The Ultrafast Dynamics of the Two Lowest Bright Excited States of Cytosine and 1-Methyl-Cytosine: A Quantum Dynamical Study

Martha Yaghoubi Jouybari,[†] Yanli Liu,[‡] Roberto Improta,[*,¶] and Fabrizio Santoro[*,†]

†*CNR–Consiglio Nazionale delle Ricerche, Istituto di Chimica dei Composti Organo Metallici (ICCOM-CNR), SS di Pisa, Area della Ricerca, via G. Moruzzi 1, I-56124 Pisa, Italy*

‡*School of Physics and Optoelectronics Engineering, Ludong University, 264025 Yantai, Shandong, PR China*

¶*CNR-Consiglio Nazionale delle Ricerche, Istituto di Biostrutture e Bioimmagini (IBB-CNR), via Mezzocannone 16, I-80136 Napoli, Italy*

E-mail: robimp@unina.it; fabrizio.santoro@pi.iccom.cnr.it

Abstract

The nonadiabatic quantum dynamics (QD) of cytosine and 1-methyl-cytosine in gas phase is simulated for 250 fs after a photoexcitation to one of the first two bright states. The nuclear wavepacket is propagated on the coupled diabatic potential energy surfaces of the lowest seven excited states, including $\pi\pi^*$, $n\pi^*$ and Rydberg states along all the vibrational degrees of freedom. We focus in particular on the interplay between the bright and the dark $n\pi^*$ states, not considering the decay to the ground electronic state. To run these simulations we implemented an automatic general procedure to parametrize linear vibronic coupling (LVC) models with time-dependent density functional theory (DFT) computations, and interfaced it with Gaussian package. The wavepacket was propagated with the multilayer version of the multiconfigurational time dependent Hartree method. Two different density functionals, PBE0 and CAM-B3LYP, which provide a different description of the relative stability of the lowest energy dark states, were used to parametrize the LVC Hamiltonian. Part of the photoexcited population on lowest HOMO-LUMO transition ($\pi_H\pi_L^*$) decays within less than 100 fs to a $n\pi^*$ state which mainly involves a promotion of an electron from the oxygen lone pair to the LUMO ($n_O\pi_L^*$). The population of the second $\pi\pi^*$ state decays almost completely, in < 100 fs, not only to $\pi_H\pi_L^*$ and to $n_O\pi_L^*$ state, but also to another $n\pi_L^*$ states involving the nitrogen lone pair. The efficiency of the adopted protocol allowed us to check the accuracy of the predictions by repeating the QD simulations with different LVC Hamiltonians parametrized either at the ground state minimum or at stationary structures of different relevant excited states.

1 Introduction

The photoactivated dynamics of DNA bases (nucleobases) is the object of an ongoing research activity, not only due its biological relevance,[1–3] but also because their potentialities as *molecular laboratories* for basic methodological and modellistic advances.[4,5] On the one hand, time-resolved (TR) experiments have shown that photoexcited nucleobases undergo effective ultrafast nonradiative decay paths.[1,2,6] They decrease photodamage in DNA and, therefore, act as a self-protecting mechanisms against UV light irradiation.[7] On the other end, the same experiments show that excited state dynamics of nucleobases is extremely complex, being characterized by different time-components spanning several order of magnitudes, from a few dozen of fs up to the ns.[1,2] As a matter of fact, experimental lifetimes depend on many different chemical physical effects, e.g. substituents of the nucleobases, solvent, temperature, excitation wavelength etc.[1,2,6] These results strongly suggest the involvement of many excited states and different decay paths.[4,5] Attaining an accurate static and dynamical theoretical description of these phenomena is thus a very difficult task and, in this respect, cytosine (Cyt) is an ideal example of the many challenges to be tackled.

In the last twenty years, Cyt and its derivatives have been thoroughly studied by different TR spectroscopy[8–14] and computational techniques,[12,13,15–26] both in the gas and in the condensed phase,[27–32] indicating the coexistence of multiple decay paths. A detailed review of these studies falls outside the scope of the present paper and those more relevant to our study will be analysed in the Discussion section. Most of the TR experimental studies agree on the presence of an ultrafast component ($\leq$ 100 fs),[8–10] together with at least one on the ps time-scale (in the range 0.5$\sim$3 ps)[8–10,33] and, finally, a slower one of a few ns.[10,11,34]There is general agreement that the lowest energy bright $\pi\pi^*$ state decays in sub-ps to

ps timescale.[4,5] Both in the gas phase and in solution,[35–38] the signature of an excited state decaying on an intermediate time-scale (from tens to hundreds of picoseconds) is found. The first obstacle for interpreting these results is that Cyt can exist in different tautomers: the enol, the keto-amino, and the keto-imino[39–42] species. In the gas phase, though both the enol and the keto-amino tautomers are almost isoenergetic, the enol form is slightly more stable ($\sim$0.03 eV).[43–45] In solution and when the N1 atom bears a sugar (as in DNA) or simply a methyl substituent the keto-amino is instead largely the predominant tautomer.[31,46] Our study is thus focused on this species and, if not otherwise specified, we shall refer to the Cyt keto-amino tautomer (Figure 1). Steady state quantum-mechanical (QM) calculations on Cyt show that several excited states, both singlet and triplets, with different nature lay within 1 eV. Considering only singlet valence transitions, five different excited states, corresponding to two bright $\pi\pi^*$ and three dark nπ^* transitions (involving either the lone pair on the oxygen $n_O\pi^*$ or on the nitrogen $n_N\pi^*$), fall within this energy range, i.e. in the region of the two lowest-energy UV absorption bands.[4,5,15,31] Several non-radiative decay pathways have been mapped, with at least three potentially accessible conical intersections (CoI) with the ground state (S_0).[4,5,26]

Few dynamical simulations of photoexcited Cyt have been reported based on trajectory surface hopping (TSH)[18,20,21,47] or ab initio multiple spawning (AIMS)[22] in combination with different levels of CASSCF and they agree in predicting an ultrafast involvement ($<$ 100 fs) of states with nπ^* character. TSH simulations with OM2/MRCI semiempirical method predicts on the contrary a direct decay from the bright state to the ground state with 350 fs.[47]

Notwithstanding the many studies carried out in the last decade, time evolution of Cyt remains unclear. While there is a general agreement on the assignment of

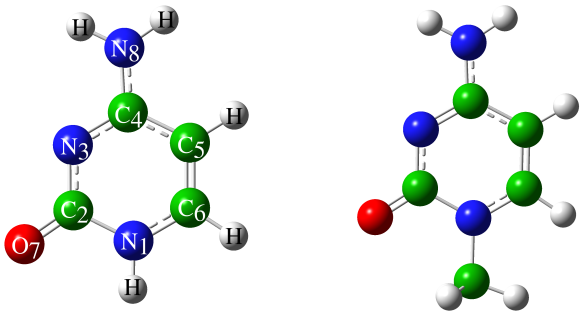
the ns time components to triplet states, many key points concerning the ps and sub-ps time components are still lively debated. For example, there is disagreement on the preferential nonradiative decay route to the ground state via the different possible CoIs and on the assignment of the dark state decaying on the dozens of ps time-scale.[4] There are indications that this state can be populated also within polynucleotides, increasing its biological relevance.[48] Actually the relative energy of the lowest energy excited states has not been definitively assessed, since it strongly depends on the computational method adopted and on its details (basis set, active space etc).[4,17]

In this scenario, full quantum dynamical (QD) studies of the decay of photoexcited Cyt would be very useful, but are still lacking. Previous comparisons on model systems have shown that QD and semiclassical predictions become different in strong-coupling regimes,[49] and, as we will show below, inter-state couplings in Cyt are remarkable. QD effects might be more and more relevant as the experimental set up reach the capability to deliver data with unprecedented time-resolution (< 20 fs[50]). As a first step in this direction, in this contribution we present a full QD investigation of the decay processes in the gas phase of Cyt and 1-methyl-cytosine (1MeCyt, Figure 1), as a minimum model for cytidine nucleoside. We use Linear Vibronic Coupling (LVC) models including all the first seven lowest electronic states and run QD simulations with the Multilayer Multiconfigurational Time Dependent Hartree (ML-MCTDH) method.[51–55] LVC models have been parametrized with time-dependent (TD) Density Functional Theory (TD-DFT). We built up an automatized diabatization interfaced with Gaussian 16[56] based on a maximum-overlap method.[57] It is similar to the one already proposed by Neugebauer et al.[58] but uses an alternative definition of the overlap of transition densities that include both excitation and de-excitation terms.

Concerning the analysis of the dynamics of Cyt, we focus on the singlet excited states; triplet pathways are investigated with TSH calculations in ref.[19,20] Additionally, the adoption of LVC harmonic models imply approximations that restrict the reliability of our approach to the ultrafast regime ($\sim$ 100 fs) in which the molecular structure remains close to planarity. This prevents the description of the complete decay to the ground-state. The time-window we monitor is nonetheless particularly meaningful since, in the presence of a very fast non radiative decay route for the bright state, the first 100 fs are the critical ones for the possible population of the dark states. On the other side a strong point of our analysis is that our LVC models allow an efficient description of the coupled QD on several excited states (7 in our case, including all the relevant $\pi\pi^*$ and $n\pi^*$ states) and considering all the nuclear degrees of freedom. Finally, to the best of our knowledge we present the first dynamical simulation of the decay after excitation of the second bright electronic state. Its role in fact deserves a closer scrutiny. The broad absorption spectrum of Cyt suggests a strong vibronic coupling between the two lowest-energy bright excited states,[59] and the radiation wavelength most commonly used in pump-probe experiments on oligonucleotides excite both states.[38] At the same time, there are computational and experimental hints that they could be the origin of different decay pathways. For example, at least in water, the dynamics strongly depends on the wavelength of the exciting radiation.[38] The excited state dynamics of Cyt and its derivatives thus appears ruled by the interplay of, at least, four close lying excited states, which are strongly vibronically coupled.

Beside the intrinsic relevance for Cyt, this study gives us the opportunity to tackle some issues of more general methodological interest. Not only we present our approach to parametrize LVC Hamiltonians from TD-DFT, but we also anal-

yse few strategies to evaluate the adequacy of the LVC models and the range of confidence of their prediction. In particular, in this contribution we build different Hamiltonians at different key structures for the photodynamics, namely the minima of the ground state and of the most involved excited states. Additionally, we also show the utility to recompute the TD-DFT states along the trajectory of the average position of the wavepacket. In fact this allows to check both the reliability of the model PES and the electronic character of the adiabatic states. Finally, thanks to the efficiency of the computational scheme, we repeated all calculations with two different DFT functionals, namely CAM-B3LYP[60] and PBE0.[61,62] They have been widely used to study, among several class of compounds, the photophysics of nucleobases and oligonucleotides.[4] In particular, CAM-B3LYP provides a picture of the excited states behavior of Cyt close to what obtained at the CC-SSD[23] and the MS-CASPT2 levels in water[31] The two functionals deliver a rather different description of the interplay between bright and dark excited states of Cyt.[4,23,63] As a consequence they enable us to test how the dynamics is affected by the shape and the position of the different PES.



(a) Cytosine (b) 1-Methyl-Cytosine

Figure 1: Molecular structures of the systems investigated in this work.

2 Parameterization of LVC Hamiltonians

3 The Model

We write our LVC Hamiltonian using ground state (S_0) dimensionless normal coordinates $\mathbf{q}$ (column vector), and conjugated momenta $\mathbf{p}$, considering a n dimensional diabatic basis $|\mathbf{d}^{(r)}\rangle$ (a row vector, as indicated by the superscript "(r)"), $|\mathbf{d}^{(r)}\rangle = (|d_1\rangle, |d_2\rangle, \cdots, |d_n\rangle)$

$$H = \sum_i \left(K + V_{ii}^{dia}(\mathbf{q})\right) |d_i\rangle\langle d_i| + \sum_{i,j>i} V_{ij}^{dia}(\mathbf{q})(|d_i\rangle\langle d_j| + |d_j\rangle\langle d_i|) \tag{1}$$

K is the kinetic term and the diagonal (V_{ii}^{dia}) potential energy surfaces (PES) are written in harmonic approximation, assuming that all the excited diabatic states have the same normal modes and frequencies of S_0. Moreover, the couplings among the diabatic states, $V_{ij}^{dia}(\mathbf{q})$ $(i \neq j)$, are taken as linear functions of $\mathbf{q}$

$$K = \frac{1}{2}\mathbf{p}^T\mathbf{\Omega}\mathbf{p} \tag{2}$$

$$V_{ii}^{dia}(\mathbf{q}) = E_i^0 + \boldsymbol{\lambda}_{ii}^T\mathbf{q} + \frac{1}{2}\mathbf{q}^T\mathbf{\Omega}\mathbf{q}, \tag{3}$$

$$V_{ij}^{dia}(\mathbf{q}) = \boldsymbol{\lambda}_{ij}^T\mathbf{q}. \tag{4}$$

In Eqs. 2-4 $\mathbf{\Omega}$ is a diagonal matrix, Ω_{ii} being the S_0 frequency of the normal mode q_i. The superscript T indicates that the transpose is taken. The time-dependent (TD) vibronic wavefunction of the system is computed solving the TD Schrödinger Equation. It is written as $|\Psi(\mathbf{q},t)\rangle = \sum_i |d_i\rangle|\Psi_i(\mathbf{q},t)\rangle$ so that electronic populations are simply $P_i(t) = \langle\Psi_i(\mathbf{q},t)|\Psi_i(\mathbf{q},t)\rangle$.

The LVC Hamiltonian is parametrized from DFT and TD-DFT calculations

with a diabatization procedure. The diabatic states are taken to be coincident with the adiabatic states $|a_i\rangle$ computed at a reference geometry. In practice, the molecular geometry is displaced along each dimensionless normal mode q_α by a quantity $\pm\Delta_\alpha$ and the diabatic states are defined as the combination of the adiabatic states that resemble as much as possible the reference states. This procedure allows to obtain the transformation

$$|\mathbf{d}^{(r)}\rangle = |\mathbf{a}^{(r)}(\Delta_\alpha)\rangle\mathbf{D}(\Delta_\alpha) \tag{5}$$

between diabatic $|\mathbf{d}^{(r)}\rangle$ and adiabatic states $|\mathbf{a}^{(r)}(\Delta_\alpha)\rangle)$, where the matrix $\mathbf{D}$ has elements $D_{ji}(\Delta_\alpha) = \langle a_j(\Delta_\alpha)|d_i\rangle$. At each geometry the matrix of the diabatic potentials $\mathbf{V}^{dia}(\Delta_\alpha)$ is computed from the diagonal adiabatic one $\mathbf{V}^{ad}(\Delta_\alpha) = diag(E_1^{ad}(\Delta_\alpha), E_2^{ad}(\Delta_\alpha), \cdot\cdot, E_n^{ad}(\Delta_\alpha))$ as

$$\mathbf{V}^{dia}(\Delta_\alpha) = \mathbf{D}^T(\Delta_\alpha)\mathbf{V}^{ad}(\Delta_\alpha)\mathbf{D}(\Delta_\alpha). \tag{6}$$

The parameters $\lambda_{ij}(\alpha)$ are obtained from Eq. 6 by numerical differentiation.[59]

3.1 Finding the diabatic-adiabatic transformation D

We define a set of reference states $|\mathbf{R}^{(r)}\rangle = \{|R_1\rangle, |R_2\rangle, \cdot\cdot, |R_n\rangle\}$, which are identical to the lowest n adiabatic states at the reference geometry $\mathbf{0}$, $|\mathbf{R}^{(r)}\rangle = |\mathbf{a}(\mathbf{0})^{(r)}\rangle$, and we consider a generic displacement Δ_α. The row vector of the n lowest adiabatic states at this geometry is $|\mathbf{a}(\Delta_\alpha)^{(r)}\rangle$ and we compute the overlap matrix with elements

$$S_{AB} = \langle R_A|a_B(\Delta_\alpha)\rangle \tag{7}$$

In section S1.1 of the Supporting Information (SI) we show that the trans-

formation matrix from adiabatic to diabatic states in Eq. 5 can be computed as

$$\mathbf{D} = \mathbf{S}^T(\mathbf{S}\mathbf{S}^T)^{-\frac{1}{2}}. \tag{8}$$

Therefore the LVC parametrization is straightforward if we are able to compute the matrix **S**. More precisely, this transformation defines pseudo-diabatic states, since an infinite basis set would be needed to define truly diabatic states. In the following for the sake of brevity we always use the term "diabatic" also to define our "pseudo-diabatic" states. In TD-DFT excited-states are actually characterized through transition densities, which in a super-operator formalism can be written as

$$|\rho\rangle\rangle = \sum_{ir} X_{ir}|\Phi_r\rangle\langle\Phi_i| + \sum_{ir} Y_{ir}|\Phi_i\rangle\langle\Phi_r| \tag{9}$$

where i, j denote occupied Kohn-Sham (KS) molecular orbitals (MOs) and r, s virtual ones. The inner product between two transition densities can be computed as

$$\begin{aligned}\langle\langle\rho^A|\rho^B\rangle\rangle &= Tr_{\mathbf{\Phi}}[\\ &\left(\sum_{ir} X_{ir}^{MO,A}|\Phi_i\rangle\langle\Phi_r| - \sum_{ir} Y_{ir}^{MO,A}|\Phi_r\rangle\langle\Phi_i|\right) \\ &\left(\sum_{js} X_{js}^{MO,B}|\Phi_s\rangle\langle\Phi_j| + \sum_{js} Y_{js}^{MO,A}|\Phi_j\rangle\langle\Phi_s|\right)]\end{aligned} \tag{10}$$

where we added a superscript MO to specify that the transition densities are written on the basis of MOs. If both transition densities are computed at the same geometry, due to the orthogonality of the MOs we have:

$$\langle\langle \rho^A | \rho^B \rangle\rangle = \sum_{ir} X_{ir}^{MO,A} X_{ir}^{MO,B} - \sum_{ir} Y_{ir}^{MO,A} Y_{ir}^{MO,B} \tag{11}$$

which is equivalent to the orthonormalization condition reported in refs.[64,65]

Consider now the transition densities relative to state "A" at the reference geometry ($|\rho^A(0)\rangle\rangle$) and state "B" at the geometry perturbed by Δ_α along mode α ($|\rho^B(\Delta\alpha)\rangle\rangle$)

$$|\rho^A(0)\rangle\rangle = \sum_{ir} X_{ir}^{MO,A,0} |\Phi_r(0)\rangle\langle\Phi_i(0)| + \sum_{ir} Y_{ir}^{MO,A,0} |\Phi_i(0)\rangle\langle\Phi_r(0)| \tag{12}$$

$$|\rho^B(\Delta_\alpha)\rangle\rangle = \sum_{js} X_{js}^{MO,B,\Delta\alpha} |\Phi_s(\Delta_\alpha)\rangle\langle\Phi_j(\Delta_\alpha)| + \sum_{js} Y_{js}^{MO,B,\Delta\alpha} |\Phi_j(\Delta_\alpha)\rangle\langle\Phi_s(\Delta_\alpha)| \tag{13}$$

By performing the trace operation in Eq. 10 the inner product $S_{AB} = \langle\langle \rho(0)^A | \rho^B(\Delta_\alpha)\rangle\rangle$ is

$$\begin{aligned} S_{AB} = & \sum_{ir,js} X_{ir}^{MO,0} X_{js}^{MO,\Delta\alpha} \langle\Phi_r(0)|\Phi_s(\Delta_\alpha)\rangle\langle\Phi_j(\Delta_\alpha)|\Phi_i(0)\rangle \\ & - \sum_{ir,js} Y_{ir}^{MO,0} Y_{js}^{MO,\Delta\alpha} \langle\Phi_i(0)|\Phi_j(\Delta_\alpha)\rangle\langle\Phi_s(\Delta_\alpha)|\Phi_r(0)\rangle \\ & + \sum_{ir,js} X_{ir}^{MO,0} Y_{js}^{MO,\Delta\alpha} \langle\Phi_r(0)|\Phi_j(\Delta_\alpha)\rangle\langle\Phi_s(\Delta_\alpha)|\Phi_i(0)\rangle \\ & - \sum_{ir,js} Y_{ir}^{MO,0} X_{js}^{MO,\Delta\alpha} \langle\Phi_i(0)|\Phi_s(\Delta_\alpha)\rangle\langle\Phi_j(\Delta_\alpha)|\Phi_r(0)\rangle \end{aligned} \tag{14}$$

where on the rhs we dropped the "A" and "B" labels since this does not generate confusion. At variance with what done in ref.,[58] our definition of the overlap S_{AB} include both excitation and de-excitations terms. We prefer this formulation because it uses the TD-DFT response vectors exactly as they are

defined. As a consequence, it allows for instance a more precise fulfillment of orthonormalization conditions. Practically speaking, for many systems results will probably be similar, but for systems where TD-DFT and Tamm Damkoff approximations show significant differences, like cyanines,[66,67] appreciable changes in the LVC parameters may be expected. For the following we follow ref[58] and move to a basis of atomic orbitals (AO), specified with greek indices, obtaining

$$
\begin{aligned}
S_{AB} \;=\; & \sum_{\mu\lambda,\sigma\nu} X^{AO,0}_{\mu\lambda} X^{AO,\Delta\alpha}_{\sigma\nu} S^{AO}_{\lambda,\nu}(0) S^{AO}_{\mu,\sigma}(0) + \\
& \sum_{\mu\lambda,\sigma\nu} Y^{AO,0}_{\mu\lambda} Y^{AO,\Delta\alpha}_{\sigma\nu} S^{AO}_{\lambda,\nu}(0) S^{AO}_{\mu,\sigma}(0) + \\
& \sum_{\mu\lambda,\sigma\nu} X^{AO,0}_{\mu\lambda} Y^{AO,\Delta\alpha}_{\sigma\nu} S^{AO}_{\mu,\nu}(0) S^{AO}_{\lambda,\sigma}(0) + \\
& \sum_{\mu\lambda,\sigma\nu} Y^{AO,0}_{\mu\lambda} X^{AO,\Delta\alpha}_{\sigma\nu} S^{AO}_{\mu,\nu}(0) S^{AO}_{\lambda,\sigma}(0)
\end{aligned}
\tag{15}
$$

To obtain the last equation, as done in ref.,[58] we approximated the the overlap matrix of the AOs at the reference and at the perturbed geometry Δ_α, with the one at the reference geometry S^{AO}_0. Further details are given in section S1.2 of SI. Eq.15 allows us to compute the element S_{AB} of the overlap matrix $\mathbf{S}$ in Eq. 7 thus completing the definition of our LVC model.

The entire procedure for LVC parametrization was implemented in a "in house" Fortran 90 code, which is available upon request, interfaced with Gaussian 16.[56]

4 Computational Details

All electronic calculations were performed with Gaussian16,[56] adopting DFT for S_0 and TD-DFT for excited states. We employed two different functionals, CAM-

B3LYP[60] and PBE0,[61,62] in combination with the 6-31+G(d,p) basis set. We focused on the keto-amino tautomer and included in our LVC model the lowest 7 excited states comprising the first two bright transitions and all the close lying states. In order to parametrize our LVC Hamiltonians we adopted reference states computed in C_s symmetry as this allows to decouple by symmetry $\pi\pi^*$ (A') and $n\pi^*$ (A") states. To that end we performed ground state optimizations imposing C_s symmetry. For Cytosine the stationary state obtained at CAM-B3LYP is a true minimum while the PBE0 structure exhibits a very small imaginary frequency corresponding to the pyramidalization of the nitrogen of the amino group. The same mode is characterized by very small imaginary frequencies also in the S_0 optimized structures of 1MeCyt in C_s symmetry. Releasing the C_s constrains we get true minima with energies only negligibly lower than the planar stationary structures. This indicates that the S_0 PES is extremely shallow along this mode so that these imaginary frequencies were simply neglected, turning them to real. In Section 4 of the SI we investigate this issue more in detail and also show that small changes in the frequency assigned to this mode or even its removal have a negligible effect on the time-evolution of the electronic populations.

In some cases LVC Hamiltonians were also parametrized at C_s stationary structures of $\pi\pi^*$ and $n\pi^*$ excited states. They are also characterized by a few imaginary frequencies reported in the SI. However, since we used S_0 normal modes, they don't affect LVC parameters. In any case, it is important to highlight that we run full-coordinate dynamics, and though our model is parametrized on planar minima, our wavepacket visits also non-planar configurations. Cartesian coordinates of all optimized structures are reported in the SI. As already done in ref.,[68] in order to obtain robust results for the LVC parameters (Eq. 6) we used a tight option for SCF convergence and a convergence threshold for TD-DFT energy equal

to 10^{-6} au, i.e. the recommended value[56] for the closely-related computation of TD-DFT numerical gradients. We performed full-dimensionality QD calculations considering all the system normal modes (33 for Cyt and 42 for 1Me-Cyt) with the ML-MCTDH method implemented in Quantics package.[69] We used a variable mean field (VMF) with a Runge-Kutta integrator of order 5 and accuracy 10^{-7}, like in the provided examples for S_2/S_1 dynamics of pyrazine with 24 normal modes.[70] As a primitive basis set we adopted an Hermite DVR representation. The same approach was already adopted for the $\pi\pi^*/n\pi^*$ decay of Thymine in gas-phase[68] and in aqueous solution[71] and for the computation of the spectrum of Cyt in gas phase.[59] Convergence tests of the ML-MCTDH propagation with respect to the dimension of the primitive basis set and the number of single particle functions (SPFs) are reported in Section S2 of the SI.

5 Results: Cytosine

5.1 Adiabatic Minima

In agreement with previous studies,[4] the lowest-energy excited states of Cyt in the Franck-Condon (FC) region (Table 1) can be conveniently described on the base of 2 Lone Pairs (LPs), 2 π bonding, 2 π^* antibonding MOs (Figures S4, S5 and S12 of the SI). The LPs are localized on N3 (n_N) or on O7 (n_O). The π bonding MOs are the HOMO (π_H), with strong bonding character with respect to the C5=C6 bond, and the second highest energy π one (π_1), with strong contribution from the amino LP. The π^* MOs are the LUMO, π_L^* with strong antibonding character with respect to the C5=C6 bond, and the second lowest energy π^*, π_2^*, with strong antibonding character with respect to the C2=O7 double bond .

Several excited states in the FC point arise from the combinations of different

Table 1: Energies (E^a_{FC}), electronic characters, oscillator strengths δ_{OPA}, and main contributions in terms of transitions among KS-MOs for the lowest excited states of keto-amino tautomer of cytosine at the ground-state minimum (FC point). The energies of the corresponding diabatic states in their minima (E^d_{min}), estimated by the LVC Hamiltonian, are also reported. CAM-B3LYP and PBE0 calculations with 6-31+G(d,p) basis set in gas phase. Energies in eV with respect to the ground state in its minimum. For comparison, literature results with other first principle methods are reported in the note to this Table

	CAM-B3LYP					PBE0				
State	E^a_{FC}	E^d_{min}	character	δ_{OPA}	Transition	E^a_{FC}	E^d_{min}	character	δ_{OPA}	Transition
S_1	5.00	4.56	$\pi_H\pi^*_L$	**0.06**	H→ L	4.78	4.29	$\pi_H\pi^*_L$	**0.04**	H→L
S_2	5.31	4.79	$n_N\pi^*_L$	0.00	H-3→L H-2→L	4.98	4.57	$n_O\pi^*_L$+ $n_N\pi^*_L$	0.00	H-1→ L
S_3	5.77	5.63	$\pi_H Ry_\sigma 1$	0.00	H→L+1	5.34	4.44	$n_O\pi^*_L$- $n_N\pi^*_L$	0.00	H-3→L
S_4	5.88	5.16	$n_O\pi^*_2$+ $n_O\pi^*_L$	0.00	H-2→L+4 H-2→L	5.59	5.45	$\pi_H Ry_\sigma 1$	0.00	H→L+1
S_5	5.94	5.56	$\pi_1\pi^*_L$	**0.13**	H-1→L	5.60	5.22	$\pi_1\pi^*_L$	**0.09**	H-2→L
S_6	6.10	5.51	$n_O\pi^*_2$-$n_O\pi^*_L$	0.00	H-2→L+4 H-2→L	5.81	5.27	$n_O\pi^*_2$	0.00	H-1→L+3
S_7	6.29	6.23	$\pi_1 Ry_\sigma 2$	0.00	H-1→L+1 H-1→L+2	6.05	5.88	$\pi_H Ry_\sigma 2$	0.00	H→L+2

MS-CASPT2/TZVP data (eV):[72] 4.68 ($\pi_H\pi^*_L$), 5.54($n_N\pi^*_L$), 5.12 ($n_O\pi^*_L$), 5.54 ($\pi_1\pi^*_L$).
CR-EOM-CCSD(T)/aug-cc-pVTZ data (eV):[73] 4.69 ($\pi_H\pi^*_L$), 5.18($n_N\pi^*_L$), 5.82 ($n_O\pi^*_L$), 5.60 ($\pi_1\pi^*_L$).

excitations (Table 1) and some of the KS-MOs themselves (Figure S4 of SI) receive contributions from the two LPs. To obtain an easier assignment, we therefore rely on natural transition orbitals (NTOs), sketched in Figures 2 and 3 for CAM-B3LYP and PBE0, respectively.

According to both functionals, the lowest energy excited state (S_1) is bright, corresponds to the HOMO→LUMO excitation and will be labeled as $\pi_H\pi^*_L$. The following bright state, S_5, is a (π_1)→LUMO excitation, will be labeled as $\pi_1\pi^*_L$. The energy gap between $\pi_H\pi^*_L$ and $\pi_1\pi^*_L$ is slightly larger according to CAM-B3LYP (0.94 vs. 0.82 eV). Both functionals predict the presence of two dark A"-symmetry Rydberg states respectively $\sim$ 0.8 eV and $\sim$ 1.2 eV less stable than $\pi_H\pi^*_L$.

The most significant disagreement between the two functionals concerns the

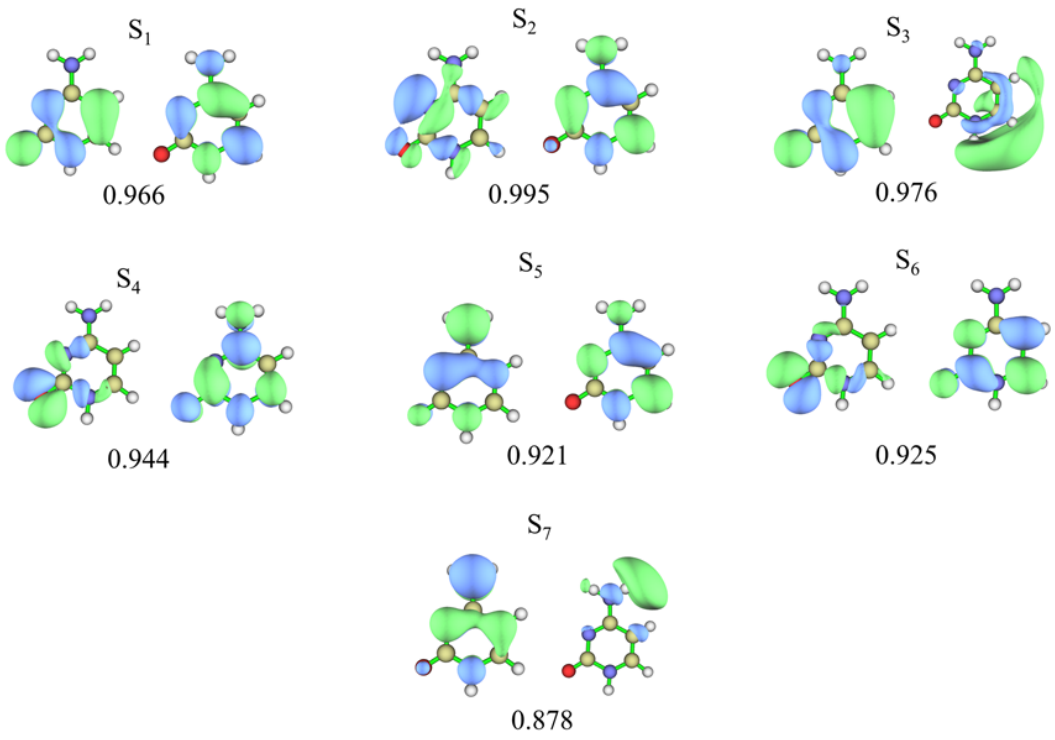


Figure 2: Natural transition orbitals (NTOs) of Cytosine in gas phase at ground state geometry using CAM-B3LYP functional and an isovalue 0.04. The weight with which each transition contributes to the corresponding excited states is also reported.

$n\pi^*$ transitions, whose assignment is made less straightforward by the strong interaction between the involved MOs. With this caveat, according to CAM-B3LYP, S_2 (hereafter labeled $n_N\pi_L^*$) mainly corresponds to the transition from n_N to the LUMO. S_4 and S_6 derive from the strong mixing of the transition from n_O to π_L^* and to π_2^*. PBE0 predicts instead that S_2 and S_3 essentially corresponds to the symmetric and antisymmetric combinations of $n_N\pi_L^*$ and $n_O\pi_L^*$, whereas S_6 is a more pure $n_O \rightarrow \pi_2^*$ transition. From the quantitative point of view, on the average, the $n\pi^*$ are $\sim$0.1 eV closer to $\pi_H\pi_L^*$ according to PBE0.

The relative energy of the excited states in the FC region provided by CAM-B3LYP is similar to that of CR-EOM-CCSD(T),[73] but for a small relative stabilization of the dark excited states, and MS-CASPT2[72] (see the note to Table

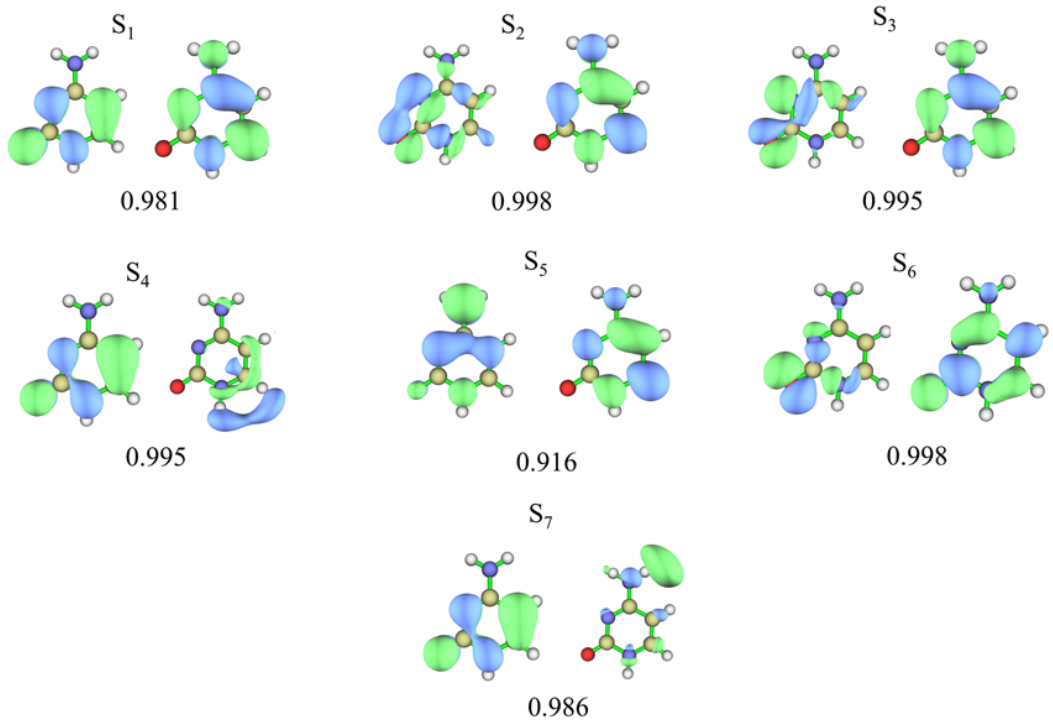


Figure 3: Natural transition orbitals (NTOs) of Cytosine in gas phase at ground state geometry using PBE0 functional and an isovalue 0.04. The weight with which each transition contributes to the corresponding excited states is also reported.

1). This latter method would also predict that $n_O\pi^*$ is more stable than $n_N\pi^*$. However, the strong mixing between these dark states makes their labeling only qualitative. Additional comparisons with other electronic methods are reported in Section S5 of the SI. It can be noticed that CASSCF calculations appear to underestimate the energy gap between $\pi_H\pi_L^*$ and the lowest $n\pi^*$ state.

Tables 2 and 3 report the energies of the adiabatic states in the $n\pi^*$ minima that we were able to locate with CAM-B3LYP and PBE0, respectively, and their electronic characters. CAM-B3LYP geometry optimizations of S_2 in C_s symmetry leads to a state with a clear-cut $n_N\pi_L^*$ character (on the S_1 PES), whereas starting from S_3 we reach a stationary point on the S_2 adiabatic PES with clear $n_O\pi_L^*$ character. When adopting PBE0, geometry optimizations of both S_2 and

S_3 lead instead to a minimum with $n_O\pi^*_L$ character lying on the S_1 PES. Further characterization of these minima can be found in the SI (sections S3.3-S3.5 for CAM-B3LYP, and S4.4-S4.5 for PBE0).

Table 2: Energies (E^a) of the adiabatic states of cytosine in the minima of the first $\pi\pi^*$ and first two $n\pi^*$ states and their electronic characters. E^d_{FC} are the energies, extrapolated at the FC point, of the diabatic states defined by LVC Hamiltonians parametrized at these minima. Energies in eV with respect to the ground state in its minimum. CAM-B3LYP/6-31+G(d,p) calculations in gas phase.

CAM-B3LYP								
First $\pi\pi^*$			First $n\pi^*$			Second $n\pi^*$		
E^a	E^d_{FC}	character	E^a	E^d_{FC}	character	E^a	E^d_{FC}	character
4.52	4.99	$\pi_H\pi^*_L$	4.67	5.35	$n_N\pi^*_L$	4.64	5.33	$\pi_H\pi^*_L$
4.96	5.79	$n_O\pi^*_L$	5.32	5.04	$\pi_H\pi^*_L$	4.80	5.94	$n_O\pi^*_L$
5.53	5.30	$n_N\pi^*_L$	6.16	5.89	$n_O\pi^*_L$	5.94	5.54	$n_N\pi^*_L$
5.90	5.78	$\pi_1\pi^*_L$	6.29	5.93	$\pi_1\pi^*_L$	6.26	6.01	$n_O\pi^*_2$
6.14	5.86	$n_O\pi^*_2$	6.59	6.64	$n_N\pi^*_2$	6.38	5.94	$\pi_1\pi^*_L$
6.16	5.83	$\pi_H Ry_\sigma 1$	6.60	5.81	$\pi_1 Ry_\sigma 1$	6.46	6.20	$\pi_H Ry_\sigma 1$
6.65	6.73	$\pi_1 Ry_\sigma 1$	6.97	6.32	$\pi_1 Ry_\sigma 2$	6.69	6.89	$n_O Ry_\sigma 1$

Table 3: Energies (E^a) of the adiabatic states of cytosine in the minima of the first $\pi\pi^*$ and $n\pi^*$ states and their electronic characters. E^d_{FC} are the energies, extrapolated at the FC point, of the diabatic states defined by LVC Hamiltonians parametrized at these minima. Energies in eV with respect to the ground state in its minimum. PBE0/6-31+G(d,p) calculations in gas phase.

PBE0					
First $\pi\pi^*$			First $n\pi^*$		
E^a	E^d_{FC}	character	E^a	E^d_{FC}	character
4.10	4.96	$n_O\pi^*_L$	4.06	5.14	$n_O\pi^*_L$
4.14	4.80	$\pi_H\pi^*_L$	4.18	5.05	$\pi_H\pi^*_L$
5.46	5.09	$n_N\pi^*_L$	5.69	5.35	$n_N\pi^*_L$
5.84	5.46	$\pi_1\pi^*_L$	5.86	5.79	$n_O\pi^*_2$
5.84	5.61	$n_O\pi^*_2$	6.07	5.97	$\pi_H Ry_\sigma 1$
5.97	5.70	$\pi_H Ry_\sigma 1$	6.08	5.77	$\pi_1\pi^*_L$
6.16	6.23	$n_O Ry_\sigma 2$	6.16	6.30	$n_O Ry_\sigma 2$

This concise analysis highlights that the different diabatic states are strongly mixed and that their coupling changes with the coordinates. This feature will

require, as detailed below, ad-hoc procedures to analyse our simulations. On the other hand, in their minima, the $n\pi^*_L$ adiabatic states acquire a well-defined electronic character, i.e. $n_O\pi^*_L$ and $n_N\pi^*_L$ states, which we will use as main 'electronic reference states' (together with $\pi_H\pi^*_L$ and $\pi_1\pi^*_L$) to interpret our QD simulations.

5.2 LVC Hamiltonian parametrized at the FC point

The adiabatic states computed at the FC point, reported in Table 1, are taken as a reference to define a set of 7 diabatic states. Truly diabatic states do not change with the coordinates. Altough this condition cannot be considered completely enforced in our pseudo-diabatic states, their electronic character does not change and can therefore be used to label them. Within a LVC framework, diabatic-states minima can be easily located by using their gradient in the FC point.

The reorganization energies E_r, i.e. the energy stabilization of the minimum (E^d_{min}) with respect to the FC point (E^a_{FC}) of the $\pi\pi^*$ and $n\pi^*$ states are rather similar, being moderately larger for $n\pi^*$ ones. Therefore, the energy order of the different minima is generally the same as the order in the FC point. The only exceptions concern the minimum of $n_O\pi^*_2$-$n_O\pi^*_L$, more stable than the minimum of $\pi_1\pi^*_L$ according to CAM-B3LYP, and the minimum of $n_O\pi^*_L$-$n_N\pi^*_L$, more stable than the one of $n_O\pi^*_L$+$n_N\pi^*_L$ according to PBE0. As expected, for states with Rydberg character E_r is very small. In their minima all the dark states considered are more stable than the FC energy of at least one of the spectroscopic states, so they can be in principle populated during the dynamics.

Analysis of the off-diagonal terms of the LVC Hamiltonian provides additional insights on the factors determining the photoactivated dynamics. The norm of the vector λ_{ij} (see Tables 4 for CAM-B3LYP and 5 for PBE0) gives a concise measure of the couplings between states i and j ($i \neq j$). The largest couplings

occur between nπ^* states, according to both functionals. Remarkable couplings also exist between $\pi\pi^*$ and nπ^* valence states while, as expected, they both have small couplings with Rydberg states. The diagonal elements of the Tables ($i = j$), are the norm of the diabatic energy gradients. On the ground of the previous discussion on the reorganization energies, it is not surprising that the largest values belong to nπ^* states and the smallest values to Rydberg ones.

Table 4: Norm of the coupling vectors between the different diabatic states (eV) for cytosine according to the LVC model parametrized at FC position. CAM-B3LYP/6-31+G(d,p) calculations in gas phase

STATE	$\pi_H\pi_L^*$	$n_N\pi_L^*$	$\pi_H Ry_\sigma 1$	$n_O\pi_2^*$+$n_O\pi_L^*$	$\pi_1\pi_L^*$	$n_O\pi_2^*$-$n_O\pi_L^*$	$\pi_1 Ry_\sigma 2$
$\pi_H\pi_L^*$	0.386						
$n_N\pi_L^*$	0.128	0.358					
$\pi_H Ry_\sigma 1$	0.075	0.018	0.231				
$n_O\pi_2^*$+$n_O\pi_L^*$	0.095	0.145	0.063	0.547			
$\pi_1\pi_L^*$	0.163	0.134	0.029	0.073	0.358		
$n_O\pi_2^*$-$n_O\pi_L^*$	0.075	0.144	0.031	0.285	0.063	0.485	
$\pi_1 Ry_\sigma 2$	0.028	0.014	0.094	0.024	0.071	0.014	0.206

Table 5: Norm of the coupling vectors between the different diabatic states (eV) for cytosine according to the LVC model parametrized at FC position. PBE0/6-31+G(d,p) calculations in gas phase.

STATE	$\pi_H\pi_L^*$	$n_O\pi_L^*$+$n_N\pi_L^*$	$n_O\pi_L^*$- $n_N\pi_L^*$	$\pi_H Ry_\sigma 1$	$\pi_1\pi_L^*$	$n_O\pi_2^*$	$\pi_H Ry_\sigma 2$
$\pi_H\pi_L^*$	0.423						
$n_O\pi_L^*$+$n_N\pi_L^*$	0.100	0.367					
$n_O\pi_L^*$-$n_N\pi_L^*$	0.122	0.329	0.546				
$\pi_H Ry_\sigma 1$	0.086	0.016	0.023	0.234			
$\pi_1\pi_L^*$	0.160	0.143	0.102	0.034	0.358		
$n_O\pi_2^*$	0.020	0.181	0.118	0.010	0.020	0.477	
$\pi_H Ry_\sigma 2$	0.096	0.010	0.009	0.130	0.059	0.010	0.261

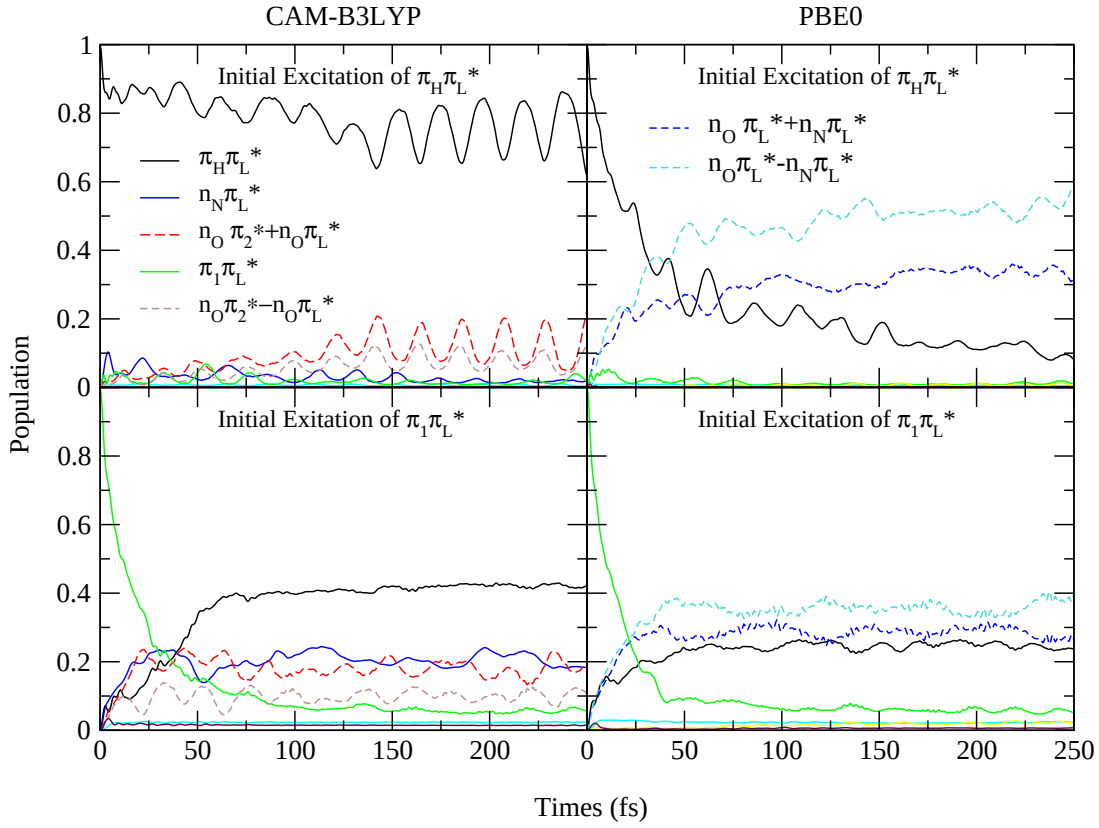


Figure 4: Nonadiabatic dynamics of electronic populations of cytosine in gas phase predicted with LVC Hamiltonians parameterized with CAM-B3LYP (left) or PBE0 (right) calculations, after an initial photoexcitation to $\pi_H\pi_L^*$ (top) or $\pi_1\pi_L^*$ (bottom). Notice that in all electronic population calculations we include 7 states but in all figures we label only the states which are significantly populated.

5.3 Electronic Population Dynamics

5.3.1 Excitation to the first bright state $\pi_H\pi_L^*$

In Figure 4 we show the time evolution of the electronic population of the excited states of Cyt in gas phase for an initial excitation on $\pi_H\pi_L^*$ (top) or $\pi_1\pi_L^*$ (bottom), according to CAM-B3LYP (left) and PBE0 (right). The time evolution of an initial population (P) of $\pi_H\pi_L^*$ drastically changes by using these two functionals. According to PBE0 both combinations of $n_O\pi_L^*$ and $n_N\pi_L^*$ are quickly populated and, at 150 fs the antisymmetric state's population is ∼0.5 and the symmetric one

~ 0.3. As a consequence, $\pi_H\pi_L^*$ has lost 80% of its initial population.

Interestingly $P(n_0\pi_L^* - n_N\pi_L^*) > P(n_0\pi_L^* + n_N\pi_L^*)$ although the former state is less stable at the FC point. On the other hand, $(n_0\pi_L^* - n_N\pi_L^*)$ has a much larger E_r and in its minimum is 0.13 eV more stable than $(n_0\pi_L^* + n_N\pi_L^*)$.

At variance, CAM-B3LYP predicts that most of the population remains on $\pi_H\pi_L^*$ since the energy gap with the other states is larger. On the other hand a significant flow to the dark states is obtained also with CAM-B3LYP. More in detail, after a small transient transfer to $n_N\pi_L^*$, in ~ 150 fs ~ 30 % of the population flows on the two strongly coupled $n_O\pi_2^*$+$n_O\pi_L^*$ and $n_O\pi_2^*$-$n_O\pi_L^*$, although they are higher in energy than $n_N\pi^*$. Furthermore their populations appear to have a very similar (in-phase) time evolution which is out-of-phase with the residual population on $n_N\pi_L^*$. Both PBE0 and CAM-B3LYP predict that none of the other electronic states included in the LVC model (like Rydberg ones) are significantly populated.

To get further insight on the interplay between the diabatic states, in Figure 5 we report as a function of time the diabatic (top) and adiabatic (bottom) energies according to the LVC Hamiltonian in the average position of the wavepacket. Since by definition the average position preserves C_s symmetry only diabatic states with the same symmetry mix to produce the adiabatic states. The diabatic energy on $\pi_H\pi_L^*$ clearly oscillates as a results of periodic vibrations on collective CC and CO stretchings.

According to PBE0 $n_O\pi_L^*$-$n_N\pi_L^*$ crosses $n_O\pi_L^*$+$n_N\pi_L^*$ and becomes more stable at t$\sim$ 25 fs, getting quite close to the $\pi_H\pi_L^*$. This explains why $P(n_O\pi_L^* - n_N\pi_L^*) > P(n_O\pi_L^* + n_N\pi_L^*)$ for $t > 25$ fs. The large splitting of the adiabatic curves however shows that these two diabatic states strongly couple for $t > 0$, so much that only one adiabatic state remains very close in energy to $\pi_H\pi_L^*$. Therefore, within an adiabatic framework, the $\pi_H\pi_L^*$ populations flows in a state that is a combination

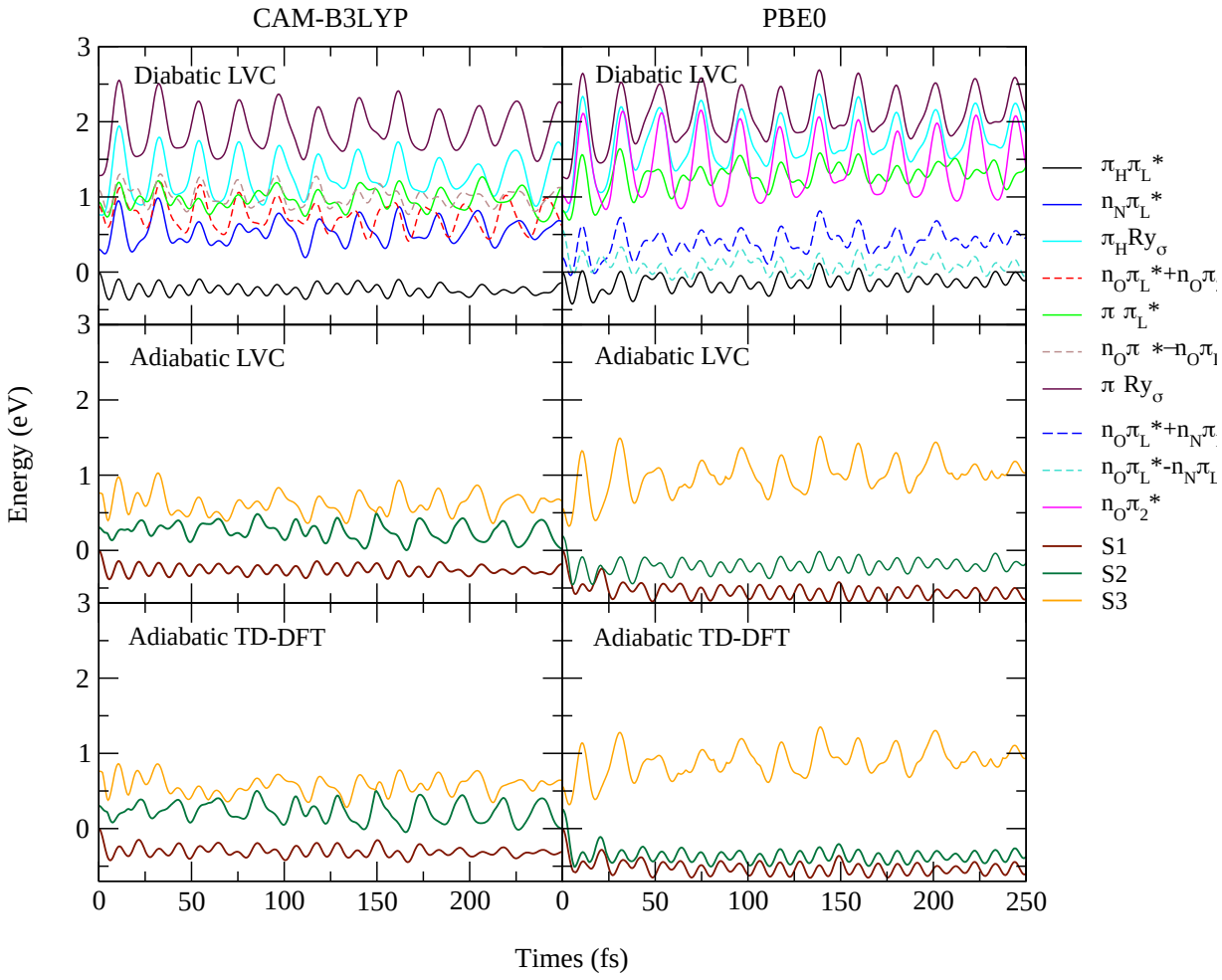


Figure 5: Diabatic (top) and adiabatic potential energies (middle, only for the three lowest states) in the average position of the wavepacket (C_s symmetry) obtained by LVC Hamiltonians parameterized with CAM-B3LYP (left) or PBE0 (right) calculations. The energy zero corresponds to the FC energy of the $\pi_H\pi_L^*$ state, i.e. respectively 5.00 eV for CAM-B3LYP and 4.78 eV for PBE0. For comparison we also report (bottom) the three lowest adiabatic energies recomputed by TD-DFT at molecular structures corresponding to the average position of the wavepacket. All 7 adiabatic potentials predicted by LVC model are reported in Figure S20 in the SI.

of these two diabatic states.

According to CAM-B3LYP all the diabatic states with nπ^* character oscillate at energies significantly higher than $\pi_H\pi_L^*$ (blue, red and brown lines). However, in terms of adiabatic states a combination of them (dark green line in the middle

panel) get rather close in energy to the lowest-energy state S_1, with $\pi_H\pi_L^*$ character after 100 fs. In particular the energy gap becomes minimum at ~ 140 and ~ 165 fs, corresponding to times where the population transfer from $\pi_H\pi_L^*$ to $n_O\pi_2^*$+$n_O\pi_L^*$ and $n_O\pi_2^*$-$n_O\pi_L^*$ is maximum (see Figure 4).

As a further analysis, in the bottom panels of Figure 5 we report the adiabatic energies at the average position of the wavepacket computed at TD-DFT level. At each time we take the average position in normal coordinates, generate the corresponding Cartesian coordinates, and repeat a TD-DFT calculation. The energies and time evolution of the true adiabatic TD-DFT states (bottom panels) and the model LVC adiabatic states (middle panels) are similar, confirming the reliability of our LVC model. This also offers us the possibility to analyse the excited states transition densities and verify what are the LPs more involved in the dynamics.

According to PBE0 results S_2 is almost degenerate with $\pi_H\pi_L^*$ and, for times t> 10-20 fs, exhibits a remarkable gap with S_3. In CAM-B3LYP dynamics S_2 is more separated from S_1, while it becomes periodically almost degenerate with S_3. We have then performed a NTO analysis of the TD-DFT electron density each 50 fs (see SI). According to PBE0, the almost degenerate S_1 and S_2 states flip between $n_O\pi_L^*$ and $\pi_H\pi_L^*$. On the contrary S_3 is always a $n\pi_L^*$ state with prevalent $n_N\pi_L^*$ character. CAM-B3LYP predicts instead that S_1 is always $\pi_H\pi_L^*$. At most of the times, S_2 and S_3 have a clear $n_O\pi_L^*$ and $n_N\pi_L^*$ character, respectively. However, when they get close in energy they mix and can switch the prevalent character.

5.3.2 Excitation to the second bright state $\pi_1\pi_L^*$

Both functionals predict that, if $\pi_1\pi_L^*$ is initially excited, in < 100 fs its population decays to less than 0.1. It flows to $\pi_H\pi_L^*$ and also to the three $n\pi^*$ states, according to CAM-B3LYP, and to the symmetric and antisymmetric combinations of $n_N\pi_L^*$

and $n_O\pi^*_L$, according to PBE0. On the other hand, the small $\pi_1\pi^*_L$ population remains almost constant up to 250 fs. Interestingly, the populations of all states after 50 fs show only small oscillations around their average values for CAM-B3LYP and no significant oscillations for PBE0. $P(\pi_H\pi^*_L)$ reaches the limit value of ~ 0.25 and 0.4, respectively for PBE0 and CAM-B3LYP, in line with the larger relative stability of $\pi_H\pi^*_L$ according to CAM-B3LYP. Interestingly, both functionals predict some differences with respect to the dynamics observed starting from $\pi_H\pi^*_L$. In PBE0, we observe that the population of $\pi_H\pi^*_L$ and of the two lowest energy $n\pi^*$ states is very similar. In particular, the long-time limit of the $\pi_H\pi^*_L$ population is larger when the dynamics is started on $\pi_1\pi^*_L$ than on $\pi_H\pi^*_L$. At variance, according to CAM-B3LYP both $P(n_O\pi^*_2 + n_O\pi^*_2)$ (limit value ~ 0.2) and $n_O\pi^*_2 - n_O\pi^*_L$ (limit value ~ 0.1) are populated, but additionally also $P(n_N\pi^*_L)$ exhibits a stable value of ~ 0.2. In other words $\pi_1\pi^*_L$ acts as door-way state to an additional dark state with respect $\pi_H\pi^*_L$. Considering that in the FC region $n_N\pi^*_L$ is the dark state with the largest energy gap with respect $\pi_1\pi^*_L$, it is clear that energetic accessibility is only one of the factors ruling the dynamics.

TD-DFT calculations along the trajectory of the average position of the wavepacket (Section S6.3 in in SI) provide, at t$\leq$ 50 fs, a picture similar to the one already seen for dynamics from the $\pi_H\pi^*_L$. According to PBE0, the quasi degenerate S_1 and S_2 states flip between the $\pi_H\pi^*_L$ and $n_O\pi^*_L$ characters whereas S_3 is the $n_N\pi^*_L$ state. CAM-B3LYP predicts instead that S_1, S_2 and S_3 are respectively $\pi_H\pi^*_L$, $n_O\pi^*_L$ and $n_N\pi^*_L$ states. However when S_2 and S_3 get close (like at t$\sim$ 150 fs), mixings or switches of the electronic character can occur.

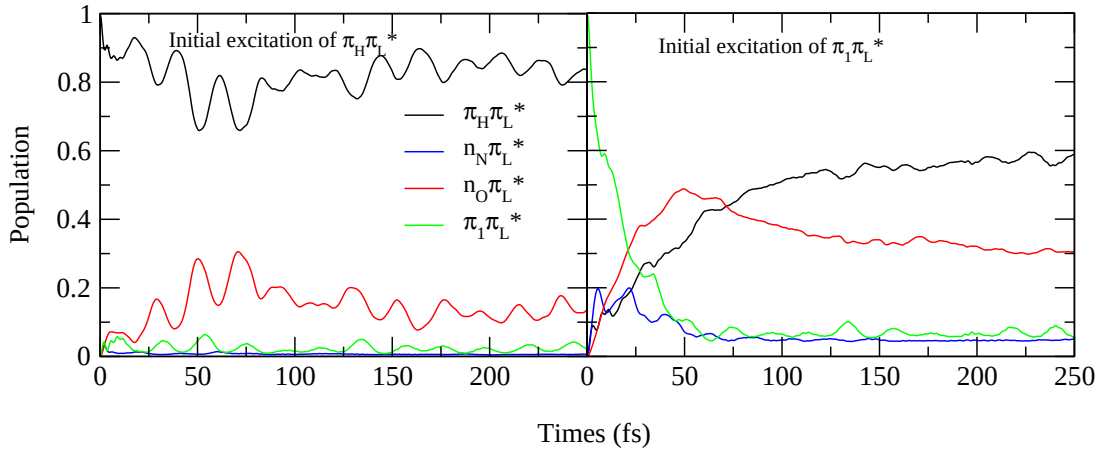


Figure 6: Nonadiabatic dynamics of electronic populations of cytosine in gas phase after an initial photoexcitation to $\pi_H\pi_L^*$ (top) or $\pi_1\pi_L^*$ (bottom) for CAM-B3LYP (left) or PBE0 (right) LVC Hamiltonians with diagonal energies corrected to reproduce CASPT2/TZVP energies at the FC point (taken from ref.[72]).

5.3.3 Predictions of a LVC Hamiltonian with CASPT2 energy gaps

CAM-B3LYP and PBE0 provide a rather different description of the coupling and the energy of the lowest energy nπ^* states. Since these and other first-principle methods (Section S4.6 in the SI) provide a different picture of the energetic of the dark states, which is not easily accessible by experiments, we have further investigated the sensitivity of QD results to the energy gaps at the FC position, fundamental parameters of the LVC model. To this aim we performed QD simulations at the CAM-B3LYP level, but shifting the vertical energies of the lowest energy states to reproduce the prediction of CASPT2. In fact, CASPT2 provides a similar description of the dark states as CAM-B3LYP, as pure $n_0\pi_L^*$ and $n_N\pi_L^*$ states.[74] However, their relative energies are rather different: $n_O\pi_L^*$ (5.12 eV) is more stable than $n_N\pi_L^*$ (5.54 eV), whose excitation energy is the same as $\pi_1\pi_L^*$.[72] $n_O\pi_L^*$ is also closer to $\pi_H\pi_L^*$ (4.68 eV) .

In this exploratory analysis we simply assumed that CASPT2 and CAM-

B3LYP PES are roughly parallel, which is confirmed by literature data at least for $\pi_H\pi_L^*$.[31]

With this LVC model, coupling the four states $\pi_H\pi_L^*$, $n_O\pi_L^*$, $n_N\pi_L^*$, $\pi_1\pi_L^*$, we run a new set of QD simulations for initial photo-excitations on $\pi_H\pi_L^*$ or $\pi_1\pi_L^*$. Figure 6 shows that with CASPT2 energies a larger population transfer to $n_O\pi_L^*$ is predicted. More specifically, for a dynamics starting $\pi_1\pi_L^*$ the decay channel toward $n_N\pi_L^*$ appears to be closed, as an effect of the destabilization of this state. Interestingly for a dynamics from $\pi_H\pi_L^*$ the population of $n_O\pi_L^*$ is faster but its long-time limit does not change remarkably. This may be a spurious effect due to the fact that in this new calculation we neglect a third $n\pi^*$ state, $n_O\pi_L^*$-$n_O\pi_2^*$, whose role in the CAM-B3LYP dynamics is not marginal.

5.4 Dynamics with LVC Hamiltonians parameterized at the different excited state minima.

LVC model PES are harmonic and therefore they are clearly an approximation of the true molecular PES and depend on the selected reference geometry. In order to evaluate the dependence of our predictions on the reference geometry, we repeated the QD simulations for LVC Hamiltonians parametrized at different excited state minima. As discussed in depth in the previous section, nπ^* states appear mixed and strongly coupled at the FC geometry. New parametrizations of the LVC Hamiltonians at the $n\pi^*$ minima, allow us also to reinvestigate the effect of these couplings, and possibly achieve a more clear chemical picture of the dynamics.

We used CAM-B3LYP and parametrized new LVC Hamiltonians as a function of the ground state normal coordinates in the minimum of the $\pi_H\pi_L^*$ state and at the minima of $n_N\pi_L^*$ and $n_O\pi_L^*$. We label these Hamiltonians LVC|Min-$\pi_H\pi_L^*$,

LVC|Min-$n_N\pi_L^*$, and LVC|Min-$n_O\pi_L^*$ whereas, for clarity, the original one will be named LVC|FC. As shown in Tables 2 and 3, extrapolation of the energies of the different diabatic states from the minima to the FC region in several cases provides rather different values with respect to those reported in Table 1. This indicates the PES can be characterized by significantly different normal modes/frequencies with respect to the ground states and/or by important anharmonicities

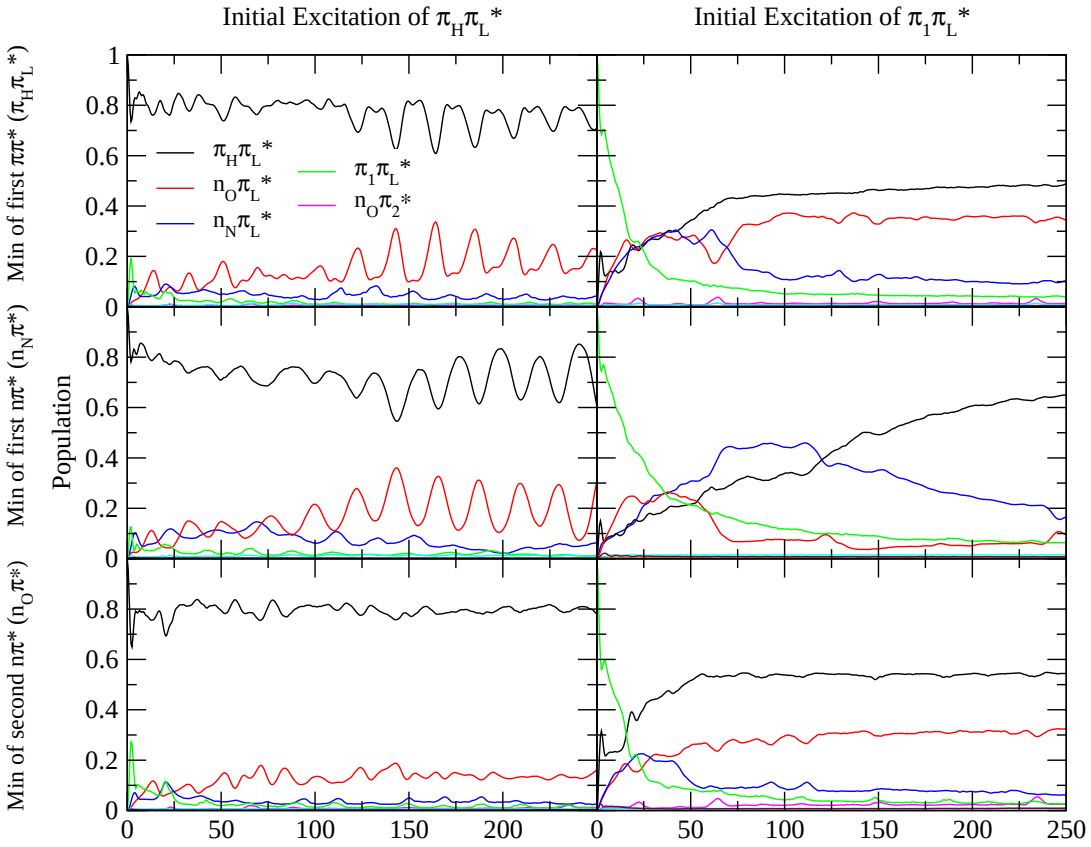


Figure 7: Nonadiabatic dynamics of electronic populations of Cyt in gas phase, as predicted by a LVC Hamiltonian parameterized with calculations at the $\pi_H\pi_L^*$ (top) $n_N\pi_L^*$ (middle) and $n_O\pi_L^*$ (bottom)C_s minimum using CAM-B3LYP functional. Initial photoexcitation to $\pi_H\pi_L^*$ (left) or $\pi_1\pi_L^*$ (right).

All the population dynamics in this section predict a spike at $t = 0$ fs corresponding to an instantaneous partial exchange of population between the two lowest bright $\pi\pi^*$ states. This is due to the fact that, according to the LVC

Hamiltonian, the diabatic $\pi\pi^*$ states are decoupled at the excited-state minimum selected for the parameterization. Therefore, they exhibit a coupling at the FC point, the initial position of the wavepacket. Such couplings evidence that the nature of the $\pi\pi^*$ states partially change from the ground-state and the excited-state minima.

Considering an initial excitation to $\pi_H\pi_L^*$, the time-evolution of the population of the bright state predicted by LVC|Min-$\pi_H\pi_L^*$, LVC|Min-$n_N\pi_L^*$ is quite similar to the one observed with LVC|FC. The main difference is that these alternative LVC models predict that most of the population lost by $\pi_H\pi_L^*$ state flows to a single $n\pi^*$ state with $n_O\pi_L^*$ character. The LVC|FC Hamiltonian also predicts localization of the excited-population in a $n_O\pi_L^*$ state, at least when moving to an adiabatic picture. Rather surprisingly, the dynamics according to LVC|Min-$n_O\pi_L^*$ shows larger differences, and namely a decrease of the population transfer to the $n_O\pi_L^*$ state. The reason is that the total coupling between $n_O\pi_L^*$ and $\pi_H\pi_L^*$ is smaller at the $n_O\pi_L^*$ minimum. For instance, it is 2.5 times smaller according to LVC|Min-$n_O\pi_L^*$ than according to LVC|Min-$n_N\pi_L^*$ one (Tables S8 and S11 in SI).

For an excitation to $\pi_1\pi_L^*$, QD with all the three new Hamiltonians predict that $P(\pi_1\pi_L^*)$ decays to <0.1 in less than 100 fs. With respect to LVC|FC, the population transfer to $\pi_H\pi_L^*$ is slightly larger with LVC|Min-$\pi_H\pi_L^*$ and LVC|Min-$n_O\pi_L^*$ (from $\sim$ 40% to $\sim$ 50%), while it increases up to 65% with LVC|Min-$n_O\pi_L^*$. The nπ^* population is dominated by $n_O\pi_L^*$ according to LVC|Min-$\pi_H\pi_L^*$ and LVC|Min-$n_O\pi_L^*$, while the fraction flowing to $n_N\pi_L^*$ is larger according to LVC|Min-$n_N\pi_L^*$.

For PBE0 we parametrized the LVC Hamiltonian around the minimum of the first nπ^* state where it has a clear $n_O\pi_L^*$ character, at variance with what happens in the FC position. This enables a more straightforward analysis of the QD results, reported in Section S6.4.2 of the SI. Shortly, the decrease of the population of $\pi_H\pi_L^*$

population is slightly lower than what predicted with LVC|FC Hamiltonian, but LVC|Min$-n_O\pi_L^*$ shows that the population is transferred almost exclusively to $n_O\pi_L^*$. An initial excitation on $\pi_1\pi_L^*$ decays to 0.05 in $\sim$ 25 fs. At later times it remains constant and an additional dark state, $n_O\pi_2^*$, is weakly populated.

6 Results: 1-Methyl-Cytosine

Table 6 shows that the electronic characters and transition energies of the excited states of 1MeCyt at the FC point are very similar to those computed for Cyt. At CAM-B3LYP level we observe only a slight stabilization of the first two $\pi\pi^*$ and a slight destabilization of the first two nπ^* states. According to PBE0 changes are even smaller and lead to a destabilization of $\mathrm{n}_O\pi_L^*$ +$\mathrm{n}_N\pi_L^*$ and a very weak stabilization of $\pi_1\pi_L^*$. Comparison of the available data indicates that methylation stabilizes the $\pi_H\pi_H^*$ state, both at CASPT2 level,[75,76] and at MS-CASPT2 level,[72,76] although calculations are not exactly with the same active spaces and basis sets. As reported in Section 6.2 of the SI, the excited states in the FC region of 1MeCyt are very close to that of cytidine, according to both CAM-B3lYP and PBE0, apart from a systematic very slight (<0.04 eV) blueshift of most of the states of cytidine. On this ground, at least from the purely electronic point of view, 1MeCyt appears a good model of the nucleoside.

Comparison of Figures 4 and 8 shows that the small changes in the energy gaps of the states due to the methyl substituent have moderate effects on the dynamics. For an excitation of the first bright state $\pi_H\pi_L^*$, according to CAM-B3LYP the population transfer to the nπ^* states is even smaller, in agreement with the relative stabilization of the $\pi_H\pi_L^*$. Interestingly, however, the behaviour is rather similar: we can notice an initial small and transient population on $\mathrm{n}_N\pi_L^*$

Table 6: Energies (E^a_{FC}), electronic characters, oscillator strengths δ_{OPA}, and main contributions in terms of transitions among Kohn-Sham orbitals for the lowest excited states of keto-amino tautomer of 1-methyl-cytosine at the FC point. The energies of the corresponding diabatic states in their minima (E^d_{min}), estimated by the LVC Hamiltonian, are also reported. CAM-B3LYP and PBE0 calculations with 6-31+G(d,p) basis set in gas phase. Energies in eV with respect to the ground state in its minimum.

	CAM-B3LYP					PBE0				
State	E^a_{FC}	E^d_{min}	character	δ_{OPA}	Transition	E^a_{FC}	E^d_{min}	character	δ_{OPA}	Transition
S_1	4.96	4.62	$\pi_H\pi^*_L$	**0.10**	H→ L	4.77	4.39	$\pi_H\pi^*_L1$	**0.07**	H→ L
S_2	5.36	4.83	$n_N\pi^*_L$	0.00	H-3→L H-2→L	5.05	4.67	$n_O\pi^*_L$ +$n_N\pi^*_L$	0.00	H-2→ L
S_3	5.79	5.66	$\pi_H Ry_\sigma 1$	0.01	H→L+1	5.35	4.43	$n_O\pi^*_L$ -$n_N\pi^*_L$	0.00	H-3→ L
S_4	5.87	5.47	$\pi_1\pi^*_L$	**0.13**	H-2→L	5.52	5.11	$\pi_1\pi^*_L$	**0.11**	H-1→ L
S_5	5.91	5.20	$n_O\pi^*_L$+$n_O\pi^*_2$	0.00	H-2→L+4 H-2→L	5.57	5.47	$\pi_H Ry_\sigma 1$	0.01	H→ L+1
S_6	6.08	5.46	$n_O\pi^*_L$-$n_O\pi^*_2$	0.00	H-2→L+4 H-2→L	5.82	5.30	$n_O\pi^*_2$	0.00	H-2→ L+3
S_7	6.21	6.16	$\pi_1 Ry_\sigma 2$	0.01	H-1→L+1	5.97	5.83	$\pi_H Ry_\sigma 2$	0.00	H→ L+2

(and $\pi_1\pi^*_L$) followed by a slight gain of population of the two states with marked $n_O\pi^*$ character that also in this case show in-phase oscillations.

PBE0 calculations predict even smaller differences for 1MeCyt and Cyt. As a matter of fact, the only small change is in the long-time limit populations: they slightly increase for $n_O\pi^*_L$+$n_N\pi^*_L$ and decrease for $\pi_H\pi^*_L$. It is noteworthy that according to ref.[12] the MS-CASPT2 gap between $\pi_H\pi^*_H$ (4.64 eV) and the first $n\pi^*$ state (5.25 eV) is $\sim$ 0.6 eV, similar to what predicted by PBE0 with the $n_O\pi^*_L$-$n_N\pi^*_L$ state that gains the largest population transfer.

For an initial excitation to the second bright state, $\pi_1\pi^*_L$, methylation in position 1 has an even smaller effect on the dynamics. According to CAM-B3LYP, as for Cyt, most of the transfers occur in the first 50 fs, but in 1MeCyt $\pi_H\pi^*_L$ population increases from $\sim$ 0.4 to $\sim$ 0.5 and, consequently, the population of $n_O\pi^*_2$+ $n_O\pi^*_L$ decreases from 0.2 to 0.1. Interestingly, for PBE0 we observe the opposite

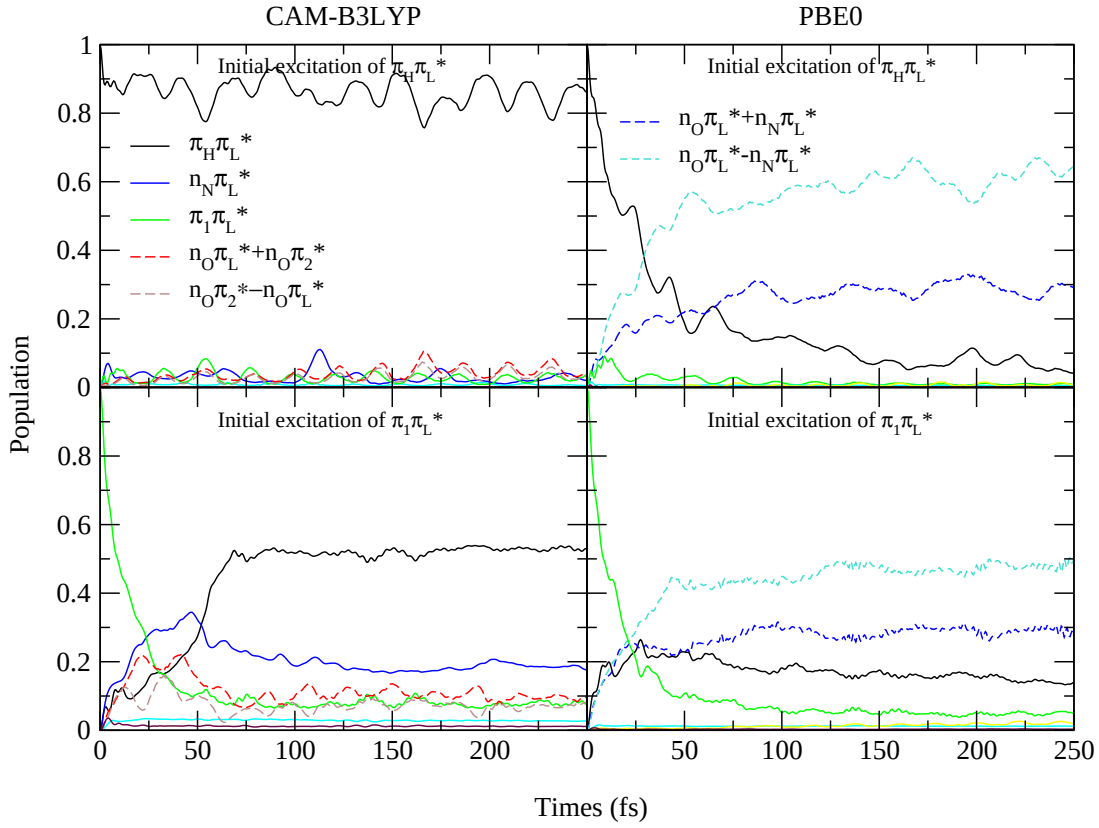


Figure 8: Nonadiabatic dynamics of electronic populations of 1-methyl-cytosine for excitations starting from two lowest bright states in gas phase.

trend. The population of $n_O\pi_L^*-n_N\pi_L^*$ increases (from ∼0.4 to ∼ 0.5) while $\pi_H\pi_L^*$ one decreases (from ∼0.25 to ∼0.15).

Also for 1MeCyt we checked the dependence of our predictions on the structure used to build the LVC models, reparameterizing with CAM-B3YP the LVC Hamiltonians taking as reference the minima of the first $\pi\pi^*$ and the first two $n\pi_L^*$ states. Results are reported in section S7.6 of the SI and also in this case provide an easier reading of dynamics, showing that after an initial excitation to $\pi_H\pi_L^*$ the population transfer is mainly to a state with $n_0\pi^*$ character. Overall, the predictions of the alternative LVC Hamiltonians in Cyt and 1MeCyt follow a similar trend with respect to LVC|FC models. LVC expanded around $n\pi^*$ minima

predict that a moderately larger fraction of the initial population flows in nπ^* dark states.

7 Discussions and Conclusions

In this contribution we used full-dimensionality quantum dynamics to study the population dynamics of Cyt and 1MeCyt in gas phase after a photoexcitation to the two lowest two bright states $\pi_H\pi_L^*$ and $\pi_1\pi_L^*$. We examined the keto-amino tautomer that is the biological relevant species, without considering the possible intersystem crossing to triplets, whose quantum yield is ≤ 0.1.[19] We adopted LVC Hamiltonians and this restricts the validity of our analysis to the ultrafast regime ($\sim$ 100 fs). In particular our methodology does not allow to describe the decay to the ground state. Therefore our study is focused on the interplay between the lowest states in the FC region and, in particular, to assess the possible ultrafast population of dark nπ^* states. Moreover, for the first time the ultrafast dynamics following excitation to the second bright excited state is simulated.

Our LVC Hamiltonians predict (Table 4 and 5) that for both Cyt and 1MeCyt some inter-state couplings are so strong to be even larger than intra-state ones (i.e. those determining the reorganization energy), a situation in which difference between quantum and semiclassical dynamics can be expected.[49]

For initial photoexcitation on the lowest bright $\pi_H\pi_L^*$ state, all calculations consistently predict that a fraction of the population flows into nπ^* states. From a quantitative point of view, we observed a remarkable sensitivity of our results on the energy gap between $\pi_H\pi_L^*$ and the nπ^* states, which depends on the electronic level of theory adopted to parametrize the LVC Hamiltonian. CAM-B3LYP destabilizes nπ^* thus disfavoring their population with respect to PBE0. A LVC

parametrized on CAM-B3LYP calculations but with vertical energies shifted to reproduce CASPT2 data gives intermediate predictions. In synthesis, at $\sim$ 150 fs $n\pi^*$ population is 20-30% according to CAM-B3LYP, is $\sim$80% for PBE0 and 30-50 % for the LVC with CASPT2 energies.

Disentangling whether the population flown to dark states has $n_N\pi^*$ or $n_O\pi^*$ character is even more challenging. Both CAM-B3LYP and PBE0 calculations indicate that these two states are strongly coupled and, in the FC position, they appear already mixed, especially according to PBE0. This complicates the interpretation with LVC|FC Hamiltonians, forcing us to adopt some non-standard tools of analysis. Calculations of the adiabatic states along the average position of the wavepacket suggest that $\pi_H\pi_L^*$ partially decays to a dark state with $n_O\pi^*$ character. This prediction is confirmed by changing the reference structure for parametrizations of the LVC Hamiltonians from the FC point to the excited state minima, where the $n\pi^*$ states acquire clear $n_O\pi_L^*$ or $n_N\pi_L^*$ characters. Interestingly, these calculations, for all the three minima ($\pi_H\pi_L^*$, $n_O\pi_L^*$ and $n_N\pi_L^*$) used in the parametrization, indicate that the electronic population flows mostly to $n_O\pi_L^*$ state. As it could be expected, parametrizations at $n\pi_L^*$ minima also predict that the transfer to $n\pi^*$ states is moderately larger than what is predicted with LVC parametrized at the FC geometry.

Interestingly, our simulations (see also the analysis of dynamics from $\pi_1\pi_L^*$) predict the existence of two characteristic time-regimes. We observe an ultrafast dynamics on a few dozens of fs time-scale, $\leq$ 50 fs, where we observe most of the changes with respect to the initial time, and a plateau, lasting until the end of our simulation window, where the populations mainly oscillate around their average value.

Our findings are fully consistent with the available experimental results. The

photophysics of Cyt in gas phase has been deeply investigated experimentally.[8,10,33,34,46,77,78] All experiments with a sufficient time-resolution has evidenced both decay constants in the ∼100 femtosecond (sometimes even < 50 fs)[77,78] and picosecond time scales. The intepretation of these data is made very complicated by the existence of several tautomers in gas phase. Adopting an excitation at 280-290 nm, the authors of ref.[10] were able to selectively excite the keto-amino tautomer, observing two short decay times respectively < 100 fs and 1.1-1.2 ps, and a longer lifetime that is 55 ps at λ=280 nm, growing to ≥ 150 ps for λ=290 nm. By comparing Cyt and 1MeCyt results, Ho et al,[46] found that the keto-amino tautomer decays with a sub picosecond lifetime depending on the excitation wavelenght (200 fs at 260 nm increasing progressively up to 1.2 ps at 300 nm). They noticed that possible population transfer to nπ^* states occurring on a < 100 fs timescale could not be investigated for the time-resolution of the experiment. A long-time component of 10 ns was assigned to the formation to triplet states.

Our simulations indicate that the fastest component should be assigned to the motion of the wavepacket out from the FC region, with vibrational relaxation on $\pi_H\pi_L^*$ and population of the dark $n\pi^*$ states. Such transfer is not complete; part of the wavepacket remains on $\pi_H\pi_L^*$ and can decay directly to the ground state, via the C6-puckered CoI, in agreement with the pictures provided by the most accurate quantum-chemical calculations.[4,14] A small, but not-zero energy barrier is present on this path,[14] giving account of ∼ ps component. Finally, the decay to S_0 on the 50∼150 ps time-scale is assigned to the dark state. Our calculations indicate that the $n\pi^*$ state has a prevalent $n_O\pi^*$ character. To the best of our knowledge gas phase experiments did not provide indications on the assignment of $n\pi^*$ population. However, experiments on cytidine in low-polarity solvents, like chloroform,[79] found the spectral signature of a dark state, populated on the subpicosecond timescale.

This state is characterized by a transient infrared spectrum very similar to that observed in water, and assigned to $n_O\pi^*_L$.[31,36] Considering that solvent could affect the infrared spectra of the different states, for the moment, experimental results seem consistent with our assignment.

As also shown by our calculations, the results of any dynamical study dramatically depends on the underlying electronic method. In our case, for example, the conclusion of this study are thus meaningful to interpret Cyt photophysics only if we admit that ground-state recovery takes least few hundreds of femtoseconds and that no 'exceptionally fast' (few fs) deactivation route to S_0 is available for the dark states. Actually, as discussed above, both conditions are met. TR experiments show that a significant part of the excited state population decays on the ps time-scale. Moreover the most recent joint experimental/computational studies of ultracold Cyt shows that the main sub-ps deactivation route is the ethylenic CoI associated to $\pi_H\pi^*_L$.[14] On this ground, the effectiveness of the other few-fs decay routes associated to the dark states (see below) are confirmed to be due to an artefact due to the use of CASSCF. More in general, the comparison between our results and those of CASSCF semi-classical dynamical simulations is not straightforward, since (i) they are made in an adiabatic base, whose electronic character can change from point to point. (ii) they follows the dynamics until the decay to S_0, the results depend on the preferential deactivation paths (iii) which are determined by the adopted electronic method. For example, at the CASSCF level, the lowest energy dark states are, on the average, closer in energy to $\pi_H\pi^*_L$ than at the CASPT2 (and TD-DFT) level and have access to low energy CoI with the ground state.[19] As a consequence, extremely fast disappearance of $\pi_H\pi^*_L$ populations is predicted to the dark $n\pi^*_L$ states, which then decay to S_0 via the sofa-like CoI (usually associated to $n_N\pi^*$) and the three-state $S_0/n_O\pi^*/\pi_H\pi^*_L$ CoI.[19,21,22,80]

In some of these studies, several trajectories were found to decay to S_0 within 10 fs. Despite these considerations, it is noteworthy that these different dynamical approaches agree on the strong coupling between $\pi_H\pi_L^*$ and the close-lying dark $n\pi^*$ states.

We also investigated, for the first time, the dynamics after an initial excitation of $\pi_H\pi_L^*$ in 1MeCyt and of $\pi_1\pi_L^*$ for both Cyt and 1MeCyt. For an initial excitation to $\pi_H\pi_L^*$, the dynamics of 1MeCyt is rather similar to Cyt, especially for PBE0. According to CAM-B3LYP methylation slightly decreases the population transfer to $n\pi^*$ states, due the larger energy gap in the FC region. The population flowing to dark states has also in this case $n_O\pi^*$ character. The similarity we find for the dynamics of Cyt and 1MeCyt is in line with the experimental results of Ho et al in ref.[46]

For both Cyt and 1MeCyt our calculations predict that $\pi_1\pi_L^*$ decays very quickly, reaching 10% in 100 fs or less. Interestingly, also in this process, we observe that the excited states populations reach a pseudo-plateau in <100 fs. As a consequence, our simulations suggest that population transfer from $\pi_1\pi_L^*$ also contributes to the fastest experimental time-components. Population flows to $\pi_H\pi_L^*$ and both to $n_O\pi^*$ and $n_N\pi^*$, but the exact proportion depends on the molecule and the level of calculation. At later times all populations reach a stationary value. $P(\pi_H\pi_L^*)$ is estimated to be in the range 20-60%. Within this rather broad range of incertitude the situation in 1MeCyt and Cyt seems rather similar. Confirming the results obtained for excitation of $\pi_H\pi_L^*$, also starting from $\pi_1\pi_L^*$ the exact amount of $n\pi^*$ population and its assignment depends on the functional and on the reference structure for computing LVC. Remarkably, on average, the population on $n_N\pi^*$ is larger than what obtained when exciting $\pi_H\pi^*$. Moreover, in some simulations, also a further $n_O\pi_2^*$ state is populated. $\pi_1\pi_L^*$ can thus acts as

a doorway toward additional states that can enrich the dynamics at later times.

Actually, the interest on the excited state decay from $\pi_1\pi_L^*$ is not purely academic. Indeed excitation at 260∼270 nm, the wavelength most commonly used in TR experiments, probably excite both $\pi_H\pi_L^*$ and $\pi_1\pi_L^*$ states, making the inclusion of this latter state very important to understand the dynamics.[38] In this respect many experiments clearly shows that the excited state dynamics depend on the excitation wavelength. Besides to varying the amount of energy deposited on the system, for Cyt, changing the excitation wavelength affects the amount of population excited on $\pi_H\pi_L^*$ and $\pi_1\pi_L^*$. In any case, inclusion of $\pi_1\pi_L^*$ enforces the conclusion that experiments at wavelength sufficiently short to excite both bright states (260 nm)[38] produce a significant population on dark nπ^* states in less than 100 fs.

Our simulations provide some interesting general insights on the effects ruling the dynamics. (i) As intuitive the relative stability in the FC region plays a role on the probability of internal conversions, (ii) however the intricate interplay of the couplings of several electronic states in a small energy interval can lead also to the population of excited states that are rather distant from the spectroscopic states in the FC point. (iii) This is particularly true for dark states stabilized by the same kind of vibrations that lead the spectroscopic state toward its minimum. These effects contribute to explain the rather puzzling result that $n_O\pi^*$ is significantly populated during the dynamics, even if it is less stable than $n_N\pi^*$ in the FC region (CAM-B3LYP results).

From the methodological point of view, we present our protocol for an automatic and general parameterization of simplified LVC Hamiltonians from TD-DFT computations. In combination with powerful ML-MCTDH propagations, this gives the possibility to run full-dimensionality calculations of the ultrafast quantum dy-

namics of nucleobases, accounting for the effect of many different coupled states. For instance, it is interesting to notice that, although we included in our models states with remarkable Rydberg character, as expected they play a marginal role in the dynamics.

At variance with semiclassical methods, where trajectories are propagated on adiabatic states that change their nature with time, QD simulations are run in a diabatic coordinate-independent representation. This different perspective allows a fresh and alternative view on the complex transformation of the electronic character of the photoexcited population with time.

Thanks to the efficiency of the proposed protocol one can easily repeat the calculations with different parametrizations of the Hamiltonian, with a limited computational cost. We have exploited this capability to check the robustness of our predictions by repeating the calculations for LVC models expanded at different key molecular structures. We also show that the recomputation of the TD-DFT potentials as a function of time at the average position of the wavepacket can be very instructive. Both these strategies revealed in fact very useful to disentangle the contributions of the coupled $n_O\pi^*$ and $n_N\pi^*$ states to the decay dynamics. The same efficiency allowed us to investigate the dependence of the dynamics on different DFT functionals. While qualitatively our results are clear, from a quantitative point, remarkable forks have been observed for the predictions of the populations with PBE0 and CAM-B3LYP. Such findings indicate that, at the state of the art, it is very challenging to get accurate values for the yield of different population-transfer channels, at least when different states with similar energy exist. In this respect QD calculations on LVC Hamiltonians can also usefully integrate surface hopping semiclassical calculations, considering that the two approaches have complementary strengths and limitations. The potentialities of these combined

approach have been recently demonstrated.[81,82]

Supporting Information. The Supporting Information is available free of charge at https://pubs.acs.org/.

Cartesian coordinates of all optimized structures. Additional Tables on the energies and properties of the excited states of Cytosine and 1-methyl-cytosine at the different optimized geometries. Plots of the Kohn-Sham Molecular Orbitals and NTOs (for those cases not given in the manuscript). Convergence test for the ML-MCTDH propagation. Additional plots of the diabatic and adiabatic energies and of the NTOs along the trajectory of the wavepacket average position. Additional electronic populations dynamics obtained with PBE0 and not reported here. Dynamics of 1MeCyt driven by LVC models parameterized at excited state minima.

Acknowledgments

This work has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skodowska-Curie grant agreement No 765266 (LightDyNAmics). The authors thank Lara Martínez Fernández, Universidad Autónoma de Madrid, for carefully reading the manuscript and useful discussions and Sonia Coriani (DTU Chemistry, Technical University of Denmark) for useful discussions on TD-DFT formalism.

References

(1) Crespo-Hernandez, C. E.; Cohen, B.; Hare, P. M.; Kohler, B. Ultrafast Excited-State Dynamics in Nucleic Acids. *Chem. Rev.* **2004**, *104*, 1977–2020.

(2) Middleton, C. T.; de La Harpe, K.; Su, C.; Law, Y. K.; Crespo-Hernández, C. E.; Kohler, B. DNA Excited-State Dynamics: From Single Bases to the Double Helix. *Ann. Rev. Phys. Chem.* **2009**, *60*, 217–239.

(3) Schreier, W. J.; Gilch, P.; Zinth, W. Early Events of DNA Photodamage. *Ann. Rev. Phys. Chem.* **2015**, *66*, 497–519.

(4) Improta, R.; Santoro, F.; Blancafort, L. Quantum mechanical studies on the Photophysics and the Photochemistry of nucleic acids and nucleobases. *Chem. Rev.* **2016**, *116*, 3540–3593.

(5) Barbatti, M.; Borin, A. C.; Ullrich, S. In *Photoinduced Phenomena in Nucleic Acids I: Nucleobases in the Gas Phase and in Solvents*; Barbatti, M., Borin, C. A., Ullrich, S., Eds.; Springer International Publishing: Cham, Switzerland, 2015; Vol. 355; pp 1–32.

(6) Gustavsson, T.; Improta, R.; Markovitsi, D. DNA/RNA: Building Blocks of Life Under UV Irradiation. *J. Phys. Chem. Lett.* **2010**, *1*, 2025–2030.

(7) Cadet, J.; Douki, T.; Ravanat, J.-L. Oxidatively Generated Damage to Cellular DNA by UVB and UVA Radiation,. *Photochem. Photobiol.* **2015**, *91*, 140–155.

(8) 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. *J. Chem. Phys.* **2005**, *122*, 074316.

(9) Ullrich, S.; Schultz, T.; Zgierski, M. Z.; Stolow, A. Electronic relaxation dynamics in DNA and RNA bases studied by time-resolved photoelectron spectroscopy. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2796–2801.

(10) Kosma, K.; Schroter, C.; Samoylova, E.; Hertel, I. V.; Schultz, T. Excited-state dynamics of cytosine tautomers. *J. Am. Chem. Soc.* **2009**, *131*, 16939–16943.

(11) Lobsiger, S.; Trachsel, M. A.; Frey, H.-M.; Leutwyler, S. Excited-State Structure and Dynamics of Keto–Amino Cytosine: The $^1\pi\pi^*$ State Is Nonplanar and Its Radiationless Decay Is Not Ultrafast. *J. Phys. Chem. B* **2013**, *117*, 6106–6115.

(12) Trachsel, M. A.; Wiedmer, T.; Blaser, S.; Frey, H.-M.; Li, Q.; Ruiz-Barragan, S.; Blancafort, L.; Leutwyler, S. The excited-state structure, vibrations, lifetimes, and nonradiative dynamics of jet-cooled 1-methylcytosine. *J. Chem. Phys.* **2016**, *145*, 134307.

(13) Trachsel, M. A.; Blaser, S.; Siffert, L.; Wiedmer, T.; Leutwyler, S. Excited-state vibrations, lifetimes, and nonradiative dynamics of jet-cooled 1-ethylcytosine. *J. Chem. Phys.* **2019**, *151*, 124301.

(14) Trachsel, M. A.; Blaser, S.; Lobsiger, S.; Siffert, L.; Frey, H.-M.; Blancafort, L.; Leutwyler, S. Locating Cytosine Conical Intersections by Laser Experiments and Ab Initio Calculations. *J. Phys. Chem. Lett.* **2020**, *11*, 3203–3210.

(15) Giussani, A.; Segarra-Martí, J.; Roca-Sanjuán, D.; Merchán, M. In *Photoinduced Phenomena in Nucleic Acids I: Nucleobases in the Gas Phase and in*

Solvents; Barbatti, M., Borin, C. A., Ullrich, S., Eds.; Springer International Publishing: Cham, Switzerland, 2015; Vol. 355; pp 55–97.

(16) Mai, S.; Richter, M.; Marquetand, P.; González, L. In *Photoinduced Phenomena in Nucleic Acids I: Nucleobases in the Gas Phase and in Solvents*; Barbatti, M., Borin, C. A., Ullrich, S., Eds.; Springer International Publishing: Cham, Switzerland, 2015; Vol. 355; pp 99–153.

(17) Blancafort, L. Energetics of Cytosine Singlet Excited-State Decay Paths—A Difficult Case for CASSCF and CASPT2. *Photochem. Photobiol.* **2007**, *83*, 603–610.

(18) 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.

(19) Richter, M.; Marquetand, P.; González-Vázquez, J.; Sola, I.; González, L. Femtosecond Intersystem Crossing in the DNA Nucleobase Cytosine. *J. Phys. Chem. Lett.* **2012**, *3*, 3090–3095.

(20) Mai, S.; Marquetand, P.; Richter, M.; González-Vázquez, J.; González, L. Singlet and triplet excited-state dynamics study of the keto and enol tautomers of cytosine. *ChemPhysChem* **2013**, *14*, 2920–2931.

(21) Barbatti, M.; Aquino, A. J. A.; Szymczak, J. J.; Nachtigallová, D.; Lischka, H. Photodynamical simulations of cytosine: characterization of the ultrafast biexponential UV deactivation. *Phys. Chem. Chem. Phys.* **2011**, *13*, 6145–6155.

(22) Hudock, H. R.; Martínez, T. J. Excited-State Dynamics of Cytosine Reveal Multiple Intrinsic Subpicosecond Pathways. *ChemPhysChem* **2008**, *9*, 2486–2490.

(23) Avila Ferrer, F. J.; Santoro, F.; Improta, R. The excited state behavior of cytosine in the gas phase: A TD-DFT study. *Computat. and Theor. Chem.* **2014**, *1040-1041*, 186 – 194.

(24) Kistler, K. A.; Matsika, S. Cytosine in Context: A Theoretical Study of Substituent Effects on the Excitation Energies of 2-Pyrimidinone Derivatives. *J. Phys. Chem. A* **2007**, *111*, 8708–8716.

(25) Kistler, K. A.; Matsika, S. Radiationless decay mechanism of cytosine: an ab initio study with comparisons to the fluorescent analogue 5-methyl-2-pyrimidinone. *J. Phys. Chem. A* **2007**, *111*, 2650.

(26) Kistler, K. A.; Matsika, S. Three-state conical intersections in cytosine and pyrimidinone bases. *J. Chem. Phys.* **2008**, *128*, 215102.

(27) Pepino, A. J.; Segarra-Martí, J.; Nenov, A.; Rivalta, I.; Improta, R.; Garavelli, M. UV-induced long-lived decays in solvated pyrimidine nucleosides resolved at the MS-CASPT2/MM level. *Phys. Chem. Chem. Phys.* **2018**, *20*, 6877–6890.

(28) Pepino, A. J.; Segarra-Martí, J.; Nenov, A.; Improta, R.; Garavelli, M. Resolving Ultrafast Photoinduced Deactivations in Water-Solvated Pyrimidine Nucleosides. *J. Phys. Chem. Lett.* **2017**, *8*, 1777–1783.

(29) Szabla, R.; Kruse, H.; Šponer, J.; Góra, R. W. Water–chromophore electron transfer determines the photochemistry of cytosine and cytidine. *Phys. Chem. Chem. Phys.* **2017**, *19*, 17531–17537.

(30) Martínez-Fernández, L.; Pepino, A. J.; Segarra-Martí, J.; Banyasz, A.; Garavelli, M.; Improta, R. Computing the Absorption and Emission Spectra of

5-Methylcytidine in Different Solvents: A Test-Case for Different Solvation Models. *J. Chem. Theory Comput.* **2016**, *12*, 4430–4439.

(31) Martínez-Fernández, L. et al. Photophysics of Deoxycytidine and 5-Methyldeoxycytidine in Solution: A Comprehensive Picture by Quantum Mechanical Calculations and Femtosecond Fluorescence Spectroscopy. *J. Am. Chem. Soc.* **2017**, *139*, 7780–7791.

(32) Kistler, K. A.; Matsika, S. Solvatochromic shifts of uracil and cytosine using a combined multireference configuration interaction/molecular dynamics approach and the fragment molecular orbital method. *J. Phys. Chem. A* **2009**, *113*, 12396.

(33) Kang, H.; Lee, K. T.; Jung, B.; Ko, Y. J.; Kim, S. K. Intrinsic lifetimes of the excited state of DNA and RNA bases. *J. Am. Chem. Soc.* **2002**, *124*, 12958–12959.

(34) Nir, E.; Müller, M.; Grace, L.; De Vries, M. REMPI spectroscopy of cytosine. *Chem. Phys. Lett.* **2002**, *355*, 59–64.

(35) Keane, P. M.; Wojdyla, M.; Doorley, G. W.; Watson, G. W.; Clark, I. P.; Greetham, G. M.; Parker, A. W.; Towrie, M.; Kelly, J. M.; Quinn, S. J. A comparative picosecond transient infrared study of 1-methylcytosine and 5'-dCMP that sheds further light on the excited states of cytosine derivatives. *J. Am. Chem. Soc.* **2011**, *133*, 4212.

(36) 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.

(37) 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. *Phys. Chem. Chem. Phys.* **2015**, *17*, 19045.

(38) Wang, X.; Zhou, Z.; Tang, Y.; Chen, J.; Zhong, D.; Xu, J. Excited State Decay Pathways of 2-Deoxy-5-methylcytidine and Deoxycytidine Revisited in Solution: A Comprehensive Kinetic Study by Femtosecond Transient Absorption. *J. Phys. Chem. B* **2018**, *122*, 7027–7037.

(39) Brown, R. D.; Godfrey, P. D.; McNaughton, D.; Pierlot, A. P. Tautomers of cytosine by microwave spectroscopy. *J. Am. Chem. Soc.* **1989**, *111*, 2308–2310.

(40) Lapinski, L.; Reva, I.; Nowak, M. J.; Fausto, R. Five isomers of monomeric cytosine and their interconversions induced by tunable UV laser light. *Phys. Chem. Chem. Phys.* **2011**, *13*, 9676–9684.

(41) Wolken, J. K.; Yao, C.; Tureček, F.; Polce, M. J.; Wesdemiotis, C. Cytosine neutral molecules and cation–radicals in the gas-phase: Structures, energetics, ion chemistry, and neutralization–reionization mass spectrometry. *Int. J. Mass Spectrom.* **2007**, *267*, 30–42.

(42) Feyer, V. et al. Tautomerism in cytosine and uracil: An experimental and theoretical core level spectroscopic study. *J. Phys. Chem. A* **2009**, *113*, 5736–5742.

(43) Kobayashi, R. A CCSD (T) study of the relative stabilities of cytosine tautomers. *J. Phys. Chem. A* **1998**, *102*, 10813–10817.

(44) Trygubenko, S. A.; Bogdan, T. V.; Rueda, M.; Orozco, M.; Luque, F. J.; Šponer, J.; Slavíček, P.; Hobza, P. Correlated ab initio study of nucleic acid bases and their tautomers in the gas phase, in a microhydrated environment and in aqueous solution Part 1. Cytosine. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4192–4203.

(45) Tomić, K.; Tatchen, J.; Marian, C. M. Quantum chemical investigation of the electronic spectra of the keto, enol, and keto- imine tautomers of cytosine. *J. Phys. Chem. A* **2005**, *109*, 8410–8418.

(46) Ho, J.-W.; Yen, H.-C.; Chou, W.-K.; Weng, C.-N.; Cheng, L.-H.; Shi, H.-Q.; Lai, S.-H.; Cheng, P.-Y. Disentangling Intrinsic Ultrafast Excited-State Dynamics of Cytosine Tautomers. *J. Phys. Chem. A* **2011**, *115*, 8406–8418.

(47) Lan, Z.; Fabiano, E.; Thiel, W. Photoinduced nonadiabatic dynamics of pyrimidine nucleobases: On-the-fly surface-hopping study with semiempirical methods. *J. Phys. Chem. B* **2009**, *113*, 3548–3555.

(48) Doorley, G.; McGovern, D.; George, M.; Towrie, M.; Parker, A.; Kelly, J.; Quinn, S. Picosecond Transient Infrared Study of the Ultrafast Deactivation Processes of Electronically Excited B-DNA and Z-DNA Forms of [poly(dG-dC)]2. *Angew. Chem., Int. Ed.* **2009**, *48*, 123–127.

(49) Ferretti, A.; Granucci, G.; Lami, A.; Persico, M.; Villani, G. Quantum mechanical and semiclassical dynamics at a conical intersection. *J. Chem. Phys.* **1996**, *104*, 5517–5527.

(50) Borrego-Varillas, R.; Teles-Ferreira, D. C.; Nenov, A.; Conti, I.; Ganzer, L.; Manzoni, C.; Garavelli, M.; Maria de Paula, A.; Cerullo, G. Observation of

the sub-100 femtosecond population of a dark state in a thiobase mediating intersystem crossing. *J. Am. Chem. Soc.* **2018**, *140*, 16087–16093.

(51) Beck, M. H.; Jäckle, A.; Worth, G. A.; Meyer, H.-D. The Multiconfiguration Time-Dependent Hartree Method: A highly efficient algorithm for propagating wavepackets. *Phys. Rep.* **2000**, *324*, 1–105.

(52) Wang, H.; Thoss, M. Multilayer formulation of the multiconfiguration time-dependent Hartree theory. *J. Chem. Phys.* **2003**, *119*, 1289–1299.

(53) Manthe, U. A multilayer multiconfigurational time-dependent Hartree approach for quantum dynamics on general potential energy surfaces. *J. Chem. Phys.* **2008**, *128*, 164116.

(54) Manthe, U. Layered discrete variable representations and their application within the multiconfigurational time-dependent Hartree approach. *J. Chem. Phys.* **2009**, *130*, 054109.

(55) Vendrell, O.; Meyer, H.-D. Multilayer multiconfiguration time-dependent Hartree method: Implementation and applications to a Henon-Heiles Hamiltonian and to pyrazine. *J. Chem. Phys.* **2011**, *134*, 044135.

(56) Frisch, M. J. et al. Gaussian˜16 Revision B.01. 2016; Gaussian Inc. Wallingford CT.

(57) Cimiraglia, R.; Malrieu, J.-P.; Maurizio, P.; Fernand, S. Quasi-Diabatic States and Dynamical Couplings from Ab Initio CI Calculations: a New Proposal. *J. Phys. B: At. Mol. Phys.* **1985**, *18*, 3073–3084.

(58) Neugebauer, J.; Baerends, E. J.; Nooijen, M. Vibronic coupling and double

excitations in linear response time-dependent density functional calculations: Dipole-allowed states of N2. *J. Chem. Phys.* **2004**, *121*, 6155–6166.

(59) Liu, Y.; Martínez Fernández, L.; Cerezo, J.; Prampolini, G.; Improta, R.; Santoro, F. Multistate coupled quantum dynamics of photoexcited cytosine in gas-phase: Nonadiabatic absorption spectrum and ultrafast internal conversions. *Chem. Phys.* **2018**, *515*, 452–463.

(60) Yanai, T.; Tew, D.; Handy, N. A new hybrid exchange−correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393*, 51.

(61) Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110*, 6158.

(62) Enzerhof, M.; Scuseria, G. E. Assessment of the Perdew–Burke–Ernzerhof exchange-correlation functional. *J. Chem. Phys.* **1999**, *110*, 5029.

(63) Improta, R.; Barone, V. In *Photoinduced Phenomena in Nucleic Acids I: Nucleobases in the Gas Phase and in Solvents*; Barbatti, M., Borin, C. A., Ullrich, S., Eds.; Springer International Publishing: Cham, Switzerland, 2015; Vol. 355; pp 329–358.

(64) Stratmann, R. E.; Scuseria, G. E.; Frisch, M. J. An efficient implementation of time-dependent density-functional theory for the calculation of excitation energies of large molecules. *J. Chem. Phys.* **1998**, *109*, 8218–8224.

(65) Casida, M.; Huix-Rotllant, M. Progress in Time-Dependent Density-Functional Theory. *Annu. Rev. Phys. Chem.* **2012**, *63*, 287–323.

(66) Send, R.; Valsson, O.; Filippi, C. Electronic Excitations of Simple Cyanine Dyes: Reconciling Density Functional and Wave Function Methods. *J. Chem. Theory Comput.* **2011**, *7*, 444–455.

(67) Moore, B.; Autschbach, J. Longest-Wavelength Electronic Excitations of Linear Cyanines: The Role of Electron Delocalization and of Approximations in Time-Dependent Density Functional Theory. *J. Chem. Theory Comput.* **2013**, *9*, 4991–5003.

(68) Liu, Y.; Cerezo, J.; Lin, N.; Zhao, X.; Improta, R.; Santoro, F. Comparison of the results of a mean-field mixed quantum/classical method with full quantum predictions for nonadiabatic dynamics: application to the $\pi\pi^*/n\pi^*$ decay of thymine. *Theor. Chem. Acc.* **2018**, *137*, 40.

(69) Worth, G. A.; Giri, K.; Richings, G. W.; Beck, M. H.; Jäckle, A.; Meyer, H.-D. The QUANTICS Package, Version 1.1, (2015), University of Birmingham, Birmingham, U.K.

(70) Raab, A.; Worth, G. A.; Meyer, H.-D.; Cederbaum, L. S. Molecular dynamics of pyrazine after excitation to the S2 electronic state using a realistic 24-mode model Hamiltonian. *J. Chem. Phys.* **1999**, *110*, 936–946.

(71) Cerezo, J.; Liu, Y.; Lin, N.; Zhao, X.; Improta, R.; Santoro, F. Mixed Quantum/Classical Method for Nonadiabatic Quantum Dynamics in Explicit Solvent Models: The $\pi\pi^*/\mathrm{n}\pi^*$ Decay of Thymine in Water as a Test Case. *J. Chem. Theory Comput.* **2018**, *14*, 820–832.

(72) Schreiber, M.; Silva-Junior, M. R.; Sauer, S. P. A.; Thiel, W. Benchmarks for electronically excited states: CASPT2, CC2, CCSD, and CC3. *J. Chem. Phys.* **2008**, *128*, 134110.

(73) Szalay, P. G.; Watson, T.; Perera, A.; Lotrich, V. F.; Bartlett, R. J. Benchmark Studies on the Building Blocks of DNA. 1. Superiority of Coupled Cluster Methods in Describing the Excited States of Nucleobases in the Franck–Condon Region. *J. Phys. Chem. A* **2012**, *116*, 6702–6710.

(74) Nakayama, A.; Harabuchi, Y.; Yamazaki, S.; Taketsugu, T. Photophysics of cytosine tautomers: new insights into the nonradiative decay mechanisms from MS-CASPT2 potential energy calculations and excited-state molecular dynamics simulations. *Phys. Chem. Chem. Phys.* **2013**, *15*, 12322–12339.

(75) Blancafort, L.; Cohen, B.; Hare, P. M.; Kohler, B.; Robb, M. A. Singlet Excited-State Dynamics of 5-Fluorocytosine and Cytosine: An Experimental and Computational Study. *The Journal of Physical Chemistry A* **2005**, *109*, 4431–4436.

(76) Li, Q.; Blancafort, L. Photochemistry and photophysics of the amino and imino tautomers of 1-methylcytosine: tautomerisation as a side product of the radiationless decay. *Photochem. Photobiol. Sci.* **2013**, *12*, 1401–1408.

(77) Ullrich, S.; Schultz, T.; Zgierski, M. Z.; Stolow, A. Electronic relaxation dynamics in DNA and RNA bases studied by time-resolved photoelectron spectroscopy. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2796–2801.

(78) Kotur, M.; Weinacht, T. C.; Zhou, C.; Kistler, K. A.; Matsika, S. Distinguishing between relaxation pathways by combining dissociative ionization pump probe spectroscopy and ab initio calculations: A case study of cytosine. *J. Chem. Phys.* **2011**, *134*, 184309.

(79) Röttger, K.; Marroux, H. J. B.; Böhnke, H.; Morris, D. T. J.; Voice, A. T.; Temps, F.; Roberts, G. M.; Orr-Ewing, A. J. Probing the excited state re-

laxation dynamics of pyrimidine nucleosides in chloroform solution. *Faraday Discuss.* **2016**, *194*, 683–708.

(80) 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.

(81) Plasser, F.; Gómez, S.; Menger, M. F. S. J.; Mai, S.; González, L. Highly efficient surface hopping dynamics using a linear vibronic coupling model. *Phys. Chem. Chem. Phys.* **2019**, *21*, 57–69.

(82) Mai, S.; González, L. Identification of important normal modes in nonadiabatic dynamics simulations by coherence, correlation, and frequency analyses. *J. Chem. Phys.* **2019**, *151*, 244115.

Graphical TOC Entry