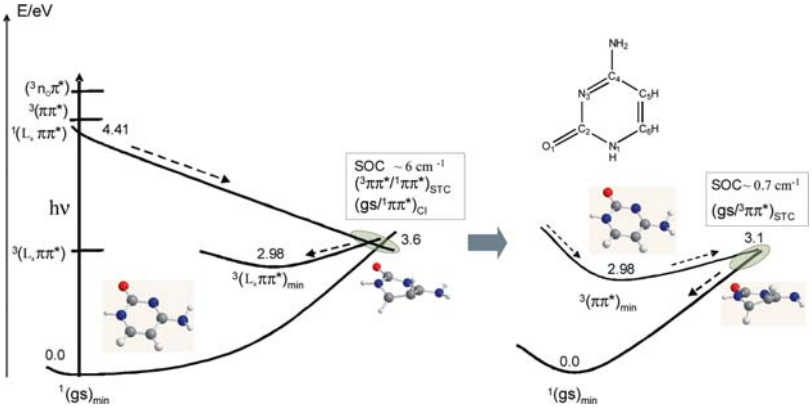


Figure 9. Scheme, based on CASPT2 results, of the photochemistry of cytosine focused on the population of the lowest-energy triplet state.

dependence of the triplet formation quantum yield reported in these three molecules is suggested to be related to the activation of the three (both at high and low excitation energies) or only one (at low energies) ISC channels. On the other hand, guanine and cytosine, having a much lower spectroscopic $^1(\pi\pi^*)$ singlet excited state below the nπ^*-type states, are not expected to display the nπ^*-mediated ISC mechanisms in regions close to the main MEP and may have only efficient ISC funnels at low energies, close to the singlet CI, a feature common to all nucleobases. The present results explain the fact that guanine and cytosine contribute much less to the phosphorescence of DNA, as it has been established.[1−3] It is noteworthy to indicate that the phosphorescence spectrum of RNA was also reported,[77] and it was shown, first, to be determined mainly by the individual properties of the ribonucleotides' π-electron systems, and second, to be composed by triplet signals of adenosine groups and centers of an unknown nature with structureless long-wavelength phosphorescence different from that in DNA. The present results would indicate that adenine and, in this case, uracil nucleobases should be preferably considered as sources of phosphorescence in RNA, as adenine and thymine are in DNA. It is clear that the present results for the isolated systems cannot be directly extrapolated to polymeric DNA/RNA. As already explained before, the described properties should be, however, considered intrinsic features of the nucleobases that, even if they may change in condensed phases or, in general, in the biological environment for the single monomers, are expected to maintain their basic characteristics, as occurs for the singlet states properties and it seems also for triplet states.[77]

Acknowledgment. This research was supported by projects CTQ2007-61260, CTQ2010-14892, and CSD2007-0010 Consolider-Ingenio in Molecular Nanoscience of the Spanish MEC/FEDER and UV-EQUIP09-5764 of the Universitat de València.

Supporting Information Available: Additional computational details, reaction paths, and Cartesian coordinates

of the singular points. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Imakubo, K. *J. Phys. Soc. Jpn.* **1968**, *24*, 1124.

(2) Szerenyi, P.; Dearman, H. H. *Chem. Phys. Lett.* **1972**, *15*, 81.

(3) Arce, R.; Rodríguez, G. *J. Photochem.* **1986**, *33*, 89.

(4) Gueron, M.; Eisinger, J.; Lamola, A. A. Excited States of Nucleic Acids. In *Basic Principles in Nucleic Acid Chemistry*; Tso, P. O. P., Ed.; Academic Press: New York, 1974; Vol. 1, pp 311−398.

(5) Daniels, M. In *Photochemistry and Photobiology of Nucleic Acids*; Wang, S. Y., Ed.; Academic Press: New York, 1976; Vol. 1, pp 23−108.

(6) Görner, H. *J. Photochem. Photobiol. B: Biol.* **1990**, *5*, 359.

(7) Cadet, J.; Vigny, P. In *Bioorganic Photochemistry*; Morrison, H., Ed.; John Wiley & Sons: New York, 1990; Vol. 1, pp 1−272.

(8) Gut, I. G.; Wood, P. D.; Redmond, R. W. *J. Am. Chem. Soc.* **1996**, *118*, 2366.

(9) Wood, P. D.; Redmond, R. W. *J. Am. Chem. Soc.* **1996**, *118*, 4256.

(10) Bosca, F.; Lhiaubet-Vallet, V.; Cuquerella, M. C.; Castell, J. V.; Miranda, M. A. *J. Am. Chem. Soc.* **2006**, *128*, 6318.

(11) Kang, H.; Lee, K. T.; Jung, B.; Ko, Y. J.; Kim, S. K. *J. Am. Chem. Soc.* **2002**, *124*, 12958–12959.

(12) Klessinger, M. In *Theoretical Organic Chemistry - Theoretical and Computational Chemistry*; Párkányi, C., Ed.; Elsevier: Amsterdam, 1998; p 581.

(13) Serrano-Pérez, J. J.; Merchán, M.; Serrano-Andrés, L. *J. Phys. Chem. B.* **2008**, *112*, 14002.

(14) Brown, I. H.; Johns, H. E. *Photochem. Photobiol.* **1968**, *8*, 273.

(15) Schreier, W. J.; Schrader, T. E.; Koller, F. O.; Gilch, P.; Crespo-Hernández, C. E.; Swaminathan, V. N.; Carell, T.; Zinth, W.; Kohler, B. *Science* **2007**, *315*, 625.

(16) Roca-Sanjuán, D.; Olaso-González, G.; González-Ramírez, I.; Serrano-Andrés, L.; Merchán, M. *J. Am. Chem. Soc.* **2008**, *130*, 10768.

(17) Crespo-Hernández, C. E.; Cohen, B.; Hare, P. M.; Kohler, B. *Chem. Rev.* **2004**, *104*, 1977–2019.

(18) Serrano-Andrés, L.; Merchán, M. In *Radiation Induced Molecular Phenomena in Nucleic Acid: A Comprehensive Theoretical and Experimental Analysis*; Shukla, M. K., Leszczynski, J., Eds.; Springer: The Netherlands, 2008; pp 435−472.

(19) Middleton, C. T.; De La Harpe, K.; Su, C.; Law, Y. K.; Crespo-Hernández, C. E.; Kohler, B. *Annu. Rev. Phys. Chem.* **2009**, *60*, 217.

(20) Serrano-Andrés, L.; Merchán, M. *J. Photochem. Photobiol. C: Photochem. Rev.* **2009**, *10*, 21.

(21) Conti, I.; Garavelli, M.; Orlandi, G. *J. Am. Chem. Soc.* **2009**, *131*, 16108.

(22) Conti, I.; Altoè, P.; Stenta, M.; Garavelli, M.; Orlandi, G. *Phys. Chem. Chem. Phys.* **2010**, DOI: 10.1039/b926608a (accessed May 4, 2010).

(23) Merchán, M.; Serrano-Andrés, L.; Robb, M. A.; Blancafort, L. *J. Am. Chem. Soc.* **2005**, *127*, 1820.

(24) Serrano-Pérez, J. J.; González-Luque, R.; Merchán, M.; Serrano-Andrés, L. *J. Phys. Chem. B.* **2007**, *111*, 11880.

(25) Climent, T.; González-Luque, R.; Merchán, M.; Serrano-Andrés, L. *Chem. Phys. Lett.* **2007**, *441*, 327.

(26) Etinski, M.; Fleig, T.; Marian, C. M. *J. Phys. Chem. A* **2009**, *113*, 11809.

(27) Klessinger, M.; Michl, J. *Excited States and Photochemistry of Organic Molecules*; VCH Publishers, Inc.: New York, 1995.

(28) *Computational Photochemistry*; Olivucci, Ed.; Elsevier: Amsterdam, 2005.

(29) Serrano-Andrés, L.; Merchán, M.; Borin, A. C. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 8691.

(30) Tatchen, J.; Gilka, N.; Marian, C. M. *Phys. Chem. Chem. Phys.* **2007**, *9*, 5209.

(31) Salzmann, S.; Tatchen, J.; Marian, C. M. *J. Photochem. Photobio. A* **2008**, *198*, 221.

(32) Carpenter, B. K. *Chem. Soc. Rev.* **2006**, *35*, 736.

(33) Harvey, J. N.; Poli, R.; Smith, K. M. *Coord. Chem. Rev.* **2003**, *238−239*, 347.

(34) González-Navarrete, P.; Coto, P. B.; Polo, V.; Andrés, J. *Phys. Chem. Chem. Phys.* **2009**, *11*, 7189.

(35) Atchity, G. J.; Xantheas, S. S.; Ruedenberg, K. *J. Chem. Phys.* **1991**, *95*, 1862.

(36) Ben-Nun, M.; Molnar, F.; Schulten, K.; Martinez, T. J. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 1769.

(37) Merchán, M.; González-Luque, R.; Climent, T.; Serrano-Andrés, L.; Rodríguez, E.; Reguero, M.; Peláez, D. *J. Phys. Chem.B* **2006**, *110*, 26471.

(38) Serrano-Andrés, L.; Merchán, M.; Borin, A. C. *Chem.—Eur. J.* **2006**, *12*, 6559.

(39) Serrano-Andrés, L.; Merchán, M.; Borin, A. C. *J. Am. Chem. Soc.* **2008**, *130*, 2473.

(40) Blancafort, L. *J. Am. Chem. Soc.* **2006**, *128*, 210.

(41) Perun, S.; Sobolewski, A. L.; Domcke, W. *J. Am. Chem. Soc.* **2005**, *127*, 6257.

(42) Perun, S.; Sobolewski, A. L.; Domcke, W. *J. Phys. Chem. A* **2006**, *110*, 13238.

(43) Marian, C. M. *J. Chem. Phys.* **2005**, *122*, 104314.

(44) Chen, H.; Li, S. H. *J. Phys. Chem. A.* **2005**, *109*, 8443.

(45) Barbatti, M.; Lischka, H. *J. Am. Chem. Soc.* **2008**, *130*, 6831.

(46) Hudock, H. R.; Martinez, T. J. *ChemPhysChem* **2008**, *9*, 2486.

(47) Canuel, C.; Mons, M.; Pluzzi, F.; Tardivel, B.; Dimicoli, I.; Elhanine, M. *J. Chem. Phys.* **2005**, *122*, 074316.

(48) Bernardi, F.; Olivucci, M.; Robb, M. A. *Pure Appl. Chem.* **1995**, *67*, 17.

(49) Takaya, T.; Su, C.; De La Harpe, K.; Crespo-Hernández, C. E.; Kohler, B. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 10285.

(50) Olaso-González, G.; Merchán, M.; Serrano-Andrés, L. *J. Am. Chem. Soc.* **2009**, *131*, 4368.

(51) Strickler, S. J.; Berg, R. A. *J. Chem. Phys.* **1962**, *37*, 814.

(52) Rubio-Pons, O.; Serrano-Andrés, L.; Merchán, M. *J. Phys. Chem. A* **2001**, *105*, 9664.

(53) Veryazov, V.; Widmark, P.-O.; Serrano-Andrés, L.; Lindh, R.; Roos, B. O. *Int. J. Quantum Chem.* **2004**, *100*, 626.

(54) Aquilante, F.; De Vico, L.; Ferré, N.; Ghigo, G.; Malmqvist, P.-Å.; Pedersen, T.; Pitonak, M.; Reiher, M.; Roos, B. O.; Serrano-Andrés, L.; Urban, M.; Veryazov, V.; Lindh, R. *J. Comput. Chem.* **2010**, *31*, 224.

(55) Noguera, M.; Blancafort, L.; Sodupe, M.; Bertran, J. *Mol. Phys.* **2006**, *104*, 925.

(56) Fleig, T.; Knecht, S.; Hättig, C. *J. Phys. Chem. A* **2007**, *111*, 5482.

(57) Zgierski, M. Z.; Patchkovskii, S.; Lim, E. C. *Can. J. Chem.* **2007**, *85*, 124.

(58) Cohen, B. J.; Goodman, L. *J. Am. Chem. Soc.* **1965**, *87*, 5487.

(59) Lavík, J.; Jelínek, O.; Plášek, J. *Photochem. Photobiol.* **1979**, *29*, 491.

(60) Sobolewski, A. L.; Domcke, W. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2763.

(61) Fülscher, M. P.; Serrano-Andrés, L.; Roos, B. O. *J. Am. Chem. Soc.* **1997**, *119*, 6168.

(62) Lower, S. K.; El-Sayed, M. A. *Chem. Rev.* **1966**, *66*, 199.

(63) Ullrich, S.; Schultz, T.; Zgierski, M. Z.; Stolow, A. *J. Am. Chem. Soc.* **2004**, *126*, 2262.

(64) Ullrich, S.; Schultz, T.; Zgierski, M. Z.; Stolow, A. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2796.

(65) Climent, T.; González-Luque, R.; Merchán, M.; Serrano-Andrés, L. *J. Phys. Chem. A* **2006**, *110*, 13584.

(66) Serrano-Pérez, J. J.; Merchán, M.; Serrano-Andrés, L. *Chem. Phys. Lett.* **2007**, *434*, 107.

(67) Gustavsson, T.; Bányász, A.; Lazzarotto, E.; Markovitsi, D.; Scalamani, G.; Frisch, M. J.; Barone, V.; Improta, R. *J. Am. Chem. Soc.* **2006**, *128*, 607.

(68) Hare, P. M.; Crespo-Hernández, C. E.; Kohler, B. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 435.

(69) Barenboim, G. M.; Domanskii, A. N. *Luminescence of biopolymers and cells*; Plenum Press: New York, 1969; p 85.

(70) Snipes, W. *Electron spin resonance and the effects of radiation on biological systems*; National Academy of Sciences: Washington, DC, 1966, p55.

(71) Gupron, M.; Shulman, R. G.; Eisinger, J. *Proc. Natl. Acad. Sci. U.S.A.* **1966**, *56*, 814.

(72) Marian, C. M. *J. Phys. Chem. A* **2007**, *111*, 1545.

(73) Nikogosyan, D. N.; Letokhov, L. S. *Riv. Nuov. Cim.* **1983**, *6*, 1.

(74) Nguyen, M. T.; Zhang, R.; Nam, P.-C.; Ceulemans, A. *J. Phys. Chem. A* **2004**, *108*, 6554.

(75) Blancafort, L. *Photochem. Photobiol.* **2007**, *83*, 603.

(76) Merchán, M.; Serrano-Andrés, L. *J. Am. Chem. Soc.* **2003**, *125*, 8108.

(77) Kudrya, V. Y.; Yashchuk, V. M.; Levchenko, S. M.; Melnik, V. I.; Zaika, L. A.; Govorun, D. M. *Mol. Cryst. Liq. Cryst.* **2008**, *497*, 425.

CT100164M

Singlet-Triplet States Interaction Regions in DNA/RNA Nucleobase Hypersurfaces

Remedios González-Luque, Teresa Climent, Israel González-Ramírez,
Manuela Merchán, Luis Serrano-Andrés*

Supporting Information

Prof. Remedios González-Luque, Prof. M. Merchán, Prof. L. Serrano-Andrés
Instituto de Ciencia Molecular, Universitat de València
Apartado 22085, ES-46071 Valencia, Spain
Fax: (+34) 96-3544427
E-mail: Luis.Serrano@uv.es

Additional computational details

All calculations reported in the present contribution have been performed using the CASPT2//CASSCF protocol, in which geometry optimizations, including minima, surface crossings, and minimum energy paths (MEP) were carried at the multiconfigurational CASSCF level, whereas electronic energy calculations use the second-order multiconfigurational perturbation approach, CASPT2.[1,2,3,4] Different active spaces have been employed in order to ensure the converged character of the results. For the pyrimidine bases, thymine, uracil, and cytosine, the final calculations used the ANO-type one-electron basis set (ANO-S) contracted to C, N, O [3s2p1d] / H [2s1p].[5] MEP optimizations employed a full ππ* active space plus three additional extra correlating π* MOs in order to balance dynamic correlation effects. The nature of the orbitals does not essentially change along the optimization process. The size of the active space is then 10 electrons distributed in 11 orbitals. Initially, geometry optimizations to obtain planar minima were performed with a valence π active space. However, addition of the extra correlating orbitals proved not to change the outcome and provide such local minima also at the highest level of calculation. The final calculations represented in the figures, at each of the CASSSF optimized structures, used the ANO-S-type one-electron basis set contracted to C, N, O [3s2p1d] / H [2s1p] (hereafter ANO-S) and an active space comprising 14 electrons distributed in 10 orbitals, including also lone pairs, and can be therefore labeled CASPT2//CASSCF(14,10)/ANO-S C, N, O [3s2p1d] / H [2s1p]. For purine nucleobases, adenine and guanine, initially, a π active space of 12 electrons and 10 orbitals (the lowest nodeless π orbital was kept inactive) was used to optimize ππ*-type states, whereas for nπ* states the space was increased to 14 electrons and 12 orbitals, adding lone-pair in-plane orbitals. Additional extra-valence orbitals were included in the active space to assure the proper behavior of the computed paths, confirming the selection performed. Calculations used a double-zeta plus polarization 6-31G(d,p) one-electron basis set in all atoms. Test calculations were performed by using a larger ANO-S type basis set contracted to C,N,O [3s2p1d] / H [2s1p] to analyze the effect of increasing the quality of the basis set. In all cases, the standard zeroth-order Hamiltonian was employed in the CASPT2 calculations, which include an imaginary level-shift correction

of 0.2 au in order to avoid the presence of intruder states.[6] All calculations used the MOLCAS-7.0 set of programs.[7]

The MEPs have been built as steepest descendent paths in a procedure[8] which is based on a modification of the projected constrained optimization (PCO) algorithm of Anglada and Bofill[9] and follows the Müller-Brown approach.[10] Each step requires the minimization of the PES on a hyperspherical cross section of the PES centered on the initial geometry and characterized by a predefined radius. The optimized structure is taken as the center of a new hypersphere of the same radius, and the procedure is iterated until the bottom of the energy surface is reached. Mass-weighted coordinates are used, therefore the MEP coordinate corresponds to the so-called Intrinsic Reaction Coordinate (IRC), measured in au, that is, bohr·(amu)$^{1/2}$. Each step in the abscissa coordinate in the figures corresponds to a step of the MEP computed with a hypersphere radius constrained to 1.06 au. The full procedure is currently implemented in the MOLCAS-7.0 package[7] and its technical description has been published elsewhere.[8] As mentioned above, at each optimized point of the MEP, CASPT2//CASSCF(14,10)/ANO-S and CASPT2//CASSCF(14,12)/6-31G(d,p) calculations were performed for the states of interest for pyrimidines and purines nucleobases, respectively. The MEP search showed to be insensitive to different selections of the step length, larger and shorter than the one reported.

In this contribution conical intersection searches were performed using the restricted Lagrange multipliers technique as included in the MOLCAS-7.0 package[7] in which the lowest-energy point was obtained under the restriction of degeneracy between the two considered states.[8] No non-adiabatic coupling elements were computed.

From the calculated CASSCF transition dipole moments (TDM) and the CASPT2 excitation energies, the radiative lifetimes have been estimated by using the Strickler-Berg relationship,[11] as explained elsewhere.[12] The spin-orbit coupling (SOC) strength between selected states was computed within the AMFI framework obtaining as the length of the spin-orbit coupling vector with the algorithms implemented in the MOLCAS-7.0 quantum-chemistry program, as described previously.[13]

Additional MEPs and molecular geometries of critical points

Figures S1 to S12 include the additional minimum energy paths mentioned in the text and that strengthen the consistency of the study, linking the crossing regions with the final lowest triplet state. Those for uracil can be found in a previous publication.[14] Regarding cytosine, the MEP on Figure S12 that links the (gs/$^3\pi\pi^*$) STC with the $^3(\pi\pi^*)$ minimum is performed using the 6-31G(d,p) instead of the ANO basis employed in the paper. We have include it here just for consistency, because no changes in the conclusions can be expected with the change of the basis set.

Table SI1 compiles the Cartesian coordinates of the optimized minima and surface crossings obtained in the present contribution for thymine, cytosine, adenine, and guanine at the described CASSCF level of theory. All remaining structures mentioned in the paper but not included here (singlet or triplet states or state crossings), have been already published and can be found in the respective papers: adenine,[15] guanine, [16] thymine, [17] uracil, [14] and cytosine, [13] together with further details, if needed. Total energies at the CASSCF(14,12)/6-31G(d,p) (purines) or CASSCF(14,10)/ANO (pyrimidines) level have been also included.

References

1 K. Andersson, P.-Å. Malmqvist and B. O. Roos, *J. Chem. Phys.* 1992, **96**, 1218.

2 B. O. Roos, M. P. Fülscher, P.-Å. Malmqvist, L. Serrano-Andrés, K. Pierloot and M. Merchán, *Adv. Chem. Phys.* 1996, ***93***, 219.

3 M. Merchán, L. Serrano-Andrés, M. P. Fülscher and B. O. Roos, in *Recent Advances in Multireference Methods*, ed. K. Hirao, World Scientific Publishing, Singapore, 1999.

4 M. Merchán and L. Serrano-Andrés, in *Computational Photochemistry*, ed. M. Olivucci, Elsevier Publishing, Amsterdam, 2005

5 K. Pierloot, B. Dumez, P. –O. Widmark and B. O. Roos, *Theor. Chim. Acta.* 1995, **90**, 87.

6 N. Forsberg and P.-Å. Malmqvist, *Chem. Phys. Lett*. 1997, **274**, 196.

7 F. Aquilante, L. De Vico, N. Ferré, G. Ghigo, P.-Å. Malmqvist, T. Pedersen, M. Pitonak, M. Reiher, B. O. Roos, L. Serrano-Andrés, M. Urban, V. Veryazov and R. Lindh, *J. Comp. Chem*. 2009, in press.

8 L. De Vico, M. Olivucci and R. Lindh, *J. Chem. Theory Comp.* 2005, **1**, 1029-1037.

9 J. M. Anglada and J. M. Bofill, *J. Comput. Chem.* 1997, **18**, 992-1003.

10 K. Müller and L. D. Brown, *Theor. Chim. Acta.* 1979, **53**, 75-93.

11 S. J. Strickler and R. A. Berg, *J. Chem. Phys.* 1962, **37**, 814.

12 O. Rubio-Pons, L. Serrano-Andrés and M. Merchán, *J. Phys. Chem. A* 2001, **105**, 9664-9676.

13 M. Merchán, L. Serrano-Andrés, M. A. Robb and L. Blancafort, *J. Am. Chem. Soc.* 2005, **127**, 1820.

14 T. Climent, R. González-Luque, M. Merchán and L. Serrano-Andrés, *Chem. Phys. Lett.* 2007, **441**, 327-331.

15 L. Serrano-Andrés, M. Merchán and A. C. Borin, *Chem. Eur. J.* 2006, **12**, 6559.

16 L. Serrano-Andrés, M. Merchán and A. C. Borin, *J. Am. Chem. Soc.* 2008, **130**, 2473-2484.

17 J. J. Serrano-Pérez, R. González-Luque, M. Merchán and L. Serrano-Andrés, *J. Phys. Chem. B*. 2007, **111**, 11880-11883.

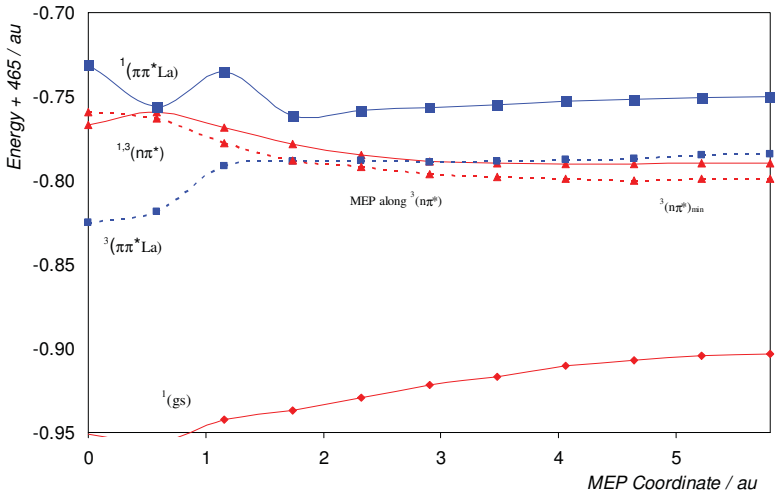


Figure S1. *Evolution of the lowest singlet and triplet states for adenine from the* $(^3n\pi^*/^1\pi\pi^*L_a)_{STC}$ *geometry along the* $^3(n\pi^*)$ *MEP.*

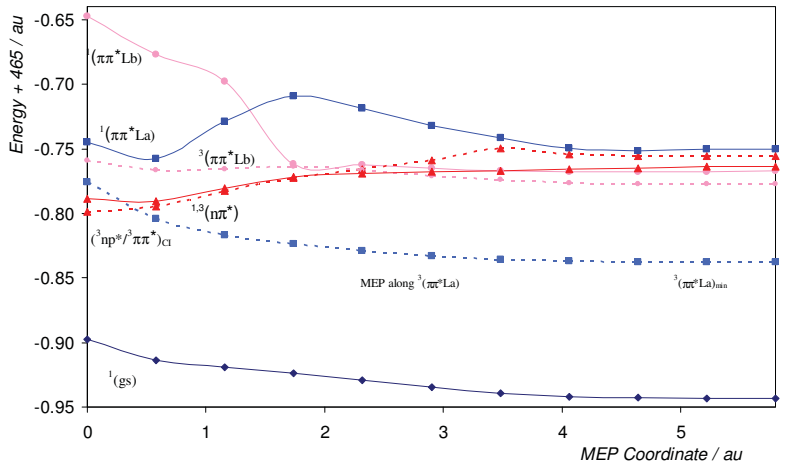


Figure S2. *Evolution of the lowest singlet and triplet states for adenine from the* $(^3n\pi^*/^3\pi\pi^*)_{CI}$ *geometry along the* $^3(\pi\pi^*)$ *MEP.*

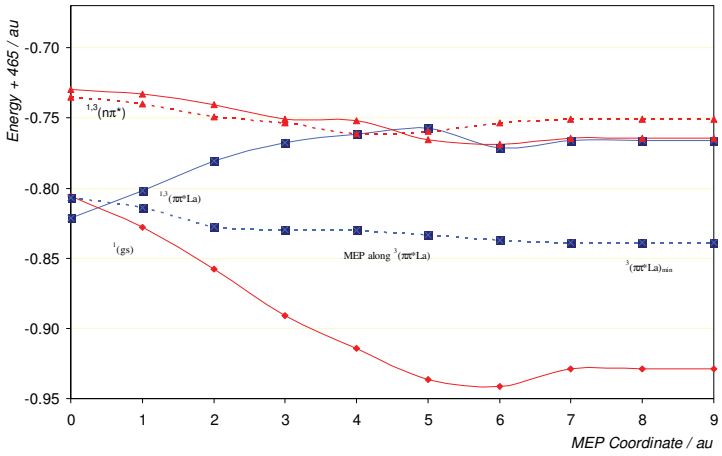


Figure S3. *Evolution of the lowest singlet and triplet states for adenine from the* $(^3\pi\pi^*/^1\pi\pi^*L_a)_{STC}$ *geometry along the* $^3(\pi\pi^*)$ *MEP.*

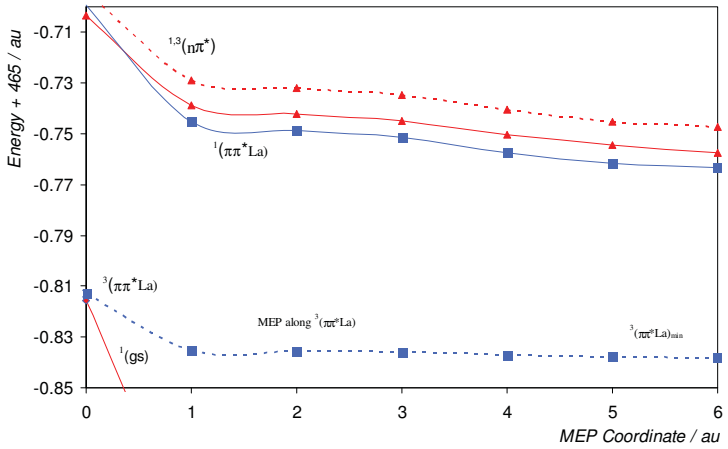


Figure S4. *Evolution of the lowest singlet and triplet states for adenine from the* $(^1gs/^3\pi\pi^*)_{STC}$ *geometry along the* $^3(\pi\pi^*)$ *MEP.*

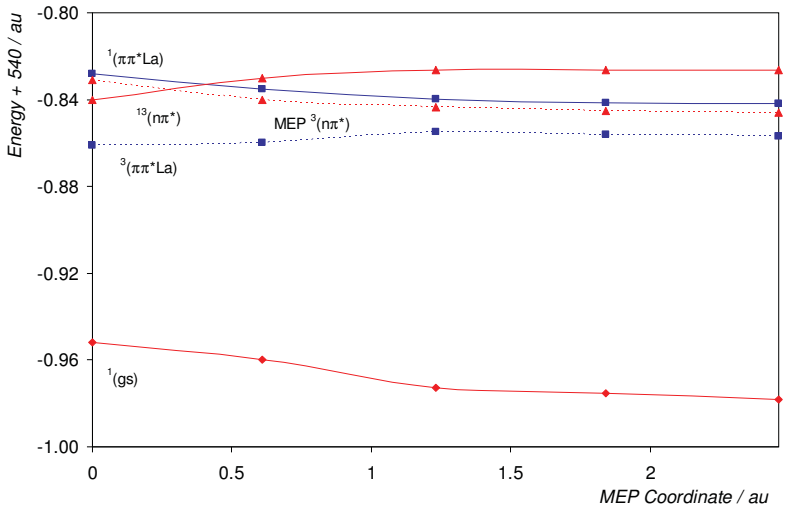


Figure S5. *Evolution of the lowest singlet and triplet states for guanine from the* $(^3n\pi^*/^1\pi\pi^*L_a)_{STC}$ *geometry along the* $^3(n\pi^*)$ *MEP.*

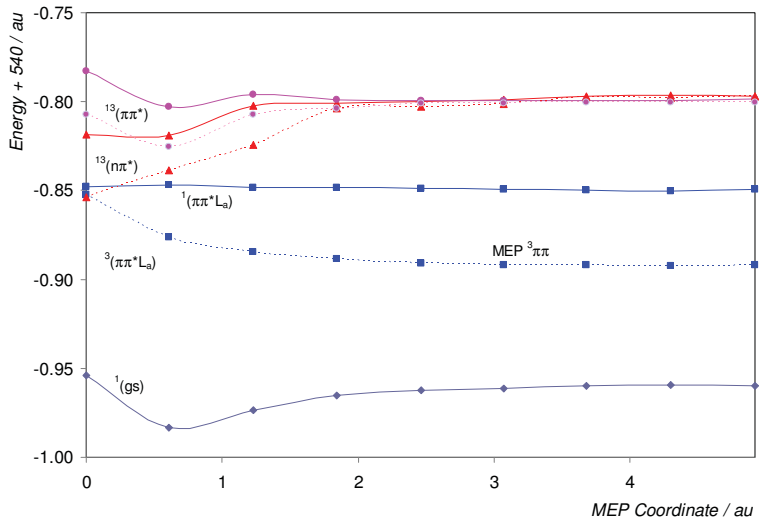


Figure S6. *Evolution of the lowest singlet and triplet states for guanine from the* $(^3n\pi^*/^3\pi\pi^*)_{CI}$ *geometry along the* $^3(\pi\pi^*)$ *MEP.*

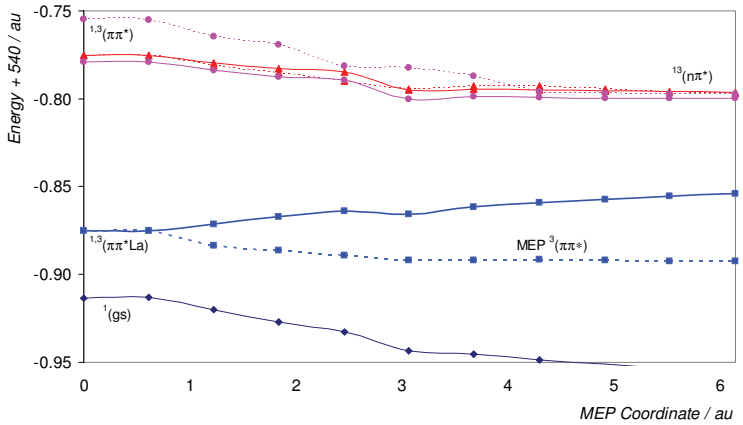


Figure S7. *Evolution of the lowest singlet and triplet states for guanine from the* $(^3\pi\pi^*/^1\pi\pi^*L_a)_{STC}$ *geometry along the* $^3(\pi\pi^*)$ *MEP.*

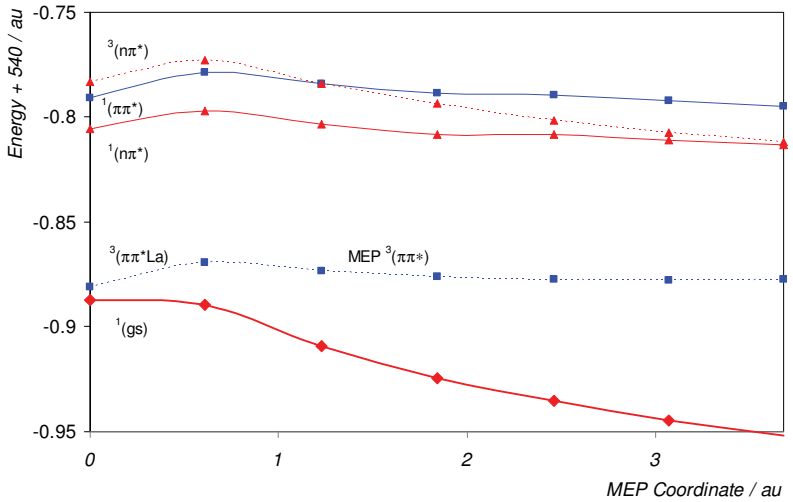


Figure S8. *Evolution of the lowest singlet and triplet states for guanine from the* $(^1gs/^3\pi\pi^*)_{STC}$ *geometry along the* $^3(\pi\pi^*)$ *MEP.*

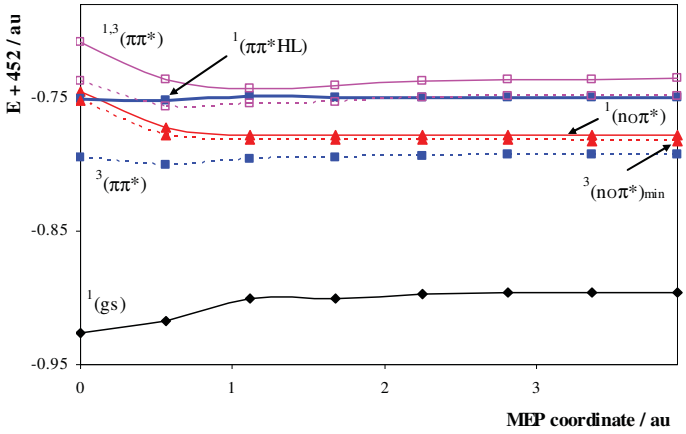


Figure S9. *Evolution of the lowest singlet and triplet states for thymine from the* $(^3n\pi^*/^1\pi\pi^*L_a)_{STC}$ *geometry along the* $^3(n\pi^*)$ *MEP.*

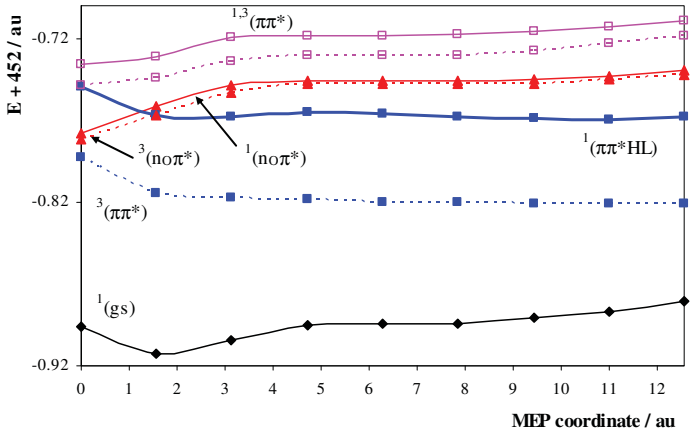


Figure S10. *Evolution of the lowest singlet and triplet states for thymine from the* $(^3n\pi^*/^3\pi\pi^*)_{CI}$ *geometry along the* $^3(\pi\pi^*)$ *MEP.*

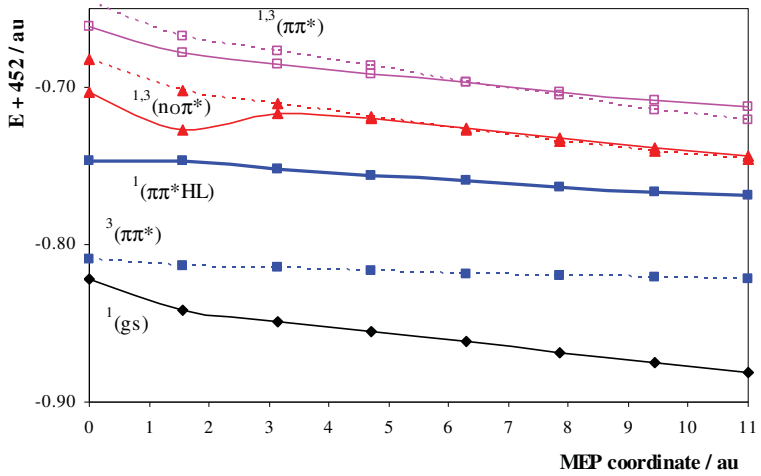


Figure S11. *Evolution of the lowest singlet and triplet states for thymine from the* $(^1gs/^3\pi\pi^*)_{STC}$ *geometry along the* $^3(\pi\pi^*)$ *MEP.*

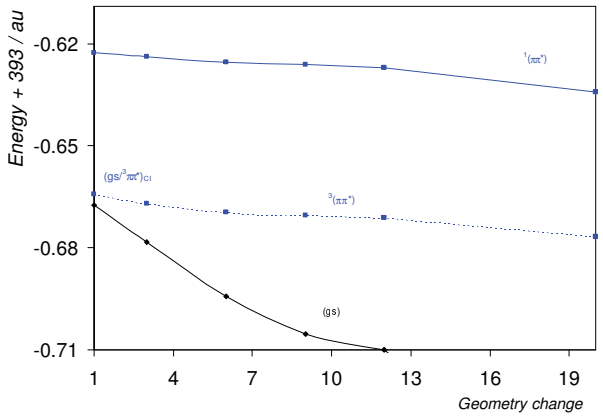


Figure S12. *Evolution of the lowest singlet and triplet states for cytosine from the* $(^1gs/^3\pi\pi^*)_{STC}$ *geometry along the* $^3(\pi\pi^*)$ *MEP. Basis set 6-31G(d,p).*

Table S1. Cartesian coordinates (Å) of the CASSCF optimized structures obtained in the paper. Total energies at the CASSCF level.

Atom	x	y	z
		Adenine $^3(\pi\pi)_{min}$	
		-464.518933 au	
N	-1.873106	-0.436581	-0.180080
N	-0.797017	0.147208	1.781233
N	1.576275	0.506736	1.311084
N	1.636597	-0.003888	-0.867996
N	-1.021581	-0.347982	-2.445799
C	-1.879446	-0.164359	1.141585
C	0.304490	0.143556	0.958971
C	0.363581	-0.184856	-0.375378
C	-0.811989	-0.649077	-1.089756
C	2.335005	0.407513	0.181384
H	-0.202088	0.049138	-2.856490
H	-1.295531	-1.150542	-2.973696
H	-2.824944	-0.214278	1.645697
H	1.871981	0.800276	2.210945
H	3.379295	0.638066	0.167407
		Adenine $(^3n\pi^*)_{min}$	
		-464.470424 au	
N	-1.873106	-0.436581	-0.180080
N	-0.797017	0.147208	1.781233
N	1.576275	0.506736	1.311084
N	1.636597	-0.003888	-0.867996
N	-1.021581	-0.347982	-2.445799
C	-1.879446	-0.164359	1.141585
C	0.304490	0.143556	0.958971
C	0.363581	-0.184856	-0.375378
C	-0.811989	-0.649077	-1.089756
C	2.335005	0.407513	0.181384
H	-0.202088	0.049138	-2.856490
H	-1.295531	-1.150542	-2.973696
H	-2.824944	-0.214278	1.645697
H	1.871981	0.800276	2.210945
H	3.379295	0.638066	0.167407

Table S1. Cartesian coordinates (Å) of the CASSCF optimized structures obtained in the paper. Total energies at the CASSCF level (cont.)

Atom	x	y	z
		Adenine $(gs/^3\pi\pi^*)_{STC}$	
		-464.329168 au (gs)	
N	-1.85770	-0.20390	-0.33861
N	-0.62700	0.08001	1.71151
N	1.79375	-0.09384	1.44644
N	1.73865	-0.27256	-0.88297
N	-0.84533	-0.47300	-2.40604
C	-1.75043	-0.00356	0.98338
C	0.46510	-0.04755	1.00481
C	0.51322	-0.17234	-0.42451
C	-0.76211	-0.30056	-1.06767
C	2.38760	0.48814	0.26439
H	-0.02489	-0.74321	-2.89654
H	-1.72446	-0.74313	-2.78107
H	-2.67474	0.08090	1.52141
H	1.95824	0.39089	2.30520
H	2.17161	1.55463	0.16937
		Adenine $(^3\pi\pi^*/^1\pi\pi^*)_{STC}$	
		-464.481417 au $^3(\pi\pi^*)$	
N	-2.09327	-0.16070	-0.40123
N	-0.93620	0.63986	1.48677
N	1.54931	0.65545	1.34155
N	1.61520	-0.18663	-0.74804
N	-0.90284	-0.37271	-2.35096
C	-1.80841	-0.27423	0.96355
C	0.27055	0.26929	1.04223
C	0.32382	-0.25622	-0.25479
C	-0.96590	-0.28246	-1.01751
C	2.31763	0.31460	0.25131
H	-0.06054	-0.66054	-2.78606
H	-1.76192	-0.43895	-2.84744
H	-2.01257	-1.19269	1.47959
H	1.84351	1.11753	2.16848
H	3.38314	0.36929	0.28166

Table S1. Cartesian coordinates (Å) of the CASSCF optimized structures obtained in the paper. Total energies at the CASSCF level (cont.)

Atom	x	y	z
	Adenine $(^3n\pi^*/^1\pi\pi^*)_{STC}$		
	-464.409053 au $^3(n\pi^*)$		
N	-2.513550	-0.188019	0.203735
N	-1.058960	-2.203888	0.457779
N	1.281876	-1.637695	0.243362
N	1.065247	0.570084	-0.200424
N	-1.617493	1.934082	0.322235
C	-2.173538	-1.522827	-0.100037
C	-0.052628	-1.402863	0.236023
C	-0.162506	0.007889	-0.038528
C	-1.500671	0.606567	0.129684
C	1.913483	-0.435226	-0.051117
H	-0.892792	2.520767	-0.026138
H	-2.545029	2.297897	0.291641
H	-2.975052	-2.139984	-0.457779
H	1.705460	-2.520752	0.395111
H	2.975052	-0.358673	-0.125992
	Adenine $(^3\pi\pi^*/^3n\pi^*)_{CI}$		
	-464.469526 au $^3(\pi\pi^*)$		
N	-1.877914	-0.429385	-0.174221
N	-0.795711	0.153925	1.782894
N	1.577913	0.497135	1.307418
N	1.631707	-0.022015	-0.870076
N	-1.040473	-0.380338	-2.446262
C	-1.881657	-0.150572	1.145588
C	0.304301	0.137143	0.959494
C	0.358868	-0.198058	-0.374194
C	-0.822041	-0.651225	-1.086237
C	2.333072	0.393241	0.175655
H	-0.305591	0.172248	-2.837161
H	-1.146208	-1.213789	-2.985955
H	-2.827126	-0.190633	1.650639
H	1.874419	0.802283	2.203156
H	3.377964	0.620968	0.158373

Table S1. Cartesian coordinates (Å) of the CASSCF optimized structures obtained in the paper. Total energies at the CASSCF level (cont.)

Atom	x	y	z
		Guanine $^3(\pi\pi^*)_{min}$	
		-539.414314 au	
O	-1.254791	2.664404	-0.214125
C	-2.324959	-0.714905	0.216380
C	-0.066355	-0.759401	0.060575
C	0.019163	0.681925	0.118925
C	-1.206601	1.467802	-0.082821
C	2.014718	-0.011693	0.119009
N	-2.300766	0.646304	-0.196869
N	-1.122776	-1.482457	0.037299
N	1.305363	1.089117	0.153065
N	1.254360	-1.152987	0.037836
N	-3.475241	-1.400161	-0.205309
H	-3.166074	1.087491	-0.412524
H	-4.285932	-1.158736	0.327257
H	-3.327762	-2.386081	-0.165228
H	1.580888	-2.089641	0.054588
H	3.083422	-0.044564	0.151943
		Guanine $(^3n_O\pi^*)_{min}$	
		-539.358824 au	
O	-1.250893	2.668777	-0.192798
C	-2.330550	-0.714693	0.231242
C	-0.067406	-0.752839	0.064253
C	0.023182	0.689564	0.089058
C	-1.203824	1.468557	-0.066292
C	2.017056	-0.011357	0.105901
N	-2.320647	0.656811	-0.126475
N	-1.125043	-1.476151	0.051134
N	1.315949	1.092715	0.105805
N	1.251780	-1.151813	0.059055
N	-3.488830	-1.383871	-0.203285
H	-3.142244	1.043031	-0.531794
H	-4.277998	-1.185223	0.377437
H	-3.332980	-2.369286	-0.225164
H	1.572982	-2.088247	0.126926
H	3.086144	-0.049535	0.134986

Table S1. Cartesian coordinates (Å) of the CASSCF optimized structures obtained in the paper. Total energies at the CASSCF level (cont.)

Atom	x	y	z
	Guanine $(gs/^3\pi\pi^*)_{STC}$		
	-539.349874 au (gs)		
O	-1.330657	2.652379	0.020296
C	-2.223970	-0.772992	0.037585
C	0.003356	-0.684734	-0.025513
C	0.077565	0.715582	-0.001489
C	-1.188867	1.459462	-0.000540
C	2.018048	-0.058502	0.271497
N	-2.289591	0.599049	-0.008511
N	-1.110755	-1.439812	0.045589
N	1.315547	1.144320	-0.207084
N	1.295600	-1.156305	-0.316555
N	-3.399047	-1.448842	0.044368
H	-3.179564	1.049786	0.004234
H	-4.287538	-0.996420	-0.004181
H	-3.353674	-2.438512	0.030649
H	1.500083	-2.057464	0.056075
H	3.071967	-0.054716	-0.113372
	Guanine $(^1\pi\pi^*/^3n_O\pi^*)_{STC}$		
	-359.308477 au $^1(\pi\pi^*)$		
O	-1.297016	2.699949	-0.186928
C	-2.317487	-0.727539	0.183721
C	-0.080077	-0.736447	0.062421
C	0.012558	0.707578	0.063136
C	-1.150559	1.381171	-0.057954
C	2.011397	-0.002261	0.114076
N	-2.339690	0.682867	-0.062192
N	-1.161107	-1.442106	0.083497
N	1.352277	1.104260	0.111586
N	1.237534	-1.145061	0.063792
N	-3.529887	-1.349836	-0.167057
H	-3.063320	1.000154	-0.672904
H	-4.215206	-1.276163	0.558306
H	-3.371443	-2.315858	-0.366263
H	1.557734	-2.081335	0.123327
H	3.080970	-0.062931	0.149426

Table S1. Cartesian coordinates (Å) of the CASSCF optimized structures obtained in the paper. Total energies at the CASSCF level (cont.)

Atom	x	y	z
	Guanine ($^3\pi\pi^*/^3n_O\pi^*)_{CI}$		
	-539.370830 au $^3(\pi\pi^*)$		
O	-1.254790	2.664400	-0.214125
C	-2.324960	-0.714905	0.216380
C	-0.066355	-0.759401	0.060575
C	0.019163	0.681925	0.118925
C	-1.206600	1.467800	-0.082821
C	2.014720	-0.011693	0.119009
N	-2.300770	0.646304	-0.196869
N	-1.122780	-1.482460	0.037299
N	1.305360	1.089120	0.153065
N	1.254360	-1.152990	0.037836
N	-3.475240	-1.400160	-0.205309
H	-3.166070	1.087490	-0.412524
H	-4.285930	-1.158740	0.327257
H	-3.327760	-2.386080	-0.165228
H	1.580890	-2.089640	0.054588
H	3.083420	-0.044564	0.151943
	Guanine ($^3\pi\pi^*/^1\pi\pi^*)_{STC}$		
	-539.349100 au $^3(\pi\pi^*)$		
O	-1.275460	2.587850	-0.125651
C	-2.242100	-0.678681	-0.043965
C	-0.080214	-0.835105	-0.028527
C	-0.113403	0.558307	0.380942
C	-1.221380	1.386820	-0.100093
C	1.945100	0.004726	0.283542
N	-2.288250	0.589502	-0.574946
N	-1.104990	-1.497530	-0.444614
N	1.187650	1.019210	0.535111
N	1.249240	-1.146310	-0.064511
N	-3.407470	-1.364610	0.120906
H	-3.128980	1.039630	-0.861409
H	-4.148480	-0.901769	0.598100
H	-3.293320	-2.326080	0.351795
H	1.635610	-2.001400	-0.386457
H	3.013110	0.001856	0.359777

Table S1. Cartesian coordinates (Å) of the CASSCF optimized structures obtained in the paper. Total energies at the CASSCF level (cont.)

Atom	x	y	z
		Cytosine $(^3\pi\pi^*)_{min}$	
		-392.644072 au	
N	-0.226042	-2.574862	-0.123304
N	-0.288399	1.558698	0.013524
N	0.890976	-0.530866	-0.004219
C	-0.343920	-1.181978	-0.018274
C	-1.525725	-0.507206	0.009710
C	0.893529	0.855654	0.003077
C	-1.531452	0.942034	0.043701
O	1.959979	1.437323	0.001530
H	0.628890	-2.901650	0.290654
H	-2.465771	-1.037385	0.003320
H	-2.409802	1.554571	-0.017541
H	-1.011627	-3.066740	0.260002
H	-0.207839	2.554389	0.022512
		Cytosine $(^3n_O\pi^*)_{min}$	
		-392.629247 au	
N	-0.212738	-2.573594	-0.145783
N	-0.282566	1.547120	0.061685
N	0.892830	-0.498535	-0.045531
C	-0.335253	-1.180024	-0.047256
C	-1.536808	-0.488907	-0.040602
C	0.833519	0.759464	-0.011373
C	-1.539369	0.909067	-0.054955
O	1.978400	1.478321	0.004220
H	0.615403	-2.898390	0.320731
H	-2.477411	-1.017954	-0.052893
H	-2.407789	1.521045	0.106887
H	-1.017697	-3.058446	0.203991
H	-0.195121	2.524530	-0.119234

Table S1. Cartesian coordinates (Å) of the CASSCF optimized structures obtained in the paper. Total energies at the CASSCF level (cont.)

Atom	x	y	z
		Cytosine (gs/$^3\pi\pi^*)_{STC}$	
		-392.634838 au (gs)	
N	-0.276338	1.241327	-0.088207
C	-0.107299	2.457091	-0.753487
N	0.157679	2.436770	-2.137625
C	0.063031	1.337980	-2.790218
C	-0.260664	0.060993	-2.122484
C	0.139273	0.049769	-0.701339
O	-0.189006	3.492722	-0.159253
N	0.156283	1.343443	-4.156370
H	0.474510	2.202961	-4.550116
H	-1.182815	-0.412907	-2.416296
H	1.107122	-0.346967	-0.426985
H	0.527704	0.526645	-4.587773
H	-0.209484	1.326057	0.902943
		Cytosine ($^3\pi\pi^*/^1\pi\pi^*)_{STC}$	
		-392.630231 au $^3(\pi\pi^*)$	
N	0.82400524	-2.37265803	0.03969720
N	-0.92826155	1.29673462	-0.07625723
N	1.02640659	-0.10939108	0.17850470
C	0.24988218	-1.14952144	0.02989101
C	-1.18372560	-1.01525374	-0.21835132
C	0.50897374	1.14641534	-0.00949887
C	-1.67078253	0.23012805	0.24549471
O	1.15123308	2.15536549	-0.08084835
H	1.80962662	-2.42154253	-0.08998982
H	-1.41001662	-1.26755082	-1.24670856
H	-2.53518961	0.38746828	0.86767798
H	0.27237214	-3.15679750	-0.21805146
H	-1.22779595	2.24209073	0.05503281

6.5 Paper V

Cyclobutane pyrimidine photodimerization of DNA/RNA nucleobases in the triplet state.

T. Climent, I. González-Ramírez, R. González-Luque, M. Merchán and L. Serrano-Andrés

J. Phys. Chem. Lett., 1, 2072-2076 (2010).

THE JOURNAL OF PHYSICAL CHEMISTRY Letters

pubs.acs.org/JPCL

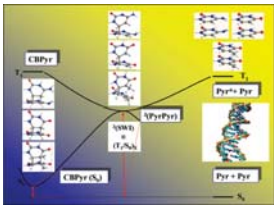
Cyclobutane Pyrimidine Photodimerization of DNA/RNA Nucleobases in the Triplet State

Teresa Climent, Israel González-Ramírez, Remedios González-Luque, Manuela Merchán, and Luis Serrano-Andrés*

Instituto de Ciencia Molecular, Universitat de València Apartado 22085, ES-46071 Valencia, Spain

ABSTRACT The photoinduced formation of cyclobutane pyrimidine dimers in the triplet excited state of the DNA/RNA pyrimidine nucleobases pairs has been studied at the CASPT2 level of theory. A stepwise mechanism through the triplet state of the homodimer is proposed for the pairs of nucleobases cytosine, thymine, and uracil involving a singlet−triplet crossing intermediary structure of biradical character representing the most favorable triplet state conformation of the nucleobases as found in the DNA environment. The efficiency of the mechanism will be modulated by two factors: the effectiveness of the triplet−triplet energy transfer process from a donor photosensitizer molecule, which relates to the relative position of the intermediate in the three acceptor systems, determined here to be lower in energy in the thymine and uracil dimers than in the cytosine pairs, and that of the intersystem crossing process toward the ground state of the photoproduct.

SECTION Dynamics, Clusters, Excited States



One of the most notorious examples of evidence of the photosensitivity of the genetic material to the action of ultraviolet (UV) light is the photoinduced formation of cyclobutane pyrimidine dimers (CBPyr or Pyr < > Pyr, see Figure 1) by pairs of DNA/RNA intrastrand adjacent pyrimidine nucleobases.[1] Those adducts constitute a major source of photoinduced DNA/RNA lesions, leading even to photomutagenesis and photocarcinogenesis, particularly in cellular DNA.[2] The process takes place in biological environments, solvents, and in the solid phase.[1] Femtosecond spectroscopy has proved that thymine (T) dimerization is an ultrafast photoreaction in which cyclobutane thymine dimers (CBT) are fully formed ∼1 ps after UV illumination.[3] Theoretical determinations have confirmed in thymine and cytosine (C) dimers a mechanism for a corresponding ultrafast nonadiabatic photoreaction mediated by the presence of a conical intersection (CI), an energy-degenerate structure between the low-lying singlet excited (S_1) and the ground state (S_0).[4−7] In those quantum-chemical CASPT2 studies, the [2 + 2] photocycloaddition reaction leading to the formation of CBT and cyclobutane cytosine (CBC) dimers was characterized in the singlet manifold.[4−7] Barrierless relaxation paths from favorable conformations of the nucleobases were shown to lead from an initially irradiated singlet state to a shearing-type CI structure, in which the nucleobases ethylenic C_5-C_6 and $C_{5'}-C_{6'}$ bonds laid parallel (parallelogram-type) and elongated, connecting the S_1 and S_0 states and allowing an efficient internal conversion process. Intrastrand nucleobase sequence and relative orientations were also proven to be essential for an efficient photoreaction to take place. This is particularly true for those conformations maximizing the overlap between the π structures of stacked nucleobases, which formed favorable excimer arrangements.[4−7] They were shown to yield the most stable structures leading to the photoreactive arrangements, in agreement with the higher yields obtained for photoproducts with cis-type parallel face-to-face conformations for the base pairs.[1,4] We attributed the low yield measured for the CBC formation as compared to CBT to the stable excimer conformations, found in CC combinations lower in energy than the CI structure and opposite to what occurs in TT pairs.[5]

Apart from the nonadiabatic photoinduced reaction in the singlet manifold taking place after irradiation in the middle or far-UV, formation of the photoproducts mediated by the nucleobases triplet state is also conceivable from the initially excited singlet state of the monomer nucleobase to the ground state of the dimer photoproduct through successive intersystem crossing (ISC) processes.[8] As it has been documented in detail experimentally[9] and computationally for cytosine,[10] uracil,[11,12] and thymine,[12,13] the lowest triplet state of the nucleobase can be populated along the ultrafast singlet-mediated internal conversion by different ISC mechanisms. Regarding dimers or oligomers, the experimental conclusions are not clear. Whereas recent time-resolved studies of CBT formation after irradiation of $(dT)_{20}$ oligonucleotides were unsuccessful at detecting the presence of the triplet state within a 200 ns time frame,[14] there has also been

Received Date: May 10, 2010
Accepted Date: June 15, 2010
Published on Web Date: June 21, 2010

DOI: 10.1021/jz100601p | J. Phys. Chem. Lett. 2010, 1, 2072–2076

THE JOURNAL OF PHYSICAL CHEMISTRY Letters

pubs.acs.org/JPCL

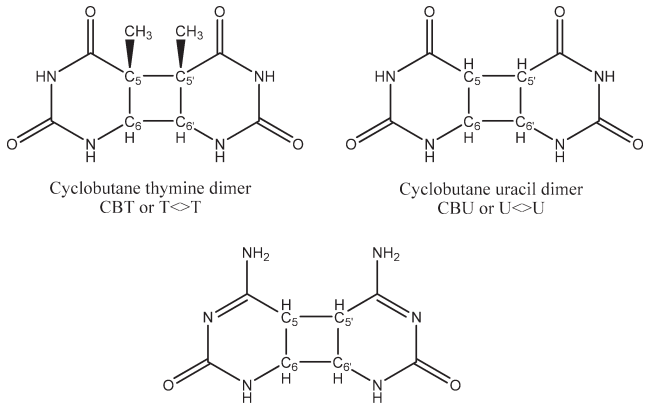


Figure 1. Structures and labeling of the DNA/RNA CBPyr homodimer photoproducts.

evidence of the ultrafast participation of the triplet state in the formation of the TT dimer.[8] Even though the role of the triplet state of DNA oligomers in the dimerization process after direct singlet irradiation has still to be clarified, the participation of the triplet states in DNA photochemistry is uncontroversial. CBPyr adducts are notoriously formed upon nucleobase photosensitization by triplet−triplet energy transfer (TET) from favorable donors, such as different ketone or phthalimidine derivatives.[1,2,15−18] The role of the nucleobases triplet states therefore seems to be linked to the ability of stacked DNA/RNA nucleobases to behave as acceptors under favorable triplet photosensitization conditions. As the efficiency of TET processes is strongly related to the relative triplet excited state energies of the donor and acceptor chromophores, the characterization of the nucleobases triplet intermediate responsible for the photosensitization action shall provide relevant insight into the photoreactive process. Great effort has been devoted to the determination of the nucleobases triplet (T_1) energy. Recent experimental studies on the TET formation of CBT dimers restrict the threshold observed for a given photosensitizer to become a potential DNA photodamager to substances with triplet state energies higher than 2.6−2.8 eV.[18,19] Such limit was considered to represent the triplet energy of thymine in DNA, known to be much lower than that of the isolated nucleobase. In order to gain insight into the mechanisms of CBPyr adducts formation on the triplet state for the three nucleobase homodimers, CC, UU, and TT, this contribution reports calculations whose results are summarized in Table 1 and Figure 2, which also contain some of the data obtained in our previous study on CBC.[4] Multiconfigurational CASPT2(14/10) and CASPT2(12/12) calculations for the monomers and the dimers, respectively, and ANO-S basis sets contracted to C,N,O[3s2p1d]/H[2s1p] were used, including CASSCF geometry optimizations of the singlet and triplet states minima, and minimum energy path (MEP) and minimum energy crossing point (MECP, singlet−triplet)

Table 1. Experimental and Theoretical CASPT2 Energies (eV) for the Lowest-Lying Triplet Excited State (T_1) of the DNA/RNA Pyrimidine Nucleobases[a]

	monomer		dimer	monomer
	E_{VA}[b]	T_e[c]	T_e $(T_1/S_0)_X$[d]	Exp[e]
cytosine	3.66	2.98	2.70	3.50
uracil	3.80	3.15	2.47	3.65
thymine	3.59	2.87	2.36	3.60

[a] All energies referred to the singlet ground state of two isolated nucleobases. See Figure 2. [b] Vertical singlet−triplet excitation energy (E_{VA}) at the monomer ground-state (S_0) Franck−Condon minimum. [c] Adiabatic S_0−T_1 monomer minimum to minimum energy (T_e): electronic excitation band origin. [d] Adiabatic S_0−T_1 homodimer minimum to minimum energy, corresponding to the lowest-energy singlet−triplet crossing $(T_1/S_0)_X$ and ^{3}(SWI) intermediate. See text and Figure 2. [e] Experimental monomer singlet−triplet band maxima, approximately corresponding to E_{VA}. Gas-phase electron energy-loss spectroscopy (refs 12, 21, and 22).

determinations in the potential energy hypersurfaces (PEH) of a system of two nucleobases. Energies were corrected for the effect of the basis set superposition error using the counterpoise procedure. The MOLCAS-7 quantum-chemistry code was employed.[20] See further details in the Supporting Information (SI) file.

Results in Table 1 determine the location of the lowest-energy triplet state in DNA/RNA pyrimidine nucleobases, both for the isolated monomer and in the context of a stacked dimer. In the monomer, the three pyrimidine nucleobases have close-by experimental triplet state signals at 3.5−3.6 eV.[12,21,22] Computation predicts, both for singlet−triplet vertical and adiabatic (band origin) transitions, that uracil lies higher in energy, followed by cytosine and thymine. Adiabatic energies cannot, however, properly represent triplet energies of the nucleobases in the biological environment, where many other effects have to be considered. We anticipate that the

DOI: 10.1021/jz100601p | J. Phys. Chem. Lett. 2010, 1, 2072–2076

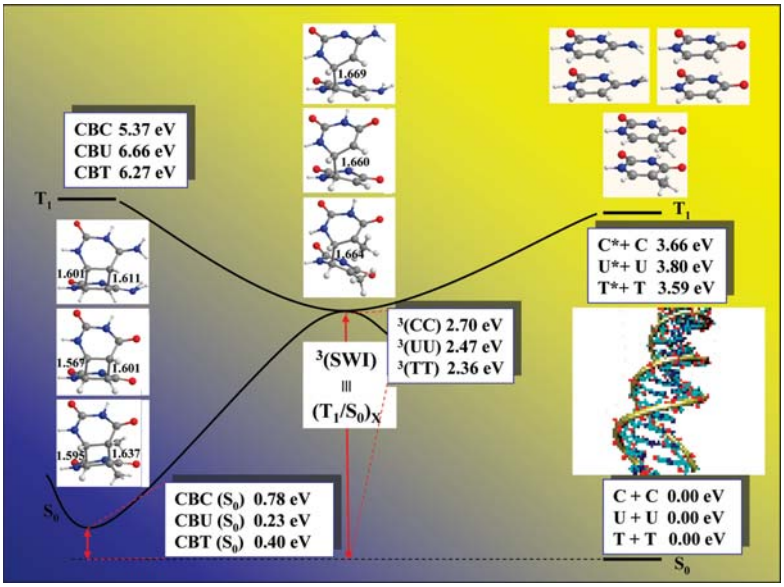


Figure 2. Scheme based on CASPT2 results of the triplet-mediated formation of CBPyr dimers (left) in isolated DNA/RNA nucleobase pairs. Center: singlet−triplet intermediate, ^{3}(SWI); right: separated nucleobases stacked pairs. Bond lengths (Å) indicated when relevant: $C_6-C_{6'}$ in ^{3}(SWI); $C_6-C_{6'}$ and $C_5-C_{5'}$ in the adducts.

major effect to take into account is the excimer stabilization undergone by the system in its excited state because of the interaction of pairs of stacked nucleobases. This would be followed, in turn, by the formation of a biradical stepwise intermediate ^{3}(SWI), in which a covalent bond is actually established between the $C_6-C_{6'}$ carbons of the adjacent nucleobases (see Figure 2). MEP calculations prove (see Figures S5 to S7 in the SI) that such structures are reached in a barrierless way from the minimum of the triplet excimer arrangements formed in the parallel stacking of the nucleobases. The ^{3}(SWI) intermediate has a biradical character with the unpaired electrons and the spin density on the two other ethylenic carbon atoms not forming the bond,[4] i.e., C_5 and $C_{5'}$. The intermediate is an actual minimum in the T_1 PEH, and also corresponds to a singlet−triplet crossing $(T_1/S_0)_X$ structure that is expected to mediate the formation of the CBPyr photoproduct. As we have demonstrated previously,[4] in order to obtain a degenerated biradical singlet−triplet crossing situation in a [2 + 2] cycloaddition between ethylenic bonds, an asymmetric charge distribution at the carbon centers is required. This takes place in the stacked nucleobases, where the C_6-type atoms are more electron deficient than the C_5-type atoms because the electron affinity of the C_6 neighbor nitrogen is higher than that of the C_5 carbon atom of the pyrimidine ring. The substituted $C_6-C_{6'}$ carbon atoms compensate for the lack of electron density by getting closer to the other monomer and forming a new bond.

It is known that ultraviolet A (UVA) radiation (3.10−3.87 eV), which comprises a range of energies inaccessible for direct singlet state population but adequate to promote TET processes from certain photosensitizers, preferentially induces the production of cyclobutane dimers at TT sites without any detectable formation of Pyr(6−4)Pyr photoproducts.[23] Once the triplet excimer in the nucleobase dimer has accepted population from the donor triplet photosensitizer, the system will evolve favorably toward the formation of the biradical triplet intermediate, representing the most stable triplet state conformation in the dimer. We propose the ^{3}(SWI) intermediate structure as the best candidate to trigger the cyclobutane photoadduct formation after ISC and evolution in the ground state PEH of the CBPyr adduct region, processes that have been suggested to have quasi unit efficiency.[14] The actual efficiency of the decay process along the triplet manifold will also rely on the enhancement of the spin−orbit coupling terms, estimated here to be just a few inverse centimeters at the in vacuo ^{3}(SWI) structures. In this respect, the solvent (or, in general, the environment) is expected to play a crucial role in the ISC process.[24] The location of the ^{3}(SWI) intermediate has been found to be the lowest in energy in the thymine homodimers (2.36 eV), followed by uracil (2.47 eV) and cytosine (2.70 eV), and it can be also related to the minimum energy required to populate the triplet state in the DNA/RNA environment. These values are slightly lower than, but in relatively good agreement with, the experimental

DOI: 10.1021/jz100601p | J. Phys. Chem. Lett. 2010, 1, 2072–2076

estimation of 2.6−2.8 eV as a minimum bound for the occurrence of TET CBT formation in solvated environments.[18,19] The present calculations therefore predict the production of thymine (CBT)and uracil (CBU) cyclobutane photoadducts in the triplet state taking place at similar energies and conditions. Higher energies would be, however, required to give rise to the cytosine adduct (CBC), which, as it occurred in the singlet manifold, might be also less favored by the formation of very stable cytosine excimers.[4] Such energy ordering has also been deduced in aqueous solution and rigid media experiments.[25] Since the efficiency of the triplet state formation has been determined to be larger in isolated thymine than in cytosine,[10,13] it is possible to rationalize the preference of the TT sites, and to a lesser extent CT, TC, and CC, to generate the cyclobutane photoproduct. CBT and CBU adducts have been found to dimerize faster and more efficiently than CBC in the presence of acetone and acetophenone.[1] Despite being obtained in lower yields, CC sequences are more relevant in the study of the damaging effects in DNA, because, unlike TT sequences, they are sites of relatively frequent CC-to-TT tandem mutations.[2]

The ^{3}(SWI) structure connects, on one hand, with the S_0 state minimum of the CBPyr adduct, but also with the ground state of the separated monomers. From the adduct geometry, the intermediate can be reached in a barrierless way along the T_1 state PEH (see the corresponding MEPs in the SI file, Figures S5 and S6). The same type of behavior was found for the S_1 state with respect to the S_0/S_1 CI in CBC,[4] and it is related with the observed photoreversibility of the adduct formation process after ultaviolet C (UVC) irradiation.[1] At the CBPyr structure the triplet state lays vertically at 4.59 (CBC), 6.43 (CBU), and 6.15 eV (CBT), whereas their respective S_1 states are placed (vertically) at 4.95, 6.28, and 5.35 eV, just in the UVC (middle to far-UV) range. It is worth mentioning that the nature of the T_1 state is different in CBC ($\pi\pi^*$), where the spin density is distributed on the C═N bond atoms in one of the monomers, and in the CBU and CBT pairs, in which the state, of nπ^* character, lies at considerably higher energy, and its spin density localizes on the C═O bond atoms neighbor to C_5. In these cases, the nature of the state changes adiabatically along the MEP, becoming of $\pi\pi^*$ in all cases at the ^{3}(SWI) structure. It is worth mentioning that the intrinsic mechanism of CBPyr formation does not seem to largely change in the biological environment, as shown in our CASPT2/MM study on the reversibility of the cytosine dimer (CBC) formation in a biological environment ($dC_{18}xdG_{18}$ was simulated).[26] Although further studies are required in all cases, the computational evidence is strong enough to illustrate that the basics of the mechanism are properly described in the isolated dimer.

In summary, once the presence of precursor singlet excimer structures favoring ultrafast internal conversion through an S_0/S_1 CI was previously proposed[3−7] as the major source of the formation of CBPyr dimers in the biopolymer nucleobases or in frozen solutions,[1] the present scheme tries to rationalize the mechanism for triplet−triplet photosensitized photoadducts from appropriate donors, an energy transfer process expected to be more favored in solution.[1] The overall lower yields of formation in cytosine dimers versus those in thymine (and uracil) can be related first to the relative location on the former of stable singlet excimer structures lower in energy than the singlet CI triggering the internal conversion. Additionally, the triplet intermediate mediating the photoadduct formation by means of energy transfer processes is in cytosine higher (and less accessible) than those in the other pyrimidine nucleobases. Determining, along the lines followed in the present research, the nature and relative location of the triplet intermediates in different nucleobase duplexes will help to rationalize the phototoxicity of different photosentisizers, driven by the energy of their reactive triplet states, and to design health care phototherapeutic nucleobase-based drugs addressed to either enhance or decrease pyrimidine dimer formation by using, for instance, appropriate nucleobase derivatives.

SUPPORTING INFORMATION AVAILABLE Computational details, MEPs and scans on the PEHs, and Cartesian coordinates for the optimized conformations. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author:

*To whom correspondence should be addressed. E-mail: Luis.Serrano@uv.es.

ACKNOWLEDGMENT The research reported in this report has been supported by the Spanish MICINN/FEDER projects CTQ2007-61260, CTQ2010-14892, and CSD2007-0010 Consolider-Ingenio in Molecular Nanoscience and by the Generalitat Valenciana.

REFERENCES

(1) Cadet, J.; Vigny, P. The Photochemistry of Nucleic Acids. In *Bioorganic Photochemistry*; Morrison, H., Ed.; John Wiley & Sons: New York, 1990; Vol. 1, pp 1−272.
(2) Douki, T.; Cadet, J. Individual Determination of the Yield of the Main UV-Induced Dimeric Pyrimidine Photoproducts in DNA Suggests a High Mutagenicity of CC Photolesions. *Biochemistry* **2001,** *40*, 2495−2501.
(3) Schreier, W. J.; Schrader, T. E.; Soller, F. O.; Gilch, P.; Crespo-Hernández, C. E.; Swaminathan, V. N.; Carell, T.; Zinth, W.; Kohler, B. Thymine Dimerization in DNA is an Ultrafast Photoreaction. *Science* **2007,** *315*, 625−629.
(4) Roca-Sanjuán, D.; Olaso-González, G.; González-Ramírez, I.; Serrano-Andrés, L.; Merchán, M. Molecular Basis of DNA Photodimerization: Intrinsic Production of Cyclobutane Cytosine Dimers. *J. Am. Chem. Soc.* **2008,** *130*, 10768−10779.
(5) Serrano-Pérez, J. J.; González-Ramírez, I.; Coto, P. B.; Merchán, M.; Serrano-Andrés, L. Theoretical Insight into the Intrinsic Ultrafast Formation of Cyclobutane Pyrimidine Dimers in UV-Irradiated DNA: Thymine versus Cytosine. *J. Phys. Chem. B* **2008,** *112*, 14096−14098.
(6) Boggio-Pasqua, M.; Groenhof, G.; Schäfer, L. V.; Grubmüller, H.; Robb, M. A. Ultrafast Deactivation Channel for Thymine Dimerization. *J. Am. Chem. Soc.* **2007,** *129*, 10996−10997.
(7) Blancafort, L.; Migani, A. Modeling Thymine Photodimerizations in DNA: Mechanism and Correlation Diagrams. *J. Am. Chem. Soc.* **2007,** *129*, 14540−14541.
(8) Kwok, W. M.; Ma, C.; Phillips, D. L. A Doorway State Leads to Photostability or Triplet Photodamage in Thymine DNA. *J. Am. Chem. Soc.* **2008,** *130*, 5131−5139.

(9) Hare, P. M.; Crespo-Hernandez, C. E.; Kohler, B. Internal Conversion to the Electronic Ground State Occurs via Two Distinct Pathways for Pyrimidine Bases in Aqueous Solution. *Proc. Natl. Acad. Sci. U.S.A.* **2007,** *104*, 435–440.
(10) Merchán, M.; Serrano-Andrés, L.; Robb, M. A.; Blancafort, L. Triplet-State Formation Along the Ultrafast Decay of Excited Singlet Cytosine. *J. Am. Chem. Soc.* **2005,** *127*, 1820–1825.
(11) Climent, T.; González-Luque, R.; Merchán, M.; Serrano-Andrés, L. On the Intrinsic Population of the Lowest Triplet State of Uracil. *Chem. Phys. Lett.* **2007,** *441*, 327–331.
(12) Etinski, M.; Fleig, T.; Marian, C. M. Intersystem Crossing and Characterization of Dark States in the Pyrimidine Nucleobases Uracil, Thymine, and 1-Methylthymine. *J. Phys. Chem. A* **2009,** *113*, 11809–11816.
(13) Serrano-Pérez, J. J.; González-Luque, R.; Merchán, M.; Serrano-Andrés, L. On the Intrinsic Population of the Lowest Triplet State of Thymine. *J. Phys. Chem. B* **2007,** *111*, 11880–11883.
(14) Marguet, S.; Markovitsi, D. Time-Resolved Study of Thymine Dimer Formation. *J. Am. Chem. Soc.* **2005,** *127*, 5780–5781.
(15) Gut, I. G.; Wood, P. D.; Redmond, R. W. Interaction of Triplet Photosensitizers with Nucleotides and DNA in Aqueous Solution at Room Temperature. *J. Am. Chem. Soc.* **1996,** *118*, 2366–2373.
(16) Cadet, J.; Sage, E.; Douki, T. Ultraviolet Radiation-Mediated Damage to Cellular DNA. *Mutat. Res.* **2005,** *571*, 3–17.
(17) Nguyen, M. T.; Zhang, R.; Nam, P.-C.; Ceulemans, A. Singlet-Triplet Energy Gaps of Gas-Phase RNA and DNA Bases. A Quantum Chemical Study. *J. Phys. Chem. A* **2004,** *108*, 6554–6561.
(18) Bosca, F.; Lhiaubet-Vallet, V.; Cuquerella, M. C.; Castell, J. V.; Miranda, M. A. The Triplet Energy of Thymine in DNA. *J. Am. Chem. Soc.* **2006,** *128*, 6318–6319.
(19) Albini, A.; Monti, S. Photophysics and Photochemistry of Fluoroquinolones. *Chem. Soc. Rev.* **2003,** *32*, 238–250.
(20) Aquilante, F.; De Vico, L.; Ferré, N.; Ghigo, G.; Malmqvist, P.- Å.; Pedersen, T.; Pitonak, M.; Reiher, M.; Roos, B. O.; Serrano-Andrés, L.; Urban, M.; Veryazov, V.; Lindh, R. MOLCAS 7: The Next Generation. *J. Comput. Chem.* **2010,** *31*, 224–247.
(21) Abouaf, R.; Pommier, J.; Dunet, H.; Quan, P.; Nam, P.- C.; Nguyen, M. T. The Triplet State of Cytosine and its Derivatives: Electron Impact and Quantum Chemical Study. *J. Chem. Phys.* **2004,** *121*, 11668–11674.
(22) Abouaf, R.; Pommier, J.; Dunet, H. Electronic and Vibrational Excitation in Gas Phase Thymine and 5-Bromouracil by Electron Impact. *Chem. Phys. Lett.* **2003,** *381*, 486–494.
(23) Mouret, S.; Baudouin, C.; Charveron, M.; Favier, A.; Cadet, J.; Douki, T. Cyclobutane Pyrimidine Dimers are Predominant DNA Lesions in Whole Human Skin Exposed to UVA Radiation. *Proc. Natl. Acad. Sci. U.S.A.* **2006,** *103*, 13765–13770.
(24) Hare, P. M.; Crespo-Hernández, C. E.; Kohler, B. Solvent-Dependent Photophysics of 1-Cyclohexyluracil: Ultrafast Branching in the Initial Bright State Leads Nonradiatively to the Electronic Ground State and a Long-Lived $^1n\pi^*$ State. *J. Phys. Chem. B* **2006,** *110*, 18641–18650.
(25) Wood, P. D.; Redmond, R. W. Triplet State Interactions between Nucleic Acid Bases in Solution at Room Temperature: Intermolecular Energy and Electron Transfer. *J. Am. Chem. Soc.* **1996,** *118*, 4256–4263.
(26) Roca-Sanjuán, D.; Olaso-González, G.; Rubio, M.; Coto, P. B.; Merchán, M.; Ferré, N.; Ludwig, V.; Serrano-Andrés, L. DNA Nucleobase Properties and Photoreactivity: Modeling Environmental Effects. *Pure Appl. Chem.* **2009,** *81*, 743–754.

DOI: 10.1021/jz100601p | *J. Phys. Chem. Lett.* **2010,** *1*, 2072–2076

Cyclobutane Pyrimidine Photodimerization of DNA/RNA Nucleobases in the Triplet State

Teresa Climent, Israel González-Ramírez, Remedios González-Luque,, Manuela Merchán, and Luis Serrano-Andrés[*]

Instituto de Ciencia Molecular, Universitat de València
Apartado 22085, ES-46071 Valencia, Spain

Luis.Serrano@uv.es

Supporting Information

Computational details

a. CASSCF and CASPT2 calculations.

Geometry optimizations of the ground-state Pyr<>Pyr dimers and the delocalized excimer $^3(TT)_{exc}$ have been studied by using the complete-active-space self-consistent-field (CASSCF) method comprising as active space a total of 16 π electrons distributed among 14 π molecular orbitals (MOs). The totally symmetric π MO of each pyrimidine was kept inactive. In order to mimic the actual interaction of pyrimidines in DNA, geometry optimization has been carried out initially within C_s symmetry constraints, allowing so for an effective and natural interaction of the two pyrimidine molecules in the biologically relevant cis-syn stereoisomer. Further optimization released the symmetry restrictions (C_1 symmetry). For the computations in C_1 symmetry, two additional π MOs were also kept inactive, since the occupation number of the corresponding natural orbitals when they were treated as active was practically 2.0. A CASSCF wave function of 12 active π electrons and 12 active π MOs was therefore employed, hereafter denoted as CASSCF(12,12).

Irrespectively of the geometrical symmetry at which the calculation was being performed, the final computations did not imposed any symmetry restriction to allow for wave function localization. Therefore, using the C_1 state-average CASSCF(12,12) wave functions for three roots, dynamic electron correlation has been subsequently taken into account perturbatively at the second-order level through the CASPT2 method,[1,2,3] labeled as CASPT2(12,12). In order to minimize weakly interacting intruder states, the imaginary level-shift technique, with IMAG=0.2 au, has been employed.[4] The reported CASPT2 (T_1/S_0) singlet-triplet crossing was obtained as an actual minimum in the T_1 PEH.[5] The obtained energies were corrected fro, the effects of the Basis Set Superposition Error (BSSE) when required (see below).

MEPs have been built as steepest descendent paths in a procedure[6] which is based on a modification of the projected constrained optimization (PCO) algorithm of Anglada and Bofill[7] and follows the Müller-Brown approach.[8] Each step requires the minimization of the PES on a hyperspherical cross section of the PES centered on the initial geometry and characterized by a predefined radius. The optimized structure is taken as the center of a

new hypersphere of the same radius, and the procedure is iterated until the bottom of the energy surface is reached. Mass-weighted coordinates are used, therefore the MEP coordinate corresponds to the so-called Intrinsic Reaction Coordinates (IRC). Each step in the abscissa coordinate in the Figures below corresponds to a step of the MEP computed with a hypersphere radius constrained to 0.1 au. The full procedure is currently implemented in the MOLCAS package[9] and its technical description has been published elsewhere.[6] Regarding the conical intersection and surface crossing searches, they were performed using the restricted Lagrange multipliers technique as included in the MOLCAS package[9] in which the lowest-energy point was obtained under the restriction of degeneracy between the two considered states, so actually they should be denominated Minimum Energy Crossing Points (MECPs).[6]

The basis set of Atomic Natural Orbital (ANO) type with the contraction scheme C,N,O[3s2p1d]/H[2s1p] was used throughout.[10] Basis set superposition error (BSSE) was taken into account by using a modified counterpoise (CP) approach based on localized molecular orbitals, specifically designed for correlated approaches (see next sextion).[11]

All the computations have been carried out by using the MOLCAS 7 quantum-chemical software.[9]

b. Basis Set Superposition Error.

The inclusion of the basis set superposition error (BSSE) is crucial to accurately describe binding energies.[12] Here the effect was taken into account by using the counterpoise correction (CP).[13]

The binding energy (E_b) was obtained as follows:

$$E_b\,(PyrPyr^*) = E_{Pyr} + E_{Pyr^*} - E_{Pyr^*Pyr} \qquad (1)$$

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

$$\text{CP-}E_b\,(\text{Pyr*Pyr}) = E_b\,(\text{Pyr*Pyr}) - [\text{CP-BSSE(Pyr*Pyr)}] \quad (2)$$

where

$$\text{CP-BSSE}\,(\text{Pyr*Pyr}) = E_{Pyr^*}\,(\text{Pyr, R=}\infty) - E_{Pyr^*}\,(\text{Pyr, R=Pyr*Pyr}) + E_{Pyr}\,(\text{Pyr*, R=}\infty) - E_{Pyr}\,(\text{Pyr*, R=Pyr*Pyr}). \quad (3)$$

In [CP-BSSE (Pyr*Pyr)] the geometry of the monomers is kept to that of the dimer. Thus, the notation E_{Pyr^*} (Pyr, R=Pyr*Pyr) indicates the energy of Pyr* computed in the ghost orbitals of Pyr at the geometry of Pyr*Pyr, whereas in E_{Pyr^*} (Pyr, R= ∞) the ghost MOs of Pyr are at infinity distance of Pyr*. In this manner the influence of the variation of geometry is accounted for in the BSSE treatment. The findings discussed correspond to CASPT2 results with inclusion of BSSE.

c. Additional figures

We have included here Figures S1 to S7 to illustrate some of the conclusions stated in the paper for the different nucleobase dimers. Figures S1 to S4 inform of the stabilization taking place in the excited state of stacked syn-cis face-to-face homodimers (here the three nucleobases are shown: cytosine, uracil, and thymine) when decreasing the intermonomer distance. The phenomenon takes place in both the singlet and triplet excited states of the dimers, but not on the ground state. This is made self-evident only when the energies are corrected of the BSSE effect. The BSSE error leads to an overestimation of the correlation energies because of basis sets borrowing effects from neighbor atoms, an error which decreases along the coordinate separating some fragments, like here both monomers. Therefore removing the BSSE reduces the binding energies, leaving the ground state of the dimer as practically dissociative, unlike the excited states. This is just the definition of an excimer.

Figure S5, on the other hand, displays the potential energy curves of the ground singlet and two triplet states in the cytosine dimer along a MEP that, starting from the minimum of the triplet T_1 excimer, known as ^{3}(LE) (see Ref. 5), connects without any barrier with the ^{3}(SWI) intermediate, also corresponding to the $(T_1/S_0)_X$ singlet-triplet crossing leading to the cyclobutane pyrimidine dimer. Notice that no actual restriction is forced on

a MEP, that represents the lowest-energy path going downhill and guarantees the absence of energy barriers along the computed connected path. In this case the MEP proves the accessibility of the ^{3}(SWI) intermediate from an excimer conformation, but, as actual calibration calculations pointed out, all type of barely stacked conformations are favorable arrangements from which the intermediate may be easily reached.

Figure S6 represents, on the other hand, the low-lying singlet states of the cytosine dimer along the MEP on S_1 starting from the CBC ground state geometry, and shows that the S_1 state connects CBC with the S_0/S_1 CI, from which it is possible to reach both the photoadduct or the separated monomers in the S_0 state, but also the CC* excimer in the S_1 state. The profile of this MEP can be easily related to the known ability of the cyclobutane pyrimidine dimer to undergo photoreversibility (and then lead to separated nucleobases) upon irradiation with UVC light, exactly that required to populate the CBPyr lowest singlet state, as it is discussed in the paper.

Figure S7 displays a similar profile, but this time representing the MEP connecting the T_1 state of the CBU dimer at the ground state of the adduct with the triplet intermediate and the minimum of the triplet excimer in the uracil dimer, in a similar fashion as it takes place for the singlet of the cytosine dimer in Figure S6. It is worth mentioning, however, that the charge distribution undergoes a large reorganization in the triplet state of the uracil dimer. At the CBU structure the lowest-energy triplet state is a $n\pi^*$ state localized on a C=O bond (as shown by the spin density and the nature of the MOs, see below), whereas on the ^{3}(SWI) intermediate the charge distribution and spin density localizes basically on the non-bound C_5 and $C_{5'}$ atoms.

Finally Figures S8 and S9 display the MOs related with the T_1 state ion CBC and CBT, which can be represented as a single one-electron promotion involving both orbitals. Notice that whereas in CBC the T_1 state is of $\pi\pi^*$ type, lower in energy and localized on the C=N bond neighbor to C_5 in one of the monomers (also the spin density localizes on those two C and N atoms), in CBT (as well as in CBU) the T_1 state has $n\pi^*$ character, and involves the C=O bond atoms neighbor to C_5 in one of the thymine monomers. The spin density localizes on C and O in this case. It is known that thymine and uracil display lowest-lying $n\pi^*$ states as compared with cytosine.[14,15,16]

d. Geometries

Table S1 compiles the optimized geometries and their computed CASSCF energies.

References

[1] Andersson, K., Malmqvist, P. -Å.; Roos, B. O. *J. Chem. Phys.* **1992,** *96*, 1218-1226.
[2] Roos, B. O., Andersson, K., Fülscher, M. P., Malmqvist, P.-Å., Serrano-Andrés, L., Pierloot, K., Merchán, M. *Adv. Chem. Phys.* **1996**, *93*, 219-331.
[3] Merchán, M.; Serrano-Andrés, L. In *Computational Photochemistry,* ed. Olivucci, M., Elsevier, Amsterdam, 2005.
[4] Forsberg, N.; Malmqvist, P.-Å. *Chem. Phys. Lett.* **1997**, *274*, 196-204.
[5] Roca-Sanjuán, D.; Olaso-González, G.; González-Ramírez, I.; Serrano-Andrés, L. *J. Am. Chem. Soc.* **2008**, *130*, 10768-10779.
[6] De Vico, L.; Olivucci, M.; Lindh, R. *J. Chem. Theory Comp.* **2005**, *1*, 1029-1037.
[7] Anglada, J. M.; Bofill, J. M. *J. Comput. Chem.*, **1997**, *18*, 992-1003.
[8] Müller, K.; Brown, L. D. *Theor. Chim. Acta.* **1979**, *53*, 75-93.
[9] Andersson, K., Barysz, M., Bernhardsson, A., Blomberg, M. R. A., Carissan, Y., Cooper, D. L., Cossi, M., Fülscher, M. P., Gagliardi, L., de Graaf, C., Hess, B., Hagberg, G., Karlström, G., Lindh, R., Malmqvist, P.-Å. , Nakajima, T., Neogrády, P., Olsen, J., Raab, J., Roos, B. O., Ryde, U., Schimmelpfennig, B., Schütz, M., Seijo, L., Serrano-Andrés, L., Siegbahn, P. E. M., Stålring, J., Thorsteinsson, T., Veryazov, V., Widmark, P. -O. *MOLCAS, version 7.0*; Department of Theoretical Chemistry, Chemical Centre, University of Lund, P.O.B. 124, S-221 00 Lund: Sweden, 2009.
[10] Pierloot, K.; Dumez, B.; Widmark, P.-O.; Roos, B. O. *Theor. Chim. Acta* **1995**, *90*, 87-114.
[11] Van Duijneveldt, F. B., Van Duijneveldt-Van de Rijdt, J. G. C. M.; Van Lenthe, J. H. *Chem. Rev.* **2004**, ***94***, 1873-1885.
[12] Olaso-González, G.; Roca-Sanjuán, D.; Serrano-Andrés, L.; Merchán, M. *J. Chem. Phys.* **2006**, *125*, 231102.
[13] Boys, S. F.; Bernardi, F. *Mol. Phys.* **2002**, *100*, 65-73.
[14] Merchán, M.; Serrano-Andrés, L.; Robb, M. A.; Blancafort, L. *J. Am. Chem. Soc.* **2005**, *127*, 1820-1825.
[15] Climent, T.; González-Luque, R.; Merchán, M.; Serrano-Andrés, L. *Chem. Phys. Lett.* **2007**, *441*, 327-331.
[16] Serrano-Pérez, J. J.; González-Luque, R.; Merchán, M.; Serrano-Andrés, L. *J. Phys. Chem. B*. **2007**, *111*, 11880-11883.

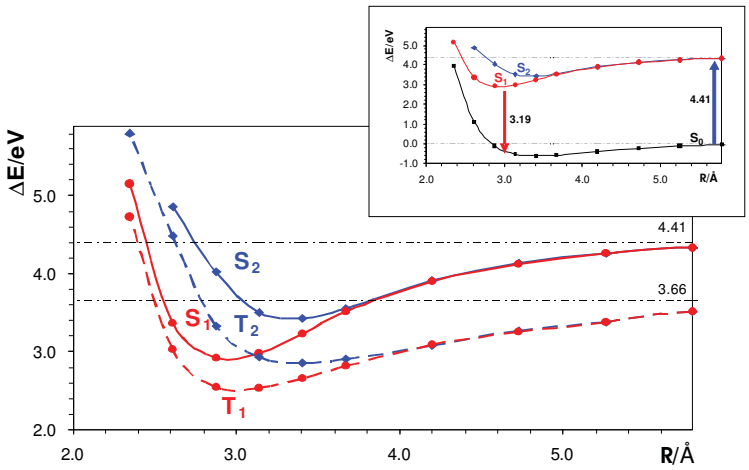


Figure S1. *CASPT2 potential energy curves built with respect to the intermolecular distance $R(C_5\text{-}C_{5'})$ of two face-to-face π-stacked cytosine molecules involving the ground and the lowest two triplet and two singlet excited states (BSSE uncorrected). The inset includes the singlet states obtained at the same level of theory.*

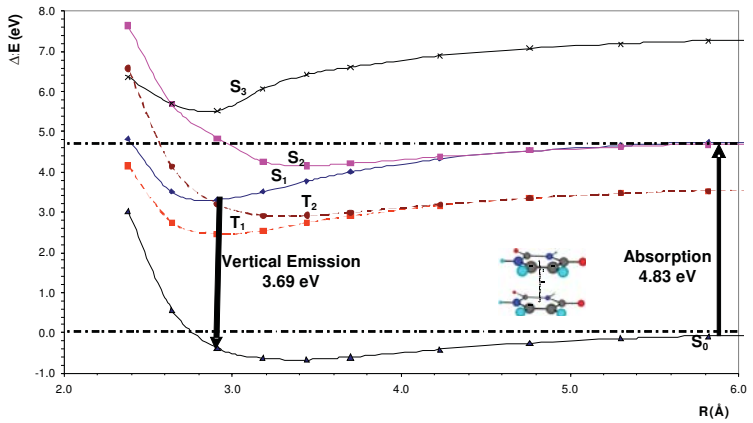


Figure S2. *CASPT2(12,12)/ANO-S C,N,O[3s2p1d]/H[2s1p] potential energy curves built with respect to the intermolecular distance considering the center of mass of two face-to-face π-stacked uracil molecules.*

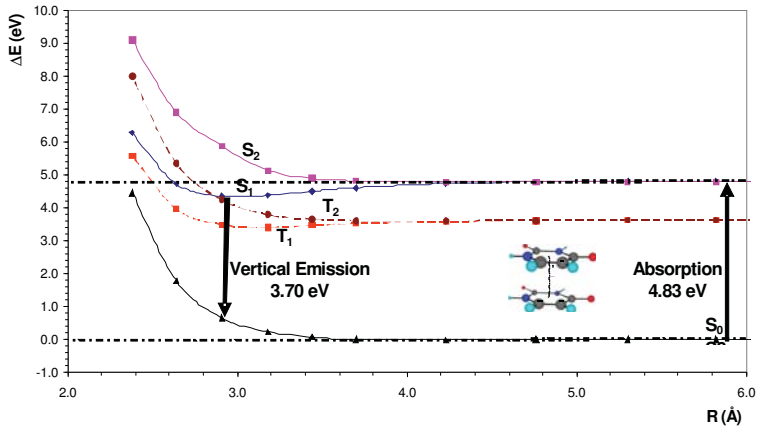


Figure S3. *Corrected BSSE-CASPT2(12,12)/ANO-S C,N,O[3s2p1d]/H[2s1p] potential energy curves built with respect to the intermolecular distance considering the center of mass of two face-to-face π-stacked uracil molecules.*

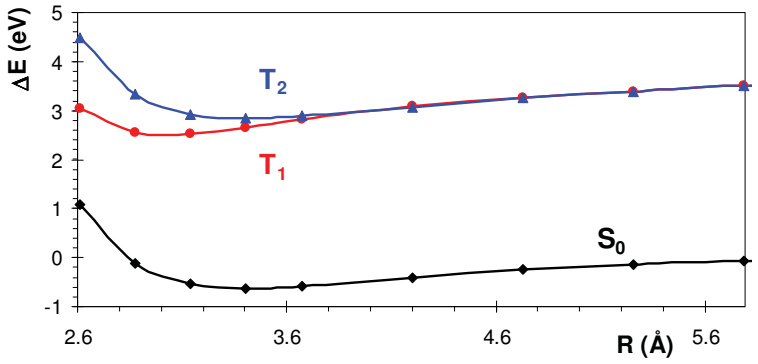


Figure S4. *Corrected BSSE-CASPT2(12,12)/ANO-S C,N,O[3s2p1d]/H[2s1p] potential energy curves built with respect to the intermolecular distance considering the center of mass of two face-to-face π-stacked thymine molecules.*

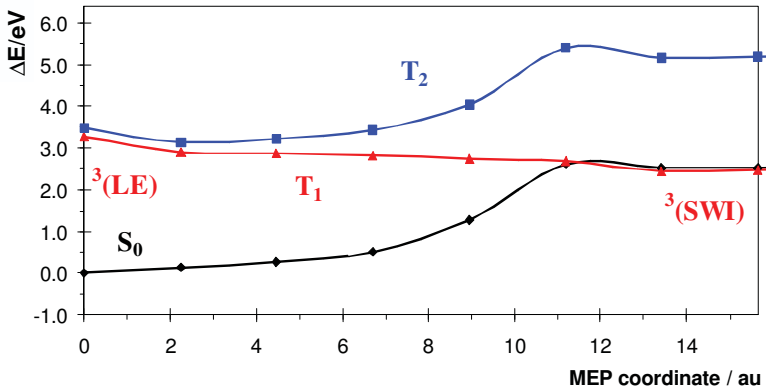


Figure S5. *Low-lying triplet excited states of the cytosine dimer computed at the CASPT2//CASSCF level along the Minimum Energy Path (MEP) of the T_1 state from the geometry of the triplet locally excited state ^{3}(LE). The T_1-MEP ends at the step-wise intermediate ^{3}(SWI) and it is isoenergetic with the ground state (S_0).*

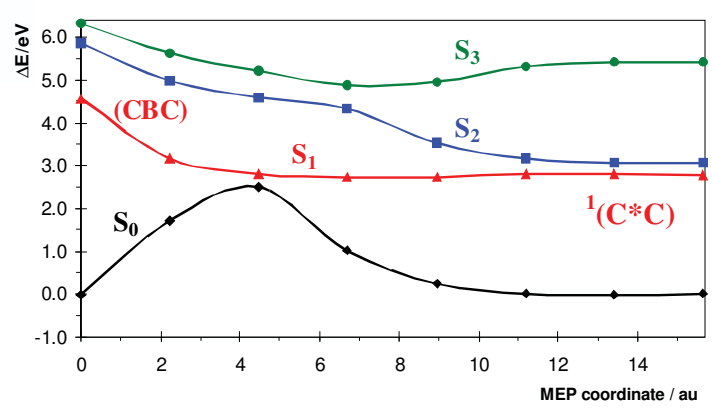


Figure S6. *Low-lying singlet excited states of the cytosine dimer computed at the CASPT2//CASSCF level along the Minimum Energy Path (MEP) of the S_1 state from the cyclobutane cytosine (CBC) dimer at its ground-state equilibrium geometry. The S_1-MEP ends at the relaxed excimer ^{1}(C*C) (see Figure 6).*

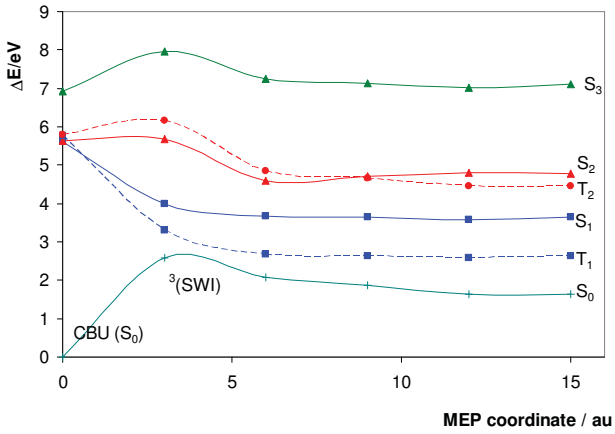


Figure S7. *Low-lying triplet and singlet states of the uracil dimer computed at the CASPT2//CASSCF level along the Minimum Energy Path (MEP) of the T_1 state from the ground-state geometry of the cyclobutane uracil dimer. The T_1-MEP goes through the step-wise intermediate ^{3}(SWI), connecting the CBU structure with that region and the separated monomers.*

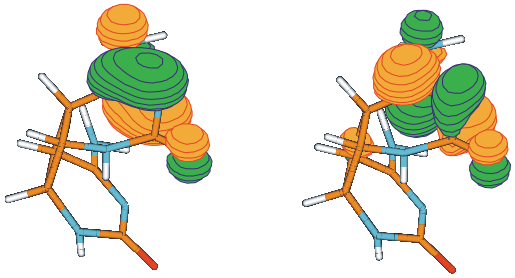


Figure S8. *At the ground state geometry of the CBC adduct the lowest-energy triplet T_1 state can be described as a single excitation from the π (left) and π^* (right) molecular orbitals, represented in the plot. They basically relate with the C=N bond (neighbor to C_5) atoms of one of the cytosine monomers. The spin density also concentrates in these two atoms.*

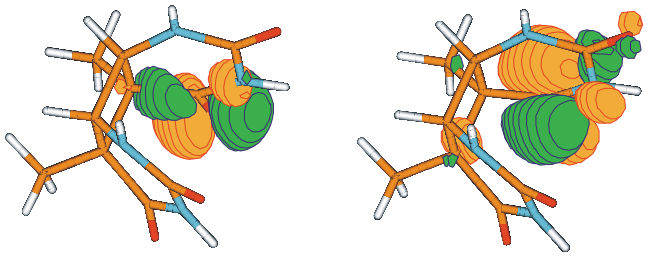


Figure S9. *At the ground state geometry of the CBT adduct the lowest-energy triplet T_1 state can be described as a single excitation from the n, oxygen lone pair (left) and π^* (right) molecular orbitals, represented in the plot. They basically relate with the C=O bond (neighbor to C_5) atoms of one of the cytosine monomers. The spin density also concentrates in these two atoms.*

Table S1. Cartesian coordinates x, y, z (in Å) of the stationary points optimized for the different dimers. CASSCF(12,12) total energy (Et) is also included.

S_0 Cytosine CASSCF(8e,7MOs)/ANO-S N,C,O[3s2p1d]/H[2s1p]
C_s symmetry
symmetry distinct nuclear coordinates
Et: -785.528173 au

N	−1.516732	−0.309951	5.302381
N	0.544315	0.867374	5.295448
N	2.529617	−0.269525	5.216023
C	−0.845807	0.906780	5.297600
C	1.154242	−0.267140	5.279228
C	0.474159	−1.556649	5.274839
C	−0.874191	−1.516454	5.288717
O	−1.464901	1.934123	5.303139
H	−2.514863	−0.261380	5.305898
H	2.945682	0.613548	5.443847
H	2.986034	−1.055067	5.633929
H	1.019989	−2.484251	5.249586
H	−1.498588	−2.396429	5.285647

S_0 CBC CASSCF(16e,14MOs)/ANO-S N,C,O[3s2p1d]/H[2s1p]
C_s symmetry ($1^1A'$)
symmetry distinct nuclear coordinates
Et: -785.492565 au

N	−1.174345	−1.029715	1.490393
N	0.129599	0.934492	1.915796
N	2.406920	0.893383	1.803909
C	−1.130331	0.320690	1.789649
C	1.177985	0.332582	1.520261
C	1.209686	−0.991614	0.805464
C	−0.139671	−1.732355	0.800324
O	−2.129488	0.935668	2.015623
H	−2.110021	−1.362973	1.372321
H	2.362412	1.864277	2.051934
H	3.160813	0.664517	1.193323
H	2.027307	−1.585923	1.203890
H	−0.067738	−2.746890	1.188110

Table S1 (continuation).

^{3}(SWI) Cytosine CASSCF(16e,14MOs)/ANO-S N,C,O[3s2p1d]/H[2s1p]
C_s symmetry ($1^3A''$)
symmetry distinct nuclear coordinates
Et: -785.424241 au

N	−1.171893	−0.988398	1.414216
N	0.054851	0.922932	2.181740
N	2.340724	0.827457	2.429911
C	−1.175869	0.305164	1.894268
C	1.147527	0.258783	1.999938
C	1.196629	−1.053376	1.421808
C	−0.038664	−1.643466	0.834312
O	−2.204127	0.879234	2.108651
H	−2.085204	−1.320023	1.185379
H	2.238904	1.802557	2.644810
H	3.130652	0.647126	1.842686
H	2.137971	−1.568375	1.326759
H	−0.081143	−2.700196	1.099079

^{3}(LE) Cytosine dimer CASSCF(16e,14MOs)/ANO-S N,C,O[3s2p1d]/H[2s1p]
C_s symmetry ($1^3A''$)
symmetry distinct nuclear coordinates
Et: -785.421156 au

N	−1.516732	−0.309951	1.679591
N	0.544315	0.867374	1.672658
N	2.529617	−0.269525	1.593233
C	−0.845807	0.906780	1.674810
C	1.154242	−0.267140	1.656438
C	0.474159	−1.556649	1.652049
C	−0.874191	−1.516454	1.665927
O	−1.464901	1.934123	1.680349
H	−2.514863	−0.261380	1.683108
H	2.945682	0.613548	1.821057
H	2.986034	−1.055067	2.011139
H	1.019989	−2.484251	1.626796
H	−1.498588	−2.396429	1.662857

Table S1 (continuation).

S_0 Uracil CASSCF(8e,7MOs)/ANO-S N,C,O[3s2p1d]/H[2s1p]
C_s symmetry
symmetry distinct nuclear coordinates
Et: -825.226095 au

N	-0.696865	1.030837	5.291772
N	-1.079768	-1.252938	5.291772
C	-1.624258	0.010406	5.291772
C	0.282947	-1.487236	5.291772
C	1.177074	-0.481320	5.291772
C	0.689855	0.906931	5.291772
O	1.391035	1.880122	5.291772
O	-2.804806	0.208124	5.291772
H	0.573817	-2.525976	5.291772
H	-1.070547	1.960861	5.291772
H	-1.730617	-2.010109	5.291772
H	2.239020	-0.654587	5.291772

S_0 CBU CASSCF(16e,14MOs)/ANO-S N,C,O[3s2p1d]/H[2s1p]
C_s symmetry ($1^1A'$)
symmetry distinct nuclear coordinates
Et: -825.154640 au

N	1.080908	0.598971	1.787935
N	-1.152615	1.215053	1.480819
C	0.099744	1.575187	1.892179
C	-1.510413	0.010395	0.800764
C	-0.412958	-1.075032	0.783986
C	0.889849	-0.730518	1.472563
O	1.729509	-1.552172	1.697257
O	0.350971	2.653561	2.345811
H	-2.457617	-0.353999	1.192367
H	1.949001	0.833749	2.232975
H	-1.816265	1.960762	1.529678
H	-0.736776	-2.028218	1.186936

Table S1 (continuation).

^{3}(SWI) uracil CASSCF(16e,14MOs)/ANO-S N,C,O[3s2p1d]/H[2s1p]
C_s symmetry (1^3A")
symmetry distinct nuclear coordinates
Et: -825.133711 au

N1	−0.695961	0.989329	1.924225
N2	−1.052570	−1.258956	1.414432
C1	−1.596751	−0.054015	1.790927
C2	0.252474	−1.424905	0.829899
C3	1.183917	−0.420852	1.427964
C5	0.683242	0.853238	1.933236
O1	1.401429	1.743010	2.308746
O2	−2.762374	0.090098	2.012786
H1	0.605618	−2.418928	1.099993
H2	−1.075009	1.838309	2.299328
H3	−1.746795	−1.955504	1.242096
H4	2.251936	−0.545727	1.371368

^{3}(LE) Uracil CASSCF(16e,14MOs)/ANO-S N,C,O[3s2p1d]/H[2s1p]
C_s symmetry (1^3A")
symmetry distinct nuclear coordinates
Et: -824.928485 au

N1	−0.695961	0.989329	1.924225
N2	−1.052570	−1.258956	1.414432
C1	−1.596751	−0.054015	1.790927
C2	0.252474	−1.424905	0.829899
C3	1.183917	−0.420852	1.427964
C5	0.683242	0.853238	1.933236
O1	1.401429	1.743010	2.308746
O2	−2.762374	0.090098	2.012786
H1	0.605618	−2.418928	1.099993
H2	−1.075009	1.838309	2.299328
H3	−1.746795	−1.955504	1.242096
H4	2.251936	−0.545727	1.371368

Table S1 (continuation).

S_0 Thymine CASSCF(8e,7MOs)/ANO-S N,C,O[3s2p1d]/H[2s1p]
C_s symmetry
distinct nuclear coordinates
Et: -905.842746 au

N	−0.732356	1.031478	11.024335
N	−1.118134	−1.240959	11.024539
C	−1.663092	0.015776	11.025098
C	0.247451	−1.477747	11.025582
C	1.154751	−0.484077	11.024699
C	2.641891	−0.702205	11.025394
C	0.648709	0.902545	11.024633
O	1.352132	1.876155	11.025800
O	−2.843758	0.219643	11.025989
H	0.529155	−2.517058	11.026998
H	−1.101588	1.961794	11.024912
H	−1.768089	−1.997177	11.023663
H	2.871350	−1.765157	11.019251
H	3.095037	−0.250930	11.905840
H	3.096774	−0.240693	10.151239

S_0 CBT CASSCF(16e,14MOs)/ANO-S N,C,O[3s2p1d]/H[2s1p]
C_s symmetry ($1^1A'$)
symmetry distinct nuclear coordinates
Et: -905.871903 au

N	−0.663032	1.034929	1.687450
N	−1.060511	−1.250889	1.493713
C	−1.557050	−0.016821	1.808581
C	0.165186	−1.500364	0.797693
C	1.199282	−0.351630	0.818750
C	2.523897	−0.707456	1.493337
C	0.689844	0.949306	1.423407
O	1.409690	1.882208	1.629656
O	−2.676235	0.161052	2.191298
H	0.589352	−2.426619	1.176232
H	−0.991511	1.910444	2.049657
H	−1.757302	−1.967895	1.502049
H	2.913191	−1.650289	1.121174
H	2.366217	−0.807822	2.566969
H	3.259216	0.073234	1.327812

Table S1 (continuation).

^{3}(SWI) thymine CASSCF(16e,14MOs)/ANO-S N,C,O[3s2p1d]/H[2s1p]
C_s symmetry ($1^3A''$)
symmetry distinct nuclear coordinates
Et: -905.791414 au

N	−0.688141	1.001310	1.796848
N	−1.032193	−1.265908	1.408710
C	−1.589139	−0.047013	1.713325
C	0.279780	−1.429363	0.831764
C	1.217760	−0.441175	1.461466
C	2.654105	−0.764035	1.746773
C	0.686700	0.848912	1.902467
O	1.379270	1.744092	2.313609
O	−2.758078	0.105525	1.911032
H	0.619089	−2.427075	1.100500
H	−1.070415	1.866369	2.129661
H	−1.720403	−1.968213	1.234455
H	2.964024	−1.667210	1.229182
H	2.796206	−0.921198	2.818184
H	3.301548	0.056975	1.449931

^{3}(LE) Thymine CASSCF(16e,14MOs)/ANO-S N,C,O[3s2p1d]/H[2s1p]
C_s symmetry ($1^3A''$)
symmetry distinct nuclear coordinates
Et: -905.740966 au

N	−0.732356	1.031478	1.699145
N	−1.118134	−1.240959	1.699349
C	−1.663092	0.015776	1.699908
C	0.247451	−1.477747	1.700391
C	1.154751	−0.484077	1.699508
C	2.641891	−0.702205	1.700204
C	0.648709	0.902545	1.699442
O	1.352132	1.876155	1.700610
O	−2.843758	0.219643	1.700798
H	0.529155	−2.517058	1.701808
H	−1.101588	1.961794	1.699721
H	−1.768089	−1.997177	1.698473
H	2.871350	−1.765157	1.694060
H	3.095037	−0.250930	2.580649
H	3.096774	−0.240693	0.826048

6.6 Paper VI

On the photoproduction of DNA/RNA cyclobutane pyrimidine dimers.

I.González-Ramírez, D. Roca-Sanjuán, T. Climent, J. J. Serrano-Pérez, M. Merchán and L. Serrano-Andrés

Theor. Chem. Acc., 128, 705-711 (2011).

Theor Chem Acc (2011) 128:705–711
DOI 10.1007/s00214-010-0854-z

REGULAR ARTICLE

On the photoproduction of DNA/RNA cyclobutane pyrimidine dimers

Israel González-Ramírez · Daniel Roca-Sanjuán · Teresa Climent · Juan José Serrano-Pérez · Manuela Merchán · Luis Serrano-Andrés

Received: 31 August 2010 / Accepted: 3 November 2010 / Published online: 23 November 2010

Abstract The UV photoreactivity of different pyrimidine DNA/RNA nucleobases along the singlet manifold leading to the formation of cyclobutane pyrimidine dimers has been studied by using the CASPT2 level of theory. The initially irradiated singlet state promotes the formation of excimers between pairs of properly oriented nucleobases through the overlap between the π structures of two stacked nucleobases. The system evolves then to the formation of cyclobutane pyrimidine dimers via a shearing-type conical intersection activating a [2 + 2] photocycloaddition mechanism. The relative location of stable excimer conformations or alternative decay channels with respect to the reactive degeneracy region explains the differences in the photoproduction efficiency observed in the experiments for different nucleobases sequences. A comparative analysis of the main structural parameters and energetic profiles in the singlet manifold is carried out for thymine, uracil, cytosine, and 5-methylcytosine homodimers. Thymine and uracil dimers display the most favorable paths, in contrast to cytosine. Methylation of the nucleobases seems to increase the probability for dimerization.

Keywords CASPT2 · Photochemisty · Conical Intersection · DNA/RNA · Cyclobutane Pyrimidine Dimers

Published as part of the special issue celebrating theoretical and computational chemistry in Spain.

I. González-Ramírez · T. Climent · M. Merchán · L. Serrano-Andrés
Instituto de Ciencia Molecular, Universitat de València, P. B. Box 22085, 46071 Valencia, Spain

D. Roca-Sanjuán (✉)
Department of Quantum Chemistry, Uppsala University, Box 518, 75120 Uppsala, Sweden
e-mail: Daniel.Roca@kvac.uu.se

J. J. Serrano-Pérez
Department of Chemistry, Computational and Structural Research Group, Imperial College London, SW7 2AZ London, UK

1 Introduction

The UV radiation directly absorbed by the nucleic acids can produce a large number of lesions [1, 2], the most common corresponding to dimer or adduct formations involving adjacent pyrimidine bases of the DNA/RNA strand [3–7]. Two types of dimerized products are frequently found: the cyclobutane pyrimidine dimers (CPDs), in which the adduct is formed via a [2 + 2] photocycloaddition involving the C=C double bonds of the pyrimidines, and the pyrimidine (6-4) pirimidone dimers, usually referred to as (6-4) photoproducts, (6-4)PPs, in which the cycloaddition takes place between C=C and C=O double bonds of the adjacent molecules. CPD lesions are more frequent than (6-4)PPs [3, 4]. Under UV-C (254-nm) radiation conditions, between two and ten CPDs per million bases are formed [5]. Two more orders of magnitude of lower-energy UV-B radiation are required to obtain the same result. Even though, in theory, various CPD diastereoisomers are possible, only the *cis-syn* conformer is found in the double helix. The *trans-syn* configuration can be also obtained, with lower yields, in single or double strands of denaturalized DNA, where the tertiary and secondary structure are lost to some extent with respect to the standard double helix [3]. Although the pyrimidine bases can react to form CPDs in many combinations, we will focus here on the homodimers of the canonical DNA/RNA bases thymine (T), uracil (U), cytosine (C) (and their adducts

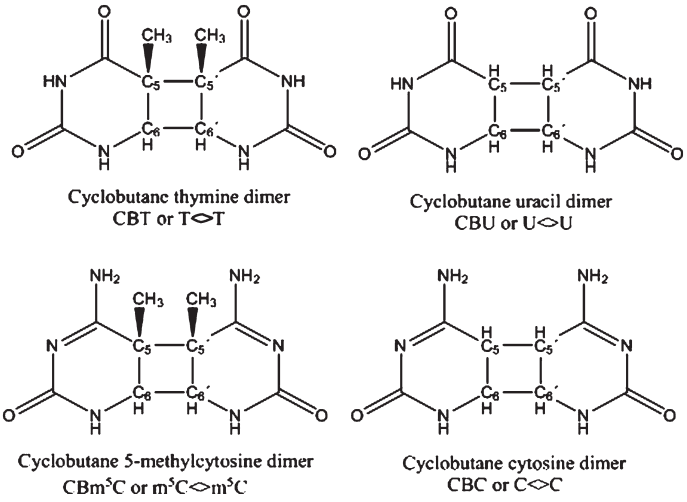


Fig. 1 Structures and labeling of the DNA/RNA cyclobutane pyrimidine homodimer photoproducts

CBT, CBU, and CBC), and on the derivative 5-methylcytosine (m^5C and CBm^5C), as displayed in Fig. 1.

Despite all CPDs can be considered DNA/RNA lesions, not all of them are actually mutagenic hotspots, something that depends on the rate of the enzymatic repair mechanisms or of the transitions from one pyrimidine to another, for instance by deamination plus ketonization in C and m^5C to U and T, respectively. Therefore, despite CBT dimers are more frequently found among the photodimers detected in DNA in different conditions [6], CC and m^5Cm^5C sites are potentially more damaging, because they might efficiently give rise to an actual mutation. Additionally, and in general, CPDs seem to undergo more rapid deamination than individual nucleobases [8–11]. The important role of the noncanonical m^5C derivative [12], which is found in significant amounts in the DNA of many eukaryotic organisms (5% in human and calf thymus DNA and 31% in wheat DNA [13, 14]), is recognized since a decade ago. Whereas earlier works were unsuccessful in detecting significant amounts of m^5C-containing CPDs [15, 16], it was found later that cyclobutane dimers of m^5C are formed when irradiated with either UV-C or UV-B [12, 17, 18] and that methylation increases the photoproduction with respect to the canonical nucleobase [17]. Tommasi et al. [18] analyzed the CPDs formation in different combinations of pyrimidine nucleobases irradiated with UV-C, UV-B, and sunlight. The methylated DNA base m^5C was the preferred target for CPD production when the natural sunlight was used.

Femtosecond spectroscopy has proved that T-dimerization is an ultrafast photoreaction in which CBTs are fully formed $\sim$1 ps after UV illumination [19]. From a theoretical standpoint, a few studies have confirmed for CBC [20] and CBT [21–24] dimers an ultrafast nonadiabatic photoreaction involving a barrierless path along the low-lying singlet excited (S_1) state. The concerted mechanism for the [2 + 2] photocycloaddition of two C- or T-molecules is mediated by the presence of a conical intersection (CI), an energy-degenerate structure between the low-lying singlet excited (S_1) and the ground state (S_0). The shearing-type CI structure—in which the nucleobases ethylenic C_5–C_6 and C_5'–C_6' bonds laid parallel (parallelogram type) and elongated—, connects the S_1 and S_0 states and allows an efficient internal conversion process [20]. Intrastrand nucleobase sequence and relative orientations were also proved to be essential for an efficient photoreaction to take place. In previous works on CC, TT, and UU pairs [20, 23, 25], we showed that those conformations maximizing the overlap between the π structures of stacked nucleobases formed favorable excimer arrangements, being the most stable structures leading to the photoreactive arrangements, in agreement with the higher yields obtained for photoproducts with *cis*-type parallel face-to-face conformations for the base pairs. Additionally, we determined that the formation of CPDs can be also obtained in the triplet manifold through a biradical intermediate involving a singlet–triplet crossing $(S_0/T_1)_X$ relating the ground (S_0) and low-lying triplet state (T_1) [20, 26, 27], and explaining the high yields of CPDs detected in solution in presence of external photogenotoxic substances acting as triplet–triplet photosensitizers [3, 28, 29].

Obtaining an accurate mapping of the relative energies of the excimer and CI structures in the different dimers is crucial to understand the formation mechanisms and the observed photoreaction yields. Earlier, we attributed the low yield measured for the CBC formation when compared to CBT to the competitive presence of stable excimer conformations and the CI responsible for nucleobase monomer deactivation (CI_{mon}) at energies similar or lower than the [2 + 2] photocycloaddition sheared-type CI (CI_{dim}) in CC. In contrast, for TT CI_{dim} is the most stable structure [23, 26], favoring the reactive process. The goal of the present contribution is to complement the theoretical study on the CPDs formation in cytosine and thymine homodimers with the analysis of the photodimerization in the other pyrimidine nucleobase, uracil, and the noncanonical C_5-methylated cytosine base, providing a wider overview of the photodimerization process in pyrimidines. It is intended to give a rationale on the predominant presence of some specific photodimers over others in terms of the differences found among the potential energy hypersurfaces of the ground state and the lowest excited state related to the [2 + 2] photocycloaddition. Concerning U and m^5C, and except for a previous DFT study on m^5C [30], to our knowledge no reliable study at the required multiconfigurational level has been reported. Therefore, the present work is the first ab initio theoretical determination of the mechanism of photodimerization in uracil and the noncanonical m^5C nucleobase, allowing an overall comparison of the CPDs formation mechanisms for the most important nucleobases homodimers.

2 Methods and computational details

In order to provide a comparative study of the cyclobutane dimer formation in pyrimidines, the same methodology as that previously employed in the cytosine and thymine homodimers has been used in the present work for the theoretical determination of the mechanism of photodimerization of two uracil and two 5-methylcytosine dimers [20, 23, 26]. The ANO-S basis set contracted to C,N,O[3s2p1d]/H[2s1p] was used throughout. Multiconfigurational CASPT2(14/10) and CASPT2(12/12) calculations [31–34] were performed, respectively, for the monomers and the dimers. This included CASSCF geometry optimizations of the singlet states minima, minimum energy paths (MEPs), and minimum energy crossing points (MECPs) determinations in the potential energy hypersurfaces (PEHs) of a system of two nucleobases, corrected at the CASPT2 level using point-wise calculations [20, 35]. Energies were also corrected on the effect of the basis set superposition error (BSSE) using the counterpoise (CP) procedure [20, 36]. As shown previously [20, 25, 37], the inclusion of BSSE is crucial to accurately describe the binding energies and compare the different mechanisms. The MOLCAS quantum-chemistry code was employed throughout [38]. In order to minimize weakly interacting intruder states, the imaginary level-shift technique with a parameter 0.2 au has been employed [39]. In order to mimic the interaction of pyrimidines in the biologically relevant *cis-syn* diastereoisomer, geometry optimizations were initially performed within the constraints of the C_s symmetry. Since the CASSCF structures of the dimer on the MECPs do not represent a crossing at the CASPT2 level, MECPs were finally obtained with the CASPT2 methodology. Further technical details can be found in previous publications [20, 23].

3 Results and discussion

As mentioned in the introduction, previous experimental and theoretical studies [19–23, 40, 41] determined a general mechanism for the ultrafast photoproduction of CPDs in nucleobases oligomers strands. It is suggested that after initial radiative population of delocalized exciton states on the nucleobases multimers, the system evolves in an ultrafast manner to either a localized excited state of the nucleobase monomer or a fluorescent excimer/exciplex state, depending on the smaller or larger degree of stacking, respectively [37]. Relaxation along the monomer path should be ultrafast ($\tau < 2$ ps), as known in the isolated systems [34, 42, 43], whereas from the long-lived excimer/exciplex state the system is expected to decay to the ground state slowly ($\tau > 10$ ps). The slower relaxation paths for the stacked nucleobases, expanding from 10 to 200 ps, have been found dominant in the decay dynamics of dinucleotides after excitation at 267 nm (4.96 eV) [19], and it can be assigned to the formation of more or less stable excimer/exciplex structures. There are additional accessible evolution paths for the pairs of bases, in particular for pyrimidine nucleobases, in which the excimer/exciplex behave as precursor for the formation of cycloadducts like CPDs or (6-4)PPs dimerized structures. As shown, there is always a $(S_0/S_1)_{CI}$ degeneracy region responsible for the photoprocess to take place. In order to elucidate the basics of the relaxation mechanisms in DNA photochemistry, it is required to provide a common framework in which all basic structures—CI of the monomer, excimers, and CI of the photoreaction—are determined at the same level of theory. The comparison between the mechanisms of photodimerization in the considered nucleobases will help to explain the distinct efficiency for photodimers production found in the experiments. Table 1 contains the relative energies, with respect to two isolated nucleobases in the ground state and selected

Table 1 Selected structural distances (R/Å) and energies (ΔE/eV) of the relevant structures along the singlet manifold in the photoproduction mechanism of cyclobutane pyrimidine dimers

Monomer/homodimer	R (C_5–C_5')	R (C_6–C_6')	ΔE
Mon (S_1)			
C* + C	–	–	4.41[a]
m^5C^* + m^5C	–	–	4.31[b]
U* + U	–	–	5.02[a]
T* + T	–	–	4.89[a]
CI_{mon} $(S_1/S_0)_{CI}$			
C	–	–	3.60[a]
m^5C	–	–	3.64[b]
U	–	–	3.90[a]
T	–	–	4.00[a]
Exc (S_1)			
C*C	3.427	3.219	3.31[c]
$m^5C^*m^5C$	3.594	3.346	3.46
U*U	2.503	2.315	3.68
T*T	2.650	2.380	3.64[d]
CI_{dim} $(S_1/S_0)_{CI}$			
CC	2.258	2.170	3.51[c]
m^5Cm^5C	2.491	2.056	3.56
UU	2.218	2.170	3.47
TT	2.350	2.220	3.26[d]
$(CPD)_{S0}$			
CBC	1.611	1.601	0.78[c]
CBm^5C	1.648	1.593	0.71
CBU	1.601	1.568	0.23
CBT	1.637	1.595	0.40[d]
$(CPD)_{S1}$			
CBC	1.611	1.601	5.35[c]
CBm^5C	1.648	1.593	5.63
CBU	1.601	1.568	5.87
CBT	1.637	1.595	5.88[d]

All energies are referred to the singlet ground state of two isolated nucleobases (see Figs. 2 and 3)

[a] Ref. [34]

[b] This work

[c] Ref. [20]

[d] Ref. [23]

geometry parameters for the five relevant structures in the photodimerization mechanism in the singlet manifold: the monomeric UV absorption (Mon), the monomer CI (CI_{mon}), the excimer minimum (Exc), the dimer CI (CI_{dim}), and the photodimer (CPD) in the four homodimers studied here: thymine, uracil, cytosine, and 5-methylcytosine. Figure 2 summarizes the basics of the mechanism.

UV solar radiation extends from the edge of the visible light at 400 nm (3.1 eV) to the far UV range (200 nm, 6.2 eV) and beyond. As mentioned previously, the yield of

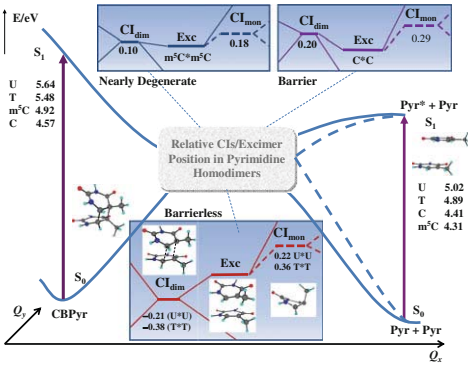


Fig. 2 Proposed scheme, based on actual CASPT2 results, for the decay path of the lowest singlet excited state S_1 of the U, T, C, and m^5C dimers involving the relaxed excimer and the conical intersection $(S_1/S_0)_{CI}$ leading to ground state cyclobutane pyrimidines (CBPyr or CPDs). Values inside the *boxes* correspond to the energies of the conical intersection of the dimer (CI_{dim}) and the monomer (CI_{mon}) structures with respect to the lowest-energy excimer (Exc). TT structures shown as an illustration

photoproduction of CPDs is wavelength dependent and increases with the energy of the absorbed radiation, because the available excess energy required to surmount energy barriers is larger for high-energy irradiations. In any case, the photoreactive process takes place ultimately in the S_1 excited state. The vertical excitation energy computed for the S_1 (HOMO → LUMO) state ranges from 4.3 eV in m^5C to 5.0 in U for the four studied nucleobases. Independently from the procedure that the S_1 is reached—directly or from decay from higher-lying singlet states—the nucleobases strand will find many arrangements in which two of the monomers will overlap their π structures, yielding very favorable conformations ready to evolve toward stable excimer/exciplex (excimer here since we are dealing with homodimers) minima in the excited state. Even when for weakly stacked pairs the excitation may localize in the monomer and decay to the ground state of the nucleobase through the CI_{mon} $(S_0/S_1)_{CI}$ (see Table 1), many arrangements will be favorable for the formation of excimers. Even in their most common biological conformation, B-DNA, it is considered that nucleobases form weakly interacting or static excimers [37]. These two types of situations, decay in the monomer through localization with access to the monomer CI and formation of excimers, can be considered responsible for the ultrafast (>2 ps) and fast (>10 ps) decays observed in femtosecond transient absorption experiments, especially in purine strands [19].

Although different excimer arrangements of distinct stability are possible, we have obtained the most stable

excimer in the four homodimers studied here as a face-to-face quasi-parallel conformation that maximizes the π overlap (see Fig. 3 as an example). At the minimum (see Table 1), the intramolecular distances between the analogous C=C double bond atoms of the adjacent monomers ($\sim$2.3–2.6 Å) are almost one Å shorter for T*T and U*U when compared to the other homodimers. T*T and U*U seem to display a stronger interaction (probably because of the two carbonyl groups), showing bonding energies with respect to the isolated monomers near 1.3 eV. In contrast, C*C and m^5C*m^5C have larger intermonomer distances ($\sim$3.2–3.6 Å) and lower binding energies ($\sim$0.9–1.1 eV). The four homodimers have also a CI degeneracy region connecting the excited S_1 and the ground S_0 state in which the system displays very similar structures, with the two monomers arranged in a quasi-parallel shearing-type conformation (see Fig. 3). This CI_{dim} $(S_0/S_1)_{CI}$ structure is responsible for the nonadiabatic [2 + 2] cyclophotoaddition reaction leading from the pair of stacked nucleobases to the final CPD photoproduct in the ground state.

The relative position of the most stable excimer conformation and the CI_{dim} $(S_0/S_1)_{CI}$ region is different in the four homodimers studied. For TT and UU pairs, CI_{dim} is the lowest-energy structure of all studied S_1 hypersurface, lying near 0.4 and 0.2 eV, respectively, below the stable excimer. It is not surprising that all the computed minimum energy paths (MEPs) in TT leads in a barrierless way from different excimer arrangements to the CI_{dim} structure [21–23], as it is the case here with UU. Additionally, the CI_{dim} structure is near 0.7 (TT) and 0.4 eV (UU) lower in energy than the CI of the respective monomers, T and U. Therefore, there is no decay process that can efficiently compete with the access to CI_{dim} and subsequent nonadiabatic transfer of energy to reach CBT or CBU, except for those systems evolving toward the splitting of the monomers. The described PEH profile is perfectly compatible with the observed high yields of production of CBT and CBU when compared with other adducts [3].

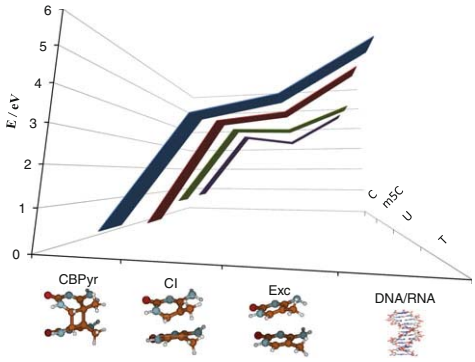


Fig. 3 Comparison of the photoreactive pathways leading to CPDs (or CBPyrs) formation along the singlet manifold for the C, m^5C, U, and T base pairs. The m^5Cm^5C structures obtained in this work are shown

Regarding CC and m^5Cm^5C, the CI_{dim} structure has been computed 0.2 and 0.1 eV, respectively, higher in energy than their relaxed excimer minima. That means that there exist conformations that will become competitive with the nonadiabatic reaction, decreasing the rate and yield of the photoprocess in the two systems, especially in CC. The same argument can be used when comparing the dimer CI with the monomer CI, which is energetically just 0.1 eV above in both cases, opening a new decay route which may compete efficiently with the formation of CBC and CBm^5C. Overall, the lower yields observed for CC tandems in contrast to TT [44] can be therefore understood by the presence in C-based dimers of several competitive structures—stable excimers and monomer decay routes—close in energy to the reactive CI. Regarding the comparison of CC and m^5Cm^5C, the former displays excimer structures somewhat more stable with respect to the CI_{dim} (0.2 eV) than in the case of the methyl derivative, in which CI_{dim} and excimer become almost degenerate. Even when the difference is small in both cases, this profile may explain the slightly higher efficiency found in the production of m^5C dimers [12, 17]. Figure 3 depicts a scheme of the relative energy levels for the studied systems.

With respect to the geometry of the dimer CI in the four systems studied (rhomboid or parallelogram type, typical of the [2 + 2] cycloaddition [20]), and although the structural discrepancies are small, they might be relevant (see Fig. 3). Whereas the C_6–C_6' distance remains similar in the various cases, C_5–C_5' shows much more noticeable differences, reflecting a major or minor distortion with respect to the ideal rhomboid geometry. Hence, the m^5C dimer presents the most distorted structure, with the largest C_5–C_5' distance, 2.491 Å, among the studied bases. This elongation is probably a consequence of the steric effect caused by the presence of both the methyl and amino groups. Thymine, also with a methyl group, comes next with a value of 2.350 Å, followed by the less distorted cytosine and uracil systems, with distances of 2.258 and 2.218 Å, respectively. As it could be expected, methylation increases the C_5–C_5' bond length in all computed structures, excimers, CIs, and adducts.

Once the CI_{dim} crossing is reached, dimers can evolve, in an ultrafast process, to form cyclobutane pyrimidines, whose structure is shown for the m^5C in Fig. 3. The singlet pathway represented in this figure and Fig. 2 allows us to better understand the photodimer production mechanism that operates after the initial excitation taking place in DNA/RNA strands. Also, photodimers can absorb UV light

to finally obtain the separated base monomers through a photoreversibility process as shown in Fig. 2. In a previous study, performed using the CASPT2/MM methodology for the $dG_{18}xdC_{18}$ system, i.e., a 18-base-pair-long double helix of poly(C)-poly(G) surrounded by water molecules [45], we determined the barrierless path connecting the CBC S_1 state with the dimer CI, which may later lead to the separation of the monomers. Although the required excitation energy in the dimer to initiate the photoreversion process (4.57–5.64 eV, within the UV-C range) is higher than the energy needed for the monomer excitation, it represents a competitive photoreaction that may reduce the yield of the photoadduct. The formation of CC photodimers was studied by Tommasi et al. [18] under UV-B and UV-C irradiations, showing an increase in the CPD production when UV-B was used, in clear agreement with the results obtained here.

4 Conclusions

The formation of cyclobutane pyrimidine dimers following UV absorption of pairs of stacked DNA/RNA nucleobases has been studied for the homodimers of thymine, uracil, cytosine, and 5-methylcytosine at the theoretical, ab initio CASPT2 multiconfigurational level. Determination of states minima, minimum reaction paths, and conical intersections in the low-lying singlet states of the dimers and monomers has led us to establish a unified mechanism for the adduct formation, which proceeds via a nonadiabatic [2 + 2] photocycloaddition reaction in the singlet manifold. A sheared-like conical intersection connecting the S_1 and S_0 states of the dimer, CI_{dim} $(S_0/S_1)_{CI}$, is the funnel controlling the reactive process. The relative position of the dimer CI with respect to stable excimer structures or the monomer decay CI region determines the efficiency of the photoreactivity. Thus, the dimer CI is the lowest-energy structure in thymine and uracil homodimers, lacking other direct competitive decay processes and favoring the higher efficiency observed in the formation of the CBT and CBU adducts. In contrast, the dimer CI computed in cytosine and 5-methylcytosine is close in energy to the most stable excimer conformation, displaying a face-to-face quasi-parallel structure, and also to the corresponding monomer CIs. Those features can easily compete with the nonadiabatic dimerization reaction decreasing its rates and yields. In fact, the stable excimer structure is lower in energy than the dimer CI in the latter systems, in which a 0.2- and 0.1-eV barrier, respectively, is found to reach the degeneracy region. Those profiles can explain the higher yields found in TT sites when compared with CC. Also, it is shown that methylation, in T and m^5C with respect to U and C, respectively, destabilizes the excimer structures with respect to the CI. Thus, in the case of cytosine, the methylation favors the probability for dimerization by decreasing energy barriers leading to the reaction. It must be highlighted that a high level of theory is required to obtain accurate profiles. CASPT2, including exhaustively the correlation energy, in contrast to simply CASSCF (or lower level approaches like TDDFT) is required to remove undesired differential correlation effects [33, 46], whereas the inclusion of the BSSE effect is essential to obtain accurate and comparable binding energies among the different systems.

Acknowledgments In memory of Luis Serrano-Andrés, he was an excellent scientist and friend. Research supported by projects CTQ2007-61260, CTQ2010-14892, and CSD2007-0010 Consolider-Ingenio in Molecular Nanoscience of the Spanish MEC/FEDER and the Generalitat Valenciana. It has also received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreements n° 255363 and n° 251955.

References

1. Olivera BM (1978) Proc Natl Acad Sci USA 75:238−242
2. Setlow RB (1974) Proc Natl Acad Sci USA 71:3363−3366
3. Cadet J, Vigny P (1990) In: Morrison H (ed) Bioorganic photochemistry. Wiley, New York, p 1
4. Friedberg EC, Walker GC, Siede W, Wood RD, Schultz RA (2006) Ellenberger T (ed) DNA Repair and mutagenesis. ASM Press, Washington, DC
5. Björn LO, McKenzie RL (2008) In: Björn LO (ed) Photobiology. The science of life and light. Springer, New York, p 503
6. Douki T, Cadet J (2001) Biochemistry 40:2495−2501
7. Schuch AP, Menck CFM (2010) J Photochem Photobiol B Biol 99:111−116
8. You YH, Li C, Pfeifer GP (1999) J Mol Biol 293:493−503
9. You YH, Pfeifer GP (2001) J Mol Biol 305:389−399
10. Ikehata H, Masuda T, Sakata H, Ono T (2003) Environ Mol Mutagen 41:280−292
11. Lee DH, Pfeifer GP (2003) J Biol Chem 278:10314−10321
12. Shetlar MD, Basus VJ, Falick AM, Mujeeb A (2004) Photochem Photobiol Sci 3:968−979
13. Friso S, Choi SW, Dolnikowski GG, Selhub J (2002) Anal Chem 74:4526–4531
14. Adams RLP, Burdon RH (1984) Molecular biology of DNA methylation. Springer, New York, pp 6–7 (Table 1.2)
15. Ehrlich M, Dove MF (1983) Photobiochem Photobiophys 6:121–126
16. Ehrlich M, Dove MF, Huang LH (1986) Photobiochem Photobiophys 11:73–79
17. Mitchell DL (2000) Photochem Photobiol 71:162–165
18. Tommasi S, Denissenko MF, Pfeifer GP (1997) Cancer Res 57:4727–4730
19. Schreier WJ, Schrader TE, Soller FO, Gilch P, Crespo-Hernández CE, Swaminathan VN, Carell T, Zinth W, Kohler B (2007) Science 315:625–629
20. Roca-Sanjuán D, Olaso-González G, González-Ramírez I, Serrano-Andrés L, Merchán M (2008) J Am Chem Soc 130:10768–10779
21. Boggio-Pasqua M, Groenhof G, Schäfer LV, Grubmüller H, Robb MA (2007) J Am Chem Soc 129:10996–10997

22. Blancafort L, Migani A (2007) J Am Chem Soc 129:14540–14541
23. Serrano-Pérez JJ, González-Ramírez I, Coto PB, Merchán M, Serrano-Andrés L (2008) J Phys Chem B 112:14096–14098
24. Zhang WY, Yuan SA, Li AY, Dou YS, Zhao JS, Fang WH (2010) J Phys Chem C 114:5594–5601
25. González-Ramírez I, Climent T, Serrano-Pérez JJ, González-Luque R, Merchán M, Serrano-Andrés L (2009) Pure Appl Chem 81:1695–1705
26. Climent T, González-Ramírez I, González-Luque R, Merchán M, Serrano-Andrés L (2010) J Phys Chem Lett 1:2072–2076
27. González-Luque R, Climent T, González-Ramírez I, Merchán M, Serrano-Andrés L (2010) J Chem Theor Comp 6:2103–2114
28. Bosca F, Lhiaubet-Vallet V, Cuquerella MC, Castell JV, Miranda MA (2006) J Am Chem Soc 128:6318–6319
29. Kwok WM, Ma C, Phillips DL (2008) J Am Chem Soc 130:5131–5139
30. Xiaoyi L, Eriksson LA (2005) Chem Phys Lett 401:99–103
31. Andersson K, Malmqvist PÅ, Roos BO (1992) J Chem Phys 96:1218–1226
32. Roos BO, Andersson K, Fülscher MP, Malmqvist PÅ, Serrano-Andrés L, Pierloot K, Merchán M (1996) Adv Chem Phys 93:219
33. Merchán M, Serrano-Andrés L (2005) In: Olivucci M (ed) Computational photochemistry. Elsevier, Amsterdam, p 35
34. Serrano-Andrés L, Merchán M (2009) J Photochem Photobiol C Photochem Rev 10:21–32
35. Serrano-Andrés L, Merchán M, Lindh R (2005) J Chem Phys 122:104107–104110
36. Boys SF, Bernardi F (2002) Mol Phys 100:65
37. Olaso-González G, Merchán M, Serrano-Andrés L (2009) J Am Chem Soc 131:4368–4377
38. Aquilante F, De Vico L, Ferré N, Ghigo G, Malmqvist PÅ, Pedersen T, Pitonak M, Reiher M, Roos BO, Serrano-Andrés L, Urban M, Veryazov V, Lindh R (2010) J Comp Chem 31:224–247
39. Forsberg N, Malmqvist PÅ (1997) Chem Phys Lett 274:196–204
40. Crespo-Hernández CE, De La Harpe K, Kohler B (2008) J Am Chem Soc 130:10844–10845
41. Takaya T, Su C, De La Harpe K, Crespo-Hernández CE, Kohler B (2008) Proc Natl Acad Sci USA 105:10285–10290
42. Crespo-Hernández CE, Cohen B, Hare PM, Kohler B (2004) Chem Rev 104:1977–2020
43. Serrano-Andrés L, Merchán M (2008) Shukla MK, Leszczynski J (eds) Photostability and photoreactivity in biomolecules: quantum chemistry of nucleic acid base monomers and dimers. In: Radiation induced molecular phenomena in nucleic acids: a comprehensive theoretical and experimental analysis. Springer, The Netherlands, p 435
44. Douki T, Cadet J (2001) Biochemistry 40:2495–2501
45. Roca-Sanjuán D, Olaso-González G, Rubio M, Coto PB, Merchán M, Ferré N, Ludwig V, Serrano-Andrés L (2009) Pure Appl Chem 81:743–754
46. Serrano-Andrés L, Merchán M (2005) J Mol Struct Theochem 729:99–108

6.7 Paper VII

On the N1-H and N3-H bond dissociation in uracil by low energy electrons: A CASSCF/CASPT2 study.

I.González-Ramírez, J. Segarra-Martí, L. Serrano-Andrés, M. Merchán, M. Rubio and D. Roca-Sanjuán

J. Chem. Theor. Comput., 8, 2769-2776 (2012).

JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

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*[,‡]

[†]Instituto de Ciencia Molecular, Universitat de València, P.O. Box 22085, 46071 València, Spain

[‡]Department of Chemistry−Ångström, Theoretical Chemistry Program, Uppsala Univeristy, P.O. Box 518, 75120 Uppsala, Sweden

Supporting Information

ABSTRACT: The dissociative electron-attachment (DEA) phenomena at the N_1−H and N_3−H bonds observed experimentally at low energies (<3 eV) in uracil are studied with the CASSCF/CASPT2 methodology. Two valence-bound π^- and two dissociative σ^- states of the uracil anionic species, together with the ground state of the neutral molecule, are proven to contribute to the shapes appearing in the experimental DEA cross sections. Conical intersections (CI) between the π^- and σ^- are established as the structures which activate the DEA processes. The N_1−H and N_3−H DEA mechanisms in uracil are described, and experimental observations are interpreted on the basis of two factors: (1) the relative energy of the $(U-H)^-$ + H fragments obtained after DEA with respect to the ground-state equilibrium structure (S_0) of the neutral molecule (threshold for DEA) and (2) the relative energy of the CIs also with respect to S_0 (band maxima). The π_1^- state is found to be mainly responsible for the N_1−H bond breaking, whereas the π_2^- state is proved to be involved in the cleavage of the N_3−H bond.

INTRODUCTION

Irradiation of nucleic acids can cause damage in several ways.[1,2] When cellular DNA/RNA is ultraviolet (UV) irradiated, the formation of lesions among adjacent pyrimidine bases (namely cyclobutane pyrimidine dimers) is the most frequent route of damage.[3,4] Nonetheless, other reactive pathways are available in DNA/RNA that take place primarily through secondary reactions involving species generated within the complex cellular system in an initial ionization step (electrons, OH, H, DNA/RNA radicals).[2] These events mostly remove electrons either from the molecules' valence orbitals in chemical bonds or from the inner core of individual atoms, subsequently causing structural lesions.[5] The remarkable work of Boudaiffa et al. unveiled the fact that low energy (3−20 eV) electron attachment in thin films of DNA could lead, through decay of temporary anion states, to single and double strand break.[6] Furthermore, data from the international commission on radiation units and measurements[7] have shown that these electrons (3−20 eV) lose the kinetic energy within picoseconds through collisions, and yet it is possible for electrons in the energy range 0.1−3 eV to cause dissociations in DNA/RNA.[5,8] Excited states of the temporary anions involving π^* electrons of the nucleobase and σ^* electrons of the sugar−phosphate C−O bond have been shown to be relevant in the mechanism of DNA strand cleavage.[2,9,10]

Nucleobases themselves can undergo decomposition reactions after attachment of low-energy electrons.[11] The most efficient reactive process in the gas phase implies the dissociative electron attachment (DEA) of the temporary anionic base which dissociates into the $(nucleobase-H)^-$ ion plus a hydrogen atom. The DEA in uracil at subexcitation energies (<3 eV) has been observed in several experiments.[5,12−17] Measurements of the total yield of $(U-H)^-$ ion as a function of electron energy show a first peak at 0.69 eV, followed by a strong and sharp feature at 1.01 eV, and next a broad band with a maximum around 1.7 eV. On the basis of theoretical G2MP2[13,17] and P2MP2[5] calculations of the dissociation energies (D) for the N−H and C−H bonds of the neutral uracil molecule and the electron affinity (EA) of the radical formed by the loss of H from the particular site, the DEA values at the N−H positions were established at low energies. In particular, the thresholds for H abstraction (E) were obtained at 0.8, 1.4, 2.2, and 2.7 eV at the N_1, N_3, C_6, and C_5 sites, respectively ($E = D - EA$)[13,17,5] (see atom labeling in Figure 1). Regarding the energy range for the different N−H

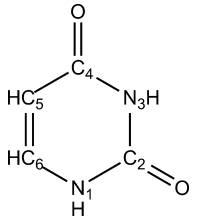


Figure 1. Numbering and atom labeling for uracil.

DEA, Ptasinska et al. measured the yields of $(U-H)^-$ production in the canonical uracil and the methylated derivative 3-methyluracil, concluding that the N_3−H channel is only responsible for the feature in the spectrum at energies above 1.4 eV, whereas the N_1−H channel is accessible also at lower energies.[17] Both theoretical and experimental results clearly

Received: February 21, 2012
Published: July 12, 2012

indicate that the two lowest peaks in the DEA experiments correspond to the N_1−H bond dissociation.

The nature of the anions involved in the mentioned DEA processes is more difficult to determine. Polar molecular systems can trap electrons in their long-range dipole field if they possess a dipole moment greater than about 2.5 D.[18−21] Such large polarity becomes a challenge both theoretically and experimentally for the correct determination of the EA[22] and the study of the low energy electron attachment behavior in a nucleic acid base such as uracil with a dipole moment of about 4.3 D.[18] The dilemma arises when this great polarity allows the existence of two different types of anions, dipole-bound (DB) and valence-bound (VB), which are expected to share the same energy region.[18,22] According to previous theoretical calculations, the vertical EA of the VB uracil anion is negative (VEA = −0.61 eV) and becomes very close to zero after vibrational relaxation (the adiabatic EA is −0.01 eV),[22] whereas the DB anion has an attractive electrostatic interaction between the electron and the molecule (EA = 0.085 eV).[18] Since the uracil nucleobase does not stabilize an electron in the valence shell, the relevance of the DB anion in the chemistry of the nucleobase is enhanced. Taking into account the coexistence of both DB and VB anions in the same range of energies and the fact that the lowest π anion appears in the electron transmission spectroscopy (ETS) data at lower energies than the threshold for the N−H DEA,[8,16] Burrow and co-workers[14−16] and Gallup et al.[23] focused on the DB anion in order to interpret the lowest two DEA signals measured in the experiments. The ETS technique, however, underestimates the energies of the anion, as shown by more recent high-level ab initio computations.[22] In the studies from Burrow and co-workers,[14−16] two-state configuration interaction calculations based on two restricted open-shell Hartree−Fock wave functions representing the DB and σ^- states were performed at different N−H distances.[15] The energies obtained, together with various empirical quantities, were then employed to build the Morse-based potential energy curves (PECs) and later to estimate the DEA signals. The peaks at 0.69 and 1.01 eV were attributed in such manner to vibrational Feshbach resonances (VFRs) of the DB anion state with vibrational levels ν = 2 and 3, respectively. Gallup et al. combined recently the finite element discrete model with the resonance *R*-matrix theory to calculate the positions of the VFRs, which were found 0.1−0.2 eV higher as compared to the experiments.[23] On the other hand, the broad band in the experimental DEA cross-section was supposed to be caused by the π_2^- VB anion,[14−16,23] according to the ETS results.[8] However, no clear evidence of this fact was provided.

The VB anion of uracil is more relevant in condensed phases, where DB anions are likely to be absent.[24−26] Although most experiments have been performed in the gas phase, there is evidence of hydrogen loss in other phases such as superfluid helium droplets kept at low temperatures[27] or even condensed-phase DNA.[28] Li et al. computed the VB DEA energy threshold for the N−H and C−H bonds in uracil with the DFT and CBS-Q methods.[29] The authors also calculated the PEC of the ground state of the anion along the N−H and C−H bonds at the DFT level. In agreement with other computations at the G2MP2 and P2MP2 levels,[5,13,17] the energy trend obtained for DEA at the N−H and C−H sites was N1 < N3 < C6 < C5.[29] In these studies, the lowest VB anion was considered as the responsible species for all the N−H and C−H DEA processes. Along the PECs, the π anion state was shown to evolve toward a σ state localized in the particular bond that dissociates. The single-reference methods employed by Li et al. did not permit them to study the role of excited states of the anion, such as the π_2^-, or to analyze properly the mixing of the π^- and σ^- states. The authors suggested further investigations able to describe the section of the PECs where states cross.

The present work is aimed to study, for the first time by means of the CASSCF/CASPT2 method,[33−36] the hydrogen dissociation process (DEA) that takes place in the uracil nucleobase at the N_1−H and N_3−H sites as a result of adding a low-energy electron into an empty π^* orbital. The role of the ground as well as the low-lying excited states of the anion in both N−H dissociations is intended. Conical intersections (CIs) between the PECs of the $\pi^-$and σ^-states along the N−H reactive coordinates will be interpreted as the points on the PECs which activate the hydrogen loss processes in the subsequent events that follow the electron attachment in the uracil nucleobase. Such crossing points have been found to be relevant features in the mechanism of several photochemical phenomena in organic molecules.

METHODS AND 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.[22] 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 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.[30] 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.[22] It was shown in previous studies of biphenyl and *p*-benzosemiquinone radical anions[31,32] 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 ($\langle r^2 \rangle$) 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.[22] No symmetry requirements (C_1 symmetry) were employed in the computations. The geometries of the neutral and the hydrogen dissociated $(U-H_1)^-$ and $(U-H_3)^-$ 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 zero-point vibrational energy correction (ZPVE) 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 PECs along the N–H reactive coordinates, including 4 additional orbitals: the σ^* orbitals related to the N_1–H and N_3–H bonds plus 2 additional diffuse orbitals required to stabilize the active space along the PECs. In total, there are 11 active electrons and 12 active orbitals [hereafter, CASSCF(11,12)]. The shape of the most relevant natural orbitals can be found in Figure S1. 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 which occur at the N–H sites. (1) The PECs between the neutral structure and the two $(U–H)^-$ 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. (2) 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. (3) 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 both the evolution of this state after electron attachment and the crossing with the dissociative σ^- excited state which ends in the fragmented nucleobase $[(U–H)^- + H]$. MEPs were built as steepest descendent paths,[37,38] 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. Test MEP calculations were also carried out with a smaller hyperspherical radius of 0.05 au and gradients computed at the CASSCF level with a small active space comprising 7 electrons distributed among 9 orbitals. Two occupied and one virtual π natural orbitals, with occupation numbers close to two and zero, respectively, were excluded from the active space.

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.[39] 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.[22]

All the calculations were performed with the CASSCF/CASPT2 method as implemented in the versions 7 of the MOLCAS quantum-chemistry package of software.[40,41]

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 (Franck–Condon transition). This state must evolve toward 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 $(U–H)^-$ fragments. Several theoretical magnitudes are valuable to describe the mechanism for DEA and to interpret the experimental observations: the vertical electron affinity (VEA) of the neutral uracil at the ground-state equilibrium structure, the relative energies between the neutral nucleobase plus the incident electron (reactants) and the H atom plus the $(U–H)^-$ ion (products), and the crossing regions which activate the DEA phenomenon. The corresponding CASSCF/CASPT2 data will be presented and discussed in the following sections.

Vertical Electron Affinities of the Uracil Neutral System. Table 1 compiles the present computed CASPT2 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[8] and reference values from a previous work.[22] 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.[22] 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.[42]

Table 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
$diff_{N1H}^-$	−0.84			
$diff^-$	−1.69			
$diff_{N3H}^-$	−1.74			
π_2^-	−2.24			−1.58

[a]This work. [b]Reference 22. [c]Reference 8.

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 ($diff_{N1H}^-$ and $diff_{N3H}^-$) 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 as reliable. The spatial extensions $\langle r^2 \rangle$ of these states are 177, 184, and 172 au^2, for $diff_{N1H}^-$, $diff^-$, and $diff_{N3H}^-$, 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.

Energy Thresholds for the Reaction U + e$^-$ → (U−H)$^-$ + H. The energy difference (E) between products and reactants of the DEA reaction has been estimated in the literature at different levels of theory by using the bond dissociation energy of the neutral uracil [D(N−H)] and the EA of the radical formed after H loss from the particular site [EA(N−H)], through the equation $E = D - EA$,[5,13,17,29] where E stands for the energy threshold of the reaction. Table 2 compiles the results obtained by Märk and co-workers with the P2MP2[5] and G2MP2[13,17] methods, data from Li et al.[29] with the DFT and CBS-Q methods, 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.[5] 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 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,[17] 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.

Table 2. Dissociation Energies (in eV) Corresponding to the N_1−H and N_3−H Bonds Computed at Different Levels of Theory with Zero-Point Vibrational Energies Included at the Same Level of Theory

	P2MP2[a]	G2MP2[b]	DFT[c]	CBS-Q[c]	CASPT2[d]
N_1H	0.8	0.8	0.71	0.84	0.61
N_3H	1.4	1.4	1.25	1.36	1.21

[a]Reference 5. [b]References 13, 17. [c]Reference 29. [d]This work.

Potential Energy Curves along the N−H Reactive Coordinates. The mapping of the PECs for the ground state of the neutral uracil molecule and the low-lying VB anionic states of the anion system was performed at the CASPT2 level from the Franck−Condon region at the ground-state equilibrium structure of the neutral nucleobase and along the reactive coordinates which lead to the U−H_1 and U−H_3 species. Figures 2 and 3 display the results obtained for the N_1−H and N_3−H DEA processes, respectively. Four anionic states are represented in the figures: 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. In agreement with Li et al.,[29] the DEA channel in which the N_1−H is dissociated appears at lower energies with respect to the N_3−H. The σ^- states have intrinsic dissociative character, although it is not clear from Figures 2 and 3. 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 2, the σ_{N1H}^- state shows initially spatial extensions, $\langle r^2 \rangle$, around 172 au^2 (dashed line), becoming localized at larger N_1−H distances with an average $\langle r^2 \rangle$ of 138 au^2. This latter value is close to the result found for the dissociated anion (133 au^2), which is a stable anion (positive EA), and also close to the $\langle r^2 \rangle$ extensions computed for the π_1^- (128−134 au^2), π_2^- (124−125 au^2), and S_0 (95−111 au^2) states. In Figure 3, similar findings are obtained. The first points of the σ_{N3H}^- state have $\langle r^2 \rangle$ around 169 au^2, and next, the electron becomes localized in the valence space, with an averaged $\langle r^2 \rangle$ value of 129 au^2, in line with the results obtained for the dissociated anion and the π_1^-, π_2^-, and S_0 states (126, 126−133, 124−125, and 105−109 au^2, respectively). Hence, descriptions on the starting region of the σ_{N1H}^- and σ_{N3H}^- 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.

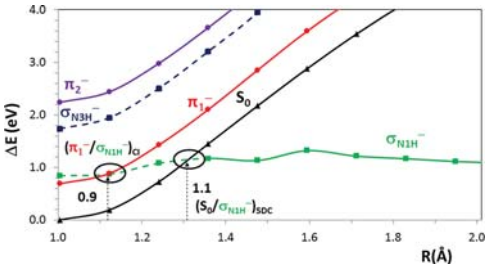


Figure 2. CASPT2 potential energy curves (PECs) for the ground state of the uracil neutral molecule (S_0) and the π_1^-, π_2^-, σ_{N1H}^-, and σ_{N3H}^- states of the anionic species along the N_1−H reactive coordinate. Dashed lines indicate points with mixed valence and dipole-bound states (see text). The corresponding CASSCF PECs can be found in Figure S2.

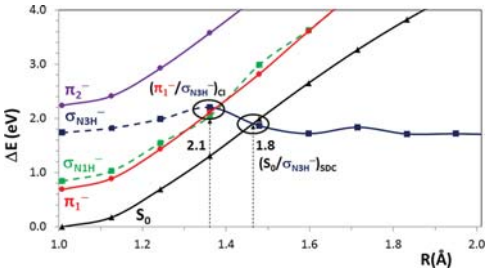


Figure 3. CASPT2 potential energy curves (PECs) for the ground state of the uracil neutral molecule (S_0) and the π_1^-, π_2^-, σ_{N1H}^-, and σ_{N3H}^- states of the anionic species along the N_3−H reactive coordinate. Dashed lines indicate points with mixed valence and dipole-bound states (see text). The corresponding CASSCF PECs can be found in Figure S3.

Two types of crossings can be distinguished in Figures 2 and 3, which are responsible for the activation of the hydrogen loss phenomena: (1) CIs between the π^- and σ^- states and (2) singlet−doublet crossings (SDCs) between the ground state of the neutral system and the σ^- states of the anion. While the role of the first crossing points in modern photochemistry is widely recognized, SDCs are not so well-known. 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. SDCs were suggested in previous

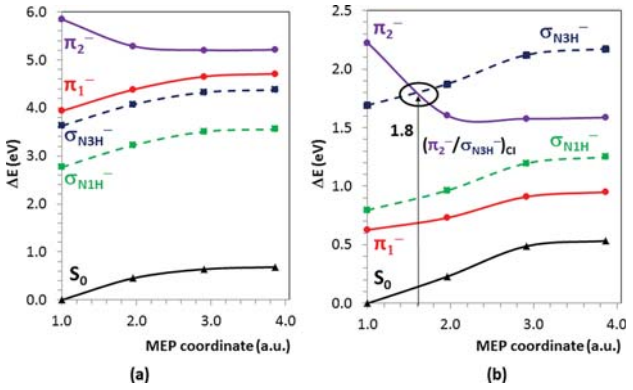


Figure 4. CASSCF (a) and CASPT2 (b) energies for the ground state of the uracil neutral molecule (S_0) and the π_1^-, π_2^-, σ_{N1H}^-, and σ_{N3H}^- states of the anionic species along the minimum energy path (MEP) of the π_2^- state. Dashed lines indicate points with valence and dipole-bound mixing (see text).

works as relevant structures to the charge transport phenomena in DNA/RNA.[43−45]

Figure 2 displays two PEC crossings involving the σ_{N1H}^- state of the anion: $(\pi_1^-/\sigma_{N1H}^-)_{CI}$ and $(S_0/\sigma_{N1H}^-)_{SDC}$. The $(\pi_1^-/\sigma_{N1H}^-)_{CI}$ connects the π_1^- state initially populated after electron-attachment with the state that drives the system toward the fragmented uracil anion $[(U-H_1)^- + H]$. Such structure, $(\pi_1^-/\sigma_{N1H}^-)_{CI}$, has an energy of 0.9 eV and therefore might mediate the DEA process at energies below 1 eV. Two points must be considered here: (1) the diffuse nature of the solutions of the σ_{N1H}^- state at this region avoids an accurate determination of the crossing points, which are expected higher in energy taking into account the dissociative nature of the σ^- states, and (2) tunneling effects will be important in this process as discussed by Scheer et al.,[15] and will operate in the other direction, decreasing the barrier to reach the σ_{N1H}^- state. The approximate $(S_0/\sigma_{N1H}^-)_{SDC}$ point appears around 1.1 eV. Apart from the relevance of such SDC in the DEA phenomenon, the energy of this point can be used as an estimate of the CI between the DB and the σ_{N1H}^- anion states. In fact, the electron is placed far from the molecule in these DB anions, and an attractive electrostatic interaction exists between both fragments; the EA of the DB anion in uracil is 0.085 eV.[18] Therefore, the equilibrium geometries of the DB anions and the neutral system are similar, and the corresponding PECs are close in energy, as explained by Probst et al.[11] In addition, the dipole moment of uracil is not changing significantly along the PEC toward the SDC region; a range of 4.2−4.7 D is obtained for the dipole moment of the neutral molecule among the first four points. Hence, the PEC for the DB anion is expected to be less than 0.1 eV far from the S_0 state (the canonical nucleobases, for example, with a large range of dipole moments of 2.56−6.55 D,[46] have small dipole-bound EAs in all the cases lower than 0.1 eV).[47] Accurate treatments of DB states require very time-demanding approaches[18] which are out of the scope of this work.

A sharp peak at 1.01 eV is observed in the cross sections measured in the DEA experiments, in contrast to the other bands which show less intense and broader features. Gallup and co-workers interpreted the shapes in the spectrum at 0.69 and 1.01 eV as VFRs of the DB anion with vibrational levels $\nu = 2$ and 3, respectively.[14−16,23] On the basis of our findings, which are in agreement with the investigations reported by Li et al.,[29] we set an alert to the conclusions obtained by the former authors:[14−16] the lowest VB π^- anion can also participate in the lowest peaks, in addition to the DB anion. The CI is interpreted in the present study as the point activating the DEA process. The SDC might also be relevant. In addition, assuming that the actual PEC of the DB-anionic state behaves approximately parallel to that related to the neutral (S_0), the DB-like state seems to activate the DEA process at similar energies than the VB π_1^- anion (see Figure 2). An accurate assignment of the peaks at 0.69 and 1.01 eV is not possible at the present because of the intrinsic limitations of the computational strategies employed. Nevertheless, a clear conclusion is obtained: not only the DB anion but also the VB π_1^- anion are responsible for the DEA processes at energies below and around 1 eV.

The CASSCF/CASPT2 results obtained for the PECs along the N_3−H reaction coordinate show also crossings between the π_1^- and S_0 states with the σ_{N3H}^- state (see Figure 3). The approximate energies obtained for these structures are around 2 eV, which locates the related processes in the broad band region of the spectrum. Hence, the electron attached in the π_1^- orbital can also undergo the cleavage of the N_3−H bond, in addition to the participation of the π_1^- state in the DEA reaction at the N_1−H site. Li et al. found similar conclusions by means of DFT constrained optimizations between the ground-state equilibrium structure of this state and the dissociated $(U-H_3)^-$ plus H products.[29] In contrast, on the basis of ETS measurements,[8] Burrow and co-workers[14−16] suggested the π_2^- state as responsible for the broad band in the spectrum, with maximum at 1.7 eV, being the vertical electron attachment energy of the second anionic state measured at −1.58 eV (see Table 1). Similar to the dissociation of the N_1−H bond, a contribution of the DB anion can be expected also here, not suggested previously. The dipole moment of the neutral system changes only slightly (3.7−4.2 D) between the Franck−Condon and the SDC crossing point; therefore, the PECs of the DB anion and the S_0 can be estimated parallel and within an energy separation of 0.1 eV, as in the equilibrium structure of the neutral uracil (0.085 eV).[18]

Figure 3 shows that the PECs of the π_2^- and σ_{N3H}^- states are indeed relatively close around the Franck−Condon region. Considering the aforementioned underestimation of the σ_{N3H}^- state in these points, both states might cross in the surroundings. A constrained optimization of the π_2^- state with a fixed N_3−H bond distance of 1.00 Å gives rise to an inversion of the PECs, bringing the π_2^- state below the energy of the σ_{N3H}^- state by 0.4 eV at the CASPT2 level. In any case, since the VEA of the π_2^- state is higher than the threshold for DEA at the N_3−H bond, unlike the situation found for the π_1^- state and the N_1−H bond, the initial evolution of the π_2^- state after electron attachment will be important to establish the thresholds for any contribution of this state to the DEA processes. Therefore, we carried out MEP calculations on the π_2^- PEC from the Franck−Condon structure. Figure 4 displays the CASSCF and CASPT2 energies for the relevant neutral and anionic states of uracil in the DEA process along the 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, the dynamical 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 σ_{N3H}^- states, $(\pi_2^-/\sigma_{N3H}^-)_{CI}$, 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 $(\pi_2^-/\sigma_{N3H}^-)_{CI}$, although the results for the crossing should be taken with caution due to the difficulties of the method to properly describe the σ^- state. Both the C═O as well as the N_1−C_2 and C_4−C_5 bonds are elongated in the region of the crossing point, as expected due to the antibonding character of the π_2^* orbital at those sites (see Figure S1). But the structure is still planar, and the N_3−H bond length is close to the Franck−Condon geometry. The N_3−H elongation will be driven by the σ_{N3H}^- dissociative state which is populated via the $(\pi_2^-/\sigma_{N3H}^-)_{CI}$. A more restrictive MEP calculation, with a shorter hyperspherical radius (0.05 au), was performed obtaining the same conclusions (cf. Figures S4 and S5). 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 $(\pi_2^-/\sigma_{N3H}^-)_{CI}$ 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.

CONCLUSIONS

The CASSCF/CASPT2 method, together with different computational strategies (LIIC, constrained optimizations, and MEPs), 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,[5,13,17,29] 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 toward the $(U−H)^-$ 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 σ_{N1H}^- state at 0.9 eV. Still, it can participate in the N_3−H DEA via a second CI, $(\pi_1^-/\sigma_{N3H}^-)_{CI}$, 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 σ_{N3H}^- state, $(\pi_2^-/\sigma_{N3H}^-)_{CI}$, in this region.

The present findings complement previous studies focused on the role of the DB states in the DEA phenomena.[14−16] The lowest peak and the sharp shape at 1 eV in the DEA cross section were assigned in those investigations to VFRs of the DB anion with ν = 2 and 3, respectively. The π_2^- was suggested to be involved in the broad band at higher energies. On the basis of our present study, the VB anion of uracil is predicted to be also responsible for the shapes at the lower region of the spectrum, and more insights are provided on the correspondence between the π_2^- state and the broad band.

ASSOCIATED CONTENT

Supporting Information

Additional figures and tables. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: Mercedes.Rubio@uv.es (M.R.), Daniel.Roca@kvac.uu.se (D.R.-S.).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

In memory of Luis Serrano-Andrés, who was an excellent scientist, mentor, and friend. Research supported by projects CTQ2010-14892 of the MICINN, the Consolider-Ingenio in Molecular Nanoscience of the Spanish MEC/FEDER CSD2007-0010, and the Generalitat Valenciana. It has also received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement 255363.

REFERENCES

(1) Connell, P.; Kron, S. J.; Weichselbaum, R. R. Relevance and Irrelevance of DNA Damage Response to Radiotherapy. *DNA Repair* **2004**, *3*, 1245−1251.

(2) Simons, J. How Do Low-Energy (0.1−2 eV) Electrons Cause DNA-Strand Breaks? *Acc. Chem. Res.* **2006**, *39*, 772−779.

(3) Cadet, J.; Vigny, P. The Photochemistry of Nucleic Acids. In *Bioorganic Photochemistry*, 1st ed.; Morrison, H., Ed.; John Wiley & Sons: New York, 1990; Vol. *1*, pp 1−273.

(4) Roca-Sanjuán, D.; Olaso-González, G.; González Ramírez, I.; Serrano-Andrés, L.; Merchán, M. Molecular Basis of DNA Photodimerization: Intrinsic Production of Cyclobutane Cytosine Dimers. *J. Am. Chem. Soc.* **2008**, *130*, 10768−10779.
(5) Hanel, G.; Gstir, B.; Denifl, S.; Scheier, P.; Probst, M.; Farizon, B.; Farizon, M.; Illenberger, E.; Märk, T. D. Electron Attachment to Uracil: Effective Destruction at Subexcitation Energies. *Phys. Rev. Lett.* **2003**, *90*, 188104−188107.
(6) Boudaiffa, B.; Cloutier, P.; Hunting, D.; Huels, M. A.; Sanche, L. Resonant Formation of DNA Strand Breaks by Low-Energy (3 to 20 eV) Electrons. *Science* **2000**, *287*, 1658−1660.
(7) *International Comission on Radiation Units and Measurements*; IRCU Report No. 31; 1979.
(8) Aflatooni, K.; Gallup, G. A.; Burrow, P. D. Electron Attachment Energies of the DNA Bases. *J. Phys. Chem. A* **1998**, *102*, 6205−6207.
(9) Kumar, A.; Sevilla, M. D. The Role of $\pi\sigma^*$ Excited States in Electron-Induced DNA Strand Break Formation: A Time-Dependent Density Functional Theory Study. *J. Am. Chem. Soc.* **2008**, *130*, 2130−2131.
(10) Kumar, A.; Sevilla, M. D. Role of Excited States in Low-Energy Electron (LEE) Induced Strand Breaks in DNA Model Systems: Influence of Aqueous Environment. *ChemPhysChem* **2009**, *10*, 1426−1430.
(11) Probst, M.; Injan, N.; Denifl, S.; Zappa, F.; Ingo, M.; Beikircher, M.; Ptasinska, S.; Limtrakul, J.; Märk, T. D.; Mauracher, A.; Scheier, P. Calculation of Processes Relevant to Reactions between Nucleic Acids and Free Electrons. *Chem. Eng. Commun.* **2008**, *195*, 1371−1381.
(12) Ptasinska, S.; Denifl, S.; Grill, V.; Märk, T. D.; Illenberger, E.; Scheier, P. Bond- and Site-Selective Loss of H^- from Pyrimidine Bases. *Phys. Rev. Lett.* **2005**, *95*, 093201−093204.
(13) Denifl, S.; Ptasinska, S.; Hanel, G.; Gstir, B.; Probst, M.; Scheier, P.; Märk, T. D. Electron Attachment to Gas-Phase Uracil. *J. Chem. Phys.* **2004**, *120*, 6557−6565.
(14) Burrow, P. D.; Gallup, G. A.; Scheer, A. M.; Denifl, S.; Ptasinska, S.; Märk, T. D.; Scheier, P. Vibrational Feshbach Resonances in Uracil and Thymine. *J. Chem. Phys.* **2006**, *124*, 124310−124316.
(15) Scheer, A. M.; Silvernail, C.; Belot, J. A.; Aflatooni, K.; Gallup, G. A.; Burrow, P. D. Dissociative Electron Attachment to Uracil Deuterated at the N1 and N3 Positions. *Chem. Phys. Lett.* **2005**, *411*, 46−50.
(16) Scheer, A. M.; Aflatooni, K.; Gallup, G. A.; Burrow, P. D. Bond Breaking and Temporary Anion States in Uracil and Halouracils: Implications for the DNA Bases. *Phys. Rev. Lett.* **2004**, *92*, 068102−068105.
(17) Ptasinska, S.; Denifl, S.; Scheier, P.; Illenberger, E.; Märk, T. D. Bond- and Site-Selective Loss of H Atoms from Nucleobases by Very-Low-Energy Electrons (< 3 eV). *Angew. Chem., Int. Ed.* **2005**, *44*, 6941−6943.
(18) Desfrançois, C.; Periquet, V.; Bouteiller, Y.; Schermann, J. P. Valence and Dipole Binding of Electrons to Uracil. *J. Phys. Chem. A* **1998**, *102*, 1274−1278.
(19) Fermi, E.; Teller, E. The Capture of Negative Mesotrons in Matter. *Phys. Rev.* **1947**, 399−408.
(20) Crawford, O. H.; Garrett, W. R. Electron Affinities of Polar Molecules. *J. Chem. Phys.* **1977**, *66*, 4968−4970.
(21) Desfrançois, C.; Abdoul-Carime, H.; Khelifa, N.; Schermann, J. P. From $1/r$ to $1/r^2$ Potentials: Electron Exchange between Rydberg Atoms and Polar Molecules. *Phys. Rev. Lett.* **1994**, *73*, 2436−2439.
(22) Roca-Sanjuán, D.; Merchán, M.; Serrano-Andrés, L.; Rubio, M. *Ab initio* Determination of the Electron Affinities of DNA and RNA Nucleobases. *J. Chem. Phys.* **2008**, *129*, 095104−095114.
(23) Gallup, G. A.; Fabrikant, I. I. Vibrational Feshbach Resonances in Dissociative Electron Attachment to Uracil. *Phys. Rev. A* **2011**, *83*, 012706−012712.
(24) Sanche, L. Nanoscopic Aspects of Radiobiological Damage: Fragmentation Induced by Secondary Low-Energy Electrons. *Mass Spectrom. Rev.* **2002**, *21*, 349−369.
(25) Sanche, L. Interactions of Low-Energy Electrons with Atomic and Molecular Solids. *Scanning Microsc.* **1995**, *9*, 619−656.
(26) Xi, L.; Cai, Z.; Sevilla, M. D. DFT Calculations of the Electron Affinities of Nucleic Acid Bases: Dealing with Negative Electron Affinities. *J. Phys. Chem. A* **2002**, *106*, 1596−1603.
(27) Denifl, S.; Zappa, F.; Mauracher, A.; Ferreira da Silva, F.; Bacher, A.; Echt, O.; Märk, T. D.; Bohme, D. K.; Scheier, P. Dissociative Electron Attachment to DNA Bases Near Absolute Zero Temperature: Freezing Dissociation Intermediates. *ChemPhysChem* **2008**, *9*, 1387−1389.
(28) Pan, X.; Cloutier, P.; Hunting, D.; Sanche, L. Dissociative Electron Attachment to DNA. *Phys. Rev. Lett.* **2003**, *90*, 208102−208105.
(29) Li, X.; Sanche, L.; Sevilla, M. D. Low Energy Electro Interactions with Uracil: The Energetics Predicted by Theory. *J. Phys. Chem. B* **2004**, *108*, 5472−5476.
(30) Simons, J.; Jordan, K. D. Ab Initio Electronic Structure of Anions. *Chem. Rev.* **1987**, *87*, 535−555.
(31) Rubio, M.; Merchán, M.; Ortí, E.; Roos, B. O. Theoretical Study of the Electronic Spectra of the Biphenyl Cation and Anion. *J. Phys. Chem.* **1995**, *99*, 14980−14987.
(32) Pou-Amérigo, R.; Serrano-Andrés, L.; Merchán, M.; Ortí, E.; Forsberg, N. A Theoretical Determination of the Low-Lying Electronic States of the *p*-Benzosemiquinone Radical Anion. *J. Am. Chem. Soc.* **2000**, *122*, 6067−6077.
(33) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. Second-Order Perturbation Theory with a Complete Active Space Self-Consistent Field Reference Function. *J. Chem. Phys.* **1992**, *96*, 1218−1226.
(34) Serrano-Andrés, L.; Merchán, M.; Nebot-Gil, I.; Lindh, R.; Roos, B. O. Towards an Accurate Molecular Orbital Theory for Excited States: Ethene, Butadiene, and Hexatriene. *J. Chem. Phys.* **1993**, *98*, 3151−3162.
(35) Roos, B. O.; Andersson, K.; Fülscher, M. P.; Malmqvist, P.-Å.; Serrano-Andrés, L.; Pierloot, K.; Merchán, M. Multiconfigurational Perturbation Theory: Applications in Electronic Spectroscopy. In *Advances in Chemical Physics, New Methods in Computational Quantum Mechanics*, 1st ed.; Prigogine, I., Rice, S. A., Eds.; John Wiley & Sons: New York, 1996; Vol. 93, pp 219.
(36) Serrano-Andrés, L.; Merchán, M.; Borin, A. C. Adenine and 2-Aminopurine: Paradigms of Modern Theoretical Photochemistry. *Proc. Natl. Acad. Sci.* **2006**, *103*, 8691−8696.
(37) De Vico, L.; Olivucci, M.; Lindh, R. New General Tools for Constrained Geometry Optimizations. *J. Chem. Theory Comput.* **2005**, *1*, 1029−1037.
(38) Anglada, J. M.; Bofill, J. M. A Reduced-Restricted-Quasi-Newton−Raphson Method for Locating and Optimizing Energy Crossing Points between Two Potential Energy Surfaces. *J. Comput. Chem.* **1997**, *18*, 992−1003.
(39) Forsberg, N.; Malmqvist, P. Å. Multiconfiguration Perturbation Theory with Imaginary Level Shift. *Chem. Phys. Lett.* **1997**, *274*, 196−204.
(40) Karlström, G.; Lindh, R.; Malmqvist, P.-Å.; Roos, B. O.; Ryde, U.; Veryazov, V.; Widmark, P.-O.; Cossi, M.; Schimmelpfennig, B.; Neogrady, P.; Seijo, L. MOLCAS: A Program Package for Computational Chemistry. *Comput. Mater. Sci.* **2003**, *28*, 222−239.
(41) Veryazov, V.; Widmark, P.-O.; Serrano-Andrés, L.; Lindh, R.; Roos, B. O. 2MOLCAS as a Development Platform for Quantum Chemistry Software. *Int. J. Quantum Chem.* **2004**, *100*, 626−635.
(42) Periquet, V.; Moreau, A.; Carles, S.; Schermann, J.; Desfrançois, C. Cluster Size Effects upon Anion Solvation of N-Heterocyclic Molecules and Nucleic Acid Bases. *J. Electron Spectrosc. Relat. Phenom.* **2000**, *106*, 141−151.
(43) Roca-Sanjuán, D.; Merchán, M.; Serrano-Andrés, L. Modeling Hole Transfer in DNA: Low-Lying Excited States of Oxidized Cytosine Homodimer and Cytosine-Adenine Heterodimer. *Chem. Phys.* **2008**, *349*, 188−196.
(44) Serrano-Andrés, L.; Merchán, M.; Roca-Sanjuán, D.; Olaso-González, G.; Rubio, M. Bioexcimers as Precursors of Charge Transfer and Reactivity in Photobiology. *AIP Conf. Proc.* **2007**, *963*, 526−532.
(45) Roca-Sanjuán, D.; Olaso-González, G.; Coto, P. B.; Merchán, M.; Serrano-Andrés, L. Modeling Hole Transfer in DNA. II. Molecular

Basis of Charge Transport in the DNA Chain. *Theor. Chem. Acc.* **2010**, *126*, 177−183.
(46) Sponer, J.; Leszczynski, J.; Hobza, P. Electron Properties, Hydrogen Bonding, Stacking, and Cation Binding of DNA and RNA Nucleobases. *Biopolymers* **2001**, *61*, 3−31.
(47) Svozil, D.; Jungwirth, P.; Havlas, Z. Electron Binding to Nucleic Acid Bases. Experimental and Theoretical Studies. A Review. *Collect. Czech. Chem. Commun.* **2004**, *69*, 1395−1428.

Supporting Information

(Total of 6 pages)

for

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*[‡,*]

[†]Instituto de Ciencia Molecular, Universitat de València, P.B. Box 22085, 46071, València, Spain

[‡]Department of Chemistry – Ångström, Theoretical Chemistry Program, Uppsala University, Box 518, 75120, Uppsala, Sweden

Mercedes.Rubio@uv.es

Daniel.Roca@kvac.uu.se

Table of Contents

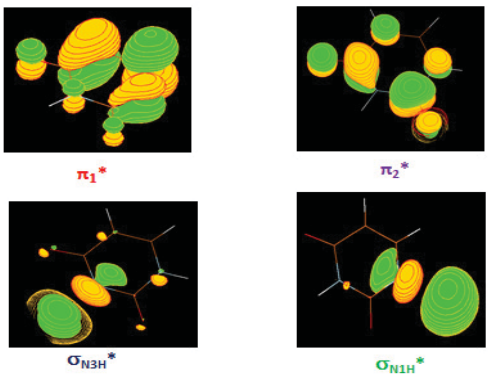


Figure S1. Set of the natural orbitals (NOs) relevant to the DEA phenomena in uracil at the N_1–H and N_3–H sites within the CASSCF (11-in-12) active space. The π_1*, π_2*, σ_{N1H}*, and σ_{N3H}* valence orbitals are depicted.

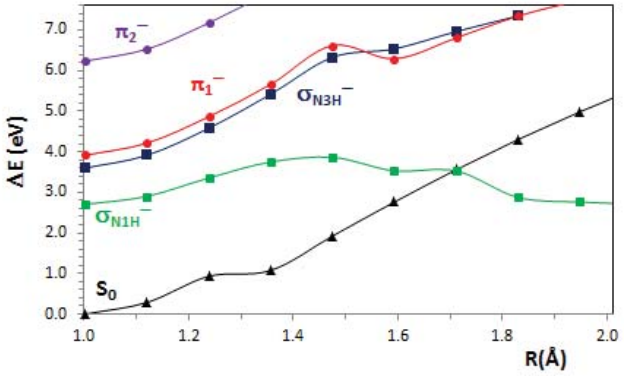


Figure S2 CASSCF potential energy curves (PECs) for the ground state of the uracil neutral molecule (S_0) and the π_1^-, π_3^-, σ_{N1H}^-, and σ_{N3H}^- states of the anionic species along the N_1–H reactive coordinate.

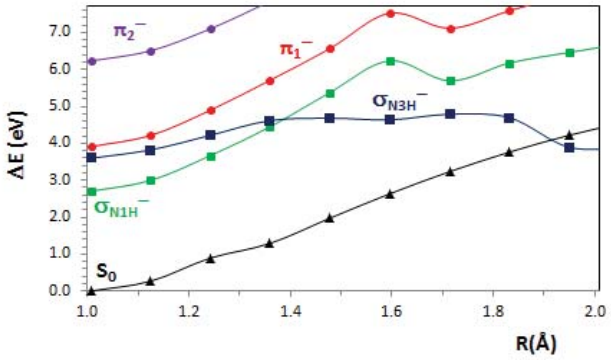


Figure S3 CASSCF potential energy curves (PECs) for the ground state of the uracil neutral molecule (S_0) and the π_1^-, π_3^-, σ_{N1H}^-, and σ_{N3H}^- states of the anionic species along the N_3–H reactive coordinate.

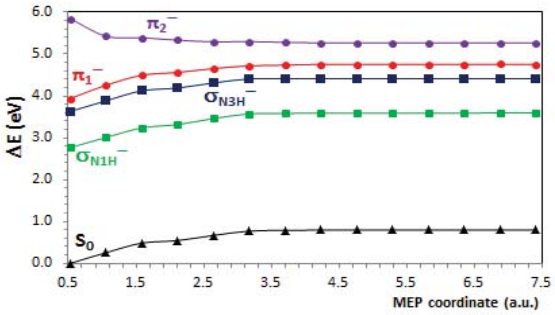


Figure S4 CASSCF energies for the ground state of the uracil neutral molecule (S_0) and the π_1^-, π_2^-, σ_{N1H}^-, and σ_{N3H}^- states of the anionic species along the minimum energy path (MEP) of the π_2^- state. Radius of hypersphere was set to 0.5 a.u.

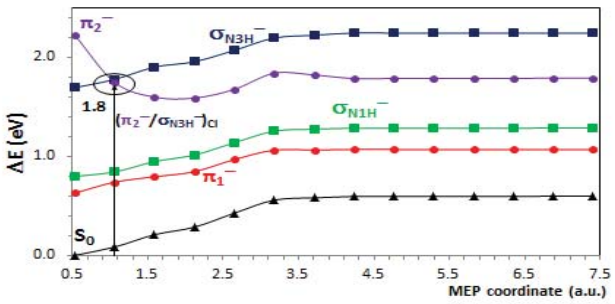


Figure S5 CASPT2 energies for the ground state of the uracil neutral molecule (S_0) and the π_1^-, π_2^-, σ_{N1H}^-, and σ_{N3H}^- states of the anionic species along the minimum energy path (MEP) of the π_2^- state. Radius of hypersphere was set to 0.5 a.u. A conical intersection among the π_2^- and σ_{N3H}^-, namely $(\pi_2^-/\sigma_{N3H}^-)_{CI}$ is observed at 1.8 eV.

Table S1. Cartesian coordinates (in Å) at the CASSCF(10,8)/ANO–L 431/21 for the ground state of the uracil neutral molecule.

C	-0.39682045	-1.20493629	-0.00000497
C	-0.35386177	1.26923815	0.00004254
C	1.10799995	1.19512603	-0.00009951
C	1.70191513	-0.00551814	0.00002770
N	0.97376935	-1.17697479	0.00004109
N	-0.98877837	0.03525842	-0.00009426
O	-1.02564421	-2.22249812	-0.00003169
O	-0.98992064	2.28596585	0.00002596
H	1.42758524	-2.05989388	0.00015568
H	-1.98555392	0.03815955	0.00057708
H	1.66493209	2.10880192	-0.00015802
H	2.76722126	-0.12625343	0.00004876

Table S2. Cartesian coordinates (in Å) at the CASSCF(10,8)/ANO–L 431/21 for the $(U–H_1)^-$ ion formed after the dissociation process.

C	-0.45888589	-1.13018281	0.00074247
C	-0.23746672	1.34658188	-0.00219044
C	1.17670324	1.12130956	-0.00353828
C	1.61552975	-0.17941555	-0.00037605
N	0.88341505	-1.29079384	0.00327000
N	-0.96084506	0.17386386	-0.00331186
O	-1.26859513	-2.04082300	-0.00034172
O	-0.81625801	2.42224544	0.00451652
H	-1.95206489	0.25689730	-0.00628145
H	1.84950483	1.95601920	0.00066840
H	2.68399256	-0.34686688	-0.00195567

Table S3. Cartesian coordinates (in Å) at the CASSCF(10,8)/ANO–L 431/21 for the $(U–H_3)^-$ ion formed after the dissociation process.

C	-0.40488116	-1.13764736	-0.00006119
C	-0.33780809	1.20667522	-0.00004229
C	1.14511998	1.16630249	-0.00028952
C	1.75470543	-0.02752545	-0.00007171
N	1.00250207	-1.16973255	0.00043223
N	-1.02245026	0.04397918	0.00009209
O	-0.95562823	-2.22949453	-0.00021857
O	-0.88446825	2.30418959	0.00013087
H	1.41856780	-2.06858183	0.00047030
H	1.69571576	2.08723309	-0.00041328
H	2.82311315	-0.15271319	-0.00031135

7 Discussion

"It is scientific only to say what it is more likely and what less likely, and not to be proving all the time the possible and impossible."
Richard Feynman (1918-1988)
Theoretical Physicist

Understanding the mechanisms of DNA/RNA damage induced by UV light and under the conditions of reductive stress was the main objective of this Thesis. The CASSCF/CASPT2 method and accurate reaction-path computational strategies have been employed in nucleobase monomers and dimers for such purposes. Regarding the DNA/RNA damage induced by UV light, we have studied the photoproduction of CPDs in π-stacked nucleobases, which can take place via the singlet (Papers I, II, and VI) or triplet (Paper V) manifolds. These photodimerization mechanisms are initiated by the absorption of light by a single pyrimidine, which occurs initially into a singlet excited state. After this localized excitation event, evolution towards dimer production can take place in those arrangements between two adjacent pyrimidines that have an effective π-stacking interaction. Hence, key structures in the mechanism of CPD formation are the bioexcimers, which are studied in Paper III for cytosine, thymine, and uracil. On the other hand, triplet population in the isolated nucleobases is also possible along the main decay channel of the bright state, as it is described in Paper IV for the five canonical bases. Then, excimer formation between the triplet state and a π-stacked nucleobase in its ground state may also give rise to the CPDs. Finally, regarding DNA/RNA damage by ionization processes, the DEA mechanism of uracil caused by low energy electrons is studied in Paper VII. Next, we discuss in more detail the results of the Papers compiled in the previous chapter.

7.1 Mechanism of photo-induced production of cyclobutane cytosine dimers via the singlet and triplet excited electronic states

In this study, our main objective was to better understand the photochemistry of cytosine oligomers in a π-stacked orientation once they are irradiated with UV photons. The results gathered allowed us to comprehend the role played by excimers (or excited dimers) of cytosine in the photoproduction of one of the most frequent lesions in DNA/RNA induced by UV light, the CPDs. The excimers were characterized by placing two cytosine monomers in a face-to-face arrangement to allow for a maximum overlap of their π electronic clouds and by calculating the energy profile of the lowest-lying singlet and triplet excited states along the separation distance of both monomers at their optimal ground-state geometry. The obtained relative minima were denoted by ^{1}(LE) and ^{3}(LE) (see Figure 1 in section 6.1).

The singlet manifold was further studied and found to be of great importance in a two-fold fashion, to offer a rationale on the pronounced wavelength dependence of the observed fluorescence on the concentration of cytosine in solutions as well as to highlight its role on the CPD production.[31,103] First, starting from the ^{1}(LE) excimer, the barrierless pathway along the lowest singlet PEH yielded the relaxed excimer $^1(CC)_{exc}$ (see Figure 6 in section 6.1). This relaxed structure exhibited a binding energy (CP-E_b) of 1.10 eV and was found to be stabilized by about ~0.5 eV with respect to the ^{1}(LE) structure. The calculated fluorescence of 3.40 eV (365 nm) for ^{1}(LE) and 2.76 eV (449 nm) for $^1(CC)_{exc}$ appears at lower energies than the one computed for the monomer, 3.96 eV (313 nm). This finding clearly renders the importance of excimers to explain the red-shifted fluorescence observed experimentally for polymers.[31,103]

Moving along the PEH of the lowest singlet state of the dimer in the direction in which the two cytosine molecules becomes closer, we were able to determine the path towards the CPD photoproduction (see Figure 6 in section 6.1). Our calculations revealed the presence of a *shearing-type* conical intersection $(S_1/S_0)_{CI}$ involving the lowest singlet (S_1) and the ground state (S_0), which might transfer, in an ultrafast manner, the population from the excited state to the ground state. This CI

point is connected on the S_0 manifold with the equilibrium structure of the photodimer. To reach the CI point and subsequently the CPD structure, the system placed at the $^1(CC)_{exc}$ excited minimum has to overcome a barrier of 0.2 eV. Nevertheless, the existence of a barrier does not imply a complete inability for the overall process of formation of CBC dimers to take place. Instead, analysis of both the theoretical and experimental data suggests that geometry orientations of the monomers close but energetically above the shearing-type CI at the time of light irradiation might be *reactive* and prone to produce CBC directly. Meanwhile, those orientations of the monomers around the region of the relaxed excimer might not be so reactive. Hence, the presence of stable excimers may decrease the effectiveness of CBC photoproduct formation along the singlet manifold. The photoreaction process is in competition with fluorescence or monomeric radiationless decay.

On the other hand, when starting from the ^{3}(LE) structure, along the triplet manifold (see Figure 3 in section 6.1), the formation of the cytosine photodimer (CBC) is possible in a barrierless fashion by means of a STC, $(T_1/S_0)_X$. These results provide the grounds for understanding the potential photogenotoxicity of endogenous and exogenous compounds via triplet-triplet sensitization and give support to the traditional view of the primary role of triplet excited states in the photochemistry of DNA.

7.2 Comparison of the photo-dimerization mechanisms of thymine and cytosine on the singlet manifold

In this work, with the fine-tuned methodological strategy employed in the previous study to determine the intrinsic photochemical and photophysical properties of the π-stacked dimer of cytosine, our next task was to provide a sound rationale, from the theoretical standpoint, to the question of why there is a higher yield for the CBX photoproduction in thymine as compared to cytosine.[44] Therefore, we extended our calculations to the processes taking place in thymine in order to grasp the main differences that give thymine the relatively higher quantum yield of CBT photoproduction. To better mimic the *in vivo* geometrical orientations, the excited states of both thymine and cytosine on this study were calculated at the geometrical arrangements of the B-DNA, namely, the $^1(TT)_B$ and $^1(CC)_B$, respectively. For thymine, we encountered the

lowest excited singlet of the $^1(TT)_B$ to be slightly red-shifted by 4.60 eV (~270 nm) with respect to the value obtained for the monomer, 4.89 eV (~254 nm). Moreover, we observed that the path from the $^1(TT)_B$ arrangement to the shearing type $(S_1/S_0)_{CI}$ structure can be accessed in a barrierless manner (see Figure 1 in section 6.2). Such accessibility into the CI in a barrierless manner is also available for cytosine when departing from the $^1(CC)_B$ structure. As seen in section 7.1, this is not the case for the evolution from the optimized dimer in the singlet excited state, $^1(CC)_{exc}$. Therefore, the geometrical arrangement of the B-DNA for the tandem base pairs provides enough flexibility to make them prone to create photodimer lesions within the strand. The route followed by the π-stacking of the bases in the face-to-face sandwich type was explored as well for thymine, as we did in cytosine, since this is the most favorable orientation for excimer formation. At the $^1(TT)_{exc}$ structure, we calculated the binding energy to be 1.25 eV. When compared to the value of 0.29 eV obtained for the $^1(TT)_B$ arrangement, we can conclude that the relaxed excimer, $^1(TT)_{exc}$, provides the optimum overlap among the monomers. Interestingly, the $^1(TT)_{exc}$ in thymine can connect in a barrierless manner with the $(S_1/S_0)_{CI}$ that lead to the formation of the CBT photolesion as well. The availability of a barrierless photoproduct formation route from the thymine excimer is in clear contrast with the results obtained for cytosine where a stable structure, $^1(CC)_{exc}$, was formed. This relaxed excimer in cytosine provides a certain degree of stability to the system since a small barrier of about 0.2 eV has to be surmounted to reach the $(S_1/S_0)_{CI}$ crossing. Therefore, the lack of a stable excimer in thymine suggests that in the TT dimer any possible arrangement attained may be susceptible to photodimerization. Moreover, the instability of the thymine excimer leaves the system without an efficient mechanism to fluoresce. Hence, the experimentally observed fluorescence for the consecutive thymine 18-mers[104] can only be attributed to individual fluorescence of the monomers. Meanwhile, in CC, the funnel toward CBC production becomes competitive with fluorescence and with the IC of the excited cytosine monomer, therefore providing a lower number of photolesions than in thymine. Overall, in this study we were able to establish that the relative stability of the formed excimers with respect to the placement of the $(S_1/S_0)_{CI}$ is the main factor responsible at the molecular level for the different efficiency observed experimentally in the production of CBT *vs.* CBC.

7.3 Excimer formation in excited pyrimidine dimers

Once the importance of the role played by the π-stacking interaction in the excimer formation of both cytosine and thymine was well established, our next goal was to prove that the excimer formation, by means of the π-stacking mechanism, is an intrinsic property of all three pyrimidine nucleobases, also including uracil. For this purpose, potential energy curves (PEC) of the lowest-lying singlet excited states were computed along the intermolecular separation of two uracil monomers in a face-to-face arrangement (see Figures 1 and 2 in section 6.3). The $R(C_5–C_{5'})$ distance was used to monitor the process since it was shown in our previous studies for cytosine and thymine (see sections 7.1 and 7.2), that it is especially relevant for the formation of CBC and CBT photolesions. Inspection of the face-to-face uracil PECs portrays a similar behavior as that observed for the previously studied pyrimidines. That is, after excitation, the low-lying excited singlet state in the uracil monomers becomes bound by 0.48 eV (see Table 1 in section 6.3). The relative minimum for the uracil excimer was located within the S_1 state at an intermolecular distance $R(C_5–C_{5'})$ of 2.910 Å. These results are in agreement with those of the excimer states obtained for both the cytosine and the thymine dimers. Thus, a binding energy of 0.58 eV was obtained for the cytosine pair at an intermolecular distance (R_{min}) of 3.076 Å, and a binding energy of 0.36 eV was calculated for the thymine pair at an R_{min} of 3.703 Å. The fact that thymine shows a lower stabilization than its counterparts and reaches a minimum at higher intermonomer distances can be accounted for by considering the fact that thymine possesses a higher steric impediment due to the presence of the methyl group. In general, the homodimers of all three pyrimidine bases give rise to the formation of excimers.

Overall, the evidence of the excimer origin of the emissive properties of all the pyrimidine homodimers highlights the importance of π-stacking and excimer formation in the modulation of the relaxation dynamics in DNA/RNA, which will combine emissive decaying paths, with nonradiative decays through accessible CIs connecting the low-lying singlet excited state with the ground state.

7.4 Triplet population mechanism in isolated DNA/RNA nucleobases

In this study, our objective was to determine the intrinsic mechanisms responsible for the population of the triplet manifold, through a STC crossing, in DNA/RNA nucleobases. This was mainly motivated by two factors. Firstly, the studied cytosine dimers showed an accessible channel for photo-production of CPDs on the triplet manifold (see section 7.1). However, the mechanism for populating the triplet state after light irradiation was not well-known. Secondly, we wanted to answer the question of why the experimental DNA phosphorescence spectra at low temperatures reported in the literature only consists of a main signal from thymine and to a lesser extent from adenine, without any contribution from the other bases.[123,124,125] In order to do so, the decay pathways on the bright excited states after near-UV irradiation of the five natural DNA/RNA bases, thymine, uracil, cytosine, adenine, and guanine were determined with the multiconfigurational CASPT2//CASSCF protocol and MEP calculations. The lowest-lying triplet excited states and the SOCs were also computed along the decay paths of the bright state of singlet nature. It is worth recalling here that the energy degeneracy among the singlet and triplet states along with a high SOC value are necessary (although not sufficient) conditions that indicate the possibility of an efficient ISC process. Under these criteria, it was observed that there were singlet-triplet interacting regions accessible along the MEP of the initially populated singlet bright $^1\pi\pi^*$ state. Interestingly, by following the decay route for this $^1\pi\pi^*$ singlet state, we found that all five natural DNA/RNA nucleobases have, at the end, a low energy, easily accessible, STC region directly connecting the lowest singlet and triplet $\pi\pi^*$ excited states (see Figures 3, 5, 7, and 9 in section 6.4). This STC region is encountered, for all five bases, prior and very close to the CI between the $^1\pi\pi^*$ and ground states. In addition, adenine, thymine, and uracil display two other crossing regions at higher energies related to the presence of low-lying singlet and triplet nπ^* states (see Figures 3 and 7 in section 6.4). Our results indicate that three STC regions can be easily accessed from the singlet main decay pathway in adenine, thymine, and uracil. These regions, where the electronic population gets trapped long enough to allow the ISC process to take place may become competitive with their IC decay mechanism, making them prone to yield a phosphorescence signal when irradiated. These two additional high-

energy competitive funnels found in adenine, thymine, and uracil are absent in guanine and cytosine (see Figures 5 and 9 in section 6.4). Therefore, guanine and cytosine are not expected to contribute to a phosphorescence signal since their bright $^1\pi\pi^*$ state appears lower in energy and less accessible toward their respective singlet and triplet $n\pi^*$ states.

Thus, the results for this study allow interpreting the fact that guanine and cytosine contribute much less to the phosphorescence of DNA than adenine and thymine. Furthermore, the findings indicate that adenine and thymine should be preferably considered as sources of phosphorescence in DNA, whereas in RNA, adenine and uracil might be the bases considered responsible for the phosphorescence. These properties should be regarded as intrinsic features of the nucleobases. Even if they may change in condensed phases or in the biological environment, for the single monomers, they are expected to maintain their main characteristics.

7.5 Cyclobutane pyrimidine photodimerization of DNA/RNA nucleobases in the triplet state

The computations and theoretical analysis performed in the study discussed in section 7.4 allowed us to have a better understanding about the underlying mechanisms responsible for the triplet excited state population in the individual DNA/RNA nucleobases through STC points. Meanwhile, from the work of section 7.1, we knew that a STC mediates the [2+2] photocycloaddition in cytosine dimers. Nevertheless, the general role of the triplet state of DNA oligomers in the dimerization process after direct irradiation of pyrimidines was still to be clarified. For instance, while some studies have shown evidence of the ultrafast participation of the triplet state in the formation of the CBT,[126] time-resolved studies of CBT formation after irradiation of $(dT)_{20}$ oligonucleotides were unsuccessful at detecting the presence of the triplet state within a 200 ns time frame.[127] Consequently, we decided to study, from a theoretical standpoint, the involvement of the triplet state within the oligomeric frame in all pyrimidines. From the literature, we knew that CPDs are notoriously formed upon nucleobase photosensitization by triplet-triplet energy transfer (TTET) from favorable donors, such as

different ketone or phthalimidine derivatives.[34,44,128,129,130,131] The role of the nucleobases triplet states therefore seems to be linked to the ability of stacked DNA/RNA nucleobases to behave as acceptors under favorable triplet photosensitization conditions. As the efficiency of TTET processes is strongly related to the relative triplet excited state energies of the donor and acceptor chromophores, we understood that the characterization of the nucleobases triplet intermediate responsible for the photosensitization action would provide relevant insight into the photoreactive process. Hence, for this study we focused our attention on modeling the CPD formation for cytosine, thymine, and uracil along the triplet manifold.

The three pyrimidine nucleobases have close-by experimental triplet state signals at 3.5–3.6 eV.[132,133,134] When the homo-bases are coming closer to each other, the effective overlap between the electron clouds of the two bases give rise to a stabilization of the triplet excited state similarly to what happens for the singlet excited states. In the triplet manifold, this excimer-type interaction leads to the formation of a biradical stepwise intermediate ^{3}(SWI), in which a covalent bond is actually formed between the C_6 and $C_{6'}$ atoms of the adjacent nucleobases (see Figure 2 in section 6.5). MEP calculations proved that such structures are reached in a barrierless manner from two face-to-face nucleobases. The ^{3}(SWI) intermediate has biradical character with the unpaired electrons and the spin density on the two other ethylenic C_5 and $C_{5'}$ atoms of the bases. The intermediate is actually a minimum on the T_1 PEH, and also corresponds to a STC crossing $(T_1/S_0)_X$ structure that is expected to mediate the formation of the CPD photoproduct. The ^{3}(SWI) intermediate was found to have the lowest energy in the thymine homodimers (2.36 eV), followed by uracil (2.47 eV) and cytosine (2.70 eV) dimers. The energy for thymine can be established as the minimum energy required to populate the triplet state in the DNA/RNA environment. Therefore, our calculations predict that the production of thymine (CBT) and uracil (CBU) photoproducts in the triplet state take place at similar energies. In contrast, higher energies are required to give rise to the CBC photodimer. Experimental studies on the TTET formation of CBT dimers restrict the threshold observed for a given photosensitizer to become a potential DNA photodamager to substances with triplet state energies higher than 2.6–2.8 eV.[131] Therefore, our results are in relatively good agreement with the experimental values. In summary, this study allows a sound rationalization of the CPD formation in the triplet manifold, provided that

an efficient energy transfer mechanism that permits the triplet to be populated takes place.

7.6 Comparison of the photoproduction mechanisms of DNA/RNA cyclobutane pyrimidine dimers in cytosine, thymine, uracil, and 5-methylcytosine via the singlet excited state

In the studies discussed in sections 7.1 and 7.2, the mechanisms for the non-adiabatic photoreaction in thymine and cytosine dimers were found to be mediated by the presence of a CI along the singlet manifold. These CIs represent energy degeneracies between the low-lying singlet excited (S_1) and the ground state (S_0) that can be accessed through barrierless relaxation paths from favorable conformations of the nucleobases. In those quantum-chemical CASPT2 studies, the [2+2] photocycloaddition reaction leading to the formation of CBC and CBT dimers along the singlet manifold was determined to take place with the nucleobases' ethylenic C_5–C_6 and $C_{5'}$–$C_{6'}$ bonds laying parallel and elongated. Additionally, it was established that the presence of a methyl group in thymine, versus cytosine, decreases the excimer stabilization, thus explaining the higher yield of photoproducts observed in the former (section 7.3).

On the grounds of these properties and with the goal of fully understanding the [2+2] photocycloadditions in DNA/RNA pyrimidines, we extended the study to uracil, which showed similar excimer properties (see section 7.3). Moreover, we also considered the non-canonical methylated cytosine (5-methyl-cytosine, m^5C), which is present in human DNA in an amount of around 5%.[135] Experimentally, all these nucleobases are known to produce CPDs when irradiated with UV-C or UV-B light.[136] In this study, the analyses were mainly focused on comparing, in cytosine, thymine, uracil, and 5-methyl-cytosine, two possible non-radiative decays that might occur on the singlet manifold of the dimers, the non-adiabatic process localized in each monomer and the one delocalized over the dimer. While the former drives the system back to the initial ground state equilibrium structure of the separated monomers, the latter produces the CPD lesions via the [2+2] photocycloaddition. The key structures that characterize these two

mechanisms are the ethylene-like CI of the isolated nucleobases (CI_{mon}) for the localized decay and the CI of the dimer (CI_{dim}) for the delocalized non-adiabatic process (see Figure 2 in section 6.6). The excimers, Exc (S_1) are also important points for the comparative analysis of the decay channels. In order to establish a comparison between the monomer- and dimer-based mechanisms, the adiabatic energies of the CI_{mon} $(S_1/S_0)_{CI}$ points were computed at the CASPT2 level for the monomers. The obtained values were 3.60, 3.64, 3.90, and 4.00 eV for cytosine, 5-methyl-cytosine, uracil, and thymine, respectively. The energy values for the excimer structures, Exc (S_1), were found to be 3.31, 3.46, 3.68, and 3.64 eV for $^1(CC)_{exc}$, $^1(m^5C\ m^5C)_{exc}$, $^1(UU)_{exc}$, and $^1(TT)_{exc}$, respectively. For the CI_{dim}, the energies of 3.51, 3.56, 3.47, and 3.26 eV were determined for the CC, m^5Cm^5C, UU, and TT dimers, respectively.

On the basis of results obtained in this study, the following scenario arises. After excitation of the two molecules of the dimer in a B-DNA orientation, the most probable decay is the relaxation of each monomer via a barrierless path along the PEH of the lowest singlet excited state towards the CI_{mon} (see Figure 2 in section 6.6). At this point, the energy is funneled to the ground state of the monomers in a radiationless manner. Even though this mechanism is the main route for energy decay, the nucleobases may adopt different arrangements within the strand that make them prone towards excimer formation by enhancing the π-π electronic overlap. In the face-to-face arrangements, excimer formation is highly plausible. Concurrently, along the excimer formation path, it is observed that the four homodimers also have a CI degeneracy region, CI_{dim}, which leads the pair of stacked nucleobases to the CPD formation. For CC and m^5Cm^5C, the CI_{dim} is found to be higher in energy than the excimer structure Exc(S_1) by 0.2 and 0.1 eV, respectively. Consequently, these conformations become somewhat stable with respect to the [2+2] non-adiabatic reaction, diminishing both the rate and the yield of the photoproducts formation. The excimer stability is higher for CC than for m^5Cm^5C which can be attributed to the lack of the methyl group in cytosine at the C5 position. On the other hand, the paths from the Exc (S_1) point toward the CI_{dim} can be accessed in a barrierless manner for UU and TT. Thus, the lower CPD production observed experimentally[44] in CC as compared to TT and UU can be interpreted by the fact that C-based dimers possess a higher number of competitive routes to deactivate the system towards the ground state of the monomer

while thymine/uracil dimers can access the CI_{dim} with fewer competitive routes and in a barrierless manner.

7.7 Theoretical study of the dissociative electron-attachment process in uracil caused by low energy electrons

In this last work, our goal was to study the effect of low energy electrons in uracil that cause bond dissociation at the N_1–H and N_3–H sites. These low energy electrons may be generated inside the cell after an initial ionization step and then cause secondary reactions leading to the formation of new ionic species. They may ultimately produce strand breaks (SSBs or DSBs), as described in chapter 1. In the case of uracil, this molecule possesses a high dipole moment of about 4.5 D and therefore it is able to trap low energy electrons (< 3 eV) which produce a dissociative electron-attachment (DEA) phenomena at the N_1–H and N_3–H bonds. This high polarity of the base can make it difficult to study the anion formation since it can give rise to the existence of two different types of anions that can share the same energy region, namely, the dipole bound (DB) and the valence bound (VB) anions. In the DB, the electron is electrostatically trapped by the nucleobase dipole moment in a Rydberg-like state, whereas in the VB the extra electron occupies one of the empty π^* orbitals of the molecule. In uracil, the DB is the only anion observed experimentally in the gas phase since the nucleobase cannot stabilize an electron in its valence shell. However, the VB anion is more relevant in condensed phases (i.e. biological medium) where DB anions are likely to be absent. Therefore, our aim for this study was to characterize, for the first time, the lowest lying VB states of the uracil molecule by means of the CASSCF/CASPT2 method in order to explore the involvement of the valence states in the hydrogen dissociation at the N_1–H and N_3–H sites.

During the electron attachment process in uracil, the system is initially ionized producing a temporary anionic state. The DEA takes place via the population transfer from the initial π^- to a dissociative σ^- state that drives the anion to the bond breaking. The dissociation energies at the CASPT2 level (taking into account the ZPVE) for the $U + e^- \rightarrow (U–H)^- + H$ reaction were calculated to be 0.61 and 1.21 eV for the N_1–H and N_3–H bonds, respectively. These values, along with those available in

the literature,[137,138,139,140] helped interpreting the DEA cross sections recorded experimentally.[137,138,139,141,142,143,144] Thus, the obtained values indicate that N_1–H cleavage may take place with an energy of the incident electron of at least 0.6 eV, whereas energies higher than 1.2 eV are required to give rise to an N_3–H dissociation. These numbers are in good agreement with the experimental observations. Additionally, the exploration of the PECs of the neutral and lowest-lying anionic states along the N_1–H and N_3–H bond distances shows the presence of CIs among the π_1^- and the σ^- states at energies 0.9 eV and 2.1 eV for N_1–H and N_3–H, respectively (see Figures 2 and 3 in section 6.7). These CIs were interpreted as the points where the DEA process is activated and, therefore, they can be compared with the band maxima of the DEA spectra.

Besides the lowest-lying π_1^- state, other higher-energy anionic states might also contribute to the DEA signals. Burrow and co-workers[142,143,144] suggested, by using electron transmission spectroscopy (ETS), that in the N_3–H dissociation, the π_2^- might be the one responsible for the broad band in the spectrum, located at 1.7 eV. By performing MEP computations following the π_2^- state, we obtained a $(\pi_2^-/\sigma_{N3H}^-)_{CI}$ crossing point at 1.8 eV confirming for the first time, using *ab initio* methods, the involvement of the π_2^- state in the N_3–H dissociation (see Figure 4 in section 6.7). At the end of the MEP, the equilibrium structure of the π_2^- state was located at an adiabatic energy of 1.59 eV. This value indicates the minimum energy that is needed for any participation of the π_2^- state in the DEA processes, i.e., no possible contributions to the DEA process can take place from this state at lower energies.

Overall, in this study we were able to establish that not only DB anions are responsible for the DEA process. The VB anions are also relevant and greatly involved in this phenomenon. The findings shed light on the DEA mechanism and allowed for a better assignment of the bands observed experimentally in the DEA cross section.

8 Conclusions

"Science, my boy, is made up of mistakes,
but they are mistakes which it is useful to make,
because they lead little by little to the truth"
Jules Verne (1828-1905)
Author

The quantum-chemical studies carried out in the present Doctoral Thesis allow determining the molecular basis of DNA/RNA photo-reactive phenomena. The findings are of interest in the field of DNA/RNA damage induced by UV light and under conditions of reductive stress. In the following lines, we list the main conclusions obtained in the works performed in the Thesis:

- The mechanism of CPD photo-production in cytosine dimers involves two key structures in the singlet manifold. First, a shearing-type CI between the lowest-lying excited state of the homodimers and the ground state connects the excited state of the dimer with the ground state of the CPD lesion. Secondly, the π-stacking interaction between the two nucleobases of the dimers gives rise to the formation of excimers which were shown to be the precursors of the CPDs. The existence of stable singlet excimers indicates the efficiency of dimer formation in cytosine.

- Three important aspects of the CPD photoproduction mechanism allow interpreting the higher experimental yield of formation of the CBT lesion as compared to the yield measured for cytosine. Firstly, thymine homodimers have more reactive orientations than those of cytosine. Thus, whereas some cytosine excimers have lower energies than the CI points that leads the system to the

lesion, in thymine, all the π-stacked arrangements are more energetic than the CI. Secondly, photoreversibility of the CPDs by direct absorption requires more energy in thymine homodimers than in those formed by cytosine. Therefore, the lesions are expected to revert less efficiently in thymine than in cytosine upon solar irradiation. Thirdly, in the dimers of cytosine, the funnel toward CBC production becomes competitive with the funnel that mediates the IC of the excited-cytosine monomer, whereas in thymine dimers, the decay of the excited monomer becomes relevant only for unstacked thymine bases.

- Homodimers of the three natural pyrimidine bases display a similar behavior with respect to their binding properties in the excited state caused by the π-stacking interaction. The formation of excimers plays a key role in the photophysical outcome and the photochemical properties of pyrimidine-containing biopolymers.

- The isolated DNA/RNA nucleobases possess accessible routes for triplet population along the main decay path of the bright electronic excited state. Adenine, thymine, and uracil present three regions of STC with significant probability of singlet to triplet population transfer. In contrast, guanine and cytosine only have one region of STC at low energies. Hence, our results indicate that adenine and thymine are the responsible sources of phosphorescence in DNA. Likewise, in RNA, adenine and uracil can be considered as the responsible nucleobases for the phosphorescence signals. On the other hand, guanine and cytosine are not predicted to contribute that much to the phosphorescence.

- The formation of the CPD lesions is also possible via the triplet excited states of the excimers. Two mechanisms are possible. First, the singlet to triplet population transfer can take place in the isolated nucleobases and then a triplet excimer can be formed via the π-stacking interaction with an adjacent nucleobase. Second, a photosensitizer can induce the formation of the triplet excimer via a triplet energy transfer process. As for the singlet photo-reactivity, the triplet excimer is a precursor of the CPDs. A key structure in this process is a biradical triplet intermediate that corresponds to a STC crossing point between the PEHs of the

ground and triplet state of the dimer. The relative energy position of these intermediate indicates that the photosensitization in cytosine homodimers requires triplet donors with higher triplet energies than in the case of uracil and thymine.

- A unified mechanism for the photoinduced CPD formation in the singlet manifold was obtained in the comparative study of the process in cytosine, thymine, uracil, and 5-methyl-cytosine. The dimerization corresponds to a nonadiabatic [2+2] photocycloaddition reaction, in which the sheared-like CI connecting the S_1 and S_0 states of the dimer is the funnel controlling the reactive process. The relative position of this CI with respect to the stable excimer structures and the CI related to the monomeric decay determines the efficiency of the photoreactivity.

- The energy threshold for the DEA process at the N_1–H and N_3–H bonds in uracil is 0.6–0.8 and 1.2–1.4 eV, respectively, which agrees with the experimental observations. The DEA process takes place via a crossing between the PECs of the initially formed π^- temporary bound states and the σ^- dissociative state. The energies of these CIs can be related to the band maxima of the cross sections recorded experimentally. The π_1^- state contributes mainly to the sharp peak at 1 eV, whereas the π_2^- state only participated in the broad band placed around 1.7 eV. In our study, we demonstrate that the VB anions, as well as the DB anions, contribute to the DEA phenomena.

9 Resumen

“Be sincere; be brief; be seated”
Franklin D. Roosevelt (1882-1945)
32nd U.S. President

En este capítulo se resumen los contenidos de la Tesis Doctoral a través de cinco secciones de introducción, métodos, estrategias computacionales, objetivos, resultados, discusión de los trabajos realizados y conclusiones.

Introducción

La investigación científica realizada a lo largo de los últimos dos siglos con el objetivo de descubrir el material que permite la transmisión de la información hereditaria presenta tres grandes hitos. El primero de ellos lo encontramos en las primeras teorías de la evolución de Charles Darwin sobre la evolución de las especies por medio de la selección natural. Tras ello, cabe destacar el descubrimiento del ADN por Friedriech Miescher. Finalmente, uno de los más importantes hallazgos fue la determinación de la estructura molecular del ADN por James Watson y Francis Crick. Este último descubrimiento permitió comprender sus propiedades de replicación y traducción y además abrió la puerta al estudio a nivel molecular del material genético en profundidad.

Una de las áreas de investigación que surgió posteriormente y que es de gran relevancia para entender la vida tal y cual la conocemos en el planeta ha sido el estudio de la interacción del ADN/ARN con la radiación solar. Al reflexionar sobre dicha interacción, surgen algunas preguntas al respecto tales como: ¿qué sucede cuando los ácidos nucleicos

son irradiados con luz, en particular, con la luz ultravioleta? ¿qué les ocurre a las bases nitrogenadas al interactuar con fuentes de radiación ionizantes o cuando se encuentran en condiciones de estrés reductor u oxidativo? El desentrañar las respuestas a éstas y otras preguntas mediante la determinación de los mecanismos de interacción luz – ADN/ARN se convirtieron en el punto de partida para esta Tesis Doctoral.

Metodología

La irradiación de los cromóforos biológicos con luz visible o UV produce estados electrónicos excitados en los mismos e induce fenómenos fotofísicos y fotoquímicos. En comparación con la reactividad en el estado fundamental, la fotofísica y fotoquímica, o en general, la química del estado excitado presenta una complejidad mayor. Así, son frecuentes las situaciones de cruces entre estados electrónicos donde existe una degeneración energética entre distintas configuraciones electrónicas. En esta Tesis, el estudio molecular de los sistemas de interés fue diseñado teniendo en mente dicha naturaleza electrónica multiconfiguracional. Para ello se seleccionaron los métodos multiconfiguracionales CASSCF y CASPT2. La combinación de estos métodos a través del protocolo CASPT2//CASSCF es probablemente una de las herramientas más útiles y prácticas para el estudio tanto espectroscópico como fotoquímico. En este protocolo las geometrías son optimizadas a nivel CASSCF, que en muchos de los casos tiene una precisión suficiente para determinar los parámetros geométricos, y luego las energías, las cuales requieren de mayor precisión, son calculadas con el método CASPT2. La ventaja de dicha combinación se debe a que puede ser aplicada a sistemas con tamaños moleculares medianos y relativamente grandes. Además, este protocolo carece de restricciones y por ello puede ser utilizado para describir todo tipo de estados y degeneraciones energéticas de forma balanceada. Finalmente, la metodología CASPT2//CASSCF cuenta con una precisión de 0.2–0.3 eV en estudios espectroscópicos, fotofísicos y fotoquímicos.

Una vez seleccionados los métodos de cálculo adecuados para los problemas de estudio, se utilizaron estrategias de cálculo capaces de determinar los caminos más probables de decaimiento energético. Entre

estas estrategias, la más precisa dentro de la modelización estática es sin duda el cálculo de caminos de mínima energía (MEP) y la localización de las zonas de intersección cónica (CI) accesibles a lo largo de dichos caminos MEP. También se recurrió a interpolaciones lineales de coordenadas internas (LIIC) con la finalidad de explorar zonas de interés de las hipersuperficies de energía potencial (PEHs).

Objetivos

El trabajo presentado en esta Tesis se enmarca dentro de la investigación de las propiedades fotofísicas y fotoquímicas de los ácidos nucléicos. El objetivo de la misma ha sido el de estudiar y obtener una mejor comprensión de los aspectos más relevantes del panorama fotofísico y fotoquímico exhibido por las bases nitrogenadas del ADN/ARN y sus dímeros, irradiados con luz UV o en presencia de electrones de baja energía. Los objetivos concretos de cada uno de los trabajos presentados en esta Tesis son los siguientes:

- Determinar los mecanismos de foto-producción de dímeros de ciclobutilpirimidina (CPDs) en las bases canónicas del ADN/ARN citosina, timina y uracilo, así como en la base 5-metil-citosina, la cual es de gran interés en epigenética. Realizar un estudio comparativo determinando las estructuras claves que caracterizan el mecanismo de fotodimerización. Interpretar los distintos rendimientos cuánticos de formación de fotodímeros determinados experimentalmente para las bases.

- Caracterizar la formación de excímeros en homodímeros de pirimidina y comprender su función como precursores del proceso fotorreactivo de formación de CPDs.

- Determinar los mecanismos de población del estado triplete en las bases pirimidínicas y púricas del ADN/ARN.

- Estudiar la formación de CPDs vía los estados electrónicos triplete.

- Estudiar el proceso disociativo por adición de electrones (DEA) en la molécula de uracilo. Asignar las secciones de cruce obtenidas experimentalmente para el proceso DEA con electrones de baja energía (0-3 eV).

Resultados y discusión

En el primero de los trabajos se determinaron principalmente los ***mecanismos de producción de dímeros de ciclobutilcitosina a través de los estados electrónicos excitados singlete y triplete***. Se estudió la formación de dímeros de ciclobutano debido a su importancia como uno de los procesos más comunes de daño en el ADN por irradiación UV. Las estructuras de los excímeros localmente excitados ^{3}(LE) y ^{1}(LE), se tomaron como punto de partida. Dichos excímeros fueron caracterizados calculando la energía de enlace a varias distancias de separación de dos monómeros de citosina con una configuración "cara a cara" y usando la geometría del estado fundamental de la citosina. Mediante la realización de cálculos *ab initio* de alto nivel, se demostró que ^{3}(LE) evoluciona a lo largo de la hipersuperficie triplete en un camino sin barreras a la formación de un intermediario, ^{3}(SWI), mientras que los cálculos a lo largo de la hipersuperficie singlete más baja dieron como resultado el excímero relajado $^1(CC)_{exc}$.

La estructura ^{3}(SWI) presenta un enlace covalente sencillo entre los átomos de carbono C_6–$C_{6'}$, con una mayor longitud en el enlace C_5–$C_{5'}$. La energía necesaria para llegar a la estructura ^{3}(SWI) partiendo de dos monómeros aislados de citosina en el estado fundamental es de 2.70 eV. Dicho valor se puede relacionar con el umbral observado experimentalmente para que un fotosensibilizador dado se convierta en un potencial agente dañino del ADN. La estructura ^{3}(SWI) constituye un cruce singlete-triplete $(T_1/S_0)_X$, que media la desactivación no radiativa hacia el estado fundamental a través de un mecanismo de cruce entre sistemas (ISC).

En cuanto a la reactividad involucrando los estados electrónicos singlete, se determinó una zona de CI entre el estado excitado más bajo y el fundamental que supone la piedra angular para entender la formación de fotoproductos a lo largo de la vía del singlete. Los cálculos

CASSCF/CASPT2 mostraron que es necesaria una energía de 0.2 eV para poder alcanzar el punto de degeneración entre PEHs desde el dímero de la citosina con la estructura $^1(CC)_{exc}$. La fluorescencia de la especie $^1(CC)_{exc}$ es, por tanto, una vía de desactivación del dímero. Sin embargo la presencia de una barrera energética no implica que el proceso fotorreactivo de formación de la ciclobutilcitosina (CBC) no pueda tener lugar. Así, el análisis de la información teórica y experimental[44] actual sugiere que en el momento de la irradiación de la luz aquellas estructuras que están cerca, pero con una energía por encima de la intersección cónica $(S_1/S_0)_{CI}$ son orientaciones reactivas que pueden dar lugar a la formación de CBC. En este sentido, la falta de excímeros estables representa un potencial mayor para alcanzar estructuras reactivas, que se traduce en una mayor incidencia en la formación de CPDs. Por el contrario, los excímeros estables disminuyen la eficacia de la formación de fotoproductos en la vía del singlete, ya que las orientaciones de los monómeros alrededor de la región del excímero relajado podrían no ser tan reactivos y se requeriría de un exceso de energía de vibración para superar las barreras correspondientes. Esto ocurre en citosina, y como resultado, la formación de CBC se vuelve globalmente menos eficaz. Ergo, la presencia de excímeros singlete estables ralentiza la eficiencia de la formación de dímeros, mientras que la ausencia de excímeros puede favorecer indirectamente la fotorreacción.

En el segundo trabajo se realizó una ***comparación de los mecanismos de foto-dimerización de timina y citosina por la vía singlete***. La formación fotoinducida de dímeros de ciclobutilpirimidina en la timina (CBT) y la citosina (CBC) se examinó en base a los resultados *ab initio* CASPT2 con la finalidad de obtener una interpretación de los rendimientos más altos observados experimentalmente para la formación de CBTs con respecto a los de la producción de CBCs.[44] Se eligieron como estructuras de partidas para la comparación las orientaciones en los dímeros de tipo "cara a cara" en la forma B del ADN, $^1(CC)_B$ y $^1(TT)_B$. En la timina, no se encontraron barreras energéticas en los caminos fotoquímicos desde las estructuras $^1(TT)_B$ o $^1(TT)_{exc}$ (excímero optimizado) hasta la intersección cónica $(S_1/S_0)_{CI}$ que conduce hacia la formación del fotoproducto. En cambio, en la citosina sólo se encontraron perfiles sin barrera para la ruta desde la estructura $^1(CC)_B$ hacia el punto de cruce $(S_1/S_0)_{CI}$. Como se vio en el trabajo previo, la vía desde $^1(CC)_{exc}$ hasta la CI presenta una pequeña barrera de 0.2 eV. Estos resultados indican que los dímeros TT tienen un mayor número de orientaciones

reactivas que los de CC, lo cual está de acuerdo con la mayor producción de CBTs que de CBCs observada experimentalmente.[44] Además, una vez formado el fotoproducto, la reversibilidad foto-inducida de CBT requiere de una energía más alta (5.48eV) en comparación con la energía de absorción de CBC (4.57eV). Por lo tanto, la citosina dispone de un mecanismo más eficiente en condiciones de luz solar para revertir la lesión CBC y regresar de nuevo hacia el excímero más estable. Por último, debido a la presencia de dichos excímeros estables en citosina, la ruta fotoquímica de producción de CBCs compite con la desactivación por fluorescencia del excímero y la conversión interna del monómero excitado citosina, mientras que para TT, el decaimiento del monómero excitado se vuelve relevante sólo para bases de timina no apiladas.

En el tercer estudio se caracterizó la ***formación de excímeros en los dímeros de pirimidinas***. Se analizó el papel de la interacción por apilamiento de tipo π entre todas nucleobases del ADN y ARN en la estabilidad, la dinámica y la reactividad de los dímeros de bases. Al igual que se hizo en el primer trabajo de la Tesis para la citosina, se calcularon las energía de enlace a varias distancias de separación de los monómeros con una configuración “cara a cara” y usando la geometría del estado fundamental de las nucleobases. Los resultados indican que los tres pares de pirimidinas CC, TT y UU tienen un comportamiento similar respecto a sus propiedades enlazantes en el estado electrónico excitado, dando lugar a la formación de excímeros. Dichas interacciones pueden ocurrir en varias de las orientaciones relativas entre las dos nucleobases disponibles dentro de la gama de flexibilidad de movimiento que poseen tanto el ADN como el ARN. Por otra parte, se mostró, en base a los cálculos CASPT2 realizados, que las interacciones por apilamiento de tipo π permiten la existencia de nuevas fuentes de fluorescencia relacionadas con la asociación de los monómeros, apoyando así la conexión entre la presencia del excímero y el desplazamiento batocrómico de la fluorescencia en oligonucleótidos de pirimidinas. Finalmente, teniendo en cuenta de forma conjunta los análisis de este trabajo junto con la determinación del mecanismo de formación de CPDs, se puede afirmar que los bioexcímeros juegan un papel fundamental en la Fotobiología modulando la redistribución de la carga en el estado excitado más bajo. Esto es crucial para entender la reactividad intrínseca de las nucleobases de ADN hacia la dimerización de pirimidinas.

En el cuarto estudio, se determinaron los ***mecanismos de población del triplete en las nucleobases aisladas***. Para ello se llevaron a cabo cálculos MEP desde el estado electrónico brillante, de tipo singlete, empezando en la geometría Franck-Condon y usando el protocolo CASPT2//CASSCF. A lo largo del camino de desactivación energética marcado por el MEP, se calcularon las energías de los estados electrónicos singlete y triplete más bajos. En el caso de la adenina, timina y uracilo, se encontraron tres regiones de cruce singlete-triplete (STC) fácilmente accesibles desde la vía principal de decaimiento del singlete. Dos de ellos se obtuvieron a altas energías y están mediados por la existencia de un estado singlete de baja energía y un triplete, ambos de tipo $n\pi^*$. La tercera vía se encontró a bajas energías cercanas al final del MEP del estado excitado $^1(\pi\pi^*)$ singlete, donde se cruzan mediante CIs este estado singlete con el estado fundamental. Nuestro estudio sugiere que la dependencia entre la longitud de onda empleada experimentalmente[31,103] para irradiar estas tres moléculas y el rendimiento cuántico de formación del triplete observado puede estar relacionada con la activación de cualquiera de los tres canales ISC disponibles o tan sólo del canal situado a bajas energías. Por el contrario, en guanina y citosina, no se obtuvieron cruces accesibles entre estados singlete y triplete de tipo $n\pi^*$ al principio del MEP. Sólo se encontraron los cruces STC a bajas energías, cerca de la CI con el estado fundamental. Estos resultados indican que en general tanto la guanina como la citosina contribuyen mucho menos a la fosforescencia del ADN que el resto de bases. En el ARN, adenina y uracilo deben ser consideradas responsables de su fosforescencia mientras que en el ADN, adenina y timina son los principales contribuyentes a este fenómeno.

En el quinto trabajo, se abordó el estudio de la ***foto-dimerización de ciclobutilpirimidinas en el estado triplete***. Se calculó y se examinó la formación de dímeros fotoinducidos de CPD en el estado triplete excitado a partir de los pares de nucleobases pirimidínicas del ADN/ARN. Se propuso un mecanismo por pasos a través del estado triplete de los homodímeros para la formación de CPDs en las nucleobases citosina, timina y uracilo. Los cálculos MEP realizados en uracilo y timina desde el excímero triplete ^{3}(LE) llevaron a la formación de un intermedio ^{3}(SWI), al igual que en el caso de la citosina presentado en el primer estudio de la Tesis. Se encontró que este intermedio es el mejor candidato para desencadenar la formación del fotoaducto del ciclobutano en la vía triplete, ya que coincide con un punto de cruce $(T_1/S_0)_X$ que conecta con

la estructura de equilibrio en el estado fundamental del CPD. Esta estructura establece, además, un límite de energía necesario para las especies fotosensibilizadoras dadoras de energía a través del proceso de transferencia de energía de tipo triplete-triplete (TTET). Así, fotosensibilizadores con energías más bajas que la del intermedio ^{3}(SWI) no pueden inducir la formación de CPDs. En nuestro estudio, la energía CASPT2 del ^{3}(SWI) de la citosina resultó ser más alta que la de timina y la de uracilo. Por tanto, la citosina será menos accesible a la fotosensibilización.

En el sexto trabajo se llevó a cabo la ***comparación de los mecanismos de fotoproducción en el ADN/ARN de dímeros de ciclobutilpirimidina en las nucleobases citosina, timina, uracilo y 5-metilcitosina a través del estado excitado singlete***. Mientras que las primeras tres moléculas constituyen las bases canónicas de tipo pirimidina del ADN/ARN, la última es un derivado de la citosina que también se encuentra en el material genético y tiene un papel crucial en los procesos de epigenética. En el estudio, se establecieron en primer lugar las principales estructuras que caracterizan la fotoquímica de los homodímeros de timina, uracilo, citosina y 5-metilcitosina, lo cual permitió establecer un mecanismo unificado y comparativo de la formación de CPDs. Estas estructuras clave son, en primer lugar, el excímero, Exc, correspondiente a la geometría optimizada en el estado excitado singlete de más baja energía con una orientación “cara a cara” de las bases. La segunda de las estructuras importantes es la CI deslocalizada en el dímero, CI_{dim}, responsable del proceso reactivo. El último punto clave es la CI_{mon} localizada en los monómeros, la cual permite la desactivación no radiativa ultrarrápida de las bases aisladas. Los análisis comparativos demostraron que la posición relativa del punto CI_{dim} con respecto a las estructuras Exc y CI_{mon} es un buen indicador de la eficiencia de la fotorreactividad. En los homodímeros de timina y uracilo, los cálculos CASPT2 mostraron que el punto de más baja energía de entre las tres estructuras es la CI_{dim}, favoreciendo así la ruta fotoquímica de formación de CBT y CBU. En cambio, en los dímeros de la citosina y la 5-metilcitosina, los mínimos Exc son los puntos de menor energía. Además, los cruces CI_{dim} y CI_{mon} tienen energías similares. Por tanto, los decaimientos radiativos desde Exc y no radiativos localizados en los monómeros (a través de CI_{mon}) compiten con las reacciones de dimerización. Estos resultados están de acuerdo con los menores rendimientos cuánticos de formación de CPDs observados

experimentalmente para citosina.[44] Por otra parte, la comparación entre citosina y 5-metilcitosina o entre uracilo y timina indica una desestabilización de las estructuras Exc con respecto a los cruces CI_{dim}, que se puede atribuir al efecto estérico del grupo metilo. La metilación aumenta entonces la probabilidad del proceso de dimerización debido a la disminución de las barreras de energía que conducen a la reacción de producción del fotoproducto.

Finalmente, en el séptimo estudio, se realizó un ***análisis teórico del proceso de disociación por adición de electrones de baja energía en el uracilo***. Los fenómenos disociativos DEA en los enlaces N_1–H y N_3–H observados experimentalmente[137,138,139,141,142,143,144] a bajas energías (< 3 eV) en uracilo se estudiaron al nivel de teoría CASSCF/CASPT2 utilizando diferentes estrategias de cálculo. El estudio se centró en las contribuciones de los aniones de valencia (VB) para completar estudios previos donde la participación de los aniones enlazados por dipolo (DB) ya fue demostrada.[142,143,144,145] Para ello, se determinaron, en nuestro trabajo, varias magnitudes teóricas que son de gran ayuda para asignar las señales de las secciones de cruce medidas experimentalmente. Estas magnitudes fueron, primero, las afinidades electrónicas de los estados aniónicos más bajos del uracilo. En segundo lugar, las energías del proceso de formación de los aniones producto de las roturas, $(U–H)^-$, más los radicales hidrógeno, ·H, relativas a la energía del uracilo neutro. Otra magnitud de interés fue la energía de los cruces entre las PECs del anión π_1^- y las del estado disociativo σ^-. La posición energética de los mínimos del estado π_2^- y de su cruce con el estado σ^- fueron también dos características del mecanismo que se determinaron para analizar la participación del estado π_2^- en los procesos DEA. Los resultados obtenidos para las energías del proceso de disociación DEA concuerdan con los umbrales de energía de 0.6–0.8 y 1.2–1.4 eV obtenidos experimentalmente para las disociaciones en N_1–H y N_3–H, respectivamente. Por otra parte, se determinó que los puntos de cruce entre las PECs de los estados π^- y σ^- aparecen a energías en torno a la posición energética de los máximos del pico más bajo y la banda ancha observada experimentalmente en la sección transversal del DEA. Por tanto, dichas CIs se interpretaron como los puntos que activan la conversión interna de los estados disociativos σ^- y que conducen al sistema hacia la formación del anión $(U–H)^-$ y un átomo de H. El estado π_1^- cruza con el σ_{N1H}^- a una energía de alrededor de 0.9 eV, apuntando,

por tanto, a la participación del anión π_1^- en la escisión del enlace N_1–H. Por otro lado, el estado π_2^- se predijo como el principal estado responsable de la señal experimental medida en la región energética en torno a 1.7 eV, debido al hecho de que se obtuvo un punto de cruce con el estado σ_{N3H}^-, $(\pi_2^-/\sigma_{N3H}^-)_{CI}$ en esta región. Nuestros resultados predicen que el anión VB del uracilo también es responsable de las señales en la región inferior del espectro y que proporcionan información adicional sobre la correspondencia entre el estado π_2^- y la banda ancha del espectro experimental.

Conclusiones

Los trabajos de investigación en química teórica llevados a cabo en la presente Tesis Doctoral han permitido determinar las bases moleculares de procesos de fotorreactividad en el ADN/ARN. A continuación, se resumen las conclusiones más relevantes obtenidas, las cuales son importantes dentro del área de investigación del daño radiativo al ADN/ARN inducido por luz UV o en condiciones de estrés reductor:

- El proceso de producción de las lesiones ciclobutilpirimidina inducidas por irradiación con luz UV ocurre a través de una CI entre el estado electrónico excitado del dímero y el fundamental. La formación de excímeros, que tiene lugar en todas las pirimidínicas, es una etapa precursora de la producción de los fotodímeros CPDs. El apilamiento de tipo π y la formación del excímero juegan un papel clave en la modulación de la dinámica de relajación del ADN/ARN.

- Tres características del mecanismo de producción de CPDs y su fotorreversibilidad son claves para entender el menor rendimiento medido experimentalmente para la producción fotoinducida de fotodímeros de citosina (CBCs) con respecto a la formación del resto de fotodímeros. En primer lugar, existen estructuras exciméricas en citosina más estables que el punto de CI que conecta con la lesión CPD. Esto no ocurre en los homodímeros de timina y uracilo, donde cualquier interacción efectiva por apilamiento de tipo π entre las bases conduce mediante un camino sin barrera a la CI y a la lesión. La

posición relativa de la CI del dímero con respecto a la estructura estable del excímero o de la región CI del decaimiento del monómero determina la eficiencia de la fotorreactividad. En segundo lugar, la fotorreversibilidad por absorción de luz UV del fotodímero CPD requiere de mayor energía en el caso de la timina con lo cual se espera que este proceso de regeneración de las bases sea menos eficiente en condiciones de luz solar. En tercer lugar, la competencia del proceso de desactivación ultrarrápida no radiativa localizado en los monómeros es mayor en timina y uracilo que en citosina.

- Las bases del ADN/ARN presentan rutas de población del triplete a lo largo del camino de decaimiento principal del estado relacionado con la transición más intensa. Nuestros resultados indican que la adenina y la timina son las responsables de la fosforescencia del ADN. Del mismo modo, en el ARN, son la adenina y el uracilo las bases que pueden considerarse como las causantes de la fosforescencia. Por otro lado, la guanina y la citosina prácticamente no contribuyen a la fosforescencia en el ADN debido al menor número de canales de población eficiente del triplete presentes en su fotoquímica.

- La formación de CPDs también puede ocurrir vía el estado electrónico triplete de los homodímeros, bien tras la población del triplete en las bases aisladas y la posterior formación de excímeros, o bien, por fotosensibilización de una especie dadora de energía triplete. La estructura clave en este mecanismo es un intermedio biradical que constituye a su vez un cruce STC entre el estado triplete del excímero y el fundamental. La ubicación relativa de las estructuras intermediarias de tripletes en diferentes homodímeros de nucleobases explica la fototoxicidad de diferentes fotosensibilizadores, impulsados por la energía de sus estados triplete reactivos.

- El umbral de energía para la disociación de los enlaces N_1–H y N_3–H en uracilo por adición de electrones de baja energía (0–3 eV) es de 0.6–0.8 y 1.2–1.4 eV, respectivamente, lo cual

concuerda con las observaciones experimentales. Se predice que los aniones de valencia VB del uracilo también contribuyen al proceso DEA, además de los aniones enlazados por dipolo (DB). El anión π_1^- participa mayoritariamente en la señal en torno a 1 eV, mientras que el π_2^- contribuye sólo a la banda ancha medida experimentalmente en torno a 1.7 eV.

Acronyms and Abbreviations

A	*Adenine*
Abs	*Absorption*
AEA	*Adiabatic Electron Affinity*
AMFI	*Atomic Mean-Field Integral*
ANO	*Atomic Natural Orbital*
aug-	*Augmented*
aug-cc-pVDZ	*Augmented Correlation Consistent Polarized Valence Double Zeta*
BSSE	*Basis Set Superposition Error*
C	*Cytosine*
CAS-CI	*Complete Active Space Configuration Interaction*
CASSCF	*Complete Active Space Self-Consistent Field*
CASPT2	*Complete Active Space Perturbation Theory to Second Order*
CBC	*Cyclobutane Cytosine*
CBT	*Cyclobutane Thymine*
CBU	*Cyclobutane Uracil*
CC	*Coupled Cluster*
cc-pVDZ	*Correlation Consistent Polarized Valence Double Zeta*
CCSD	*Coupled Cluster Singles and Doubles*
CCSD(T)	*Coupled Cluster Singles and Doubles and Triples computed perturbativelly*
CGTO	*Contracted Gaussian Type Orbital*
CI	*Configuration Interaction / Conical Intersection*
CISD	*Configuration Interaction Singles and Doubles*
CP	*Counterpoise*
CPD	*Cyclobutane Pyrimidine Dimer*
CSF	*Configurational State Function*
D	*Doubly*
DB	*Dipole Bound*
dim	*Dimer*
DEA	*Dissociative Electron-Attachment*
DNA	*Deoxyrribonucleic Acid*
DSB	*Double Strand Break*
DZ	*Double Zeta*
DZP	*Double Zeta Polarization*
EA	*Electron Affinity*
ETS	*Electron Transmission Spectroscopy*
Exc	*Excimer*
F	*Fluoresecence*
FCI	*Full Configuration Interaction*
G	*Guanine*
GTO	*Gaussian Type Orbital*
HF	*Hartree-Fock*

IC	*Internal Conversion*
IMAG	*Imaginary*
IP	*Ionization Potential*
IPEA	*Ionization Potential Electron Affinity*
IRC	*Intrinsic Reactive Coordinate*
ISC	*Intersystem Crossing*
IUPAC	*International Union of Pure and Applied Chemistry*
LE	*Locally Excited*
LIIC	*Linear Interpolation of Internal Coordinates*
m^5C	*5-methylcytosine*
MBPT	*Many-Body Perturbation Theory*
MCSCF	*Multi-Configurational Self-Consistent Field*
MEP	*Minimum Energy Path*
MO	*Molecular Orbital*
MO-LCAO	*Molecular Orbitals as Linear Combination of Atomic Orbitals*
mon	*Monomer*
MP	*Møller-Plesset perturbation theory*
MP2	*Møller-Plesset perturbation theory to Second order*
MP3	*Møller-Plesset perturbation theory to Third order*
MRCC	*Multi-Reference Coupled Cluster*
MRPT	*Multi-Reference Perturbation Theory*
MRCI	*Multi-Reference Configuration Interaction*
MS	*Multistate*
NO	*Natural Orbital*
P	*Phosphorescence*
PEC	*Potential Energy Curve*
PEH	*Potential Energy Hypersurface*
PGTO	*Primitive Gaussian Type Orbital*
Pyr	*Pyrimidine*
QCEXVAL	*Quantum Chemistry of the Excited State Universitat de València*
QM	*Quantum Methods*
RHF	*Restricted Hartree-Fock*
RNA	*Ribonucleic Acid*
ROHF	*Restricted Open-shell Hartree-Fock*
ROS	*Reactive Oxygen Species*
S	*Singly / Singlet*
SCF	*Self-Consistent Field*
SDC	*Singlet-Doublet Crossing*
SDCI	*Singly and Doubly Excited Configuration Interaction*
SOC	*Spin-Orbit Coupling*
SONO	*Single-Occupied Natural Orbital*
SP	*Saddle Point*
SSB	*Single Strand Break*
STC	*Singlet-Triplet Crossing*
STO	*Slater Type Orbital*
SWI	*Step-Wise Intermediate*
T	*Thymine / Triply / Triplet*
TDM	*Transition Dipole Moment*

TS	*Transition State*
TTET	*Triplet-Triplet Energy Transfer*
TZ	*Triple Zeta*
TZP	*Triple Zeta Polarization*
U	*Uracil*
UHF	*Unrestricted Hartree-Fock*
UV	*Ultraviolet*
UV-Vis	*Ultraviolet-Visible*
VAE	*Vertical Attachment Energy*
VB	*Valence Bound*
VDZ	*Valence-Double-Zeta*
VEA	*Vertical Electron Affinity*
VTZ	*Valence-Triple-Zeta*
X	*Crossing*
ZPVE	*Zero-Point Vibrational Energy*

References

[1] University of California Museum of Paleontology. *Erasmus Darwin (1731-1802).* Online. Cited on September 21st, 2015. Available at http://www.ucmp.berkeley.edu/history/Edarwin.

[2] The British Broadcasting Corporation. *Charles Darwin: Evolution and the story of our species.* Online. Cited on September 21st, 2015. Available at http://www.bbc.co.uk/timelines/zq8gcdm

[3] John Alroy. *Erasmus Darwin (1731-1802).* Online. Cited on September 21st, 2015. Available at https://www.nceas.ucsb.edu/~alroy/lefa/ErasmusDarwin.html

[4] Indiana University of Pennsylvania. *The Charles Darwin Collection.* Online. Cited on September 21st, 2015. Available at http://www.iup.edu/archives/digital-projects-and-exhibits/the-charles-darwin-collection/

[5] John Alroy. *Johann Gregor Mendel(1822-1884).* Online. Cited on September 21st, 2015. Available at https://www.nceas.ucsb.edu/~alroy/lefa/Mendel.html

[6] Dahm, R. *Developmental Biology* **2005**, *278*, 274.

[7] Nature Education. *Gregor Mendel: A private Scientist.* Online. Cited on September 21st, 2015. Available at http://www.nature.com/scitable/topicpage/gregor-mendel-a-private-scientist-6618227

[8] George Wolf. *Friedriech Miescher, The Man Who Discovered DNA.* Online. Cited on June 18th, 2014. Available at http://www.bizgraphic.ch/miescheriana/html/the_man_who_dicovered_dna.html

[9] Steven M. Carr. *Leven's Tetranucleotide Hypothesis.* Online. Cited on October 2nd, 2015. Available at http://www.mun.ca/biology/scarr/Tetranucleotide_Hypothesis.html

[10] U.S. National Library of Science. *The Oswald T. Avery Collection. After the Discovery: The Transforming Principle's Reception by the Scientific Community.* Online. Cited on October 4th, 2015. Available at http://profiles.nlm.nih.gov/ps/retrieve/Narrative/CC/p-nid/158/p-visuals/

[11] Brian Allensby. *The People Responsible for the Discovery of DNA.* Online. Cited on October 4th, 2015. Available at http://www.ba-education.com/for/science/dnadiscovery.html

[12] Van Slyke, D. D.; Jacobs, W. Biographical Memoir of Phoebus Aaron Theodor Levene, **1945** Vol. 23, pp. 75– 86, National Academy of Sciences, Washington, D. C.

[13] Levene, P. A.; *J. Biol. Chem.* **1919**, *40*, 415.

[14] IUPAC-IUB Joint Commission on Biochemical Nomenclature *Eur. J. Biochem.* **1983**, *131*, 9.

[15] Avery, O. T.; MacLeod, C. M.; McCarty, M. *J. Exp. Med.* **1944**, *79*, 137.

[16] Katherine Nightingale. *Behind the Picture: Photo 51.* Online. Cited on October 5th, 2015. Available at http://www.insight.mrc.ac.uk/2013/04/25/behind-the-picture-photo-51

[17] Franklin, R. E.; Gosling, R. G. *Nature* **1953**, *171*, 740.

[18] Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 737.

[19] Wilkins, M. H. F.; Stokes, A. R.; Wilson, H. R. *Nature* **1953**, *171*, 738.

[20] A. Barrington Brown. *The Discoverers of the DNA structure , James Watson , and Francis Crick, with their model of a DNA molecule.* Online. Cited on October 4th, 2015. Available at https://news.wsu.edu/2012/01/11/nih-funds-scientists-work-to-unravel-cell-repair/#.VTa3O9J_Oko

[21] Rich, A.; Davies, D. R. *J. Am. Chem. Soc.* **1956**, *78*, 3548.

[22] Rich, A. *Nature Struct. Biol.* **2003**, *10(4)*, 247.

[23] Varshavsky, A. *Cell* **2006**, *127(7)*, 1295.

[24] Rich, A. *Q. Rev. Biophys.* **2009**, *42(2)*, 117.

[25] Erik Chaulk. *Alexander Rich, 1924-2015.* Online. Cited on October 5th, 2015. Available at http://www.asbmb.org/asbmbtoday/201508/InMemoriam/Rich

[26] University of Maryland. *Past and Present Distinguished Advisors.* Online., Cited on October 5th, 2015. Available at http://www.chem.umd.edu/aboutus/committeeofdistinguishedadvisors/committeeofdistinguishedadvisorspastandpresentdistinguishedadvisors

[27] Eisinger, J.; Shulman, R. G. *Science* **1968**, *161*, 1311.

[28] Crespo-Hernández, C. E.; Cohen, B.; Hare, P. M.; Kohler, B. *Chem. Rev.* **2004**, *104*, 1977.

[29] Shreier, W. J.; Schrader, T. E.; Soller, F. O.; Gilch, P.; Crespo-Hernández, C. E.; Swaminathan, V. N.; Carell, T.; Zinth, W.; Kohler, B., *Science* **2007**, *315*, 625.

[30] Callis, P. R. *Annu. Rev, Phys. Chem.* **1983**, *34*, 329.

[31] Eisinger, J.; Guéron, M.; Shulman, R. G.; Yamane, T. *Proc. Natl. Acad. Sci. U.S.A.* **1966**, *55*, 1015.

[32] Kwok, W. M.; Ma, C.; Phillips, D. L. *J. Am. Chem. Soc.* **2006**, *128*, 11894.

[33] Canuel, C.; Mons, M.; Pluzzi, F.; Tardivel, B.; Dimicoli, I.; Elhanine, M. *J. Chem. Phys.* **2005**, *122*, 074316.

[34] Cadet, J; Vigny, P. In *Bioorganic Photochemistry*; Morrison, H., Ed.; John Wiley & Sons: New York, 1990; Vol. 1, pp 1-272.

[35] Danilov, V. I.; Slyusarchuk, O. N.; Alderfer, J. L.; Stewart, J. J. P.; Callis, P. R. *Photochem. Photobiol.* **1994**, *59*, 125.

[36] Marguet, S.; Markovitsi, D. *J. Am. Chem. Soc.* **2005**, *127*, 5780.

[37] Holman, M. R.; Ito, T.; Rokita, S. E. *J. Am. Chem. Soc.* **2007**, *315*, 625.

[38] von Sonntag, C. *Free-Radical-Induced DNA Damage and Its Repair: A Chemical Perspective*; Springer-Verlag: Berlin, 2006.

[39] Lubin, D.; Jensen, E. H. *Nature* **1995**, *377 (6551)*, 710.

[40] Häder, D.-P.; Kumar, H. D.; Smith, R. C.; Worrest, R. C. *Photochem. Photobiol. Sci.* **2007**, *6(3)*, 267.

[41] Solomon, K. R. *Atmos.-Ocean* **2008**, *46(1)*, 185.

[42] Friedberg, E. C.; Walker, G. C.; Siede, W.; Wood, R. D.; Schultz, R. A. *DNA Repair and Mutagenesis*, Ellenberger, T. ed.; ASM Press: Washington DC, 2006.

[43] Björn, L. O.; McKenzie, R. L. *Photobiology. The Science of Life and Light*; Björn, L.O. ed.; Springer: New York, 2008, 503.

[44] Douki, T.; Cadet, J. *Biochemistry* **2001,** *40*, 2495

[45] Schuch, A. P.; Menck, C. F. M. *J. Photochem. Photobiol. B Biol.* **2010**, *99*, 111.

[46] Patricia Shapley. *Light and the Electromagnetic Spectrum.* Online, Cited on May 11th, 2015. Available at http://butane.chem.uiuc.edu/pshapley/GenChem2/A3/3.html

[47] Lesczczynski, J., Ed. *Computational Molecular Biology*; Elsevier: Amsterdam, 1999.

[48] Eriksson, L. A., Ed. *Theoretical Biochemistry. Processes and Properties In Biological Systems*; Elsevier: Amsterdam, 2001.
[49] Leszczynski, J; Shukla, M., Eds., *Radiation Induced Molecular Phenomena in Nucleic Acids. A Comprehensive Theoretical Analysis*; Springer: The Netherlands, 2008.
[50] Boon, P. J.; Cullis, P. M.; Symons, M. C. R.; Wren, B. W. *J. Chem. Soc. Perkin Trans II* **1984**, 1393.
[51] Sevilla, M. D.; Becker, D.; Yan, M.; Summerfield, S. R. *J. Phys. Chem.* **1991**, *95*, 3409.
[52] Yan, M.; Becker, D.; Summerfield, S. R.; Renke, P.; Sevilla, M. D. *J. Phys. Chem.* **1992**, *96*, 1983.
[53] Simons, J. *Acc. Chem. Res.* **2006**, *39(10)*, 772
[54] Jensen, F. *Introduction to Computational Chemistry*; John Wiley & Sons: Chichester, 1999.
[55] Andrés, J., Beltrán, J. *Química Teórica y Computacional*; Universitat Jaume I: Castelló, 2000.
[56] Merchán, M.; Serrano-Andrés, L. In *Computacional Photochemistry*; Olivucci, M., Ed.; Elsevier: Amsterdam, 2005.; Vol. 16, Chap. 2, p. 35.
[57] Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry. Introduction to Advanced Electronic Structure Theory*; Dover Publications Inc.: Mineola, New York, 1996.
[58] Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*; John Wiley & Sons: Chichester, 2004.
[59] Widmark, P.-O.; Roos B. O. *European Summerschool of Quantum Chemistry 2005*. *Books I, II, III*; Lund University: Lund, 2005.
[60] Hartree, D. R. *Proc. Cambridge Philos. Soc.* **1928**, *24*, 328.
[61] Fock, V. A. *Z. Phys.* **1930**, *15*, 126.
[62] Roothaan, C. C. J. *Rev. Mod. Phys.* **1951**, *23*, 69.
[63] Pople, J. A.; Nesbet, R. K. *J. Chem. Phys.* **1954**, *22*, 571.
[64] Roothaan, C. C. J. *Rev. Mod. Phys.* **1960**, *32*, 179.
[65] Meckler, A. *J. Chem. Phys.* **1953**, *21*, 1750.
[66] Čížek, J. *J. Chem. Phys.* **1966**, *45*, 4256.
[67] Čížek, J. *Adv. Chem. Phys.* **1969**, *14*, 35.
[68] Čížek, J.; Paldus, J. *Int. J. Quantum Chem.* **1971**, *5*, 359.
[69] Møller C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618.
[70] Roos, B. O.; Taylor, P. R.; Siegbahn, P. E. M. *Chem. Phys.* **1980**, *48*, 157.
[71] Roos, B. O. *Int. J. Quantum Chem. Symp.* **1980**, *14*, 175.
[72] Roos, B. O. In *Advances in Chemical Physics; Ab Initio Methods in Quantum Chemistry-II*; Lawley, K. P., Ed.; John Wiley & Sons Ltd.: 1987; p. 399.
[73] Andersson, K.; Malmqvist, P.-Å.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. *J. Phys. Chem.* **1990**, *94*, 5483.
[74] Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. *J. Chem. Phys.* **1992**, *96*, 1218.
[75] Roos, B. O.; Andersson, K. *Chem. Phys. Lett.* **1995**, *245*, 215.
[76] Roos, B.O.; Andersson, K.; Fülscher, M. P.; Serrano-Andrés, L.; Pierloot, K.; Merchán, M.; Molina, V. *J. Mol. Struct. THEOCHEM* **1996**, *388*, 257.
[77] Forsberg, N.; Malmqvist, P.-Å. *Chem. Phys. Lett.* **1997**, *274*, 196.
[78] Ghigo, G.; Roos, B. O.; Malmqvist, P.-Å. *Chem. Phys. Lett.* **2004**, *396*, 142.

[79] Finley, J.; Malmqvist, P.-Å.; Roos, B. O.; Serrano-Andrés, L. *Chem. Phys. Lett.* **1998**, *288*, 299.
[80] Slater, J. C. *Phys. Rev.* **1930**, *36*, 57.
[81] Boys, S. F. *Proc. R. Soc. (London) A* **1950**, *200*, 542.
[82] Heher, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; John Wiley & Sons: 1986.
[83] Dunning Jr, T. H. *J. Chem. Phys.* **1989**, *90*, 1007.
[84] Wilson, A. K.; Van Mourik, T.; Dunning Jr, T. H. *J. Mol. Struct.* **1996**, *388*, 339.
[85] Almlöf, J.; Taylor, P. R. *J. Chem. Phys.* **1978**, *86*, 4070.
[86] Widmark, P.-O.; Malmqvist, P.-Å.; Roos, B. O. *Theor. Chim. Acta* **1990**, *77*, 291.
[87] Widmark, P.-O.; Persson, B. J.; Roos, B. O. *Theor Chim Acta* **1991**, *79*, 291.
[88] Pou-Amérigo, R.; Merchán, M.; Nebot-Gil, I.; Widmark, P.-O.; Roos, B. O. *Theor. Chim. Acta* **1995**, *92*, 149.
[89] Pierloot, K.; Dumez, B.; Widmark, P.-O.; Roos, B. O. *Theor. Chim. Acta* **1995**, *90*, 87.
[90] Roos, B. O.; Lindh, R.; Malmqvist, P.-Å.; Veryazov. V.; Widmark, P.-O. *J. Phys. Chem. A* **2004**, *108*, 2851.
[91] Davidson, E. R.; Chakravory, S. *J. Chem. Phys. Lett.* **1994**, *217*, 49.
[92] Boys, S. F.; Bernardi, F. *Molecular Physics* **1970**, *19*, 553.
[93] Turro, N. J. *Modern Molecular Photochemistry*; University Science Books: Sausalito, 1991.
[94] Serrano-Andrés, L.; Merchán, M. In *Encyclopedia of Computational Chemistry*; Schlegel, P. V. R.; Schreiner, P. R.; Schaefer III, H. F.; Jorgensen, W. L.; Thiel, W.; Glen, R. C., Eds.; Wiley: Chichester, 2004.
[95] Robb, M. A.; Olivucci, M.; Bernardi, F. In *Encyclopedia of Computational Chemistry*; Schlegel, P. V. R. ; Schreiner, P. R.; Schaefer III, H. F.; Jorgensen, W. L.; Thiel, W.; Glen, R. C., Eds.; Wiley: Chichester, 2004.
[96] Klessinger, M.; Michl, J. *Excited States and Photochemistry of Organic Molecules*; VCH Publishers: New York, 1995.
[97] Peyerimhoff, S. In *Encyclopedia of Computational Chemistry*; Schlegel, P. V. R. ; Schreiner, P. R.; Schaefer III, H. F.; Jorgensen, W. L.; Thiel, W.; Glen, R. C., Eds.; Wiley: Chichester, 2004.
[98] Serrano-Andrés, L.; Merchán, M. *Anales de Química* **2004**, *100*, 16.
[99] Kasha, M. *Disc. Farad. Soc.* **1950**, *9*, 14.
[100] Birks, J. B. *Rep. Prog. Phys.* **1975**, *38*, 903.
[101] Stevens, B.; Hutton, E. *Nature, Lond.* **1960**, *186*, 1045.
[102] Birks, J. B. *The Exciplex*; Academic Press: New York, San Francisco, London, 1975.
[103] Plessow, R.; Brockhinke, A.; Eimer, W.; Kohse-Höinghaus, K. *J. Phys. Chem. B* **2000**, *104*, 3695.
[104] Crespo-Hernández, C. E.; Cohen, B.; Kohler, B. *Nature* **2005**, *436*, 1141.
[105] Teller, E. *J. Phys. Chem.* **1937**, *41*, 109.
[106] Klessinger, M. *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 549.
[107] Bernardi, F.; Olivucci, M.; Robb, M. A. *Chem. Soc. Rev.* **1996**, *25*, 321.
[108] Domcke, W.; Yarkony, D. R.; Köppel, H. *Conical Intersections: Electronic Structure, Dynamics and Spectroscopy*; World Scientific: Singapore, 2004.
[109] Rodríguez, A. R.; Román, J. Z. *Espectroscopía*; Pearson Prentice Hall: Madrid, 2004.
[110] Perkampus, H.-H. *Encyclopedia of Spectroscopy*; VCH Publishers: Winheim, 1993.

[111] Pilar, F. *Elementary Quantum Chemistry*; Dover Publications, Inc.: 2001.
[112] Serrano-Andrés, L.; Serrano-Pérez, J. J. In *Handbook of Computational Chemistry*, Leszczynski, J., Ed.; Springer Verlag: Berlin, Heidelberg, 2012; pp 483-560.
[113] Robb, M. A.; Garavelli, M., Olivucci, M., Bernardi, F. *Rev. Comput. Chem. 2000*, *15*, 87.
[114] Truhlar, D. G.; Steckler, R.; Gordon, M. S. *Chem. Rev.* **1987**, *87*, 217.
[115] De Vico, L.; Olivucci, M.; Lindh, R. *J. Chem. Theor. Comput.* **2005**, *1*, 1029.
[116] Anglada, J. M.; Bofill, J. M. *J. Comput. Chem.* **1997**, *18*, 992.
[117] Müller, K.; Brown, L. D. *Theor. Chem. Acc.*, **1979**, *53*, 75.
[118] Karlström, G.; Lindh, R.; Malmqvist, P.-Å.; Roos, B.O.; Ryde, U.; Veryazov, V.; Widmark, P.-O.; Cossi, M.; Schimmelpfennig, B.; Neogrady, P.; Seijo, L. *Comp. Mat. Sci.* **2003,** *28*, 222.
[119] Veryazov, V.; Widmark, P.–O.; Serrano-Andrés, L.; Lindh, R.; Roos, B.O. *Int. J. Quantum Chem.* 2004, *100*, 626.
[120] Aquilante, F.; De Vico, L.; Ferré, N.; Ghigo, G.; Malmqvist, P.-Å.; Neogrády, P.; Pedersen, T. B.; Pitonák, M.; Reiher, M.; Roos, B. O.; Serrano-Andrés, L.; Urban, M.; Veryazov, V.; Lindh, R. *J. Comput. Chem.* **2010**, *31*, 224.
[121] Heβ, B. A.; Marian, C. M.; Wahlgren, U.; Gropen, O. *Chem. Phys. Lett.* **1996**, *251*, 365.
[122] Malmqvist, P.-Å.; Roos, B. O. *Chem. Phys. Lett.* **1989**, *155*, 189.
[123] Imakubo, K. *J. Phys. Soc. Jpn.* **1968**, *24*, 1124.
[124] Szerenyi, P.; Dearman, H. H. *Chem. Phys. Lett.* **1972**, *15*, 81.
[125] Arce, R.; Rodríguez, G. *J. Photochem.* **1986**, *33*, 89.
[126] Kwok, W. M.; Ma, C.; Phillips, D. L. *J. Am. Chem. Soc.* **2008**, *130*, 5131.
[127] Marguet, S.; Markovitsi, D. *J. Am. Chem. Soc.* **2005**, *127*, 5780.
[128] Gut, I. G.; Wood, P. D.; Redmond, R. W. *J. Am. Chem. Soc.* **1996**, *118*, 2366.
[129] Cadet, J.; Sage, E.; Douki, T. *Mutat. Res.* **2005**, *571*, 3.
[130] Nguyen, M. T.; Zhang, R.; Nam, P.-C.; Ceulemans, A. *J. Phys. Chem. A* **2004**, *108*, 6554.
[131] Bosca, F.; Lhiaubet-Vallet, V.; Cuquerella, M. C.; Castell, J. V.; Miranda, M. A. *J. Am. Chem. Soc.* **2006**, *128*, 6318.
[132] Etinski, M.; Fleig, T.; Marian, C. M. *J. Phys. Chem. A* **2009**, *113*, 11809.
[133] Abouaf, R.; Pommier, J.; Dunet, H.; Quan, P.; Nam, P.- C.; Nguyen, M. T. *J. Chem. Phys.* **2004**, *121*, 11668.
[134] Abouaf, R.; Pommier, J.; Dunet, H. *Chem. Phys. Lett.* **2003**, *381*, 486.
[135] Friso, S.; Choi, S. W.; Dolnikowski, G. G.; Selhub, J. *Anal. Chem.* **2002**, *74*, 4526.
[136] Tommasi, S.; Denissenko, M. F.; Pfeifer; G. P. *Cancer Res.* 1997, 57, 4727.
[137] Hanel, G.; Gstir, B.; Denifl, S.; Scheier, P.; Probst, M.; Farizon, B.; Farizon, M.; Illenberger, E.; Mäŕk, T. D. *Phys. Rev. Lett.* **2003**, *90*, 188104.
[138] Denifl, S.; Ptasinska, S.; Hanel, G.; Gstir, B.; Probst, M.; Scheier, P.; Märk, T. D. *J. Chem. Phys.* **2004**, *120*, 6557.
[139] Ptasinska, S.; Denifl, S.; Scheier, P.; Illenberger, E.; Märk, T. D. *Angew. Chem., Int. Ed.* **2005**, *44*, 6941.
[140] Li, X.; Sanche, L.; Sevilla, M. D. *J. Phys. Chem. B* **2004**, *108*, 5472.
[141] Ptasinska, S.; Denifl, S.; Grill, V.; Märk, T. D.; Illenberger, E.; Scheier, P. *Phys. Rev. Lett.* **2005**, *95*, 093201.

[142] Burrow, P. D.; Gallup, G. A.; Scheer, A. M.; Denifl, S.; Ptasinska, S.; Märk, T. D.; Scheier, P. *J. Chem. Phys.* **2006**, *124*, 124310.
[143] Scheer, A. M.; Silvernail, C.; Belot, J. A.; Aflatooni, K.; Gallup, G. A.; Burrow, P. D. *Chem. Phys. Lett.* **2005**, 411, 46.
[144] Scheer, A. M.; Aflatooni, K.; Gallup, G. A.; Burrow, P. D. *Phys. Rev. Lett.* **2004**, *92*, 068102.
[145] Gallup, G. A.; Fabrikant, I. I. *Phys. Rev. A* **2011**, *83*, 012706.