The final publication is available at

https://doi.org/10.1039/d0cp04557h

Additional Information

Experimental and Theoretical Studies on Thymine Photodimerization Mediated by Oxidatively Generated DNA Lesions and Epigenetic Intermediates

Mauricio Lineros-Rosa,[a] Antonio Francés-Monerris,[b,c] Antonio Monari,*[b] Miguel Angel Miranda,*[a] and Virginie Lhiaubet-Vallet*[a]

Interaction of nucleic acids with light is a scientific question of paramount relevance not only in the understanding of life functioning and evolution, but also in the insurgence of diseases such as malignant skin cancer and in the development of biomarkers and novel light-assisted therapeutic tools. This work shows that the UVA portion of sunlight, not absorbed by canonical DNA nucleobases, can be absorbed by 5-formyluracil (ForU) and 5-formylcytosine (ForC), two ubiquitous oxidatively generated lesions and epigenetic intermediates present in living beings in natural conditions. We measure the strong propensity of these molecules to populate triplet excited states able to transfer the excitation energy to thymine-thymine dyads, inducing the formation of cyclobutane pyrimidine dimers (CPDs). By using steady-state and transient absorption spectroscopy, NMR, HPLC, and theoretical calculations, we quantify the differences in the triplet-triplet energy transfer mediated by ForU and ForC, revealing that the former is much more efficient in delivering the excitation energy and producing the CPD photoproduct. Although significantly slower than ForU, ForC is also able to harm DNA nucleobases and therefore this process has to be taken into account as a viable photosensitization mechanism. The present findings evidence a rich photochemistry crucial to understand DNA damage photobehavior.

Introduction

Nucleic acids are constantly exposed to endogenous and exogenous agents that can modify their chemical structure and therefore compromise their normal biological function.[1,2] Among such agents, solar light and especially UVB radiation (~290-320 nm) at high doses represents a major threat to public health. As a matter of fact, ca. 5-10% of the total UV radiation that reaches Earth's surface falls in this wavelength range[3] and DNA nucleobases directly absorb UVB wavelengths, triggering complex photochemical pathways leading to the appearance of DNA lesions.[4–11] Although DNA nucleobases are known for their high photostability,[4–8] harmful DNA modifications such as cyclobutane pyrimidine dimers (CPDs) and 6-4 photoproducts (64-PP) are constantly produced and may accumulate in the skin, inducing mutation and carcinogenesis.[12,13] The unprotected exposure to UV light is nowadays recognized as one of the main causes of malignant skin cancer such as melanoma.[14–16]

The toxicity of DNA photolesions has been extensively studied: whereas CPDs are the most abundant photoproducts and are highly resistant to repair, 64-PP are less common, better repaired, but extremely mutagenic.[17] The high repair resistance of CPDs also results in their accumulation at the genome level. Furthermore, the deficiency of photolesions nucleotide excision repair (NER) enzymes is the main cause of highly debilitating pathologies, such as *xeroderma pigmentosum*, and of the exponential increase of skin cancer predisposition.

The rest of UV solar radiation, which is not filtered by the ozone layer (~90%), belongs to the less energetic UVA wavelengths (~320-380 nm).[3] In contrast to UVB photons, UVA light is weakly absorbed by DNA nucleobases (ε(UVA) of ca. 10 for a dA$_{20}$.dT$_{20}$ duplex).[18] In this region, the presence of photosensitizers, either exogenous or endogenous, located spatially close to DNA, modifies the damage distribution opening the way to indirect photochemical reactions.[19,20] These usually involve the population of triplet states of the chromophores by means of intersystem crossing (ISC) followed by several photochemical routes traditionally categorized in three classes: (*i*) triplet-triplet energy transfer (TTET) to DNA, (*ii*) activation of molecular oxygen to its singlet excited state, and (*iii*) photoinduced electron transfer processes usually involving guanine. Other mechanisms such as photodissociation or hydrogen abstraction have also been reported.[4,21] While electron transfer and singlet oxygen activation usually produce oxidatively generated DNA lesions, triplet-triplet energy transfer has been shown, both computationally and experimentally, to induce thymine dimerization.[12,22,23] Here we focus on the first class of photosensitization, i.e. the triplet-triplet energy transfer from the chromophore to DNA.

A wide range of external photosensitizers have been studied in the last years,[12,24–27] a representative example being the paradigmatic case of benzophenone.[28–30] Nevertheless, it has been recently proposed that DNA lesions, such as 6-4PP,[31] can themselves act as photosensitizers, since the modifications of their chemical structure may alter their optical and photophysical properties. It was found that the presence of the pyrimidone moiety leads to absorption in the UVA range and induces CPD production after ISC and subsequent triplet-triplet energy transfer. The possibility of triplet transfer was also confirmed using molecular modeling and simulations.[32]

Recently, the role of oxidatively generated lesions, namely 5-formyluracil (ForU)[33] and 5-formylcytosine (ForC)[34,35] (see Scheme 1), as DNA photosensitizers has been studied through experiments[36] and theoretical calculations.[37–39] It has indeed been recognized, both in model systems and in DNA, that the the presence of ForU in solution is correlated with a significant

increase in the yield of CPD damages.[36] Furthermore, molecular modeling and simulations have shown that ForU is prone to a facile ISC, leading to an energetically favorable TTET to thymine and that those conditions are maintained in the B-DNA environment.[37] Subsequently, it was shown that ISC is also possible in the case of ForC, even if non-adiabatic molecular dynamics have pointed out a less-favorable phenomenon as compared to ForU.[39]

The behavior of ForC in comparison with ForU is indeed extremely intriguing, not only from a photophysical perspective, but also due to its biological relevance.

Interestingly, ForU can be obtained both from endogenous and exogenous processes. Indeed, it is part of the lesions formed under γ-radiation and has also been observed in significant yields as a result of UVA irradiation of DNA in the presence of type I photosensitizers.[40–43] Moreover, ForU is produced as an intermediate of thymine oxidation catalyzed, in low yield, by ten eleven translocation (Tet) enzymes.[44] ForC can also be considered an oxidatively generated lesion obtained by high doses of γ-rays,[45] but also an epigenetic intermediate in the demethylation process of 5-methylcytosine (5MetC).[46] Indeed, the methylation of cytosine at 5 position, occurring in non-coding DNA regions with a high density of guanine and cytosine, the so-called CpG islands, results in the silencing of the specific gene expression. In eukaryotic cells, the regulation of the gene expression as a response to external stress is controlled by DNA methyltransferases (DNMT) that provide an oxidative route to demethylation involving the oxidation of 5MetC to ForC, and the subsequent participation of the base excision repair (BER) machinery. Deregulation in gene expression is found in many diseases including cancers, therefore, epigenetics intermediates are also recognized as most valuable biomarkers. The fact that an epigenetic intermediate, ForC, can potentially induce DNA photolesions, clearly opens fundamental questions concerning the subtle interplay between epigenetic phenomena and DNA photolesions, and hence can contribute to solve the molecular grounds of the problems related to the cellular tumorigenesis and the protective mechanisms that cells may have developed as a result of evolutionary pressure. Fortunately, ForU and ForC are efficiently repaired by glycosylases, and thus an estimated frequency of these lesions in cellular DNA is expected to be lower than a few modifications per 10^6 bases.

In this article, we rationalize the propensity of ForC to photoinduce thymine dimerization in model systems, also comparing the process with the one operative in the case of ForU. The crucial differences between the two chromophores and their effects on thymine dimerization are also duly analyzed. Indeed, the description of these processes at a molecular and electronic level is of utmost importance to comprehend the naturally occurring mechanisms of DNA damage.

Results and discussion

Triplet excited states of ForU and ForC

The photochemistry of ForU and ForC (Scheme 1) was tackled in aqueous media at room temperature in order to evaluate their potential to act as DNA photosensitizers. In this context, triplet excited states have been well-established as key intermediates in the photosensitized generation of damages, through Type I/II processes or triplet-triplet energy transfer (TETT).[12,47] Here, we will use the thymine dimer formation as an analytical tool to compare the harmfulness of ForU and ForC triplet excited states. Indeed, the photosensitization of cyclobutane thymine dimer (Thy<>Thy, Scheme 1) is a clean reaction that occurs by TTET from the photosensitizer excited state to the pyrimidine base. To be efficient, this process requires a photosensitizer with a high triplet energy (higher than that of thymine), a high ISC quantum yield to allow significant population of the triplet manifold, and a triplet lifetime long enough to maximize the likelihood of the energy transfer process. Hence, monitoring Thy<>Thy formation provides valuable information on the overall strength of photosensitizing compounds.

The first important property of an efficient DNA photosensitizer is absorption of light in a wavelength region where nucleobases do not (or barely) absorb, to assure the exploitation of the maximum number of incoming photons.
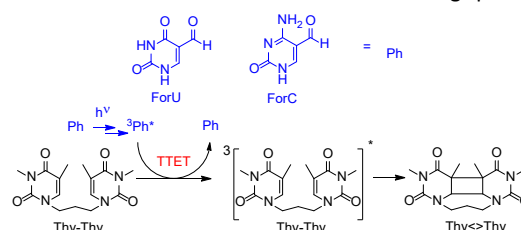
Scheme 1. Structures of both photosensitizers (Ph) under study (ForU and ForC) and the photosensitized [2+2] photocycloaddition of Thy-Thy to afford Thy<>Thy

As shown in Figure S3, in MeCN:H$_2$O (1:1, v:v), ForC and ForU share similar absorption maxima (λ_{max}) at 276-280 nm. Both compounds exhibit a band that reaches the UVA region allowing their selective excitation in the presence of canonical DNA bases.

Nonetheless, the uracil derivative has a larger band, which broadens the spectral range available for its excitation. The broader absorption band for ForU can be correlated with its larger flexibility, and hence to the coupling between electronic and vibrational degrees of freedom. Indeed, as shown by the analysis of the computed vibrational frequencies, ForU has four low frequency, large amplitude vibrations from 23 to 66 cm^{-1}, which are absent in the case of ForC. None of the compounds show fluorescence emission, pointing toward an efficient deactivation of the singlet excited state by nonradiative pathways such as internal conversion or ISC. By contrast, phosphorescence was detected in ethanol glass at 77K with a maximum at 445 and 425 nm for ForU and ForC, respectively (Figure S4). This emission informs not only on the occurrence of ISC, which populates the triplet excited state, but also on the triplet energy (E$_T$) of the uracil and cytosine derivatives determined at ca. 3.67 and 3.94 eV, respectively.[39] The triplet excited state behavior at room temperature was monitored using laser flash photolysis, since this technique allows the detection and characterization of transient species absorbing between 300 and 700 nm and whose lifetimes are

typically in the microsecond timescale. Thus, nitrogen bubbled solutions in H_2O:MeCN (1:1) of the two formyl derivatives were excited at 266 nm using the 4th harmonic of a Nd:YAG laser. The transient spectrum of ForU (Figure 1) showed a single broad band with maximum at 440-460 nm that completely disappeared after 40 μs without giving rise to any other species. This transient signal was assigned to the triplet-triplet absorption by comparison with our previously published data in PBS, but the spectrum in aqueous acetonitrile was cleaner and much less noisy.[36] The decay monitored at 440 nm can be fitted with a monoexponential function, f(t) = A exp (-t/τ), yielding a lifetime τ of ca. 12 μs (Figure 1, inset) that is much longer than the one found in PBS.

The cytosine derivative, ForC, also displayed a signal centered at 440 nm together with a shoulder at 560 nm, both bands decreasing with the same kinetics (Figure 2); this transient signal was assigned to the triplet excited state of ForC.[38]

Concerning the kinetics, a biexponential behavior was found for the 440 and 560 nm decays, which were adjusted using the equation f(t) = A_1 exp(-t/τ_1) + A_2 exp(-t/τ_2). The obtained lifetimes are τ_1= 0.07 μs (A_1= 0.25) and τ_2= 2.3 μs (A_2= 0.75) (Figure 2, inset). Both characteristic lifetimes are significantly shorter than that of [3]ForU*, and thus, the dynamic range for [3]ForC* intermolecular quenching is reduced, especially in the case of τ_1. Thus, bimolecular rate constants for [3]ForC* quenching were further determined using the changes of τ_2, only.

These intrinsic differences in the triplet lifetimes correlate very well with the nature of the emissive triplet states predicted by previous calculations for the two monomers. The emissive triplet states have been identified as [3]n,π* and [3]π,π*, for ForU
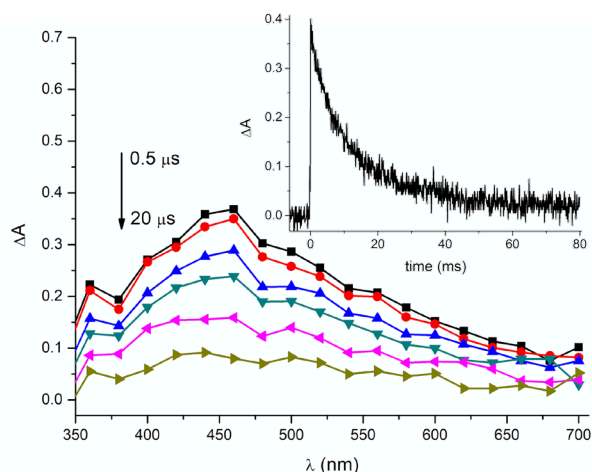

Figure 1. Transient absorption spectra of 5-formyluracil in MeCN:H_2O (1:1, v:v) under N_2 at different times after the 266 nm laser pulse. Inset: decay of ForU monitored at 440 nm.

and ForC respectively, on the basis of their vertical emissions.[39] Thus, the [3]n,π* state of ForU has a longer lifetime given the lowest transition dipole moment with the singlet ground state and hence a smaller oscillator strength, resulting, following Einstein thumb rule, in a slower radiative decay as compared to the [3]π,π* of ForC.[39] The above results establish that, upon excitation, both ForU and ForC populate triplet excited states with energies higher than that of thymine and with a lifetime

in the microsecond timescale, hence they can potentially induce energy transfer. Thus, the next step was to evaluate the ability of these excited states to photosensitize CPDs formation by means of time-resolved and steady-state experiments.

Triplet-triplet energy transfer rates to Thy-Thy measured by laser flash photolysis

First, laser flash photolysis was employed to determine the rate of the triplet-triplet energy transfer from [3]ForU* and [3]ForC* to Thy-Thy (Scheme 1), used as a model of adjacent nucleobases. As shown in Figure 3 (and Figure S5), the addition of increasing amounts of Thy-Thy resulted in the shortening of the triplet excited state decay kinetics of the formyl derivatives, together with a decrease of the end-of-pulse signal intensity, due to the filter effect produced by the thymine chromophore at the excitation wavelength. Stern-Volmer plot representation of τ_0/τ as a function of the quencher concentration, reported as inlay, provided the bimolecular TTET rate constant (k_q) for both sensitizers (see Material and Methods in ESI for more details). A value of 1.3 x 10^9 M^{-1} s^{-1} was obtained for the [3]ForU* quenching, while the process was almost three times slower for [3]ForC* with a k_q of ca. 4.6 x 10^8 M^{-1} s^{-1}. Taking into account that the model system has two thymine units, the obtained values are in agreement with those previously given in the literature for TTET processes involving thymine as energy acceptor using other photosensitizers.[12,48]
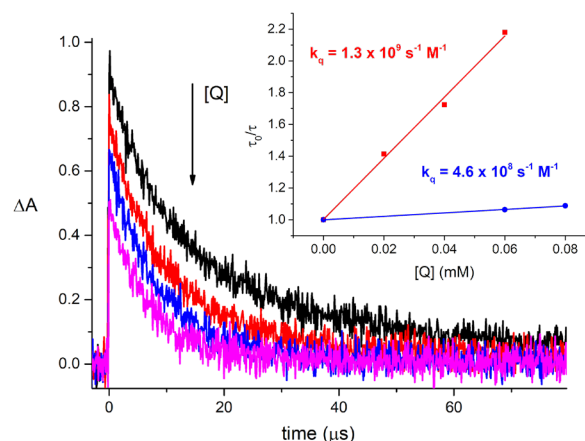

Figure 2. Transient absorption spectra of 5-formylcytosine in MeCN:H_2O (1:1, v:v) under N_2 at different times after the 266 nm laser pulse. Inset: decay of ForC monitored at 440 nm.

Curiously, an inverted order for the rate constants was expected based on the triplet energies, higher for ForC than for ForU, and on the Sandros´ equation.[49] This equation establishes that, for triplet-triplet energy transfer processes, the larger the energy difference between the donor (ie. formyl derivatives) and the acceptor (ie. the Thy-Thy dyad), the higher the rate constant. The data obtained here pointed toward the occurrence of a more complex process than a typical energy transfer, for instance the formation of triplet exciplex.
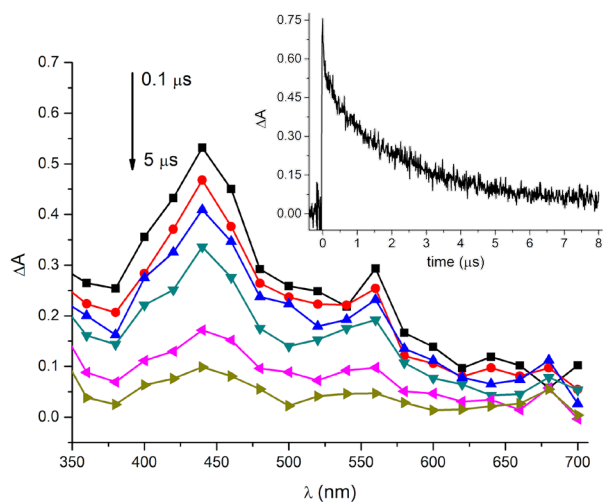
Figure 3. Decays of ForU in deaerated MeCN:H$_2$O (1:1,v:v) at 440 nm after the laser pulse at 266 nm in the presence of different Thy-Thy concentrations (from 0 to 0.08 mM). Inset: Stern-Volmer quenching of ForC (blue) and ForU (red) by Thy-Thy.

Cyclobutane pyrimidine dimer production monitored by NMR and HPLC

Next, steady-state photolysis was performed on deaerated MeCN:H$_2$O (1:1, v:v) solutions of Thy-Thy in the presence of ForU or ForC and using monochromatic light at λ_{exc} = 310 nm. All samples contained the same concentration of Thy-Thy (3.5 mM), ForC concentration was fixed at 1 mM, while in the case of ForU, the photosensitizer concentration was adjusted in order to obtain the same absorbance at λ_{exc} as for ForC.

The course of the photoreaction was followed by ^{1}H NMR and HPLC (Figures 4 and S6-S8). In both cases, a clean process was observed with the formation of the *cis-syn* cyclobutane dimer of Thy-Thy (Thy<>Thy) as the only photoproduct. This assignment was based on the comparison with the HPLC elution time and with the NMR chemical shifts of synthetized Thy<>Thy (see Figure S6 and S8, up).

The NMR spectra of irradiated samples gave relevant information on the course of the photoreaction and on the nature of the photoproducts. They show that the [2+2] photocycloaddition leading to Thy<>Thy formation results in the saturation of the C5-C6 double bond and induces characteristic changes in the chemical shifts (δ) of the protons belonging to the nucleobase and to the trimethylene bridge.
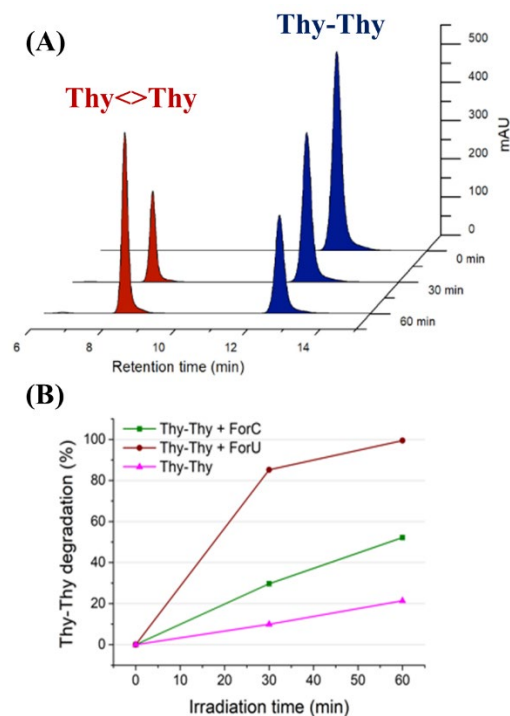

Figure 4. (A) HPLC chromatograms registered at 240 nm of ForC:Thy-Thy (1 mM: 3.5 mM) in H$_2$O:MeCN (1:1, v:v) irradiated from 0 to 60 min at 310 nm, (B) Time-dependent photodegradation of Thy-Thy, in the presence of ForU (red), ForC (green) and alone (pink) (λ_{irr} = 310 nm).

The most pronounced change was observed for the H6 proton that shifted from 7.90 to upper fields at ca. 4.48 ppm (Figure S6). Signals of the methyl group at N3 also experienced shielding passing from 3.77 to 3.51 ppm, while those of the C5 methyl moved from 2.38 to 2.00 ppm. The ratio of the integrals of these signals (taken as pairs) can be used to evaluate the reaction course. In the case of ForC irradiation, it was observed that after 1h, 50% of Thy-Thy was decomposed to yield Thy<>Thy as the sole product (Figure S7, up). Interestingly, when using ForU as a photosensitizer, the initial dyad is completely consumed under the same conditions (Figure S7, bottom).

The HPLC analysis provided similar results. Figure 4A shows the chromatograms registered for different irradiation times of ForC in presence of Thy-Thy. The peak corresponding to the Thy-Thy signal, that is eluting at 12 min, decreases concomitantly with the appearance and the increase of a new peak corresponding to Thy<>Thy (8 min). A qualitatively analogous behavior was observed for the irradiation of the mixture ForU:Thy-Thy (Figure S8, up). A control experiment performed with Thy-Thy alone irradiated at λ_{exc}= 310 nm (Figure S8, bottom) was also carried out and confirmed the lower yields for the direct CPD formation in the absence of photosensitizer. As expected from the difference in the bimolecular rate constants, Thy-Thy was more efficiently consumed when ForU is acting as photosensitizer. This is clearly illustrated in Figure 4B showing that the complete consumption of Thy-Thy was achieved after 1h of irradiation in presence of ForU, while only 50% of Thy-Thy is depleted in the

presence of ForC, and 20% without any sensitizer. However, as recently reported for 6-4PP,[50] these numbers can vary in the real system, where ForU or ForC is part of the double helix, depending on the sequence and the distance between the intrinsic photosensitizer and the Thy-Thy site.[51,52] Studies using synthetic double stranded oligonucleotides are indeed necessary to evaluate the effect of photosensitizer/dithymine site ratio decrease by comparison with the concentration used for the intermolecular TTET with our model system.

Altogether, these experimental results showed that in spite of its higher triplet excited state energy, ForC is a less efficient Thy<>Thy photosensitizer than ForU. Thus, in order to shed more light on the molecular aspects responsible for this difference, molecular modelling was performed for a mixture of each formyl derivative and thymine.

Mechanistic aspects revealed by molecular modeling

Several factors explain the different yields of triplet-triplet energy transfer and CPD formation induced by ForU and ForC. First, the population of the triplet excited state is slower for ForC than for ForU, as shown by the respective time constants, (τ_{ISC} = 1940.5 and 3559.4 fs, respectively) determined by non-adiabatic molecular dynamics.[39] Second, the triplet state of ForC decays faster, as evidenced by the time-resolved spectroscopic data shown in Figures 1 and 2, in coherence with the nature of the lowest triplet state ($^3n,\pi^*$ and $^3\pi,\pi^*$ for ForU and ForC, respectively).[39] Finally, the potential energy landscape of the electronic states involved in the photoprocess are different for ForU-Thy and ForC-Thy systems, as shown in Figure 5.

Note that Figure 5 provides an overview of the energetic level of the most important electronic states calculated for relevant points of the potential energy surfaces, namely Franck-Condon, i.e. the ground state equilibrium geometry, and the optimized structures of the lowest singlet and triplet states.

In the case of ForC (Figure 5A) and as in the case of the monomer,[39] the first excited singlet state, of $^1n,\pi^*$ nature and localized over ForC ($^1n,\pi^*_{ForC}$), lies, at its equilibrium geometry, 0.48 eV below the triplet state of $^3\pi,\pi^*_{ForC}$ nature, i.e. the other player involved in the ISC, justifying the slower population of the triplet manifold. Notably, and again coherently with non-adiabatic dynamics performed for the monomer,[39] spin-orbit couplings corroborate the role of $^3\pi,\pi^*_{ForC}$ state as the doorway to ISC (Table S1). In contrast, in ForU, the corresponding states ($^1n,\pi^*_{ForU}$ and $^3\pi,\pi^*_{ForU}$) are quasi-degenerated at the $^1n,\pi^*_{ForU}$ equilibrium geometry, leading to a much more efficient triplet population (Figure 5B).[37]
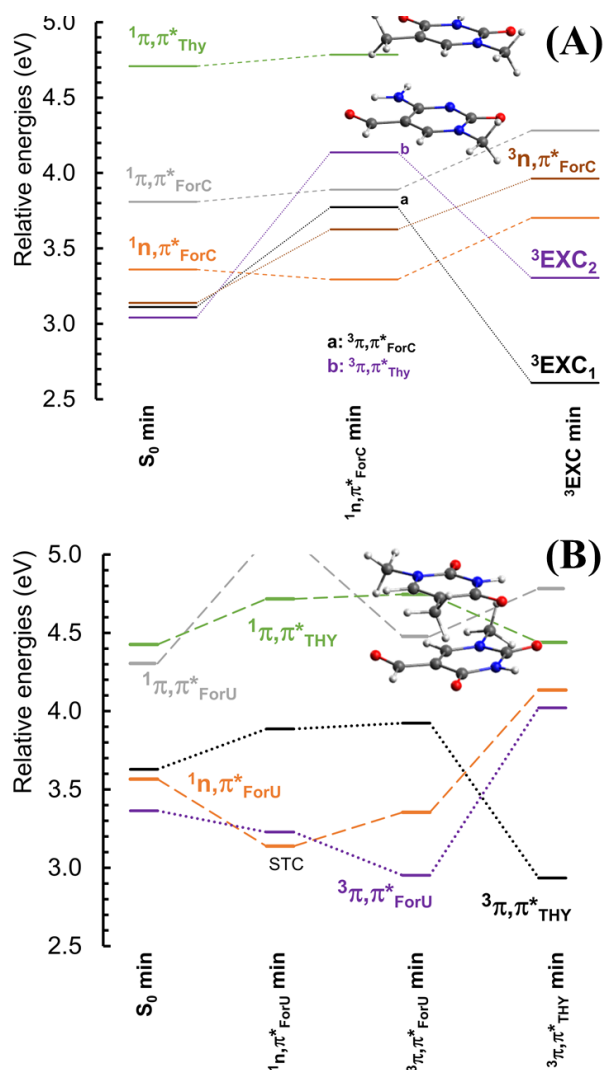
Figure 5. (A) CASPT2 energy profiles for the most relevant excited states of the ForC-Thy system. (B) CASPT2 energy profiles for the most relevant excited states of the ForU-Thy system, taken from Ref. [37]. STC = singlet-triplet crossing. The horizontal axis defines the equilibrium geometry of the most important states, obtained via geometry optimization, the dashed (singlets) and dotted (triplet) lines are used to connect the states having the same diabatic nature (correlation diagram) as a guide for the eyes. While the order of the structures presented on the x-axis reflects the temporal evolution of the global TTET process, no mechanistic insights in terms of activation energies or energy barriers can be directly inferred. Molecular geometries have been optimized using Gaussian 16,[53] whereas all multiconfigurational calculations have been conducted with the OpenMolcas software.[54]
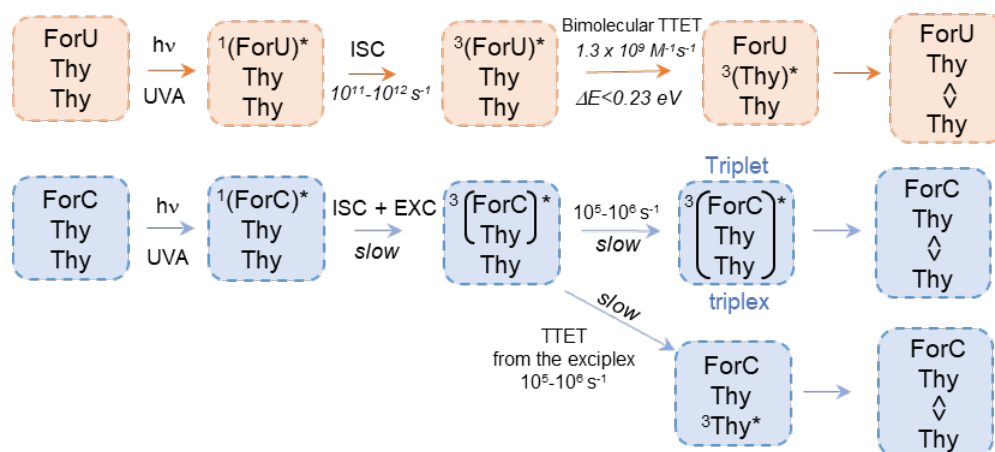
Scheme 2. Schematic representation of the proposed Thy<>Thy production from excitation of ForU (top) and ForC (bottom), according to our experimental and computational data. ISC = intersystem crossing, TTET = triplet-triplet energy transfer.
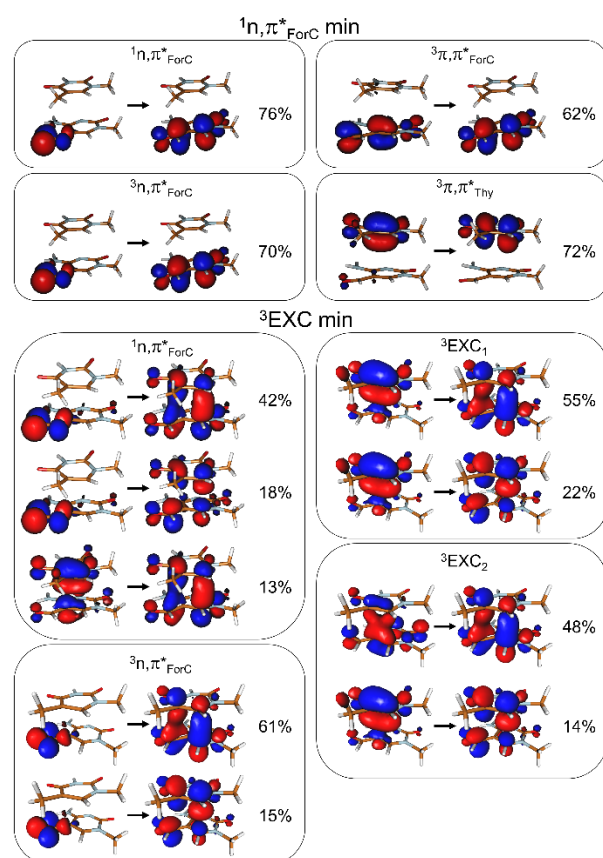

Figure 6. CASSCF orbitals corresponding to the highest weight electronic transitions that characterize the excited states of the ForC-Thy system. The weights of the configurations are also shown. Other transitions with much lower weights adding up to the total 100% are not reported. The orbitals defining the analogous electronic density reorganization in the ForU-Thy system can be found in ref [37].

Other significant differences between the two sensitizers are due to the distribution of the spin density between the couple

photosensitizer-nucleobase. The triplet states of ForC-Thy are characterized by a marked delocalization of the π clouds, describing the unpaired electrons, over both ForC and thymine moieties, especially involving the antibonding π^* orbitals (see Figure 6). Hence, the triplet states can be classified as exciplex (EXC$_1$, EXC$_2$). The delocalization is especially pronounced at the S$_0$ min and the T$_1$ min, that in diabatic notation is labeled as ^{3}EXC min. On the contrary, in the case of ForU, the spin density is much better localized over the two moieties, also giving rise to two well distinct electronic states, with a small but non-negligible driving force favoring the transfer of spin density to Thy, in a process requiring to bypass only a relatively small energy barrier. The presence of the low energy ^{3}EXC$_1$ in ForC can contribute to further explain the different observed yields. Indeed, the possibility of the formation of a delocalized state involving the sensitizer and two thymine bases attached to a rigid scaffold has been previously invoked in the case of benzophenone, another photosensitizer.[55] Different mechanisms could be at play here (see Scheme 2) for ForC, in which (*i*) the exciplex state is too stable to be dissociated, and act as the energy donor for TTET to Thy, however the decreased energy of ^{3}EXC$_2$ (2.61 eV) locates this state below of Thy triplet excited state (at 2.87 eV)[23], which makes TTET process unfavored; or (*ii*) the exciplex ^{3}EXC$_1$ can eventually further delocalize involving a second Thy unit forming a triplet triplex that evolves forming Thy<>Thy; nonetheless, such a trimolecular process can clearly be seen as slower and less efficient than the most direct route operative in the case of ForU involving localized states (Scheme 2). However, the energy gap is not the only ingredient determining the ISC efficiency. SOC, and more generally the topology of the potential energy surfaces, also play a relevant role, that in some cases may shape the photophysical outcome has also shown by previous molecular modelling and simulations works.[29,56–58]

Of note, all calculations in the present contribution as well as the non-adiabatic molecular dynamics simulations[39] have been carried on in the gas phase. While the effect of the solvent is not expected to modify significantly the qualitative picture

sketched out here, as demonstrated for the ForU-Thy system with polarizable-continuum model calculations,[37] some effects, especially for quasi-degenerate states, are susceptible to slightly alter the potential energy surfaces landscape.

Conclusions

The present work characterizes the formation of CPD lesions through triplet-triplet energy transfer from ubiquitous oxidized nucleobases and epigenetic intermediates. The differences of the photosensitization capabilities of ForU and ForC are revealed in detail by time-resolved and steady-state absorption spectroscopy, NMR, HPLC, and molecular modeling. It has been observed that the triplet-triplet energy transfer rate from ForU to thymine-thymine is three times faster than that from ForC ($k_q \sim 1.3 \times 10^9$ $M^{-1}s^{-1}$ for ^{3}ForU* and 4.6×10^8 M^{-1} s^{-1} for ^{3}ForC*), leading to a complete depletion of the thymine-thymine reactant after 1 hour of irradiation in the presence of ForU. Under the same conditions, the consumption of the reactant is only about 50% when ForC acts as photosensitizer. Molecular modeling provides further evidence supporting a slower intersystem crossing and triplet-triplet energy transfer in ForC, since the relaxation of the singlet excited state leads to a region where the triplet states are less accessible. Exciplex states of $^3\pi,\pi^*$ (^{3}EXC) nature dominate the triplet state manifold in ForC-Thy and are identified as mediators of thymine dimerization in this system. Even if our calculations have been made in gas-phase only, and hence care should be taken when drawing general conclusions, our picture offers a qualitative explanation of the involved mechanism also greatly reinforced by the good agreement between the theoretical and experimental data. Furthermore, in a biological environment the lesions studied in this work occur in the hydrophobic core of DNA, in which the influence of the polar environment is expected to be more limited.

A step forward is under progress with insertion of these formyl derivatives in short double stranded oligonucleotides to evaluate the potential of these lesions as intrinsic photosensitizers.

Conflicts of interest

There are no conflicts to declare

Acknowledgements

Notes and references

1   C. E. Crespo-Hernandez, B. Cohen, P. M. Hare and B. Kohler, *Chem. Rev.*, 2004, **104**, 1977–2019.

2   R. Improta, F. Santoro and L. Blancafort, *Chem. Rev.*, 2016, 116, 3540–3593.

3   E. Sage, P.-M. Girard and S. Francesconi, *Photochem. Photobiol. Sci.*, 2012, **11**, 74–80.

4   A. Francés-Monerris, H. Gattuso, D. Roca-Sanjuán, I. Tuñón, M. Marazzi, E. Dumont and A. Monari, *Chem. Sci.*, 2018, **9**, 7902–7911.

5   Y. Zhang, K. De La Harpe, A. A. Beckstead, R. Improta and B. Kohler, *J. Am. Chem. Soc.*, 2015, **137**, 7059–7062.

6   Y. Zhang, X.-B. Li, A. M. Fleming, J. Dood, A. A. Beckstead, A. M. Orendt, C. J. Burrows and B. Kohler, *J. Am. Chem. Soc.*, 2016, **138**, 7395–7401.

7   D. B. Bucher, A. Schlueter, T. Carell and W. Zinth, *Angew. Chem. Int. Ed.*, 2014, **53**, 11366–11369.

8   K. Rottger, H. J. B. Marroux, M. P. Grubb, P. M. Coulter, H. Bohnke, A. S. Henderson, M. C. Galan, F. Temps, A. J. Orr-Ewing and G. M. Roberts, *Angew. Chem. Int. Ed.*, 2015, **54**, 14719–14722.

9   J. J. Nogueira, F. Plasser and L. Gonzalez, *Chem. Sci.*, 2017, **8**, 5682–5691.

10 M. Barbatti, A. J. A. Aquino, J. J. Szymczak, D. Nachtigallová, P. Hobza and H. Lischka, *Proc. Natl. Acad. Sc. U.S.A.*, 2010, **107**, 21453–21458.

11 S. Reiter, D. Keefer and R. De Vivie-Riedle, *J. Am. Chem. Soc.*, 2018, **140**, 8714–8720.

12 M. C. Cuquerella, V. Lhiaubet-Vallet, F. Bosca and M. A. Miranda, *Chem. Sci.*, 2011, **2**, 1219–1232.

13 S. Mouret, C. Baudouin, M. Charveron, A. Favier, J. Cadet and T. Douki, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 13765–13770.

14 H. Ikehata, T. Mori, Y. Kamei, T. Douki, J. Cadet and M. Yamamoto, *Photochem. Photobiol.*, 2020, **96**, 94–104.

15 G. P. Pfeifer and A. Besaratinia, *Photochem. Photobiol. Sci.*, 2012, **11**, 90–97.

16 F. P. Noonan, M. R. Zaidi, A. Wolnicka-Glubisz, M. R. Anver, J. Bahn, A. Wielgus, J. Cadet, T. Douki, S. Mouret, M. A. Tucker, A. Popratiloff, G. Merlino and E. C. De Fabo, *Nat. Commun.*, 2012, **3**, 884.

17 R. P. Sinha and D.-P. Häder, *Photochem. Photobiol. Sci.*, 2002, **1**, 225–236.

18 S. Mouret, C. Philippe, J. Gracia-Chantegrel, A. Banyasz, S. Karpati, D. Markovitsi and T. Douki, *Org. Biomol. Chem.*, 2010, **8**, 1706–1711.

19 B. Epe, *Photochem. Photobiol. Sci.*, 2012, **11**, 98–106.

20 J. Cadet and T. Douki, *Photochem. Photobiol. Sci.*, 2018, **17**, 1816–1841.

21 A. Francés-Monerris, I. Tuñón and A. Monari, *J. Phys. Chem. Lett.*, 2019, **10**, 6750–6754.

22 D. Roca-Sanjuán, G. Olaso-González, I. González-Ramírez, L. Serrano-Andrés and M. Merchán, *J. Am. Chem. Soc.*, 2008, **130**, 10768–10779.

23 T. Climent, I. González-Ramírez, R. González-Luque, M. Merchán and L. Serrano-Andrés, *J. Phys. Chem. Lett.*, 2010, **1**, 2072–2076.

24 E. Dumont and A. Monari, *Front. Chem.*, 2015, **3**, 43.

25   V. Lhiaubet-Vallet, Z. Sarabia, D. Hernández, J. V. Castell and M. A. Miranda, *Toxicol. Vitr.*, , DOI:10.1016/S0887-2333(03)00108-5.

26   S. Sauvaigo, T. Douki, F. Odin, S. Caillat, J.-L. Ravanat and J. Cadet, *Photochem. Photobiol.*, 2001, **73**, 230–237.

27   V. Lhiaubet-Vallet, F. Bosca and M. A. Miranda, *Photochem. Photobiol.*, 2009, **85**, 861–868.

28   M. C. Cuquerella, V. Lhiaubet-Vallet, J. Cadet and M. A. Miranda, *Acc. Chem. Res.*, 2012, **45**, 1558–1570.

29   M. Marazzi, S. Mai, D. Roca-Sanjuán, M. G. Delcey, R. Lindh, L. González and A. Monari, *J. Phys. Chem. Lett.*, 2016, **7**, 622–626.

30   E. Dumont, M. Wibowo, D. Roca-Sanjuán, M. Garavelli, X. Assfeld and A. Monari, *J. Phys. Chem. Lett.*, 2015, **6**, 576–580.

31   V. Vendrell-Criado, G. M. Rodriguez-Muniz, M. C. Cuquerella, V. Lhiaubet-Vallet and M. A. Miranda, *Angew. Chem. Int. Ed.*, 2013, **52**, 6476–6479.

32   E. Bignon, H. Gattuso, C. Morell, E. Dumont and A. Monari, *Chem. Eur. J.*, 2015, **21**, 11509–11516.

33   D. K. Rogstad, J. Heo, N. Vaidehi, W. A. Goddard, A. Burdzy and L. C. Sowers, *Biochemistry*, 2004, **43**, 5688–5697.

34   M. Bachman, S. Uribe-Lewis, X. Yang, H. E. Burgess, M. Iurlaro, W. Reik, A. Murrell and S. Balasubramanian, *Nat. Chem. Biol.*, 2015, **11**, 555–557.

35   Y. Wang, X. Zhang, G. Zou, S. Peng, C. Liu and X. Zhou, *Acc. Chem. Res.*, 2019, **52**, 1016–1024.

36   I. Aparici-Espert, G. Garcia-Lainez, I. Andreu, M. A. Miranda and V. Lhiaubet-Vallet, *ACS Chem. Biol.*, 2018, **13**, 542–547.

37   A. Francés-Monerris, C. Hognon, M. A. Miranda, V. Lhiaubet-Vallet and A. Monari, *Phys. Chem. Chem. Phys.*, 2018, **20**, 25666–25675.

38   X. Wang, Y. Yu, Z. Zhou, Y. Liu, Y. Yang, J. Xu and J. Chen, *J. Phys. Chem. B*, 2019, **123**, 5782–5790.

39   A. Francés-Monerris, M. Lineros-Rosa, M. A. Miranda, V. Lhiaubet-Vallet and A. Monari, *Chem. Commun.*, 2020, **56**, 4404–4407.

40   J. Cadet and T. Douki., *Int. J. Radiat. Biol.*, 1999, **75**, 571–581.

41   T. Douki, T. Delatour, F. Paganon and J. Cadet, *Chem. Res. Toxicol.*, 1996, **9**, 1145–1151.

42   S. Frelon, T. Douki, J.-L. Ravanat, J.-P. Pouget, C. Tornabene and J. Cadet, *Chem. Res. Toxicol.*, 2000, **13**, 1002–1010.

43   J. Cadet, T. Douki and J.-L. Ravanat, *Acc. Chem. Res.*, 2008, **41**, 1075–1083.

44   T. Pfaffeneder, F. Spada, M. Wagner, C. Brandmayr, S. K. Laube, D. Eisen, M. Truss, J. Steinbacher, B. Hackner, O. Kotljarova, D. Schuermann, S. Michalakis, O. Kosmatchev, S. Schiesser, B. Steigenberger, N. Raddaoui, G. Kashiwazaki, U. Müller, C. G. Spruijt, M. Vermeulen, H. Leonhardt, P. Schär, M. Müller and T. Carell, *Nat. Chem. Biol.*, 2014, **10**, 574–581.

45   G. S. Madugundu, J. Cadet and J. R. Wagner, *Nucleic Acids Res.*, 2014, **42**, 7450–7460.

46   T. F. J. Kraus, D. Globisch, M. Wagner, S. Eigenbrod, D. Widmann, M. Münzel, M. Müller, T. Pfaffeneder, B. Hackner, W. Feiden, U. Schüller, T. Carell and H. A. Kretzschmar, *Int J Cancer*, 2012, **131**, 1577–90.

47   V. Lhiaubet-Vallet and M. A. Miranda, in *CRC Handbook of Organic Photochemistry and Photobiology.*, eds. F. Ghetti, A. G. Griesbeck and M. Oelgemöller, CRC Press, 2012, pp. 1541–1555.

48   O. R. Alzueta, M. C. Cuquerella and M. A. Miranda, *J. Org. Chem*, 2019, **84**, 13329–13335.

49   V. Lhiaubet-Vallet, M. C. Cuquerella, J. V. Castell, F. Bosca and M. A. Miranda, *J. Phys. Chem. B*, , DOI:10.1021/jp070167f.

50   T. Douki, *ChemPhotoChem*, 2020, **4**, 294–299.

51   A. Kuhlmann, L. Bihr and H.-A. Wagenknecht, *Angew. Chem. Int. Ed.*, 2020, **59**, 17378–17382.

52   L. Antusch, N. Gaß and H.-A. Wagenknecht, *Angew. Chem. Int. Ed.*, 2017, **56**, 1385–1389.

53   M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. a. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. a. Petersson, H. Nakatsuji, X. Li, M. Caricato, a. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, a. F. Izmaylov, J. L. Sonnenberg, Williams, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. a. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. a. Keith, R. Kobayashi, J. Normand, K. Raghavachari, a. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox, 2016, Gaussian 16, Revision C.01, Gaussian, Inc., Wallin.

54   I. Fernández Galván, M. Vacher, A. Alavi, C. Angeli, F. Aquilante, J. Autschbach, J. J. Bao, S. I. Bokarev, N. A. Bogdanov, R. K. Carlson, L. F. Chibotaru, J. Creutzberg, N. Dattani, M. G. Delcey, S. S. Dong, A. Dreuw, L. Freitag, L. M. Frutos, L. Gagliardi, F. Gendron, A. Giussani, L. Gonzalez, G. Grell, M. Guo, C. E. Hoyer, M. Johansson, S. Keller, S. Knecht, G. Kovačević, E. Källman, G. Li Manni, M. Lundberg, Y. Ma, S. Mai, J. P. Malhado, P. A. Malmqvist, P. Marquetand, S. A. Mewes, J. Norell, M. Olivucci, M. Oppel, Q. M. Phung, K. Pierloot, F. Plasser, M. Reiher, A. M. Sand, I. Schapiro, P. Sharma, C. J. Stein, L. K. Sørensen, D. G. Truhlar, M. Ugandi, L. Ungur, A. Valentini, S. Vancoillie, V. Veryazov, O. Weser, T. A. Wesolowski, P.-O. Widmark, S. Wouters, A. Zech, J. P. Zobel and R. Lindh, *J. Chem. Theory Comput.*, 2019, **15**, 5925–5964.

55   P. Miro, V. Lhiaubet-Vallet, M. L. Marin and M. A. Miranda, *Chem. Eur. J.*, 2015, **21**, 17051–17056.

56   T. Schnappinger, P. Kölle, M. Marazzi, A. Monari, L. González and R. De Vivie-Riedle, *Phys. Chem. Chem. Phys.*, 2017, **19**, 25662–25670.

57   A. Francés-Monerris, J. Segarra-Martí, M. Merchán and D. Roca-Sanjuán, *Theor. Chem. Acc.*, 2016, **135**, 1–15.

58   R. Gonzalez-Luque, T. Climent, I. Gonzalez-Ramirez, M. Merchan and L. Serrano-Andres, *J. Chem. Theory Comput.*, 2010, **6**, 2103–2114.