Review

Are the five natural DNA/RNA base monomers a good choice from natural selection?
A photochemical perspective

Luis Serrano-Andrés*, Manuela Merchán
Instituto de Ciencia Molecular, Universitat de València, Apartado 22085, ES-46071 Valencia, Spain

Abstract

In order to prevent the damaging effects of sun radiation in the genetic material, its constituent chromophores, the five natural DNA/RNA nucleobases cytosine, thymine, uracil, adenine, and guanine, should be able to efficiently dissipate absorbed radiation, UV specifically, avoiding as much as possible photoreactions leading to lesions. It has been established experimentally and theoretically that efficient internal conversion channels, still open and relevant in the oligomer-stacked strands, exist in the monomers allowing an effective waste of the initial energy. Previous evidences cannot explain, however, why minor differences in the molecular structure modify drastically the photochemistry of the systems, leading for many derivatives to slower decays, sometimes to intense fluorescence, and also to reactivity. Using the accurate CASPT2/CASSCF quantum chemical method and the Photochemical Reaction Path Approach it is determined that the five natural nucleobases display barrierless paths from the allowed excited state toward accessible conical intersection seams with the ground state. Such features are known to be the funnels for efficient energy decay and fluorescence quenching. Modified nucleobases, except the methylated ones, are predicted less photostable because they display energy barriers along lowest-energy paths and hence restricted accessibility of the internal conversion channel. This specificity speaks in favor of the choice of the biological nucleobases by natural selection based on their resistance to photochemical damage. Whereas natural and methylated nucleobases, also frequent in the genetic code, are photostable and cannot be photochemically discarded, other non-natural nucleobases may have been eliminated at early stages of the natural selection process.

© 2009 Elsevier B.V. All rights reserved.
Luis Serrano-Andrés was born in Baden, Switzerland, in 1966. In 1989 he graduated in chemistry by the University of Valencia, Spain, where he obtained in 1994 his Ph.D. degree by his studies on theoretical quantum-chemical organic photochemistry. His postdoctoral research was mainly developed from 1995 to 1998 in the University of Lund, Sweden, within the group of Björn Roos, working in particular in the development and application of quantum-chemical methods to the excited molecular state and their implementation in the MOLCAS software package. He has been invited researcher in the Universities of Lund, Sweden, Torun, Poland, and Sao Paulo, Brazil. In 1999 he joined the Quantum Chemistry of the Excited State (QCEXVAL) group in the University of Valencia, where he co-leads with Manuela Merchán the QCEXVAL group. He was a recipient of the Spanish Royal Chemical Society Young Scientist Award in 2001 and Promising Scientist Nominee Prize of the International Society of Theoretical Physical Chemistry and the Centre de Mecánica Ondulatoria Apilquelle of Paris in 2002.

Manuela Merchán was born in Castellón, Spain, in 1956. After graduating in 1980 in chemistry by the University of Valencia, Spain, she received in 1985 in this institution her Ph.D. degree for her studies in the quantum chemistry of metal–ligand compounds. After developing postdoctoral and research stays in the Universities of Toulouse, France, and Lund, Sweden, she became Associate Professor of Physical Chemistry in the University of Valencia in 1987, where she founded in 1993 the Quantum Chemistry of the Excited State (QCEXVAL) group. In 2004 she became Professor of Physical Chemistry in the University of Valencia. Her main research interest focuses on the development of computational strategies by using accurate quantum chemical ab initio methods and their application to the chemistry of the electronic excited states, in particular for relevant biomolecular systems. She was recipient of the National Prize in Physical Chemistry of the Spanish Royal Chemical Society in 2007.

1. Introduction

Are the five biologically relevant DNA and RNA base monomers cytosine (C), thymine (T), uracil (U), adenine (A), and guanine (G) the best molecules that natural selection could have chosen to build the genetic code among other close tautomers and derivatives? Although most part of the answer relates to structural considerations and its connection with the genetic function, it is possible also to focus on another aspect such as the stability of the genetic code upon external perturbations, for instance to the presence of UV radiation, especially frequent in the early stages of life on Earth [1,2]. We can get some insight into a complex problem if we encounter a manner to explain why analogous molecular structures are more prone to undergo productive photochemistry mainly leading to mutations. As it shall be discussed later, it is understandable that rapid non-radiative decay after UV-light absorption may have been highly advantageous during molecular evolution as a procedure to dissipate the energy before harmful photoproducts may be formed [3].

Experimentally it was earlier established that nucleic acid base monomers in aqueous phase largely quench their fluorescence [4–6]. Evidences of extremely efficient non-radiative decays to non-emitting states [6,7] were also put forward. Recently the presence of ultrafast internal conversion (IC) events transferring the energy from the excited to the ground state in nucleic acids has been directly detected by subpicosecond techniques as an intrinsic feature of nucleotides, nucleosides, and nucleic acid base monomers (see Fig. 1), in isolated conditions and condensed phases [8–13]. Photostability of nucleobases has been therefore an issue of a lively debate in recent years. Is DNA/RNA really so resistant to near-UV radiation? Additional long-lived excited states components have been determined in polynucleotide chains and double-stranded DNA [14–18], particularly related to the stacked-bases structure and the existence of excited state eximers [14,16,18]. Also quite recently, the mechanism of formation of mutated cyclobutane dimers between DNA pyrimidine bases after UV irradiation has been established [19–23]. As a matter of fact, the production of such lesions, considering that has as a prior step the formation of longer-lived singlet excited eximer states, is a strong proof of a more general hypothesis: that the increase in the lifetime of an excited state enhances the probability for a photoreaction, as it occurs for triplet states, even if the correlation of these magnitudes is not necessarily simple. One may then wonder if the photostability of the individual nucleobases can be after all extended to the DNA strand. It has been clearly established that the paths for ultrafast deactivation found in nucleobases are still open in the oligonucleotide strand. A large part of the initially absorbed energy will decay in an ultrafast manner by the same channels as those of the monomers, in particular from those conformations that remain basically unstacked or slightly stacked, but even – with slightly larger lifetimes – from bound excimer states [3,14–18,24,25]. In summary, the ultrafast decay processes may well predominate and most of the energy absorbed by the system will be rapidly dissipated in those systems in which the monomeric nucleobases provide efficient funnels for internal conversion to the ground state.

Independent of the degree of photostability of the biological compound, a striking evidence is put forward by experiment. The photochemical properties of many nucleobases tautomers and derivatives are radically different from those of the natural chromophores. Upon tautomerization or substitution on the nucleobase the ultrafast energy decay channels are severely affected [8,10,26,27]. In some cases, like for 7-substituted adenine derivatives, the femtosecond channel vanishes [8,10,28]; in other situations the picosecond lifetime increases in one or two orders of magnitude [8,26,27,29,30]. There are even more obvious cases, like 2-aminopurine, a close constitutional isomer of adenine (6-aminopurine) [31]. Whereas the latter efficiently quenches its fluorescence leading to very small quantum yields in water ($\phi_F \approx 10^{-4}$) [6,8], 2-aminopurine displays a strong emission ($\phi_F \approx 0.66$) [29,30], in such a way that the compound is commonly used to substitute adenine in DNA as a fluorescent probe to detect protein-induced local conformational changes [32–36]. Similar scenarios are obtained for protonated, oxidized, and reduced nucleobases [8], whereas substitution by methyl does not seem to strongly affect the photochemical outcome [8,10,11,26–28]. Even if it cannot be claimed that such type of behavior is universal for all possible modified nucleobases – and specific systems have been found to decay rapidly in solution [37] –, it is common enough to be understood as a pattern to have consequences in the biological context.

As mentioned above, a question to be raised is, therefore, why natural nucleobases display such characteristic photochemical feature, so convenient for a safe genetic behavior, and why this property can be perturbed by minor structural modifications. First we have to establish the mechanism for rapid energy dissipation in natural nucleobases. Former proposals to explain low quantum yields of fluorescence and excited-singlet-state deactivation in nucleobases by means of excited-state photoreactions or phototautomerisms were ruled out because of the absence of photoproducts and deuterium isotope effects in different solvents [8]. Modern theoretical photochemistry associates the efficiency of radiationless decay between different electronic states taking place in IC processes to the presence of crossings of different potential energy hypersurfaces (PEH) along a hyperline or seam of molecular structures, i.e., a N–2 dimensional subspace on the N-dimensional surfaces [38–42]. The points belonging to these subspaces, named conical intersections (CI) and representing energy degeneracies between two or more states, behave as energy funnels in whose proximity the probability for non-adiabatic, non-radiative, jumps is
high \cite{40,43,44}. The nature of the deactivation is certainly related with the strong vibronic coupling taking place between the interacting states when the energy gap decreases, in particular in situations such as conical intersections \cite{41}. The electronic excited state related to the transition with the largest absorption probability will in principle play the main role in the photochemistry of a given compound, together with the corresponding pathways leading to the lowest-excited state S$_1$, and from it to the conical intersections connecting such state with the ground state S$_0$ \cite{31}. Joint efforts of different groups of quantum chemists in recent years \cite{31,39,45–67} have clearly established that the ultrafast radiationless decays in natural nucleobases is directly linked to the presence of this type of (S$_0$/S$_1$)$_{CI}$ conical intersections with biradical character \cite{65–67}.

In previous studies we have proved that the existence of a low-energy CI between the initially populated allowed spectroscopic ππ$^*$ singlet (L$_a$ in Platt’s nomenclature or HOMO$\rightarrow$LUMO (HL)) state and the molecular ground state is an intrinsic property not only of the five natural nucleobases, but also of some close tautomers and derivatives, even those that do not decay non-radiatively in an efficient way \cite{31,48,49,52}. This implies that it is not merely the existence of a low-energy CI, but its accessibility which guarantees the efficiency of the decay \cite{31,39,49,52,54,56,64–68}. Our previous calculations determined that the five natural nucleobases display barrierless reaction paths from the initially excited 1(ππ$^*$ L$_a$) state – that carrying the largest intensity – at the starting Franck-Condon (FC) region toward the (S$_0$/S$_1$)$_{CI}$ \cite{31,48,49,52,68}. Any barrier that the system finds along
the most favorable relaxation profile will give rise to retardation in the reaction rate, even leading to well-defined minima from which the excited state may emit [31,68]. In the present paper we will carefully analyze the absence or presence of such energy barriers, and will relate it with the larger or smaller efficiency, respectively, for non-radiative energy transfer to the ground state. As carefully established recently [31,48,49,52], the only safe procedure to guarantee an accurate conclusion within the static model is the application of the Photochemical Reaction Path Approach. This is a theoretical strategy established by Bernadi, Olivucci, and Robb some years ago [69] based on mapping the path for the efficient transit of the energy in the excited states toward the accessible Robb some years ago [69] based on mapping the path for the efficient transit of the energy in the excited states toward the accessible minimum energy path (MEP), that is, a lowest-energy path connecting the initial and final reaction structures [31,46,48–52,60,61,70], here the Franck–Condon ground state geometry and the (S0/S1)CI. No other type of interpolation or reaction path determination can guarantee the absence or presence of energy barriers along a given profile, and connect the main relaxation path, not simply with the usually computed lowest-energy CI, but with the first accessible CI [52] which, as recently pointed out by Michl [71], is the actually relevant funnel for energy transfer. Determination of MEPs constitutes the basis for future theoretical studies on the dynamics of the system which may lead to ultimately predict the time evolution of the process. As quantum chemical method we will employ a highly accurate multiconfigurational perturbation procedure (CASSCF/CASPT2) [72–75], so far, together with the multireference configuration interaction (MRCI) methods [53–57,76,77], the only generally applicable procedure that has proved to be solid enough to compute this type of electronic structure problems.

In the present paper we will review the status of the relaxation dynamics of natural nucleobases (see Fig. 1) and a number of tautomers and derivatives (see Figs. 2 and 3) in relation to the presence or absence of energy barriers along the lowest-energy relaxation path. Focusing on the calculation of different MEPs along the \( ^1(\pi\pi^* \Sigma) \) state of the systems, we will demonstrate here by the first time that there is at least a barrierless deactivation path for the excited state energy which is present in the natural nucleobases and absent or hindered in other close molecules. This feature makes the modified nucleobases prospective mutational hot spots. The absence or presence of such barriers will be related to the larger or smaller, respectively, efficiency for non-radiative energy transfer to the ground state. By doing so we shall illustrate the importance of obtaining a quantum chemical description of PEHs based on the computation of state minima, barrier heights, conical intersections, and in particular, the proper calculation of MEPs connecting the singular points, in a way that enables us to carefully determine the most favorable paths for the energetic decay. Our conclusions establish solidly the prevalence of a model based on the conical intersection concept in modern photochemistry [31,38–40].

2. Theoretical background

2.1. Structure determinations: calculation of minimum energy paths

Optimizations of minima, PEH crossings, and MEPs have been performed at the CASSCF level of theory. MEPs have been built as steepest descendent paths in a procedure [70] which is based on a modification of the projected constrained optimization (PCO) algorithm of Anglada and Bofill [78] and follows the Müller–Brown approach [79]. Each step requires the minimization of the PEH on a hyperspherical cross section of the PEH 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 paper 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-7.0 package [80–82] and its technical description has been published elsewhere [70]. Regarding the conical intersection searches, they were performed using the restricted Lagrange multipliers technique as included in the MOLCAS package in which the lowest-energy point was obtained under the restriction of degeneracy between the two considered states. The conical intersection searches were carried out first at the CASSCF level, obtaining minimum energy crossing points (MECP). In order to locate the conical intersection at a higher level of theory, CASPT2 calculations around the CASSCF optimized structures were performed in order to find the lowest-energy structure for the highest state which had the smallest energy difference between the computed states. The CASPT2 scans were performed by computing a grid of points along distortions involving the coordinates with the smallest values for the CASSCF gradients in the region of the crossing.

2.2. Multiconfigurational calculations

Unless otherwise specified the structure determinations employed a \( \pi\pi^* \) valence active space for states of such nature, whereas the results including \( \pi\pi^* \) states added orbitals corresponding two additional lone-pair orbitals and electrons. Sometimes, even when it is technically feasible it is not always adequate to add orbitals that are almost doubly occupied along all the studied process, especially when the calculations include geometry optimizations. The reason is that inner orbitals deep in energy may easily rotate between the inactive and active spaces. As a general rule, orbitals with occupations close to two or zero should be excluded from the active space. It is not only that the convergence of the calculation is poorer, but also that undesired active orbitals enter into the active space. For instance, in adenine or 2-aminopurine to carry out geometry optimizations for \( \pi\pi^* \) states, an active space (9,10) was finally selected. By using this space, the three lone pair orbitals and the deepest occupied \( \pi \) orbital, have been excluded of the space. In general, the deeper orbitals are not strictly localized. Whereas the lowest occupied \( \pi \) orbital combines, depending on the case, contributions of the five-membered ring and the amino group, the two lowest lone pair orbital, well separated from the highest lone pair, combine contributions of the nitrogen on the five-membered ring with other from both nitrogen atoms in the six-membered ring. If these orbitals are not excluded from the calculations in geometry optimizations, \( \sigma \)-type orbitals enter into the active space and unbalance the geometry optimizations and minimum energy paths. These restrictions cannot be applied without checking constantly the obtained results with larger active spaces in order to analyze the energy and occupation of the deepest, excluded, orbitals. In any case, it is not only the number of selected orbitals in the active space which is important, but its nature. Along a reaction path the importance of an orbital for a distortion may increase, and therefore its occupation may also decrease, leaving other orbitals as the deepest ones, at it occurs with the orbital containing a large contribution of the amine group in the computation of the lowest CI in 2-aminopurine, when the out-of-plane \( \text{NH}_2 \) distortion becomes important.

The one-electron atomic basis sets ANO-S C,N,O,F[3s2p1d]/H[2s] and 6-31G(d,p) were used for pyrimidine and purine nucleobases, respectively. The extension of the 6-31G(d,p) to larger ANO-type basis was also tested for the purine and pyrimidine nucle-
obases in parallel studies. As it will be later detailed, whereas the
correct description of the reaction profiles was shown to require
the use of an ANO-type basis set in pyrimidines [48,68], the basic
features in purines were shown not to be extremely dependent on
the quality of the one-electron basis set [48,68,83].

It is crucial to emphasize that the CASSCF level of calculation
can provide relatively accurate geometries and properties, and ade-
quate reference wave functions, but not quantitative electronic
energies, simply because it includes just a small fraction of the elec-
tronic correlation energy. An additional account of such effects is
obtained in a general, reliable, and computationally efficient man-
ner by using second-order perturbation theory. At the obtained
CASSCF points and using the corresponding wave function as refer-
ce (typically a state-average CASSCF solution for several roots of
the symmetry), CASPT2 calculations using a standard zeroth-order
Hamiltonian were carried out on several singlet states in order
to include the necessary dynamical correlation effects. Spurious
intruder states effects were prevented with an imaginary level shift
of 0.2 au. The overall protocol is usually named CASPT2//CASSCF,
that is, geometry optimizations at the CASSCF level and point
energy calculations at the CASPT2 level, and has proved its accuracy
repeatedly [31,39,45–52,58–63,68,72–75,84–88]. No symmetry
restrictions were imposed during the calculations. In all cases the
MOLCAS-7 set of programs has been employed [80,81,83].

3. Photophysics of the natural nucleobases: ultrafast decays

The photochemistry of DNA and RNA nucleobases begins with
the absorption of energy by the allowed $1(\pi\pi^* L_2)$ singlet state,
computed vertically at 5.02 (U), 4.89 (T), 4.41 (C), 5.35 (A), and
4.93 (G) eV, with related oscillator strengths ranging from 0.053
for C to 0.436 for U, in near agreement with previous theoretical
and gas-phase experimental data, 5.1 (U), 4.8 (T), 4.6 eV (C), 5.2
(A), and 4.6 (G) eV [8,89–91]. In cytosine and guanine the allowed

![Fig. 3. Structure, labeling, name, and acronym used here for the purine DNA/RNA nucleobase tautomers, isomers, and derivatives studied in the present paper.](image)
Fig. 4. Computed CASPT2//CASSCF minimum energy paths (MEPs) along the allowed \(^1(\pi^*\pi_L)\) state for the natural pyrimidine nucleobases cytosine, thymine, and uracil (top to bottom) from the Franck–Condon region leading in a barrierless way to the conical intersection \((gs/\pi\pi^*)_CI\).

Fig. 5. Computed CASPT2//CASSCF minimum energy paths (MEPs) along the allowed \(^1(\pi^*\pi_L)\) state for the natural purine nucleobases adenine and guanine (top to bottom) from the Franck–Condon region leading in a barrierless way to the conical intersection \((gs/\pi\pi^*)_CI\).

The calculation of MEPS is a computationally expensive and complex task. As it was shown previously, the outcome of the MEPS calculated for uracil and thymine strongly relied on the level of theory employed [48]. The use of ANO-type instead of 6-31G(d,p) basis sets and the enlargement of the active space with three additional extra-valence orbitals was required to stabilize the CASSCF results and get a path leading toward the out-of-plane CI, and not to a planar minimum. A and G profiles did not display such a dependency (\((gs/\pi\pi^*)_CI\)). The geometrical distortion taking place during the process involves twisting of the corresponding double bond C5C6 or C2N3, with hydrogen atoms (or carbon in T and nitrogen in G) displaying a dihedral angle near 120° (see Fig. 4) at the CI, which lies adiabatically (from the ground-state minimum) at 3.9 (U), 4.0 (T), 3.6 (C), 4.1 (A), and 4.3 (G) eV. Most of the absorbed energy will decay non-radiatively to reach the \((gs/\pi\pi^*)_CI\) funnel by an ultrafast relaxation that we assign to the femtosecond component of the multi-exponential decay measured in molecular beams at 130 (U), 105 (T), 160 (C), 100 (A), and 148 (G) fs [10]. In a recent study on the guanine molecule we reported a preliminary on-the-fly \(ab\ initio\) reaction dynamics study based on classical trajectories along the \(^1(\pi^*\pi_L)\) state hypersurface in which it was shown that at least a number of trajectories reached the seam of \((gs/\pi\pi^*)_CI\) CIs in about 100 fs [52]. The total decay involves at least a second relaxation in the picosecond range. Even if getting quantitative reaction rates will have to wait until more exhaustive studies are performed, we can attribute the ultrafast decay of the excited nucleobases as an intrinsic molecular property based on the barrierless character of this main reaction path leading to the \((gs/\pi\pi^*)_CI\) states degeneration, as indicated recently by model reaction dynamics studies of Barbatti and Lischka [77].
The case of cytosine (C) is slightly more problematic [31,49,52]. Computed CASSCF MEPs for cytosine at all levels lead to a planar minimum located almost isoenergetic with the ethene-like CI, a situation which makes the profile toward such minimum competitive with that toward the CI. We therefore performed a linear interpolation in internal coordinates (LIIC) between the FC and (gs/ππ*)CI geometries (see Fig. 4), defining in this way a connected path which provides a higher bound for the reaction barriers along the studied profile. At the beginning of the LIIC a small barrier of 2.5 kcal mol\(^{-1}\) is obtained. Being a higher bound for the actual barrier, this is an indication that when dynamic correlation energy (at the CASPT2 or MRCI levels) could be included to determine the MEP points structures, the path for C will be also found barrierless toward the final CI. It might happen, however, that such small barrier is dependent of the environmental conditions. Indeed, a theoretical study by Blancafort and Migani [93] reported a decrease in the barrier in water, indicating that it is probably too small to affect the decay of the system. As shown below, we performed a similar effort in one of the tautomers to see the importance of the barriers and check the reliability of the proposed model.

In U, T, and A, a favorable surface crossing – and a nearby CI – takes place between the \(1(\pi\pi^*)L_a\) and an \(n\pi^*\) singlet excited state (the latter involving an oxygen, \(nO\pi^*\) for U and T, or nitrogen, \(nN\pi^*\) for A, electron lone-pair) along the lowest-energy MEP on \(1(\pi\pi^*L_a)\), because of the lower excitation energy of the \(n\pi^*\) state at the FC geometry [31,48–52,68]. Such feature indicates, as it was suggested previously [31,48–52,68,94] and it has been recently measured [13], that a second relaxation channel can be opened by an energy switch from the \(n\pi^*\) to the \(n\pi^*\) state. In C and G it has been also proved that there are other favorable connections between these states (\(\pi\pi^*\) and \(nO\pi^*\)) guaranteeing the transfer of the energy and therefore the presence of the second decay, whose connections with the ground state differ [68]. In A and G we proved the existence of low-energy CIs between the ground and the \(n\pi^*\) state (gs/\(n\pi^*)CI [31,49,52,68], related to other ring distortions, a feature also found for C [45,46,54,56,60,92,93]. In T and U, however, such CI seems to lie at much higher energies [59]. In any case, and as recent reaction dynamics calculations on adenine suggest [95], the \(n\pi^*\) state might be only intermediate in the ultrafast femtosecond deactivation toward the ethene- or methanamide-like (gs/\(\pi\pi^*)CI, whereas the picosecond relaxation would correspond to the final deactivation of the CI, to the ground state. Direct relaxation channels through the \(n\pi^*\) or higher \(n\pi^*\) states would probably be opened at much higher energies and involving longer lifetimes, and they might increase its importance in non-natural nucleobases, where the main relaxation channel toward (gs/\(\pi\pi^*)CI is not so favorable.

4. Photophysics of the non-natural nucleobases tautomers and derivatives: slower decays

4.1. Photophysics of modified pyrimidine nucleobases

Once the main conclusion on the photophysics of natural nucleobases has been established, that is, the existence in these systems of a barrierless path from the initially populated state toward a CI with the ground state explaining the presence of an ultrafast energy decay quenching most of the fluorescence, we can perform the same type of study for several of their tautomers and derivatives. It must be clearly emphasized that all MEPs have been computed and displayed in the figures until finding a minimum or a conical intersection, independently of the number of steps required to reach that region. Two exceptions will be found in which we have completed the scheme by representing a barrier separating the minimum and the CI. We can start by substituting positions C5 and N1 in uracil by two species with opposite characteristics, a methyl group and a fluorine atom, which are an electron donor and acceptor, respectively. The two positions of the ring are of great interest. C5 belongs to the double C=O bond involved in the out-of-plane ethene-like CI, (gs/\(\pi\pi^*)CI, whereas N1 is the atom linking to the sugar moiety in the nucleotide and DNA [3].

As displayed in Fig. 6, the fate of the \(1(\pi\pi^*L_a)\) state in 5-fluorouracil (5FU) and 1-fluorouracil (1FU) along the computed MEP is significantly different with respect to U itself. In all cases we confirm here the existence of the CI (gs/\(\pi\pi^*)CI as an intrinsic property of the systems, but its relative accessibility varies. 1FU displays a barrierless MEP at the CASSCF level, but a barrier (point 5 of the MEP) of 4.8 kcal mol\(^{-1}\) shows up when the energy profile is corrected at the CASPT2 level. One can expect that such barrier decreases somewhat in a purely CASPT2 MEP. Regarding

![Image](Image)

Fig. 6. Computed CASPT2//CASSCF minimum energy paths (MEPs) along the allowed \(1(\pi\pi^*L_a)\) state for several uracil derivatives (5- and 1-fluorouracil and 1-methyluracil) from the Franck–Condon region. Only methyl-substituted compounds lead in a barrierless way to the conical intersection (gs/\(\pi\pi^*)CI.

**Fig. 7.** Computed CASPT2/CASSCF minimum energy paths (MEPs) along the allowed $^1(\pi^*\pi_1)$ state for one thymine derivative and one tautomer (6-methylthymine and 4-hydroxythymine) from the Franck–Condon region. Only methyl-substituted compounds lead in a barrierless way to the conical intersection (gs/$^1\pi\pi^*$). 5FU, the most significant change with respect to U is the presence of a pronounced energy barrier of 13.7 kcal mol$^{-1}$ hampering an easy access to the final CI at the end of the MEP. Such CI has been computed independently of the MEP as the minimum energy crossing point connecting the ground and lowest $\pi\pi^*$ states. Contrary to substitutions with fluorine, the 1-methyluracil (1MeU) and 5-Methyluracil (5MeU) (actually, thymine) derivatives both display barrierless paths, as in U itself. In fact the lack of barrier in the path computed for 1MeU, in which a substitution of a NH with a NC bond has occurred as in the biological compound, confirms, as measured [7], the ultrafast decays obtained also for nucleotides. The computed MEPs nicely rationalize the available experimental evidence. Almost identical decay times have been measured in water for U (96 fs) and 1MeU (93 fs), and similarly for T (195 fs), the 5-methyl derivative. On the contrary, 5FU displays a seven-fold increase in the state lifetime (694 fs) [26,27]. It is clear that modifications in the accessibility of the ethene-like CI reflects in a distinct photophysics of the nucleobase derivative. Substitution in position N1 does not lead to so drastic changes, as in 1-cyclohexyluracil, which displays decays in water at 120 fs and 13 ps. [96]. This could be expected, considering that nucleosides and nucleotides have similar relaxation behavior as nucleobase monomers [8].

As shown in Fig. 7, methyl substitution in position C6 – also a carbon atom belonging to the $C=C$ bond twisted in the ethene-like CI – does not introduce relevant consequences in the expected ultrafast decay of the nucleobases. The MEP computed for 6-methylthymine (6MeT, could also named 5,6-dimethyluracil) also leads in a barrierless way from FC to the ethene-like CI. Indeed, a lifetime of 97 fs has been reported for 6-methyluracil in water [26]. We have also computed the corresponding MEP for 5MeC (not shown here), and it leads to a low-energy minimum, a situation similar to that found in the parent natural C molecule. However, the height of the barrier toward the CI in C almost vanishes in water [93], whereas in 5MeC it can be expected slightly larger, considering the slight increase of the lifetime observed for the latter in the aqueous medium [8]. More important, the decay time decreases in the corresponding nucleotide, confirming the photostability found in other alkylated nucleobases. Another example of perturbation in the ultrafast decay was reported by Kistler and Matsika [54], who computed a less accessible CI in the fluorescent analogue 5-methyl-2-pyrimidinone (5MeC without the amino group). Finally, we have calculated the MEP for one of the enol tautomers of thymine, 4OHT (see Fig. 8), which takes part in one of the non-Watson-Crick (WC) AT pairs [3], as shall be commented later. The MEP on 4-hydroxythymine (4OHT) leads to a planar minimum at the end of the MEP (see Fig. 7), at much higher energies than those of the ethene-like CI in other similar systems. It must be understood that the displayed PEHs do not correspond to a truncated MEP, but simply that the minimum has been found in few steps. The presence of such minimum guarantees also the existence of a barrier hindering the access to the final CI (which still exists, see below), and therefore suggesting a slower decay.

4.2. Photochemistry of modified purine nucleobases

Regarding purine nucleobases tautomers and derivatives, Figs. 9 and 10 compile computed MEPs for three adenine and three guanine analogs. 7H-Tautomers (see Fig. 3), both for adenine and guanine [31,49,52] have MEPs along the $^1(\pi^*\pi_1)$ surface freely leading from FC toward high-lying planar minima. As shown in previous works [31,49,52], this behavior does not imply that the CI (gs/$^1\pi\pi^*$) does not exist at low-energies; it does, but its accessibility will be restricted by the presence of the intermediate minimum. From the experimental viewpoint, no femtosecond decay has been reported for 7-derivatives in molecular beams, unlike 9-derivatives such as 9-methyladenine [8,10,28], and larger lifetimes have been obtained in condensed phases for the former. In particular 7-methyladenine displays a 20-fold increase with respect to 9-methyladenine [8]. It is worth pointing out that

![Adenine-Thymine, Watson-Crick pair](image)

![Inosinoadenine-4-Hydroxythymine, non-Watson-Crick pair](image)
The nucleobase is attached to the DNA strand through position N9, and alkylation at this spot may closely represent the situation found in nucleosides and nucleotides. Substitution in position N7, on the other hand, makes the system prone to be perturbed by the formation of intermolecular hydrogen bonds with the amine group [49]. Another tautomer, 9H-iminoadenine (IA), has a similar behavior as 7H-adenine (7HA), with a MEP yielding a planar minimum and therefore a barrier in the relaxation path. As observed in Fig. 8, 4OHT and IA form one of the non-Watson-Crick pairs obtained after a double hydrogen transfer from the Watson-Crick pair constituted by A and T [3]. Whereas the two latter and natural monomers have been shown to be highly photostable and efficient decay channels have been also found for canonical WC pairs [97,98], those two systems forming the mutated non-Watson-Crick pair are clearly not. This is a property which favors the existence of the natural WC conformation. The out-of-plane deformation leading to the (gs/ππ*) state does not strongly affect the structure of the pyrimidine and purine WC pair found in double strand DNA. This non-adiabatic relaxation process can be therefore considered also available in DNA itself, not only in the monomers, unlike those mechanisms related with ππ* type states, strongly perturbed in protic environments and in WC pairs because of the presence of hydrogen bonds [26,27].

Regarding 9H-2-aminopurine (2AP), a constitutional isomer of A (formally 6-aminopurine), the computed CASPT2 MEP displayed in Fig. 9 helps to understand its strikingly different photochemical behavior as compared with natural A. The 1(ππ* Lα) state is the lowest one at all geometries, and carries most of the intensity, while the 1(ππ* Lb) state is placed energetically too high to be of primary photophysical importance [31]. When populated, the 1(ππ* Lα) state evolves from the FC structure to an almost planar minimum, 1(ππ* Lα)min. From that region, the system has to release
the energy, either by emitting radiation or by reaching a higher-lying conical intersection, \((gs/ππ^*L_a)CI\), surmounting a barrier of 5.0 kcal mol\(^{-1}\) through an intermediate transition state. The computed adiabatic transition relating both ground and excited state minima, 3.89 eV, matches quite well the experimental measurement for absorption and emission band origins, 4.01 eV \([99,100]\). The measured fluorescence quantum yield increases from 0.02 to 0.66 when going from non-polar to polar solvents \([29,32]\), indicating that the barrier to access the CI will be enlarged with the augmented solvent polarity.

Guanine enol-tautomers show the same type of behavior as that found in enol-pyrimidine nucleobases. The MEP computed for 9H-2-amino-6-hydroxypurine (9H6OHG) also leads to a planar minimum at high energies (as it happens for the 7H-tautomer, not shown), and therefore the ultrafast fs decay observed in natural guanine will not take place (see Fig. 9) \([52]\). It is exactly what the detailed analysis of the IR-UV gas-phase absorption spectrum indicates \([57]\). Guanine enol-tautomers, more stable in the gas phase than the natural oxo-tautomer G, show structured absorption pointing out to the presence of a clearly defined minimum \([57]\). For natural G, on the other hand, a band origin could not be assigned, something that we explain by the absence of a well-defined minimum energy structure along the main, barrierless relaxation path \([52,68]\). The same behavior as in G or A was encountered for the computed MEP on \((ππ^*L_a)\) for 9-methylguanine (9MeG). Once more the profile was obtained barrierless from FC to the methanamine-like CI \((gs/ππ^*L_a)\), pointing out both that methyl substitution does not strongly modify the photochemical outcome of the nucleobases and that position N9 in purine bases is less sensitive to substitution, as supported by the experimental decay times in molecular beams and in a solvent, 110 and 220 fs, respectively, for 9-methyladenine \([10]\). This is also an indication that nucleosides and nucleotides share similar mechanisms as the monomers \([8]\). On the other hand, more drastic substitutions can be expected to further perturb the photochemical outcome. This is the case of the propanodeoxyguanosine, in which an exocyclic ring prevents the out-of-plane ring deformation and therefore the ultrafast channel becomes deactivated, yielding a strong fluorescence \([101]\).

As already mentioned for pyrimidine systems, in purine nucleobases there are other presumably slower but available relaxation channels whose importance as energy dissipation funnels will have to be evaluated in the future. They certainly do not directly con-

5. **Final remarks**

In the present contribution, we have shown how the new concepts and tools of modern theoretical, quantum–chemical, photochemistry are able to shed light on important chemical and biological issues. In particular, we emphasize the importance of the conical intersection concept, understood as the seam of degenerate structures belonging to a surface crossing of electronic states responsible for internal conversion processes releasing the energy acquired by the molecule \([38–40]\). Also, we state that timescales in the energy relaxation processes in excited molecular systems rely on the accessibility of the seam of conical intersections, that is, on the presence or absence of energy barriers along the reaction path. Computation of minimum energy paths and related transition states and conical intersections – using accurate quantum chemical methods such as the CASSCF/CASPT2 approach – along the main relaxation pathways is the only procedure which can guarantee the proper calculation of the relevant energy barriers along the lowest-energy profile \([31,49,52,68]\).

With these tools at hand we have studied the photochemical behavior of DNA/RNA natural and non-natural nucleobases. We consider that the response of the nucleobases as chromophores upon absorption of near-UV light can constitute one mechanism of natural selection of DNA/RNA composition in a quest for obtaining systems resistant to external perturbation. In particular, it is concluded that the basic photostability displayed by the five natural nucleobases upon UV-irradiation \([4–13]\) can be explained by the expedient accessibility of the energy relaxation path on the allowed excited singlet state leading to a channel for favorable internal conversion toward the ground state. This property is accounted for through the calculation of a barrierless MEP connecting the initial Franck–Condon region with a conical intersection between the excited singlet \(ππ^*\) and the ground state, where both are degenerated in energy. The extrapolation of the photostability of the monomeric nucleobases to the genetic material is not straightforward. Long-lived decay channels, related mainly to the formation of stacked excimers, have been described in oligonucleotides, buy still a large part of the initial absorbed energy dissipates by the monomer funnel in slightly stacked bases \([18]\), and even the decay of the excimers probably rely on the position of the \((gs/ππ^*L_a)CI\) of the individual nucleobase \([23]\). The mechanism of energy release described here in the natural nucleobases does still play an important role in the relative photostability of the genetic material, as it surely played in the prebiotic material, probably guaranteeing the persistence of the natural nucleobases in DNA/RNA.

Modification of the nucleobases by means of tautomeration or substitution increases the excited states lifetimes, sometimes even yielding intense emission \([8–13,26–37]\). These phenomena imply absence or restricted access to efficient internal conversion funnels for energy deactivation. Indeed, MEPs computed for different non-natural modified nucleobases, exactly at the same level that those calculated for the natural systems, yield energy barriers, and therefore local minima, along the relaxation pathways, preventing or hindering the access to the key conical intersection, \((gs/ππ^*L_a)CI\). Derivatives obtained by fluorination, tautomerization or constitutional isomerization of the nucleobases have been shown to give rise to compounds which do not release their absorbed energy so efficiently \([8–13,26–37]\) and therefore are more prone to photoreact and, in cases, to yield back natural nucleobases after enzymatic repair \([3]\). Longer state lifetimes in non-natural nucleobases mean
a larger probability to undergo mutations, not to participate in the DNA replication process, and therefore, in the long run, not to be included in the genetic code. It is possible to envisage that in early stages of life on Earth, with larger doses of UV radiation, such photochemical mechanism contributed to the natural selection of the photostable nucleobases to form the nucleic acids.

As shocking exceptions to this trend, selectively methylated (alkylated) nucleobases, especially those substituted in positions not perturbing the WC structure, display different behavior than other modified nucleobases. Our computed MEPs for the methylated species are also barrierless toward the low-lying CI, suggesting also that they are as photostable as the natural chromophores. This is also confirmed by experiment, in which some methylated pyrimidine and purine nucleic acid bases have shown basically the same ultrafast femtosecond lifetimes as those of the parent molecules °−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°−°-