Journal Name

ARTICLE

UV-Induced Long-Lived Decays in Solvated Pyrimidine Nucleosides Resolved at the MS-CASPT2/MM Level

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Ana Julieta Pepino,[a,⊥] Javier Segarra-Martí,[b,⊥,*] Artur Nenov,[a] Ivan Rivalta,[b] Roberto Improta[c,d,*] and Marco Garavelli[a,*]

The most relevant 'dark' electronic excited states in DNA/RNA pyrimidine nucleosides are mapped in water employing hybrid MS-CASPT2/MM optimisations with explicit solvation and including the sugar. Conical intersections (CIs) between initially accessed bright $^1\pi\pi^*$ and the lowest energy dark $^1n\pi^*$ excited states, involving the lone pair localised on the oxygen and/or nitrogen atoms are characterised. They are found at the vicinities of the Franck-Condon (FC) region and are shown to facilitate non-adiabatic population transfer. The excited state population of the $^1n_O\pi^*$ state, localised in the carbonyl moiety on all pyrimidine nucleosides, is predicted to rapidly evolve to its minimum, displaying non-negligible potential energy barriers along its non-radiative decay, and accounting for the ps signal registered in pump-probe experiments as well as for an efficient population of the triplet state. Cytidine displays an additional $^1n_N\pi^*$ state localised in the N3 atom and that leads to its excited state minimum displaying large potential energy barriers in the pathway connecting to the CI with the ground state. Sugar-to-base hydrogen/proton transfer processes are assessed in solution for the first time, displaying a sizable barrier along its decay and thus being competitive with other slow decay channels in the ps and ns timescales. A unified deactivation scheme for the long-lived channels of pyrimidine nucleosides is delivered, where the $^1n_O\pi^*$ state is found to mediate the long-lived decay in the singlet manifold and act as the doorway for triplet population and thus accounting for the recorded phosphorescence and, more generally, for the transient/photoelectron spectral signals registered up to the ns timescale.

Introduction

DNA nucleobases, the building blocks of the genomic material, possess an outstanding ability to dissipate the energy gained upon UV-light absorption in a harmless manner.[1-6] Beyond the intrinsic interest in the photophysics of these systems, this exceptional trait has been postulated to provide the required photostability for the hereditary genetic information to survive under strong UV-exposure in prebiotic conditions, ensuring its proper replication.[7,8] Due to the role played in the photo-protection mechanisms occurring in the DNA/RNA genomic material, these photoinduced processes have been extensively studied over the years both theoretically[5-7,9-12] and experimentally with state-of-the-art techniques,[13-20] providing a wealth of information towards elucidating the concrete pathways promoting photostability and separating them to those that lead to photo-damaging instances.

There is general consensus on the paths responsible for the sub-ps (<200 fs) ultrafast internal conversion occurring in water-solvated nucleoside derivatives deoxy-Thymidine (Thd), oxy-Uridine (Urd) and oxy-Cytidine (Cyd), the pyrimidine-based building blocks of DNA/RNA: this involves $S_1(^1\pi_H\pi^*)/S_0$ ethylenic and, as recently proposed, $S_1(^1\pi_H\sigma^*)/S_0$ ring-opening conical intersections.[21-23] *N-H and O-H bond-breaking events have been reported to occur along a σπ* state that becomes relevant and low in energy upon stretching and that leads to a conical intersection with the ground state,[24,25] being also suggested to partake in an ultrafast timescale. Other alternative scenarios have also been proposed: i) the ultrafast component of the excited state decay could account only for the initial relaxation of the wave packet away from the FC and its movement towards the excited state minimum,[23] ii) the involvement of $^1n\pi^*$ states and their ultrafast decay to the ground state mediated by $^1n\pi^*/S_0$ conical intersections in vacuo[12] and iii) this signal accounting for the population transfer between the initially accessed $^1\pi\pi^*$ and a neighbouring $^1n\pi^*$ state at the vicinities of the FC region in the gas phase[26] as well as in solution.*[18] On the other hand, the degree of involvement of dark excited states ($^1n\pi^*$ or intramolecular charge transfer transitions) in the photoactivated dynamics of nucleobases in solution is not so well understood. It remains to be completely ascertained whether dark states, such as those of charge transfer (CT) character or due to dipole forbidden

[a] *Dipartimento di Chimica Industriale "Toso Montanari", Università di Bologna, Viale del Risorgimento 4, I-40136 Bologna, Italy. E-mail: marco.garavelli@unibo.it*
[b] *Univ Lyon, ENS de Lyon, CNRS, Unversité Lyon 1, Laboratoire de Chimie UMR 5182, Allée d'Italie 46, FR-69342 Lyon, France. E-mail: Javier.segarra-marti@ens-lyon.fr*
[c] *Istituto di Biostrutture e Bioimmagini CNR, Via Mezzocannone 16, I-80134 Napoli, Italy. E-mail: robimp@unina.it*
[d] *LIDYL, CEA, CNRS, Université Paris –Saclay, F-91191 Gif-sur-Yvette, France.*
⊥ *Contributed equally*

† Footnotes relating to the title and/or authors should appear here.
Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

transitions like $^1n\pi^*$ states, can be directly populated after UV (*i.e.* 255 nm) light irradiation, and if these are indeed responsible for the long-lived decay components (from dozen/hundreds of ps to ns)[27-29] registered in pyrimidine nucleobases/sides/tides over the years as it is usually assumed. Long-lived channels may in turn enhance the likelihood of deleterious photochemical instances.[7] This has been shown, for example, in methylated DNA derivative 5-methyl-cytosine, where upon methylation the excited state lifetime increases an order of magnitude.[30] This feature has been correlated to the 5-methyl-cytosine's ability to cause DNA damage,[31] where methylated cytosine sequences are found to promote a 5- to 15-fold increase in the damage by pyrimidine photo-dimerisation damage as compared to cytosine.[32] Long lived deactivation channels, albeit less studied in the literature, are thus fundamental in order to fully understand DNA's photo-protection mechanisms and will be the focus of this work.

In the present contribution, we assess the most relevant 'optically dark' deactivation pathways in solvated pyrimidine nucleosides Thd, Urd and Cyd (i.e., the biologically relevant species in the biologically relevant environment) and their dependence on a variety of factors, such as the effect of the sugar and the explicit solvation along both singlet and triplet manifolds. These factors are considered theoretically by means of both statically and dynamically correlated potential energy surfaces. Furthermore, we provide a comparison of the so-obtained decay routes and their underlying spectroscopic fingerprints with state-of-the-art time-resolved experimental evidence. To this aim we adopt a quantum mechanics/molecular mechanics (QM/MM) scheme where both statically and dynamically correlated multistate complete active space second-order perturbation theory (MS-CASPT2)[33] is coupled to an explicit MM description of the solvent.

Computational Details

MM dynamics. MM dynamics simulations were performed using the Amber-11 suite of programs and the Parm99 force field.[34,35] A cubic solvent box of 12x12x12 Å comprising around a thousand water molecules of explicit TIP3P was initially utilised.[36] The whole system, including the nucleoside moiety, was equilibrated from 0 to 300°K for 1 ns while maintaining a constant volume and pressure (1 atm). A production run was then performed for 100 ns recording snapshots every 200 fs.

Initial conditions sampling. To select the initial geometry for QM/MM optimisations we sampled the conformational space via performed a cluster analysis based on the Root Mean Square (RMS) deviation within a 2.0 Å difference of the nucleosides coordinates over all snapshots recorded along the 100 ns molecular dynamics run using the MMTSB toolbox.[37] We obtained two different clusters, with base to sugar syn/anti populations of 73%/27%, 62%/38% and 71%/29% for Urd, Thd and Cyd, respectively.[21] The initial structure was therefore selected as the snapshot with the geometrical parameters closest to the centroid (*i.e.*, the mean position) of the structures obtained in a syn conformation, which was found to be the most representative according to the above analysis. Oxy-Uridine and oxy-Cytidine were chosen as model systems due to their presence in RNA, whereas deoxy-Thymidine was favoured due to its presence in DNA (see Figure 1).

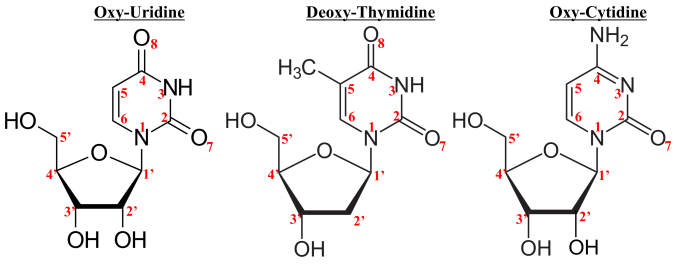

Figure 1. Chemical structure and atom labelling of oxy-uridine, deoxy-thymidine and oxy-cytidine.

It is worth noting that a single water arrangement for the different systems was considered along the diverse deactivation channels studied. The choice is based on the approach we recently tested in the related 5-methyl-cytosine system,[30] where we observed that considering different water arrangements improves the quantitative description of the excitation energies but did not change the qualitative conclusions attained from a single starting water distribution, which is what we also expect to occur in this work.

QM/MM calculations. QM/MM calculations were performed using the COBRAMM interface.[38] The cut between the QM and MM regions has been done so that it includes the nucleobase in the QM region, except for characterising the CT channels, where the whole nucleoside was included. A three-layer approach (high, medium and low) was used throughout: nucleoside/base was included in the QM region (high layer) whereas the medium layer comprises the movable MM atoms, consisting of the sugar moiety plus the water molecules within a 5 Å radius of the centre of mass of the nucleobase, the remainder of the MM system being kept frozen during all optimisation procedures in the low layer. Equilibrium geometries and photoreaction paths were determined by using fully unconstrained minima, transition state and minimum energy path (intrinsic reaction coordinate; IRC) optimizations using the Berny optimiser implemented in the Gaussian 09 program package[39] interfaced with COBRAMM. Conical Intersection (CI) optimisations were performed with the gradient projection algorithm of Bearpark *et al.*[40,41] as implemented in the COBRAMM interface. Energies in the QM region were computed making use of the complete active space second-order perturbation theory (CASPT2) method[42,43] in its multistate extension (MS)[33] as implemented in the MOLCAS package.[44] Gradients and non-adiabatic couplings (required for the CI optimisations) were computed numerically using MS-CASPT2 energies and wave functions. *It is important to highlight that all gradients employed for the different optimisations of critical points in this study (excited state minima, transition states and interstate crossings) have been carried out at the CASPT2 level making use of finite differentiation. Hessians have been computed numerically at the CASPT2 level to ascertain the nature of the different transition states optimised, and employ a mixed implementation where the first layer of the finite differentiation is carried out by MOLCAS*[44] *while the second is performed in parallel by our COBRAMM interface.*[21,30] The active space selected comprised the full π space plus the two lone-pair orbitals to provide a proper description of the $n\pi^*$ states, thus making 14 electrons in 10 orbitals. A CASSCF wave function averaged over eight singlet states was used in the

single-point CASPT2 energy corrections. An imaginary level shift of 0.2 a.u. was employed in the perturbation step to avoid intruder states, and the IPEA shift[45] was set to 0 unless specified otherwise. Two different basis sets were employed throughout: 6-31G**[46] was used in the optimizations at the CASPT2 level whereas an ANO-L-VTZP[47,48] was employed to refine the CASPT2 energies of the different key structures. Transition states were characterised by the presence of one negative frequency and IRCs were applied to verify that the correct states were connected. Spin-orbit couplings (SOCs) between singlet and triplet states were computed within the atomic mean-field integral (AMFI) framework,[49] obtaining the length of the SOC vector as implemented in MOLCAS and described in detail elsewhere.[50]

Excited state absorptions (ESAs). ESA estimates have been computed averaging over 25 states and assuming only absorption contributions arising from the populated state (the state characterising the minima in which it is computed). These therefore do not correspond to the actual pump-probe signal, as that would contain the negative contributions arising from the ground state bleach and stimulated emission.[53] It is worth noting, however, that those contributions would lie in spectral regions slightly different to those relevant and mainly discussed in the present manuscript, and would refer to ground state bleaching contributions of the optically dark ^{1}nπ* states, which are therefore not expected to change drastically the final pump-probe signal. *The present computations make use of the 14 electrons in 10 orbitals previously described. While this active space might seem insufficient to properly describe the high-lying excited states accessible in pump-probe spectroscopy, we have shown in previous work[54] how this active space suffices to obtain a qualitative description of the main signals featured in both Vis and UV probe windows. This has been recently highlighted for adenosine monomers and dimers,[55] where the main fingerprint recorded experimentally for these systems is placed in the 300-360 nm window[56] and is well reproduced even with computationally cheaper restricted active space approaches.* The spectra have been built by convoluting all states computed with a Gaussian function with FWHM = 0.3 eV, and where the intensity is based on the computed oscillator strength of each transition using Gabedit.[57]

Photoelectron signals. Cationic signals were computed as the difference between doublet and singlet states at each of the critical structures considered, the doublets being obtained by computing the cationic forms of the nucleobases. As it has been suggested previously in the literature,[58] a rescaling factor is needed for the energy of the cationic states as these deviate from their experimental counterpart in a more pronounced manner than the estimates singlet states (see ESI Table S1). *This shift is applied based on previous evidence in the literature,[59] where the description of cationic states in DNA/RNA nucleobases is shown to be more challenging and its errors larger than those registered for the often considered singlet states even for strongly correlated coupled-cluster based methods.* Given that in this case our estimates resemble remarkably well those arising from steady-state absorption maxima, the rescaling is carried out only for the cationic states in the following manner:

$$\Delta = \left(E_{D_0}^{CASPT2} - E_{S_0}^{CASPT2}\right) - IE_{S_0 \rightarrow D_0}^{Vertical,Exp,aq}$$

where $IE_{S_0 \rightarrow D_0}^{Vertical,Exp,aq}$ refer to experimental ionisation energies taken from Slavicek *et al.*[60] We assume 238 nm (5.21 eV) probe pulses analogous to those used by Buchner *et al.*,[20] the average kinetic energy (AKE) being:

$$AKE = 5.21 - \left(\left(E_{D_0}^{CASPT2} - \Delta\right)^{Rescaled} - E_{S_1}^{CASPT2}\right)$$

where energies refer to eV. A 0.25 a.u. IPEA shift was employed as it has been reported to better reproduce adiabatic cationic signals.[59,61] Despite IPEA-shifted MS-CASPT2 values being reported in the text, it is shown in the ESI (Table S2) that both MS and IPEA treatments do not affect the qualitative results obtained and therefore play no critical role in the ensuing discussion.

It is important to note that most photoelectron experiments monitor the photoelectron absorption signals, which provide better suited signals for tracking excited state dynamics and that have been considered in other computational studies.[62-64] These signals are however not computed in the present manuscript, where we rely on qualitative estimates of other spectroscopic data (ESAs and AKEs).

Table 1. Singlet and triplet vertical excitations (oscillator strength in parenthesis) for Urd, Thd and Cyd at the MS-CASPT2(14,10)/ANO-L-VTZP level at the Franck Condon region.

	Urd (eV)	Thd (eV)	Cyd (eV)	
$^1\pi_H\pi^*$ (H→L)	4.64 (0.255) *[4.64][51]*	4.70 (0.262) *[4.72][51]*	$^1\pi_H\pi^*$ (H→L)	4.60 (0.065) *[4.59][52]*
$^1n_O\pi^*$	5.14 (0.005)	5.17 (0.006)	$^1n_N\pi^*$	5.54 (0.010)
$^1\pi_2\pi^*$ (H→L+1)	5.82 (0.161)	5.88 (0.145)	$^1\pi_2\pi^*$ (H-1→L)	5.57 (0.154)
$^3\pi_H\pi^*$ (H→L)	3.70	3.73	$^1n_O\pi^*$	5.70 (0.002)
$^3n_O\pi^*$	4.71	5.06	$^3\pi_H\pi^*$ (H→L)	3.75
			$^3n_N\pi^*$	4.50

Experimental absorption maxima are provided in brackets for comparison.

Results and Discussion

^{1}nπ* decays

Pyrimidine nucleosides display low-lying bright excited ^{1}ππ* states ($^1\pi_H\pi^*$ and $^1\pi_2\pi^*$, see Table 1) that partake in the ultrafast deactivations of these systems,[21] being firstly populated upon absorption. In addition, for all the bases, optically dark excited state ^{1}nπ* transitions involving the lone par of oxygen and/or nitrogen atoms, are close in energy in the Franck-Condon (FC) region. Non-adiabatic molecular dynamics studies, mainly in the gas phase,[11,12,65-67] have reported the involvement of dark states in the excited state deactivation dynamics. As a first step, we verify the accessibility of the different ^{1}ππ*/^{1}nπ* CIs close to the FC region.

Figures 2 (Urd/Thd) and 3 (Cyd) depict the stationary points characterised along the PESs plotted in terms of their mass

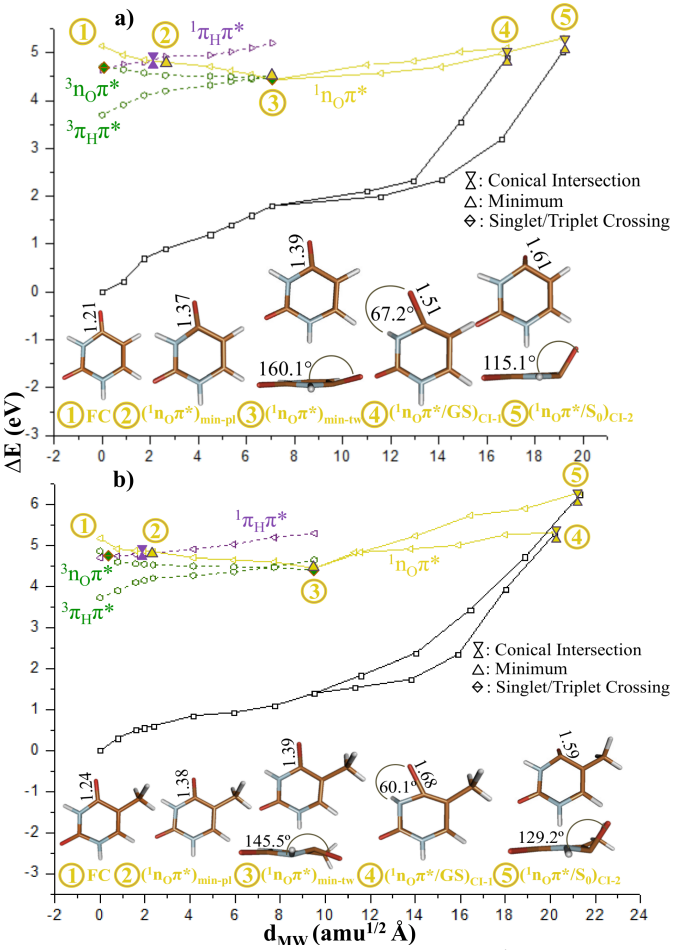

Figure 2. Evolution of the ground (black lines) and lowest-lying $^1n_O\pi^*$ (yellow lines) excited states for Urd (a) and Thd (b) along the computed $^1n_O\pi^*$ relaxation pathways plotted in mass-weighted coordinates. Dotted lines represent the evolution of the singlet ($^1\pi\pi^*$, purple line) and triplet ($^3\pi\pi^*$ and $^3n_O\pi^*$, green lines) manifold along those channels. Key points characterised along the decay routes are also depicted, the numbers along the potential energy curves referring to the structures given below with their main associated geometrical distortions.

weighted displacement from the corresponding ground state equilibrium geometry. *Energetically low-lying $(^1n\pi^*/^1\pi_H\pi^*)_{CI}$ and $(^1n\pi^*/^1\pi_2\pi^*)_{CI}$ CIs exist at the vicinities of the FC region, accessible through small molecular distortions, owing to the energetic proximity of the partaking states and that are expected to mediate $^1n\pi^*$ population, which has been suggested to occur in an ultrafast timescale.[26] This non-adiabatic population transfer can be analysed by comparing the energetic position of the absorbing $^1\pi_H\pi^*$ states for Urd, Thd and Cyd of 4.64, 4.70 and 4.60 eV, respectively (see Table 1) and the adiabatic energies of their respective $(^1n_O\pi^*/^1\pi_H\pi^*)_{CI}$ placed at 4.95, 4.90 and 5.10 eV, respectively (see Table 2). As can be seen, relatively small energy barriers of 0.31 and 0.2 eV are predicted for Urd and Thd, respectively, whereas Cyd displays a much larger 0.5 eV difference from one of the absorbing $^1\pi\pi^*$ states ($^1\pi_H\pi^*$), while a barrierless path is predicted from the more intense and higher-lying $^1\pi_2\pi^*$ state, that is also expected to be populated upon absorption.[21]*

Monitoring the key C4-O8 (C2-O7) distances in critical points close to the FC region highlights this point as well: Urd, Thd and Cyd display C-O distances of 1.21, 1.24 and 1.25 Å in their FC, respectively, 1.28, 1.25 and 1.27 Å in their $^1\pi_H\pi^*$ planar minima, and 1.26, 1.27 and 1.26 Å[21] in the twisted $^1\pi_H\pi^*$ minima; finally the CO bond in $(^1n_O\pi^*/^1\pi_H\pi^*)_{CI}$ increases up to 1.28, 1.38 and 1.34 Å for the three bases. It is thus envisaged that high-frequency vibrations, particularly due to C=O carbonyl stretching like those involved in the intramolecular vibrational dynamics of the $^1\pi\pi^*$ state, would allow for few $^1n\pi^*/^1\pi\pi^*$ re-crossings[68] and facilitate a partial non-adiabatic population transfer to the dark state before the out-of-plane C5 puckering motion draws the system towards the ultrafast $^1\pi\pi^*/S_0$ CI. *Even in early time ultrafast deactivations along the initially accessed $^1\pi_H\pi^*$ state, where a relatively large energy gap between the bright $^1\pi_H\pi^*$ and dark $^1n_O\pi^*$ states is observed,[21] a sizable C-O stretching can lead the system to populate the $^1n_O\pi^*$ state and can justify an ultrafast population branching where a large part of the initial wave packet ends in the $^1n_O\pi^*$ state.* Moreover, the $^1n\pi^*/^1\pi\pi^*$ coupling has been also linked to increase along the out-of-plane motion,[69] which would further promote $^1\pi\pi^*$ to $^1n\pi^*$ population transfer along the $^1\pi\pi^*$ main decay channel.[21] The involvement of a dark excited state in the photophysics can account for the experimental evidence from pump-probe transient absorption (fs-TA) measurements. In uracil, ~15-20% of the excited state population of water-solvated uracil and ~40% of oxy-uridine monophosphate (UMP) show a long-lived ps signal.[27] In this case, a 24 ps component was observed,[27] and a similar value (26ps) was registered for the 1-cyclohexyl derivative,[70] while oxy-UMP displays a 147 ps component[27] that rises to 240 ps when present in an oxy-UMP+deoxy-AMP mixture.[71] In thymine, around 15-20% of excited state population was measured to decay on the ps time scale, showing a lifetime of 30 ps for water-solvated thymine and 127 ps for deoxy-thymidine monophosphate (dTMP).[27] In cytosine, the recent study by Ma *et al.*[19] provides fs-TA data that shows ~25% excited state population on the ps channels for solvated cytosine, 49% for Cyd, 33% for dCyd and 39% for cytidine monophosphate (CMP), with lifetimes of 7.7, 35, 30 and 34 ps, respectively. All measurements recording this long-lived component have been carried out with fs-TA techniques, time-resolved fluorescence (fs-TRF)[19] and fluorescence up-conversion (fs-FU)[22] being unable to register them, which further supports their assignment to non-emissive/dark states such as $^1n\pi^*$ (still presenting sizeable transient absorption signals as recorded in fs-TA). Finally, a prominent phosphorescence signal has been registered in DNA/RNA nucleobases/tides/sides[72] implying a significant involvement of triplets in the overall deactivation mechanism. Addition of the sugar moiety thus strongly affects the long-living component, yielding analogous results for DNA and RNA based cytosine derivatives, which is the only case with enough data available in the literature for a meaningful comparison.[19]

Uridine and Thymidine. Figure 2 depicts the fate of the $^1n_O\pi^*$ state upon population either through $(^1n_O\pi^*/^1\pi_H\pi^*)_{CI}$ or $(^1n_O\pi^*/^1\pi_2\pi^*)_{CI}$ crossings for Urd/Thd. A swift decay is observed leading towards a planar minimum $(^1n_O\pi^*)_{min\text{-}pl}$ placed adiabatically at 4.84 and 4.87 eV for Urd and Thd, respectively (see Table 2), and featuring a pronounced C4-O8 (see Figure 1 for atom labelling) bond lengthening of ~0.07 and ~0.14. The $^1n_O\pi^*$ state further decays towards a slightly twisted minimum defined by an out-of-plane carbonyl motion $(^1n_O\pi^*)_{min\text{-}tw}$ present in all pyrimidine nucleosides and placed at 4.43 and 4.67 eV for Urd and Thd, respectively. This

excited state minimum has an associated C4-O8 bond length increase with respect to the FC of 0.17 and 0.12 Å. Inclusion of hydrogen-bonded water molecules directly linked to the carbonyl[30,73] and of dynamic electron correlation in the model[23,74-76] appear to be important for properly describing this twisted excited state minimum as observed by the inability shown by CASSCF at characterising this structure. Two sloped CIs to the ground state along the $^1n_O\pi^*$ PES involving strong C-O elongation and in-plane ($(^1n_O\pi^*/S_0)_{CI\text{-}1}$) or out-of-plane ($(^1n_O\pi^*/S_0)_{CI\text{-}2}$) bending of the carbonyl group have been characterised for Urd/Thd (see Fig. 2). They are similar to the structures already reported in the literature for Uracil and Thymine in gas-phase[6,77] and solution,[78] and show a strong increase of C4-O8 bond lengths of 0.3 and 0.35 Å, respectively. The path connecting $(^1n_O\pi^*)_{min\text{-}tw}$ and the CIs exhibit very large potential energy barriers (Urd: 0.77 and 1.33 eV, Thd: 0.83 and 1.08 eV, respectively, see Table 2), consistent with their potential involvement in the slow (several ps) component of the decay shown above. Due to the large PES barriers associated with $(^1n_O\pi^*/S_0)_{CI\text{-}2}$ in the following we will focus only on the decays mediated by the lower-lying and more accessible $(^1n_O\pi^*/S_0)_{CI\text{-}1}$. Interestingly, these large potential energy barriers obtained for Urd/Thd correlate well with the long lifetimes measured for UMP and TMP.[27] Due to the large potential energy barriers separating the minimum from the CI with S_0, it is also possible that population is back-transferred to $^1\pi_H\pi^*$, decaying along the channels associated to the sub-ps decay.[68]

Table 2. Main critical points characterised along the different potential energy surfaces considered in the present work enclosing excited state minima, conical intersections, intersystem crossing regions and transition states, together with their associated adiabatic (with respect to the FC region) energies and oscillator strengths computed at the MS-CASPT2(14,10)/ANO-L-VTZP. AE and VE stand for adiabatic (containing zero-point energy corrections for both excited and ground states) and vertical emissions, respectively.

	Urd (eV)	Thd (eV)	Cyd (eV)	
$^1\pi_H\pi^*$ (H→L) path				
$(^3n_O\pi^*/^1\pi_H\pi^*)_{ISC}$	4.7 (SOC ~ 9 cm^{-1})	5.1 (SOC ~ 1 cm^{-1})	$(^3\pi_2\pi^*/^1\pi_H\pi^*)_{ISC}$	4.55 (SOC~12 cm^{-1})
$(^3\pi_H\pi^*/^1\pi_H\pi^*)_{ISC}$	3.90 (SOC ~ 5 cm^{-1})	4.51 (SOC ~ 1 cm^{-1})	$(^3\pi_H\pi^*/^1\pi_H\pi^*)_{ISC}$	4.01 (SOC ~ 24 cm^{-1})
$(^3\pi_H\pi^*/S_0)_{ISC}$	3.51 (SOC ~ 1 cm^{-1})	3.92 (SOC ~ 1 cm^{-1})	$(^3\pi_H\pi^*/S_0)_{ISC}$	3.85 (SOC ~ 2 cm^{-1})
$^1\pi_H\pi^*_{min\text{-}pl}$	4.40 (AE: 4.24, VE: 3.75, f~0.185)	4.38 (AE: 4.14, VE: 3.78, f~0.235)	$^1\pi_H\pi^*_{min\text{-}pl}$	4.04 (AE: 3.96, VE: 3.84, f~0.077)
$^1\pi_H\pi^*_{min\text{-}tw}$	4.03 (AE: 3.82, VE: 2.59, f~0.180)	4.14 (AE: 4.01, VE: 2.43, f~0.231)	$^1\pi_H\pi^*_{min\text{-}tw}$	3.99 (AE: 3.89, VE:3.02, f~0.074)
$(^1\pi_H\pi^*_{min\text{-}tw}\rightarrow CI_{\pi H\pi^*/S0})_{TS}$	4.12 (ΔE: 0.09)	4.26 (ΔE: 0.12)	$(^1\pi_H\pi^*_{min\text{-}tw}\rightarrow CI)_{TS}$	4.14 (ΔE: 0.15)
$(^1\pi_H\pi^*/S_0)_{CI}$	3.98	4.22	$(^1\pi_H\pi^*/S_0)_{CI}$	3.90
$(^1\pi_H\pi^*_{min\text{-}pl}\rightarrow CI_{\pi H\delta^*/S0})_{TS}$	4.54 (ΔE: 0.14)	4.56 (ΔE: 0.18)	$^1\pi_H\pi^*_{min\text{-}3}$	4.28 (AE: 4.01, VE: 3.78 f~0.055)
$(^1\pi_H\sigma^*/S_0)_{CI}$	4.08	4.47	$(^1\pi_H\pi^*_{min\text{-}pl}\rightarrow CI_{\pi H\delta^*/S0})_{TS}$	4.28 (ΔE: 0.24)
			$(^1\pi_H\sigma^*/S_0)_{CI}$	4.10
$(^3\pi_H\pi^*)_{min}$	3.20	3.53	$(^3\pi_H\pi^*)_{min}$	3.45
$(^3n_O\pi^*)_{min}$	4.41	4.47	$(^3\pi_2\pi^*)_{min}$	4.34
$(^3\pi_H\pi^*/^3n_O\pi^*)_{CI}$	4.43	4.50	$(^3\pi_H\pi^*/^3n_N\pi^*)_{CI}$	4.36
$n_O\pi^*$ path				
$(^1\pi_H\pi^*/^1n_O\pi^*)_{CI}$	4.95	4.90	$(^1n_O\pi^*)_{min\text{-}pl}$	5.10
$(^1n_O\pi^*)_{min\text{-}pl}$	4.84	4.87	$(^1n_O\pi^*)_{min\text{-}tw}$	4.54
$(^1n_O\pi^*)_{min\text{-}tw}$	4.43	4.67	$(^1\pi_2\pi^*/^1n_O\pi^*)_{CI}$	5.60
$(^1n_O\pi^*/S_0)_{CI\text{-}1}$	5.20 (ΔE~0.77)	5.50 (ΔE~0.83)	$(^1n_O\pi^*/S_0)_{CI}$	6.27 (ΔE: 1.73)
$(^1n_O\pi^*/S_0)_{CI\text{-}2}$	5.73 (ΔE~1.30)	5.75 (ΔE~1.08)	$(^3\pi_H\pi^*/^1n_O\pi^*)_{ISC}$	5.10 (SOC ~42 cm^{-1})
$(^3\pi_H\pi^*/^1n_O\pi^*)_{ISC}$	4.43 (SOC ~ 42 cm^{-1})	4.67 (SOC ~ 40 cm^{-1})	$((^1n_O\pi^*)_{min\text{-}tw}\rightarrow 3st\text{-}CI)_{TS}$	4.93 (ΔE: 0.41)
			$3st_{CI}$	4.60
$\pi_2\pi^*$ (H-1 →L) path				
$(^1\pi_H\pi^*/^1\pi_2\pi^*)_{CI}$	4.75	4.78	$(^1\pi_H\pi^*/^1\pi_2\pi^*)_{CI}$	4.70
$^1\pi_2\pi^*_{min}$	4.72 (AE: 4.48, VE: 3.95, f~0.083)	4.77 (AE: 4.50, VE: 4.03, f~0.091)	$^1\pi_2\pi^*_{min}$	4.70 (AE: 4.40, VE: 4.05, f~0.085)
$(^1\pi_2\pi^*/S_0)_{CI}$	5.60 (ΔE~0.88)	5.01 (ΔE~0.24)	$(^1\pi_2\pi^*/^1n_N\pi^*)_{CI}$	5.28
			$n_N\pi^*$ path	
			$(^1n_N\pi^*)_{min}$	4.57
			$(^1\pi_H\pi^*/^1n_N\pi^*)_{CI}$	4.80
			$(^1n_N\pi^*_{min}\rightarrow CI)_{TS}$	5.23 (ΔE: 0.66 eV)
			$(^1n_N\pi^*/S_0)_{CI}$	4.45
			$(^3\pi_H\pi^*/^1n_N\pi^*)_{ISC}$	4.90 (SOC ~ 8 cm^{-1})

Cytidine. Cyd (see Fig. 3) displays an analogous qualitative profile along the $^1n_O\pi^*$ deactivation as those described above for Urd/Thd. After initial access through $(^1n_O\pi^*/^1\pi_H\pi^*)_{CI}$ or $(^1n_O\pi^*/^1\pi_2\pi^*)_{CI}$ at the vicinities of the FC, the $^1n_O\pi^*$ decays to a planar $(^1n_O\pi^*)_{min\text{-}pl}$ minimum placed 5.10 eV adiabatically and displaying a 0.09 Å lengthening of the C2-O7 carbonyl moiety. This bond length then decreases by 0.05 Å going towards $(^1n_O\pi^*)_{min\text{-}tw}$, whose adiabatic energy is 4.54 eV (see Table 2), resulting in a significant energy stabilisation with respect to its planar minimum. Two different CIs are also accessible connecting the ground and $^1n_O\pi^*$ states as it was shown for Urd/Thd: $(^1n_O\pi^*/S_0)_{CI}$, characterised with an out-of-plane bending of the carbonyl group, and a three-state CI $(^1n_O\pi^*/^1\pi_H\pi^*/S_0)_{CI}$ denoted as 3_{st}CI in Figure 3. As opposed to the former, 3_{st}CI shows lesser carbonyl out-of-plane distortion despite featuring a ~0.28 Å elongation of the C2-O7 bond with respect to the FC. This 3_{st}CI is considered the photochemically active CI as it is characterised by having the smallest C-O stretching while displaying the smallest potential energy barrier connecting the $^1n_O\pi^*$ and

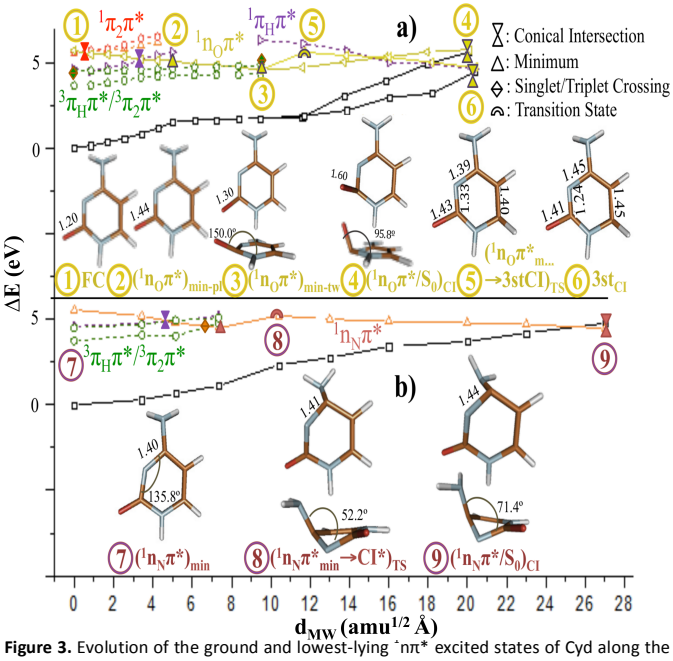

Figure 3. Evolution of the ground and lowest-lying ˚nπ* excited states of Cyd along the computed relaxation pathways of the $^1n_O\pi^*$ (yellow lines) state (a) and $^1n_N\pi^*$ (orange line) state (b) plotted in mass-weighted coordinates. Dotted lines represent the evolution of the singlet ($^1\pi\pi^*$, purple line) and triplet ($^3\pi\pi^*$, green lines) manifold along those channels. Key points characterised along the decay routes are also depicted, the numbers along the potential energy curves referring to the structures given below with their main associated geometrical distortions.

ground states at 0.41 eV, which may explain the systematically faster long-lived component observed in Cyd[19,27] as compared to Urd/Thd.[27]

Cyd also presents a low-lying $^1n_N\pi^*$ state, not present in Urd/Thd, that appears energetically between the photoactive $^1\pi_H\pi^*$ and $^1\pi_2\pi^*$ states at the FC region (see Table 1). $(^1n_N\pi^*/^1\pi_H\pi^*)_{CI}$ and $(^1n_N\pi^*/^1\pi_2\pi^*)_{CI}$ crossings have been located and are placed in this case further away from the FC compared to those previously described for the $^1n_O\pi^*$ state (see Fig. 3), requiring slightly out-of-plane carbon displacements and making its non-adiabatic population immediately after absorption less likely. Optimisation along the $^1n_N\pi^*$ state leads to a minimum, its reaction pathway continuing towards the $(^1n_N\pi^*/S_0)_{CI}$ structure featuring a strongly-puckered out of plane amino group (see Fig. 3),[79,80] and characterised by bond length elongations of C4-N3 and C4-N8 by moving from the FC to the minimum (0.06 and 0.05 Å, respectively), and a further 0.02 and 0.04 Å by moving from the minimum to the CI, triggering the progressive out-of-plane modes of the amino moiety and the N3 atom. Strong $^1\pi_H\pi^*/^1n_N\pi^*$ mixing is observed along this channel (see SI Table S4), the CI displaying large $^1\pi_H\pi^*$ contributions.[79] We characterised the transition state connecting the minimum and CI along this decay route, revealing a sizeable potential energy barrier (0.66 eV, see Table 2) and making it another potential candidate for the long-lived decay in Cyd. These $^1n\pi^*$ mediated routes have in the past been proposed to actively contribute to the ultrafast decays of nucleobases based on *in vacuo* computations.[12] However, we here show that inclusion of solvent effects and of dynamical correlation significantly destabilises the related CI ruling out the involvement of these channels in the ultrafast (sub-ps) timescale.

Nucleoside-mediated decays

N1 substitution. According to the experiments mentioned above, N1 substitution affects mainly the long-lived components by systematically increasing their lifetimes.[19,27] This is often ascribed to the effects held by the sugar/phosphate moieties on the carbonyl group describing the $^1n_O\pi^*$ excitation. As previously reported,[21] the syn nucleoside to sugar conformation here employed is the most representative structure as obtained from our molecular mechanics studies, and involves a hydrogen bond between 5' O–H of the sugar moiety and O8 (O7 in Cyd) of the nucleobase (see Fig. 1). We anticipate that this intramolecular hydrogen bond can potentially affect the C2-O7 stretching mode that is the one activated along the Cyd $^1n_O\pi^*$ decay channel thus increasing its energy barrier and lifetime.[19] This entails a hydrogen bond that is expected to hamper the out-of-plane motion thus impacting this decay channel. The recent study by Ma *et al.*[19] is the only one providing enough experimental evidence to assess this point, investigating both oxy- and deoxy-cytosine derivatives and where a clear trend can be observed, the lifetime of cytosine increasing from the bare nucleobase (7 ps) to oxy- and deoxy-cytidine (30 and 34 ps) as well as CMP (34 ps). The possibility of forming intramolecular sugar/base hydrogen bonds seem to be plausible according to our molecular dynamics, and thus a similar behaviour could be expected for uracil and thymine, displayed through the increase of their lifetimes by going from nucleobase to nucleotide,[27] even though additional experimental evidence and a purposely tailored study of the syn/anti equilibrium would be required for a definitive assessment. Sugar/phosphate substitution can thus affect intramolecular vibrational relaxation and interfere with the solute/solvent hydrogen bonding network, ultimately affecting the $^1n_O\pi^*$ lifetime. It is worth noting that in 1-cyclo-hexyl uracil (where this kind of hydrogen bond cannot be formed) the dark state decays in 26 ps[70] as in uracil (24 ps), further supporting the role played by intramolecular hydrogen bonds.[27]

Intramolecular hydrogen transfer. We next consider a recently proposed[81,82] mechanism whereby a sugar-to-base charge transfer state with an accessible CI to the ground state can be populated in purine nucleoside derivatives by an aborted proton transfer from the ribose sugar to the N3 atom of the base. Based on ADC(2) computations, this channel has been suggested to be activated on sub-ps timescales in gas-phase nucleosides,[83] and to compete with $^1\pi\pi^*$ ultrafast deactivations in the sub-ps timescale. We considered this possibility in pyrimidines and explored the intramolecular excited-state 5'-O–H···O7 H-transfer reaction using the same procedure devised by Tuna *et al.*[81,82] and including the whole nucleoside in the QM region. Figure 4a shows the potential-energy profiles of the electronic excited states of interest for Thd (Urd and Cyd display analogous behaviors, see SI Figures S5-S6). This is based on a ground-state relaxed scan, which displays a pronounced dissociative character along the hydrogen transfer coordinate that is unlikely to be overcome with thermal (vibrational) energy. The CT state (displaying a shift of electron density from the 5'-OH group to the π system of the nucleobase moiety) approaches $^1\pi_H\pi^*$ and $^1\pi_2\pi^*$ states at high energies with O-H distances of ~1.45 Å in Thd and Urd (1.35 Å for Cyd). An unconstrained optimisation of the CT state (optCT in Fig. 4b) displays an O-H distance of ~1.47 Å, and

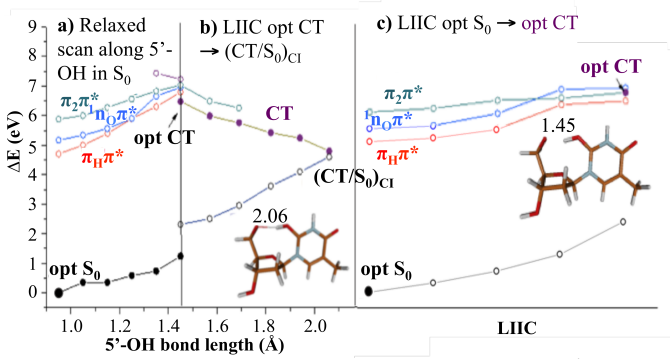

Figure 4. Energy profiles of the electronic ground state, the lowest bright $^1\pi\pi^*$ state ($^1\pi_H\pi^*$), the dark $^1n\pi^*$ state, the second $^1\pi\pi^*$ state ($^1\pi_2\pi^*$) and the charge-transfer (CT) state along the H-transfer coordinate of the 5'-OH group of the ribose moiety of Thymidine. (a) Relaxed scan along the 5'-O–H internuclear distance in the electronic ground state (full black circles) with vertical excitation energies of the $^1\pi_H\pi^*$, $^1n_O\pi^*$, $^1\pi_2\pi^*$ and CT states (empty red, blue, green and purple circles, respectively). (b) Linear interpolation between the optimised charge-transfer state (full purple circles) and the conical intersection connecting the CT to the ground state $(CT/S_0)_{CI}$. (c) Linear interpolation between the ground-state equilibrium geometry (full black circle at the lower left) and the optimised geometry of the charge-transfer state (full purple circle at the right). The $^1\pi_2\pi^*$ state correlates adiabatically to the charge-transfer state at large O–H distances. The energy profiles for Urd and Cyd are given in the SI (Figures SI5 and SI6, respectively).

upon relaxed scan of this state, leads in a barrierless manner to a CI with the ground-state at a 5'-O–H distance of about 2.1 Å (Fig. 4b). This 5'-O–H distance may be potentially activated along the ground state dynamics according to our previous molecular dynamics simulations in less than 10% of the overall population,[21] ruling out a prominent role in the photophysics. Moreover, this CT state is optically dark and should thus be accessed through $^1\pi\pi^*$ states, $^1\pi_2\pi^*$ in this case, which would act as their doorway. However, following $^1\pi\pi^*$ population we observe an ultrafast nucleobase-centred deactivation,[21] which would be more likely than surmounting the sizable potential energy barrier (~0.6 eV, Figure 4c) featured along the 5'-O-H···O7 reaction coordinate, ruling out H-transfer in the ultrafast (sub-ps) time domain. This holds even when considering the additional stabilisation experienced by CT states in polar solvents[84] as included in our QM/MM scheme, and makes us attribute this channel to the slower ps decays.[19,27]

It is also worth noting that the possibility of a 1'-H···O7 hydrogen transfer has also been recently suggested in Cyd[85] and being related to photo-anomerisation and nucleobase loss reactions characterised theoretically *in vacuo*. These reactions are important due to their involvement in the pre-biotical syntheses of pyrimidine nucleotides,[86,87] where prolonged UV-exposure results in the degradation of biologically irrelevant stereoisomers of Cyd/Urd while preserving their canonical forms. This is therefore related to how life originated and how the present nucleobases were chosen as its building blocks.[88] Despite its importance, we focus on the 5'-O-H···O7 interaction favoured in our molecular dynamics, the 1'-H···O7 channel being beyond the scope of the present study and being considered for an upcoming study. A last channel not accounted for in this work, is the CT mediated by neighbouring water molecules and has been recently described theoretically only for 7H-adenine.[89] We cannot assess this particular pathway due to treating the water molecules at the MM level. On the other hand, our contribution provides the grounds for additional studies in this direction since we can extract important insights form our molecular dynamics simulations on the composition of the solvation shell in order to include the most important water molecules to add to the QM region.

Triplet mediated decays

The last channel considered to partake in the long-lived components registered in pyrimidine nucleobases emerges from intersystem crossing (ISC) events, previously reported in gas phase[50,65,66,90] and that are here addressed in solution with dynamically correlated multireference PESs for the first time. While some results are similar to those obtained in the gas phase[50,65,66,90] (and assumed to be representative also for non-polar solvents), we also find significant differences, as detailed below. As reported in Table 3, along the initially accessed $^1\pi_H\pi^*$ state and its decay, two efficient ISC regions appear: one in the proximity of the FC region, through $(^3n_O\pi^*/^1\pi_H\pi^*)_{ISC}$, and a second one closer to the crossing region with the ground state, via $(^{1,3}\pi_H\pi^*)_{ISC}$. These two ISC regions have spin orbit coupling (SOC) terms for Cyd (~12 and 24 cm^{-1}) larger than for Urd (~9 and 5 cm^{-1}) or Thd (~1 cm^{-1}), and therefore greater singlet-triplet efficiency would be expected in Cyd and Urd in solution compared to their *in vacuo* counterparts (SOC ~1-2 cm^{-1}).[50,90-92] Those channels could in principle contribute to the global population of the triplet manifold, especially in Cyd that features the largest SOC values. On the other hand, the contribution of these crossings is likely small, since their decay is ultrafast while longer timescales are normally required for ISCs to take place.[93] However, we have found a very efficient singlet-triplet crossing region at the vicinities of the different $^1n_O\pi^*$ twisted minima $(^3\pi_H\pi^*/^1n_O\pi^*)_{ISC}$ with much larger SOCs (see Table 3) that guarantees efficient ISC rate in agreement with El-Sayed rules.[94,95] This ISC region, reported for Thymine[92] and Uracil[50] in gas-phase, is also characterised here for Cyd. This pathway, the main contributor towards triplet population in pyrimidine nucleosides, appears to be less efficient in polar than in apolar solvents/gas-phase (gas phase SOC ~ 60 cm^{-1}),[50] which may explain partially the experimental evidence that shows a decrease in the triplet yield in polar solvents (<2%).[96] Additionally, Table 3 shows larger SOCs and more ISC regions are found in Cyd with respect to Urd/Thd, which could partially account for their different triplet yields ($9x10^{-3}$ for Cyd and $3x10^{-3}$ for Urd/Thd),[97,98] even though a more accurate time-dependent treatment would be required for a more robust characterisation.[99] As shown in Table 3, the $^1n_N\pi^*$ state in Cyd was also considered and ruled out due to its negligible SOC compared to those arising from the $^1n_O\pi^*$ state. Moreover, comparison with available phosphorescence maxima in the literature for Thd and Urd in ethanol of ~340 nm (3.65 eV) and ~420 nm (2.95 eV),[72] respectively, can be clearly assigned to the phosphorescent emission from our computed $^3\pi_H\pi^*$ minima placed at 3.53 and 3.20 eV (see Table 2) in Thd and Urd, suggesting direct relaxation along $^3\pi_H\pi^*$ after the ISC events take place.

Table 3. Singlet-triplet crossing points characterised along the relaxation of both $^1\pi\pi^*$ ($^1\pi_H\pi^*$ and $^1\pi_2\pi^*$) and $^1n\pi^*$ ($^1n_O\pi^*$ and $^1n_N\pi^*$) potential energy surfaces. Adiabatic energies are given in eV, and the associated spin orbit coupling (SOC) for each of the crossing points is given in cm^{-1}.

	Urd (eV)	Thd (eV)	Cyd (eV)	
$(^3n_O\pi^*/^1\pi_H\pi^*)_{ISC}$	4.7 (SOC ~ 9 cm^{-1})	4.8 (SOC ~ 1 cm^{-1})	$(^3\pi_2\pi^*/^1\pi_H\pi^*)_{ISC}$	4.55 (SOC~12 cm^{-1})
$(^3\pi_H\pi^*/^1\pi_H\pi^*)_{ISC}$	3.90 (SOC ~ 5 cm^{-1})	4.51 (SOC ~ 1 cm^{-1})	$(^3\pi_H\pi^*/^1\pi_H\pi^*)_{ISC}$	4.01 (SOC ~ 24 cm^{-1})
$(^3\pi_H\pi^*/S_0)_{ISC}$	3.51 (SOC ~ 1 cm^{-1})	3.92 (SOC ~ 1 cm^{-1})	$(^3\pi_H\pi^*/S_0)_{ISC}$	3.85 (SOC ~ 2 cm^{-1})
$(^3\pi_H\pi^*/^1n_O\pi^*)_{ISC}$	4.43 (SOC ~ 42 cm^{-1})	4.67 (SOC ~ 40 cm^{-1})	$(^3\pi_2\pi^*/^1n_O\pi^*)_{ISC}$	4.56 (SOC ~ 42 cm^{-1})
			$(^3\pi_H\pi^*/^1n_N\pi^*)_{ISC}$	4.90 (SOC ~8 cm^{-1})

Spectral signals

To assign the available spectroscopic signals associated to the long-lived components registered in pyrimidine nucleosides, spectral fingerprints (featuring ESAs and photoelectron signals) out of the stationary points characterised in this work have been systematically computed and compared with the available experimental evidence.

Excited state absorptions. The main workhorse behind the experimental characterization of the photophysics of DNA/RNA systems has been fs-TA technique.[2] These experiments provide specific ESAs that depend on the different electronic structure of each of the photoactive state, and are in principle able to provide an indirect signature of their specific population. Our ESA computations rely on the assumption that the individual signals are dominated by the electronic structure at each excited state minimum.

Figure 5 shows a schematic picture of the different channels and associated signals computed, the bottom panels reporting the contributions to the ESA spectrum and displaying the registered experimental signals in brackets for comparison. All systems show intense ESAs in the 300-400 nm range arising from the $^1n_O\pi^*$ state, either at $(^1n_O\pi^*)_{min\text{-}pl}$ or $(^1n_O\pi^*)_{min\text{-}tw}$, its relative intensity increasing by going from planar to twisted minimum; the computed signals are in agreement with the experimental evidence recorded in that wavelength range.[27] This assignment is particularly important for Cyd, as it allows us to assign $^1n_O\pi^*$ as the main actor in the long-lived signal registered, featuring sizable contributions in the 300-400 nm range, instead of $^1n_N\pi^*$.[30] However, recent theoretical studies on gas phase Cytosine micro-hydrated by two water molecules have proposed the potential role of a charge transfer event mediated by the solvent and that could potentially show a fingerprint in the IR regime.[100] Here we note that bulk solvent effects and the presence of the outer solvation shell strongly affect the IR spectra of an excited state with partial CT character, and require careful analysis that will be thoroughly considered in a forthcoming publication where IR signals are analysed. Moreover, the precise position of the water molecules, which is ultimately selected *ad hoc* in previous studies, is expected to impact the result and potentially bias the conclusions attained.[100] Finally, techniques with higher spectral and temporal resolution than those previously used in the IR regime[101] would be required to provide a definitive assignment.[16,102] The most pronounced band in the 450-650 nm region is related to triplet $^3\pi_H\pi^*$ absorption and correlates with the conclusions of an experimental study on Thymine.[28] Considering the similarities between pyrimidine nucleosides, we predict Urd and Cyd to also possess a very similar triplet deactivation mechanism. It is worth noting that the sugar to base proton/hydrogen transfer mechanism, having been postulated theoretically[81,82] but lacking a spectroscopic fingerprint for its assignment, presents low intense bands in the 500-600 nm range, which unfortunately overlaps with the more intense $^1\pi_H\pi^*$ ESAs.[21] However, and given the steep decay along the H-transfer coordinate (Figure 4b), the contributions of the $^1\pi_2\pi^*$ state minimum are expected to be the intermediate that

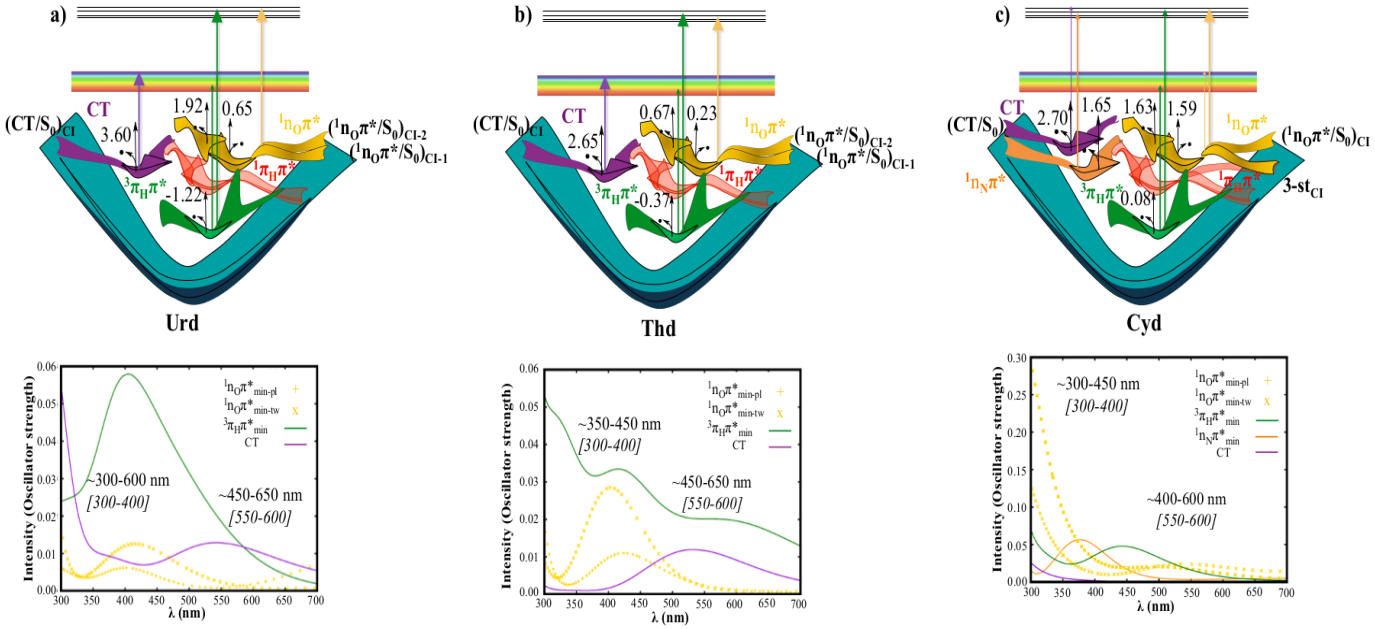

Figure 5. Diagram containing the photochemically relevant mechanisms in water-solvated Urd (a), Thd (b), and Cyd (c) involved in their long-lived decay components and their associated spectroscopic fingerprints. The coloured arrows denote ESAs, their thickness representing their relative intensity (computed values in nm). The corresponding calculated ESA spectra are displayed in nm in the bottom panels (computed ranges are reported together with the corresponding experimental values in brackets). Black arrows correspond to AKEs in eV, negative values refer to the energies required for ionization, *i.e.* the simulated 5.21 eV probe pulses are unable to ionise the molecule.

potentially allows CT population, being the potential fingerprint for this state. We expect TA will not be very efficient in disentangling properly the CT deactivation channels through ESAs as these signals overlap with those of the ultrafast channels at early times, even if it may partially be able to separate contributions at longer times,[71] and that these may require complex multidimensional[16] and X-ray[102] spectroscopies for their unequivocal disentanglement.

Average kinetic energies. Another way to track photoinduced phenomena is based on photoelectron spectroscopy, where a molecule is ionised and the ejected electron and average kinetic energy (AKE) can be monitored. A photoelectron set-up based on the one employed by Buchner *et al.*[20] has been simulated (see Computational Details), yielding AKEs for the different channels characterised in the present study (*cf.* Figure 5). Large AKE values are obtained for the $^1n\pi^*$ state initially accessed planar minima of Ura and Cyd, similar to those previously obtained at the FC region for these systems.[21] This hampers a proper separation of the contributions arising from $^1\pi\pi^*$ and $^1n\pi^*$ at early times and is likely responsible of the proposed lack of $^1n\pi^*$ population according to some recent studies.[20] Thd displays smaller values for the planar minimum yet shows the same trend as Urd by reducing their AKEs when moving from planar to twisted $^1n_O\pi^*$ minima, evidencing a larger singlet to doublet gap traceable by current state-of-the-art instruments in the long-ps timescale. This also applies to Cyd, where planar and twisted minima yield very similar values that could be monitored with the probe pulses considered and provides an additional fingerprint to separate them from the bright $^1\pi\pi^*$ states.[21] However, both planar $^1n_O\pi^*$ and $^1n_N\pi^*$ minima in Cyd depict roughly the same AKEs, not allowing to distinguish which dark state is populated. AKEs were computed on top of the triplet state minima for the first time, triplet to doublet gaps being extremely large, making the probe pulses here employed (5.2 eV) unable to ionise the systems (Cyd being the only case ionised with an AKE ~ 0.08 eV) and thus suggesting the move to even more energetic probe pulses (vacuum UV) for their characterization.[103] The CT state, on the other hand, shows smaller singlet to doublet gaps, readily accessible for photoelectron studies with less energetic probe pulses. Our estimates reveal larger AKEs to those found in the FC region for all systems, being ~3 eV and over an eV larger than those registered for FC, and even larger towards the CI with the ground state, thus being a unique fingerprint for their characterization. However, these do not appear in the available experimental evidence,[20] which suggests a negligible participation of this channel in the overall deactivation mechanism.

Conclusions

Accurate MS-CASPT2/MM energies and gradients with explicit solvation on realistic systems are here used to map the excited state decay channels in water-solvated pyrimidine nucleosides contributing to the long lifetimes (several ps to ns) observed experimentally. The underlying spectral signals (transient absorption, photoelectron signals and phosphorescence emission maxima) are modelled and compared with available observations, clarifying experimental assignments. Our study supports initial population of bright $^1\pi\pi^*$ ($^1\pi_H\pi^*$ and $^1\pi_2\pi^*$) states followed by ultrafast non-adiabatic population branching to the $^1n_O\pi^*$ state at the vicinities of the FC region, which subsequently decays to its minima. These minima act as a sink where the population remains trapped given the large energy barrier required to reach the funnel for non-radiative decay back to the ground state. This mechanism justifies lifetimes of ~100 ps in Urd/Thd and ~30 ps in Cyd, the latter displaying a comparatively smaller barrier mediating its decay through a 3-state CI with $^1\pi_H\pi^*$ and the ground state, not present in the other pyriminidine nucleosides, thus accounting for the faster decay recorded experimentally. We also predict the $^1n_O\pi^*$ minima to act as the main doorway mediating ISC events, its computed SOCs being large and correlating with the phosphorescence quantum yields observed in natural nucleosides. N1 substitution is suggested to hamper the carbonyl stretching and thus increase the lifetime and yield of $^1n_O\pi^*$ population. The possibility of a sugar to nucleobase proton/hydrogen transfer was also considered, requiring large activation energies, its contribution to the sub-ps channels being considered small as opposed to what has been proposed for similar systems *in vacuo*.[81,82] Computed ESAs display sizable contributions in the 300-400 nm range, in agreement with the experimental evidence,[27] that allow us to tentatively assign $^1n_O\pi^*$ as the main state behind the long-lived signal in Cyd, where both $^1n_O\pi^*$ and $^1n_N\pi^*$ could in principle contribute. Our model also displays the characteristic triplet to triplet signals[28] in the 300-550 nm range supporting the ISC scenario for the longer-lived (up to ns) component. Simulated photoelectron signals suggest that dark $^1n\pi^*$ states may be differentiated from the $^1\pi\pi^*$ states thus assisting their further assignment in the long-living (dozens of ps) time regime in contrast with recent reports,[20] and rule out a prominent involvement of the sugar to nucleobase proton/hydrogen transfer route.

This work provides a fundamental step to understand photoinduced events in the DNA/RNA pyrimidine nucleoside building blocks, reconciling available experimental evidence with predictions within a unified deactivation model that accounts for signals spanning from the femto- to the nano-second regime. Despite the absence of quantum dynamical (or semi-classical) simulations, we give the most accurate static picture of the photophysics of DNA/RNA nucleosides including the long-lived spectral signals, and provide the basis for more complex excited state dynamics calculations that can be directly compared with the available experimental evidence. This unified deactivation scenario, where all ultrafast[21] and long-lived events registered arise from common mechanisms among all pyrimidine bases, calls for a shared evolutionary route taken by our genomic material under extreme UV-light exposure in prebiotic conditions,[7,8] where natural selection chose the most photostable and/or suitable building blocks for information storage.[88]

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

M.G. acknowledges support by the European Research Council Advanced Grant STRATUS (ERC-2011-AdG No. 291198). I.R. thanks the French Agence National de la Recherche (FEMTO-2DNA, ANR-15-CE29-0010) and the ENS-Lyon (grants: 900/S81/BS81-14, MI-LOURD-FR15). J.S.-M. and I.R. acknowledge the use of HPC resources of the "Pôle Scientifique de Modélisation Numérique" (PSMN) at the ENS-Lyon, France. R.I. thanks Dr. L. Martínez-Fernández for useful discussions, CNR/CNRS Progetto Bilaterale 2015 and the Chaire D'Alembert program of University Paris-Saclay for financial support.

References

(1) Middleton, C. T.; de La Harpe, K.; Su, C.; Law, Y. K.; Crespo-Hernandez, C. E.; Kohler, B.: DNA Excited-State Dynamics: From Single Bases to the Double Helix. In *Annual Review of Physical Chemistry*; Annual Review of Physical Chemistry; Annual Reviews: Palo Alto, 2009; Vol. 60; pp 217-239.

(2) Crespo-Hernandez, C. E.; Cohen, B.; Hare, P. M.; Kohler, B. Ultrafast excited-state dynamics in nucleic acids. *Chemical Reviews* **2004**, *104*, 1977-2019.

(3) Kleinermanns, K.; Nachtigallova, D.; de Vries, M. S. Excited state dynamics of DNA bases. *Int. Rev. Phys. Chem.* **2013**, *32*, 308-342.

(4) Schreier, W. J.; Gilch, P.; Zinth, W. Early Events of DNA Photodamage. *Annual Review of Physical Chemistry* **2015**, *66*, 497-519.

(5) Giussani, A.; Segarra-Martí, J.; Roca-Sanjuán, D.; Merchán, M.: Excitation of Nucleobases from a Computational Perspective I: Reaction Paths. In *Photoinduced Phenomena in Nucleic Acids I*; Barbatti, M., Borin, A. C., Ullrich, S., Eds.; Topics in Current Chemistry; Springer International Publishing, 2015; Vol. 355; pp 57-97.

(6) Improta, R.; Santoro, F.; Blancafort, L. Quantum Mechanical Studies on the Photophysics and the Photochemistry of Nucleic Acids and Nucleobases. *Chemical Reviews* **2016**, *116*, 3540-3593.

(7) Serrano-Andrés, L.; Merchán, M. Are the five natural DNA/RNA base monomers a good choice from natural selection? A photochemical perspective. *J. Photochem. Photobiol. C-Photochem. Rev.* **2009**, *10*, 21-32.

(8) Beckstead, A. A.; Zhang, Y.; de Vries, M. S.; Kohler, B. Life in the light: nucleic acid photoproperties as a legacy of chemical evolution. *Physical Chemistry Chemical Physics* **2016**, *18*, 24228-24238.

(9) Marchetti, B.; Karsili, T. N. V.; Ashfold, M. N. R.; Domcke, W. A 'bottom up', ab initio computational approach to understanding fundamental photophysical processes in nitrogen containing heterocycles, DNA bases and base pairs. *Physical Chemistry Chemical Physics* **2016**, *18*, 20007-20027.

(10) Improta, R.; Barone, V.: Excited States Behavior of Nucleobases in Solution: Insights from Computational Studies. Topics in Current Chemistry; Springer Berlin Heidelberg, 2014; pp 1-29.

(11) Mai, S.; Richter, M.; Marquetand, P.; González, L.: Excitation of Nucleobases from a Computational Perspective II: Dynamics. Topics in Current Chemistry; Springer Berlin Heidelberg, 2014; pp 1-55.

(12) Barbatti, M.; Aquino, A. J. A.; Szymczak, J. J.; Nachtigallova, D.; Hobza, P.; Lischka, H. Relaxation mechanisms of UV-photoexcited DNA and RNA nucleobases. *Proceedings of the National Academy of Sciences of the United States of America* **2010**, *107*, 21453-21458.

(13) McFarland, B. K.; Farrell, J. P.; Miyabe, S.; Tarantelli, F.; Aguilar, A.; Berrah, N.; Bostedt, C.; Bozek, J. D.; Bucksbaum, P. H.; Castagna, J. C.; Coffee, R. N.; Cryan, J. P.; Fang, L.; Feifel, R.; Gaffney, K. J.; Glownia, J. M.; Martinez, T. J.; Mucke, M.; Murphy, B.; Natan, A.; Osipov, T.; Petrović, V. S.; Schorb, S.; Schultz, T.; Spector, L. S.; Swiggers, M.; Tenney, I.; Wang, S.; White, J. L.; White, W.; Gühr, M. Ultrafast X-ray Auger probing of photoexcited molecular dynamics. *Nat Commun* **2014**, *5*.

(14) Bucher, D. B.; Pilles, B. M.; Carell, T.; Zinth, W. Charge separation and charge delocalization identified in long-living states of photoexcited DNA. *Proceedings of the National Academy of Sciences* **2014**, *111*, 4369-4374.

(15) Bucher, D. B.; Schlueter, A.; Carell, T.; Zinth, W. Watson–Crick Base Pairing Controls Excited-State Decay in Natural DNA. *Angewandte Chemie International Edition* **2014**, *53*, 11366-11369.

(16) Prokhorenko, V. I.; Picchiotti, A.; Pola, M.; Dijkstra, A. G.; Miller, R. J. D. New Insights into the Photophysics of DNA Nucleobases. *The Journal of Physical Chemistry Letters* **2016**, *7*, 4445-4450.

(17) Chatterley, A. S.; West, C. W.; Stavros, V. G.; Verlet, J. R. R. Time-resolved photoelectron imaging of the isolated deprotonated nucleotides. *Chemical Science* **2014**, *5*, 3963-3975.

(18) Xue, B.; Yabushita, A.; Kobayashi, T. Ultrafast dynamics of uracil and thymine studied using a sub-10 fs deep ultraviolet laser. *Physical Chemistry Chemical Physics* **2016**, *18*, 17044-17053.

(19) Ma, C.; Cheng, C. C.-W.; Chan, C. T.-L.; Chan, R. C.-T.; Kwok, W.-M. Remarkable effects of solvent and substitution on the photo-dynamics of cytosine: a femtosecond broadband time-resolved fluorescence and transient absorption study. *Physical Chemistry Chemical Physics* **2015**, *17*, 19045-19057.

(20) Buchner, F.; Nakayama, A.; Yamazaki, S.; Ritze, H.-H.; Lübcke, A. Excited-state relaxation of hydrated thymine and thymidine measured by liquid-jet photoelectron spectroscopy: experiment and simulation. *Journal of the American Chemical Society* **2015**, *137*, 2931-2938.

(21) Pepino, A. J.; Segarra-Martí, J.; Nenov, A.; Improta, R.; Garavelli, M. Resolving Ultrafast Photoinduced Deactivations in Water-Solvated Pyrimidine Nucleosides. *The Journal of Physical Chemistry Letters* **2017**, *8*, 1777-1783.

(22) Onidas, D.; Markovitsi, D.; Marguet, S.; Sharonov, A.; Gustavsson, T. Fluorescence Properties of DNA Nucleosides and Nucleotides: A Refined Steady-State and Femtosecond Investigation. *The Journal of Physical Chemistry B* **2002**, *106*, 11367-11374.

(23) Gustavsson, T.; Banyasz, A.; Lazzarotto, E.; Markovitsi, D.; Scalmani, G.; Frisch, M. J.; Barone, V.; Improta, R. Singlet excited-state behavior of uracil and thymine in aqueous solution: A combined experimental and computational study of 11 uracil derivatives. *Journal of the American Chemical Society* **2006**, *128*, 607-619.

(24) Sobolewski, A. L.; Domcke, W.; Dedonder-Lardeux, C.; Jouvet, C. Excited-state hydrogen detachment and hydrogen transfer driven by repulsive (1)pi sigma* states: A new paradigm for nonradiative decay in aromatic biomolecules. *Physical Chemistry Chemical Physics* **2002**, *4*, 1093-1100.

(25) Sobolewski, A. L.; Domcke, W. On the mechanism of nonradiative decay of DNA bases: ab initio and TDDFT results for the excited states of 9H-adenine. *Eur. Phys. J. D* **2002**, *20*, 369-374.

(26) Canuel, C.; Mons, M.; Piuzzi, F.; Tardivel, B.; Dimicoli, I.; Elhanine, M. Excited states dynamics of DNA and RNA bases:

Characterization of a stepwise deactivation pathway in the gas phase. *Journal of Chemical Physics* **2005**, *122*, 074316.

(27) Hare, P. M.; Crespo-Hernández, C. E.; Kohler, B. Internal conversion to the electronic ground state occurs via two distinct pathways for pyrimidine bases in aqueous solution. *Proceedings of the National Academy of Sciences of the United States of America* **2007**, *104*, 435-440.

(28) Hare, P. M.; Middleton, C. T.; Mertel, K. I.; Herbert, J. M.; Kohler, B. Time-resolved infrared spectroscopy of the lowest triplet state of thymine and thymidine. *Chemical Physics* **2008**, *347*, 383-392.

(29) Vayá, I.; Gustavsson, T.; Miannay, F.-A.; Douki, T.; Markovitsi, D. Fluorescence of Natural DNA: From the Femtosecond to the Nanosecond Time Scales. *Journal of the American Chemical Society* **2010**, *132*, 11834-11835.

(30) Martinez-Fernandez, L.; Pepino, A. J.; Segarra-Martí, J.; Jovaisaite, J.; Vayá, I.; Nenov, A.; Markovitsi, D.; Gustavsson, T.; Banyasz, A.; Garavelli, M.; Improta, R. The Photophysics of Deoxycytidine and 5-methyl-deoxycytidine in Solution: a comprehensive picture by Quantum Mechanical Calculations and Femtosecond Fluorescence Spectroscopy. *Journal of the American Chemical Society* **2017**, *139*, 7780-7791.

(31) Tommasi, S.; Denissenko, M. F.; Pfeifer, G. P. Sunlight Induces Pyrimidine Dimers Preferentially at 5-Methylcytosine Bases. *Cancer Research* **1997**, *57*, 4727-4730.

(32) Esposito, L.; Banyasz, A.; Douki, T.; Perron, M.; Markovitsi, D.; Improta, R. Effect of C5-Methylation of Cytosine on the Photoreactivity of DNA: A Joint Experimental and Computational Study of TCG Trinucleotides. *Journal of the American Chemical Society* **2014**, *136*, 10838-10841.

(33) Finley, J.; Malmqvist, P. A.; Roos, B. O.; Serrano-Andrés, L. The multi-state CASPT2 method. *Chemical Physics Letters* **1998**, *288*, 299-306.

(34) Case, D.; Darden, T. A.; Cheatham, T. E.; Simmerling, C.; Wang, J.; Duke, R.; Luo, R.; Crowley, M.; Walker, R.; Zhang, W.; Merz, K. M.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Kollman, P.: *Amber 11*: University of California, San Francisco, 2011.

(35) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber biomolecular simulation programs. *Journal of Computational Chemistry* **2005**, *26*, 1668-1688.

(36) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **1983**, *79*, 926-935.

(37) Feig, M.; Karanicolas, J.; Brooks, C. L. MMTSB Tool Set: enhanced sampling and multiscale modeling methods for applications in structural biology. *Journal of Molecular Graphics and Modelling* **2004**, *22*, 377-395.

(38) Altoé, P.; Stenta, M.; Bottoni, A.; Garavelli, M. A tunable QM/MM approach to chemical reactivity, structure and physico-chemical properties prediction. *Theoretical Chemistry Accounts* **2007**, *118*, 219-240.

(39) Frisch, M. J. T., G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgometry Jr, J. A; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Coss, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J.: GAUSSIAN 09 Rev. B01. GAUSSIAN 09, Rev. B01 ed.; Gaussian, Inc.: Wallingford, CT, 2010.

(40) Bearpark, M. J.; Robb, M. A.; Bernhard Schlegel, H. A direct method for the location of the lowest energy point on a potential surface crossing. *Chemical Physics Letters* **1994**, *223*, 269-274.

(41) Bearpark, M. J.; Larkin, S. M.; Vreven, T. Searching for Conical Intersections of Potential Energy Surfaces with the ONIOM Method : Application to Previtamin D. **2008**, 7286-7295.

(42) Roca-Sanjuán, D.; Aquilante, F.; Lindh, R. Multiconfiguration second-order perturbation theory approach to strong electron correlation in chemistry and photochemistry. *Wiley Interdisciplinary Reviews-Computational Molecular Science* **2012**, *2*, 585-603.

(43) Andersson, K.; Malmqvist, P. A.; Roos, B. O. 2nd-order perturbation-theory with a complete active space self-consistent field reference function. *Journal of Chemical Physics* **1992**, *96*, 1218-1226.

(44) Aquilante, F.; Autschbach, J.; Carlson, R.; Chibotaru, L.; Delcey, M. G.; De Vico, L.; Fernández Galvan, I.; Ferré, N.; Frutos, L. M.; Gagliardi, L.; Garavelli, M.; Giussani, A.; Hoyer, C.; Li Manni, G.; Lischka, H.; Ma, D.; Malmqvist, P. A.; Müller, T.; Nenov, A.; Olivucci, M.; Pedersen, T. B.; Peng, D.; Plasser, F.; Pritchard, B.; Reiher, M.; Rivalta, I.; Schapiro, I.; Segarra-Martí, J.; Stenrup, M.; Truhlar, D. G.; Ungur, L.; Valentini, A.; Vancoillie, S.; Veryazov, V.; Vysotskiy, V.; Weingart, O.; Zapata, F.; Lindh, R. Molcas 8: New Capabilities for Multiconfigurational Quantum Chemical Calculations across the Periodic Table. *Journal of Computational Chemistry* **2016**, *37*, 506-541.

(45) Ghigo, G.; Roos, B. O.; Malmqvist, P.-Å. A modified definition of the zeroth-order Hamiltonian in multiconfigurational perturbation theory (CASPT2). *Chemical Physics Letters* **2004**, *396*, 142-149.

(46) Hehre, W. J.; Ditchfield, R.; Pople, J. A. Self—Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian—Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *The Journal of Chemical Physics* **1972**, *56*, 2257-2261.

(47) Widmark, P. O.; Malmqvist, P. A.; Roos, B. O. DENSITY-MATRIX AVERAGED ATOMIC NATURAL ORBITAL (ANO) BASIS-SETS FOR CORRELATED MOLECULAR WAVE-FUNCTIONS .1. 1ST ROW ATOMS. *Theor. Chim. Acta* **1990**, *77*, 291-306.

(48) Widmark, P. O.; Joakim, B.; Persson; Roos, B. O. DENSITY-MATRIX AVERAGED ATOMIC NATURAL ORBITAL (ANO) BASIS-SETS FOR CORRELATED MOLECULAR WAVE-FUNCTIONS .2. 2ND ROW ATOMS. *Theor. Chim. Acta* **1991**, *79*, 419-432.

(49) Iliaš, M.; Kellö, V. r.; Visscher, L.; Schimmelpfennig, B. Inclusion of mean-field spin–orbit effects based on all-electron two-component spinors: Pilot calculations on atomic and molecular properties. *The Journal of Chemical Physics* **2001**, *115*, 9667-9674.

(50) González-Luque, R.; Climent, T.; González-Ramírez, I.; Merchán, M.; Serrano-Andrés, L. Singlet-Triplet States Interaction Regions in DNA/RNA Nucleobase Hypersurfaces. *Journal of Chemical Theory and Computation* **2010**, *6*, 2103-2114.

(51) Cohen, B.; Crespo-Hernandez, C. E.; Kohler, B. Strickler-Berg analysis of excited singlet state dynamics in DNA and RNA nucleosides. *Faraday Discussions* **2004**, *127*, 137-147.

(52) Peon, J.; Zewail, A. H. DNA/RNA nucleotides and nucleosides: direct measurement of excited-state lifetimes by femtosecond fluorescence up-conversion. *Chemical Physics Letters* **2001**, *348*, 255-262.

(53) Nenov, A.; Giussani, A.; Segarra-Martí, J.; Jaiswal, V. K.; Rivalta, I.; Cerullo, G.; Mukamel, S.; Garavelli, M. Modeling the high-energy electronic state manifold of adenine: Calibration for nonlinear electronic spectroscopy. *The Journal of Chemical Physics* **2015**, *142*, 212443.

(54) Giussani, A.; Segarra-Martí, J.; Nenov, A.; Rivalta, I.; Tolomelli, A.; Mukamel, S.; Garavelli, M. Spectroscopic fingerprints of DNA/RNA pyrimidine nucleobases in third-order nonlinear electronic spectra. *Theoretical Chemistry Accounts* **2016**, *135*, 1-18.

(55) Segarra-Martí, J.; Jaiswal, V. K.; Pepino, A. J.; Giussani, A.; Nenov, A.; Mukamel, S.; Garavelli, M.; Rivalta, I. Two-dimensional electronic spectroscopy as a tool for tracking molecular conformations in DNA/RNA aggregates. *Faraday Discussions* **2017**, DOI:10.1039/C1037FD00201G.

(56) Kwok, W.-m.; Ma, C.; Phillips, D. L. Femtosecond Time- and Wavelength-Resolved Fluorescence and Absorption Spectroscopic Study of the Excited States of Adenosine and an Adenine Oligomer. *Journal of the American Chemical Society* **2006**, *128*, 11894-11905.

(57) Allouche, A.-R. Gabedit—A graphical user interface for computational chemistry softwares. *Journal of Computational Chemistry* **2011**, *32*, 174-182.

(58) Thompson, A. L.; Martinez, T. J. Time-resolved photoelectron spectroscopy from first principles: Excited state dynamics of benzene. *Faraday Discussions* **2011**, *150*, 293-311.

(59) Roca-Sanjuán, D.; Rubio, M.; Merchán, M.; Serrano-Andrés, L. Ab initio determination of the ionization potentials of DNA and RNA nucleobases. *Journal of Chemical Physics* **2006**, *125*, 084302.

(60) Slavíček, P.; Winter, B.; Faubel, M.; Bradforth, S. E.; Jungwirth, P. Ionization Energies of Aqueous Nucleic Acids: Photoelectron Spectroscopy of Pyrimidine Nucleosides and ab Initio Calculations. *Journal of the American Chemical Society* **2009**, *131*, 6460-6467.

(61) Segarra-Martí, J.; Merchán, M.; Roca-Sanjuán, D. Ab initio determination of the ionization potentials of water clusters (H2O)(n) (n=2-6). *Journal of Chemical Physics* **2012**, *136*, 244306.

(62) Werner, U.; Mitrić, R.; Bonačić-Koutecký, V. Simulation of time resolved photoelectron spectra with Stieltjes imaging illustrated on ultrafast internal conversion in pyrazine. *The Journal of Chemical Physics* **2010**, *132*, 174301.

(63) Arbelo-González, W.; Crespo-Otero, R.; Barbatti, M. Steady and Time-Resolved Photoelectron Spectra Based on Nuclear Ensembles. *Journal of Chemical Theory and Computation* **2016**, *12*, 5037-5049.

(64) Ruckenbauer, M.; Mai, S.; Marquetand, P.; González, L. Photoelectron spectra of 2-thiouracil, 4-thiouracil, and 2,4-dithiouracil. *The Journal of Chemical Physics* **2016**, *144*, 074303.

(65) Richter, M.; Mai, S.; Marquetand, P.; Gonzalez, L. Ultrafast intersystem crossing dynamics in uracil unravelled by ab initio molecular dynamics. *Physical Chemistry Chemical Physics* **2014**, *16*, 24423-24436.

(66) González-Vázquez, J.; González, L. A Time-Dependent Picture of the Ultrafast Deactivation of keto-Cytosine Including Three-State Conical Intersections. *ChemPhysChem* **2010**, *11*, 3617-3624.

(67) Hudock, H. R.; Levine, B. G.; Thompson, A. L.; Satzger, H.; Townsend, D.; Gador, N.; Ullrich, S.; Stolow, A.; Martinez, T. J. Ab initio molecular dynamics and time-resolved photoelectron spectroscopy of electronically excited uracil and thymine. *Journal of Physical Chemistry A* **2007**, *111*, 8500-8508.

(68) Mercier, Y.; Santoro, F.; Reguero, M.; Improta, R. The decay from the dark npi* excited state in uracil: an integrated CASPT2/CASSCF and PCM/TD-DFT study in the gas phase and in water. *The journal of physical chemistry. B* **2008**, *112*, 10769-10772.

(69) Improta, R.; Barone, V.; Lami, A.; Santoro, F. Quantum Dynamics of the Ultrafast ππ*/nπ* Population Transfer in Uracil and 5-Fluoro-Uracil in Water and Acetonitrile. *The Journal of Physical Chemistry B* **2009**, *113*, 14491-14503.

(70) Hare, P. M.; Crespo-herna, 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 1 n π * State. *Journal of Physical Chemistry B* **2006**, *110*, 18641-18650.

(71) Takaya, T.; Su, C.; Harpe, K. d. L.; Crespo-Hernández, C. E.; Kohler, B. UV excitation of single DNA and RNA strands produces high yields of exciplex states between two stacked bases. *Proceedings of the National Academy of Sciences of the United States of America* **2008**, *105*, 10285-10290.

(72) Görner, H. Phosphorescence of nucleic acids and DNA components at 77 K. *Journal of Photochemistry and Photobiology B: Biology* **1990**, *5*, 359-377.

(73) Li, Q.; Giussani, A.; Segarra-Martí, J.; Nenov, A.; Rivalta, I.; Voityuk, A. A.; Mukamel, S.; Roca-Sanjuán, D.; Garavelli, M.; Blancafort, L. Multiple Decay Mechanisms and 2D-UV Spectroscopic Fingerprints of Singlet Excited Solvated Adenine-Uracil Monophosphate. *Chemistry – A European Journal* **2016**, *22*, 7497-7507.

(74) Segarra-Martí, J.; Francés-Monerris, A.; Roca-Sanjuán, D.; Merchán, M. Assessment of the Potential Energy Hypersurfaces in Thymine within Multiconfigurational Theory: CASSCF vs. CASPT2. *Molecules* **2016**, *21*, 1666.

(75) Segarra-Martí, J.; Garavelli, M.; Aquilante, F. Multiconfigurational Second-Order Perturbation Theory with Frozen Natural Orbitals Extended to the Treatment of Photochemical Problems. *Journal of Chemical Theory and Computation* **2015**, *11*, 3772-3784.

(76) Yamazaki, S.; Taketsugu, T. Nonradiative Deactivation Mechanisms of Uracil, Thymine, and 5-Fluorouracil: A Comparative ab Initio Study. *Journal of Physical Chemistry A* **2012**, *116*, 491-503.

(77) Perun, S.; Sobolewski, A. L.; Domcke, W. Conical intersections in thymine. *Journal of Physical Chemistry A* **2006**, *110*, 13238-13244.

(78) Nakayama, A.; Arai, G.; Yamazaki, S.; Taketsugu, T. Solvent effects on the ultrafast nonradiative deactivation mechanisms of thymine in aqueous solution: Excited-state QM/MM molecular dynamics simulations. *The Journal of Chemical Physics* **2013**, *139*, 214304.

(79) Kistler, K. A.; Matsika, S. Photophysical pathways of cytosine in aqueous solution. *Physical Chemistry Chemical Physics* **2010**, *12*, 5024-5031.

(80) Blancafort, L. Energetics of Cytosine Singlet Excited-State Decay Paths—A Difficult Case for CASSCF and CASPT2†. *Photochemistry and Photobiology* **2007**, *83*, 603-610.

(81) Tuna, D.; Sobolewski, A. L.; Domcke, W. Mechanisms of Ultrafast Excited-State Deactivation in Adenosine. *The Journal of Physical Chemistry A* **2014**, *118*, 122-127.

(82) Tuna, D.; Domcke, W. Excited-state deactivation in 8-oxo-deoxyguanosine: comparison between anionic and neutral forms. *Physical Chemistry Chemical Physics* **2016**, *18*, 947-955.

(83) Camillis, S. D.; Miles, J.; Alexander, G.; Ghafur, O.; Williams, I. D.; Townsend, D.; Greenwood, J. B. Ultrafast non-radiative decay of gas-phase nucleosides. *Physical Chemistry Chemical Physics* **2015**, *17*, 23643-23650.

(84) Yin, H.; Ma, Y.; Mu, J.; Liu, C.; Rohlfing, M. Charge-Transfer Excited States in Aqueous DNA: Insights from Many-Body Green's Function Theory. *Physical Review Letters* **2014**, *112*, 228301-228301.

(85) Szabla, R.; Campos, J.; Sponer, J. E.; Sponer, J.; Gora, R. W.; Sutherland, J. D. Excited-state hydrogen atom abstraction initiates the photochemistry of [small beta]-2[prime or minute]-deoxycytidine. *Chemical Science* **2015**, *6*, 2035-2043.

(86) Powner, M. W.; Gerland, B.; Sutherland, J. D. Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. *Nature* **2009**, *459*, 239-242.

(87) Powner, M. W.; Anastasi, C.; Crowe, M. A.; Parkes, A. L.; Raftery, J.; Sutherland, J. D. On the Prebiotic Synthesis of Ribonucleotides: Photoanomerisation of Cytosine Nucleosides and Nucleotides Revisited. *ChemBioChem* **2007**, *8*, 1170-1179.

(88) Sutherland, J. D. The Origin of Life—Out of the Blue. *Angewandte Chemie International Edition* **2016**, *55*, 104-121.

(89) Barbatti, M. Photorelaxation Induced by Water–Chromophore Electron Transfer. *Journal of the American Chemical Society* **2014**, *136*, 10246-10249.

(90) Merchán, M.; Serrano-Andrés, L.; Robb, M. A.; Blancafort, L. Triplet-state formation along the ultrafast decay of excited singlet cytosine. *Journal of the American Chemical Society* **2005**, *127*, 1820-1825.

(91) Etinski, M.; Fleig, T.; Marian, C. A. Intersystem Crossing and Characterization of Dark States in the Pyrimidine Nucleobases Uracil, Thymine, and 1-Methylthymine. *Journal of Physical Chemistry A* **2009**, *113*, 11809-11816.

(92) Serrano-Perez, J. J.; Gonzalez-Luque, R.; Merchan, M.; Serrano-Andres, L. On the intrinsic population of the lowest triplet state of thymine. *Journal of Physical Chemistry B* **2007**, *111*, 11880-11883.

(93) Marian, C. M. Spin–orbit coupling and intersystem crossing in molecules. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2012**, *2*, 187-203.

(94) El-Sayed, M. A. Spin—Orbit Coupling and the Radiationless Processes in Nitrogen Heterocyclics. *The Journal of Chemical Physics* **1963**, *38*, 2834-2838.

(95) El-Sayed, M. A. Triplet state. Its radiative and nonradiative properties. *Accounts of Chemical Research* **1968**, *1*, 8-16.

(96) Salet, C.; Bensasson, R.; Becker, R. S. TRIPLET EXCITED STATES OF PYRIMIDINE NUCLEOSIDES AND NUCLEOTIDES. *Photochemistry and Photobiology* **1979**, *30*, 325-329.

(97) Rahn, R. O.; Shulman, R. G.; Longworth, J. W. The UV-induced triplet state in DNA. *Proceedings of the National Academy of Sciences of the United States of America* **1965**, *53*, 893-896.

(98) Rahn, R. O.; Shulman, R. G.; Longworth, J. W. Phosphorescence and Electron Spin Resonance Studies of the uv-Excited Triplet State of DNA. *The Journal of Chemical Physics* **1966**, *45*, 2955-2965.

(99) Mai, S.; Marquetand, P.; González, L. A general method to describe intersystem crossing dynamics in trajectory surface hopping. *International Journal of Quantum Chemistry* **2015**, *115*, 1215-1231.

(100) Szabla, R.; Kruse, H.; Sponer, J.; Gora, R. W. Water-chromophore electron transfer determines the photochemistry of cytosine and cytidine. *Physical Chemistry Chemical Physics* **2017**, *19*, 17531-17537.

(101) Quinn, S.; Doorley, G. W.; Watson, G. W.; Cowan, A. J.; George, M. W.; Parker, A. W.; Ronayne, K. L.; Towrie, M.; Kelly, J. M. Ultrafast IR spectroscopy of the short-lived transients formed by UV excitation of cytosine derivatives. *Chem. Commun.* **2007**, 2130-2132.

(102) Wolf, T. J. A.; Myhre, R. H.; Cryan, J. P.; Coriani, S.; Squibb, R. J.; Battistoni, A.; Berrah, N.; Bostedt, C.; Bucksbaum, P.; Coslovich, G.; Feifel, R.; Gaffney, K. J.; Grilj, J.; Martinez, T. J.; Miyabe, S.; Moeller, S. P.; Mucke, M.; Natan, A.; Obaid, R.; Osipov, T.; Plekan, O.; Wang, S.; Koch, H.; Gühr, M. Probing ultrafast ππ*/nπ* internal conversion in organic chromophores via K-edge resonant absorption. *Nature Communications* **2017**, *8*, 29.

(103) Horton, S. L.; Liu, Y.; Chakraborty, P.; Matsika, S.; Weinacht, T. Vibrationally assisted below-threshold ionization. *Physical Review A* **2017**, *95*, 063413.

1