JCTC Journal of Chemical Theory and Computation

Open Access

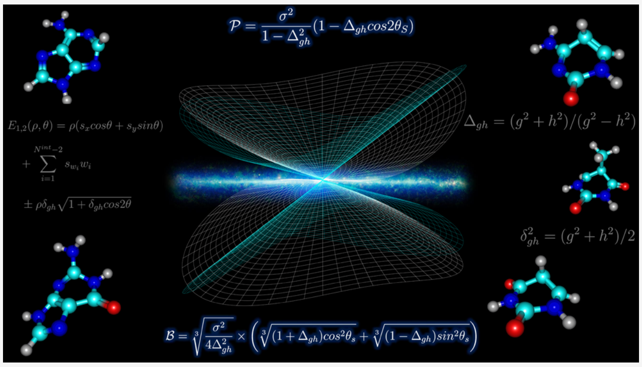
This article is licensed under CC-BY 4.0

Characterizing Conical Intersections in DNA/RNA Nucleobases with Multiconfigurational Wave Functions of Varying Active Space Size

Juliana Cuéllar-Zuquin, Ana Julieta Pepino, Ignacio Fdez. Galván,* Ivan Rivalta, Francesco Aquilante, Marco Garavelli, Roland Lindh,* and Javier Segarra-Martí*

Cite This: *J. Chem. Theory Comput.* 2023, 19, 8258–8272 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information



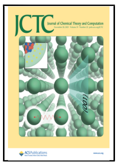
ABSTRACT: We characterize the photochemically relevant conical intersections between the lowest-lying accessible electronic excited states of the different DNA/RNA nucleobases using Cholesky decomposition-based complete active space self-consistent field (CASSCF) algorithms. We benchmark two different basis set contractions and several active spaces for each nucleobase and conical intersection type, measuring for the first time how active space size affects conical intersection topographies in these systems and the potential implications these may have toward their description of photoinduced phenomena. Our results show that conical intersection topographies are highly sensitive to the electron correlation included in the model: by changing the amount (and type) of correlated orbitals, conical intersection topographies vastly change, and the changes observed do not follow any converging pattern toward the topographies obtained with the largest and most correlated active spaces. Comparison across systems shows analogous topographies for almost all intersections mediating population transfer to the dark $^1n_{O/N}\pi^*$ states, while no similarities are observed for the "ethylene-like" conical intersection ascribed to mediate the ultrafast decay component to the ground state in all DNA/RNA nucleobases. Basis set size seems to have a minor effect, appearing to be relevant only for purine-based derivatives. We rule out structural changes as a key factor in classifying the different conical intersections, which display almost identical geometries across active space and basis set change, and we highlight instead the importance of correctly describing the electronic states involved at these crossing points. Our work shows that careful active space selection is essential to accurately describe conical intersection topographies and therefore to adequately account for their active role in molecular photochemistry.

1. INTRODUCTION

DNA/RNA nucleobases are the chromophoric units of our genetic material and have been extensively studied over the years given their prominent role in deleterious photochemical mutations as well as due to their outstanding intrinsic photostability.[1−4] Their photostability is strongly related to the ability possessed by DNA bases to dissipate the excess energy gained upon absorption in a nonradiative manner at ultrafast time scales, being mediated by a variety of conical intersections (CIs) connecting different excited electronic states with each other and with the ground state, and enabling an efficient funnelling down of excited-state population converting it into thermal energy.[5,6] This self-protecting mechanism against UV-light damage has been proposed to be key for selecting the nucleobases as the building blocks of our genetic material at prebiotic times under extreme UV-light exposure,[7−10] by choosing the most suitable (photostable) compounds and thus aiding in its photoprotective design, further securing efficient DNA replication.

The existence of conical intersections is nowadays widely recognized and their importance often highlighted in modern photochemistry,[11−13] their direct characterization being still elusive by experimental means[14] even though indirect fingerprints have been recorded.[15] Given the importance of CIs for rationalizing the outcome of a given photochemical reaction,[16−18] recent efforts have been put toward efficiently facilitating their characterization in realistic systems. This includes extending the available protocols to hybrid quantum mechanics/molecular mechanics (QM/MM) schemes,[19,20] hence including solvent effects, or locating CIs with alternative penalty functions[21] and updated branching plane techniques[22]

Received: May 30, 2023
Revised: October 13, 2023
Accepted: October 13, 2023
Published: October 26, 2023



© 2023 American Chemical Society

https://doi.org/10.1021/acs.jctc.3c00577

that do not require the evaluation of expensive nonadiabatic couplings.

Albeit difficult to describe, CIs have been reported in the literature for a number of years, being detailed first for crossings within states of different symmetry[23,24] and later on within states of the same symmetry.[25,26] Different efficient methodologies have been devised over the years for CI optimization: these range from linear-[27] to quadratic-based[28] projection techniques to the use of the nudged elastic band method,[29] and being mostly combined with multiconfigurational techniques that possess first-order nonadiabatic coupling formulations and implementations.[30−34] Characterizing the potential energy surfaces around conical intersections is important as their shape or topography is believed to influence excited-state reactivity[17,35−38] and may therefore be fundamental to fully understand photochemistry.

Here, we present a systematic study of the topography of the lowest-lying conical intersections of the different pyrimidine (uracil, thymine, and cytosine) and purine (adenine and guanine) nucleobases (see Figure 1). Despite nucleobase CIs being reported in the literature over the years,[39−48] this work represents to our knowledge the first systematic analysis of the active space dependence on the characterization of such critical points and provides also an overview of the potential impact shown by the diffuseness of the basis set employed in the optimization procedure.

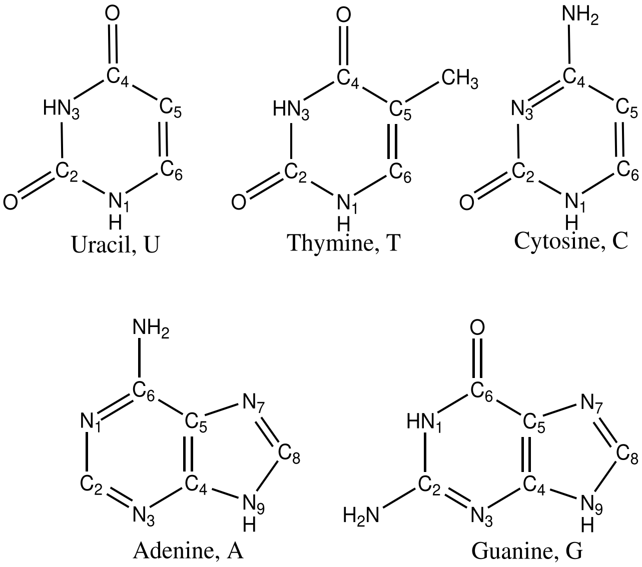


Figure 1. Chemical structure and atom labeling of the DNA and RNA pyrimidine-based (upper half) and purine-based (lower half) nucleobases.

Active space selection[49] is a complex process that can potentially affect the outcome of simulations, and our thorough analysis reveals its impact. Using multiple different active spaces for each of the five canonical nucleobases (and totaling over 300 optimized CIs), we find that the different electron correlation included through varying active spaces strongly shapes CI topography: active space strongly affects conical intersection topographies in almost all cases studied. When examining pyrimidines, we observe that the states less affected by changes in the active space were those involving $^1n_O\pi^*$, with the exception of uracil, which showed significant sensitivity to these changes. On the other hand, conical intersections involving an excited state (regardless of its nature) and the ground state in purines were more influenced by changes in the (static) electron correlation retained in the model. Here, we focus on conical intersection topography due to active space size: other important aspects to excited-state reactivity, such as the accessibility of the different conical intersections characterized, are beyond the scope of the present manuscript.

Interestingly, comparing optimized structures for a given conical intersection reveals minimal changes in the resulting geometry due to active space size. These findings suggest that the topography of conical intersections may depend primarily on the accurate description of the electronic states involved in the crossing rather than on the optimization procedure itself. These results have significant implications for modeling and understanding photoinduced phenomena in DNA and RNA nucleobases.

The manuscript is organized as follows: we first cover the computational details describing the active space selection procedure and other simulation details. The results are separated between pyrimidine and purine derivatives to simplify the analysis, considering the different relevant low-lying (and thus photochemically accessible) CIs featured in the diverse DNA/RNA building blocks. A discussion follows from the similarities that emerge among the different nucleobases, ending with Section 5 that summarizes the present findings.

2. COMPUTATIONAL DETAILS

All computations were carried out with the OpenMOLCAS package.[50−52] Different complete active space self-consistent field (CASSCF)[53] schemes were considered as shown in Figures S1−S5 in the Supporting Information (SI): a systematic procedure was employed whereby active space size was reduced from their full π (occupied and virtual) and $n_{O/N}$ (occupied) valence active space (14 electrons in 10 orbitals (or (14,10) from here onward) for pyrimidines, (16,12) for adenine, and (18,13) for guanine) by removing one by one the least contributing orbitals in terms of occupation number, i.e., those occupied closest to 2 and those unoccupied closest to 0. This leads to slightly different active spaces for the different systems and the different conical intersections studied but that are more consistent within specific systems and CIs. It is worth noting adenine and guanine feature (18,13) and (20,14) full π and $n_{O/N}$ valence active spaces, respectively: in these cases, one n_N orbital was removed due to its low contribution and to make computations more feasible (see Figures S4 and S5). Tables S1−S36 in the SI contain the specific orbitals included and information about natural orbital occupation numbers and the natural intersection orbitals[54] for each of the conical intersections for each of the nucleobases studied. To optimize conical intersections with different active spaces, we used the most correlated (i.e., CASSCF(14,10) for pyrimidines, CASSCF(16,12) for adenine, and CASSCF(18,13) for guanine) as the starting reference. An equal weights state averaging procedure was employed comprising the lowest-lying five roots for pyrimidines and seven roots for purines whenever feasible, with the smallest active spaces being averaged over three roots.

Atomic Natural Orbital basis sets with a large contraction (ANO-L)[55,56] were used in their double-ζ (VDZP) and triple-ζ (VTZP, available in the SI) polarized contractions. The atomic compact Cholesky decomposition (acCD)[57] was used throughout to speed-up the two-electron integrals[58−60] as well as for computing analytic CASSCF gradients[61] and non-adiabatic couplings.[34] Conical intersections were characterized using the method introduced by Fdez. Galván et al.[34] and available in OpenMolcas.[50,51]

All minimum energy conical intersections (referred to as conical intersections from here onward) optimized in this work were analyzed in terms of their $\mathcal{P}$ and $\mathcal{B}$ parameters, defined in previous work by Fdez. Galván et al.[34] and that are related to those originally formulated by Ruedenberg and co-workers in their seminal work on conical intersections.[26] These values are defined as

$$\mathcal{P} = \frac{\sigma^2}{1 - \Delta_{gh}^2}(1 - \Delta_{gh}\cos 2\theta_s) \tag{1}$$

$$\mathcal{B} = \sqrt[3]{\frac{\sigma^2}{4\Delta_{gh}^2}} \times \left(\sqrt[3]{(1 + \Delta_{gh})\cos^2\theta_s} + \sqrt[3]{(1 - \Delta_{gh})\sin^2\theta_s}\right) \tag{2}$$

where σ refers to the pitch, Δ_{gh} to the asymmetry, and θ_s to the tilt heading.[34] Depending on the value taken by these two parameters ($\mathcal{P}$ and $\mathcal{B}$), we have the following classification system

$$\mathcal{P}\begin{cases} >1 & \text{sloped} \\ <1 & \text{peaked} \end{cases} \text{ and } \mathcal{B}\begin{cases} >1 & \text{single-path} \\ <1 & \text{bifurcating} \end{cases}$$

A three-dimensional view of the two potential energy surfaces in the branching space for each of the conical intersections resulting from the combination of the $\mathcal{P}$ and $\mathcal{B}$ parameters described above can be found in Figure 2.

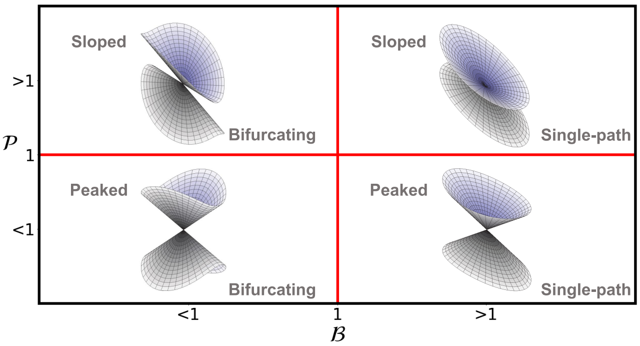


Figure 2. Classification of the conical intersections in terms of $\mathcal{P}$ and $\mathcal{B}$ parameters with a three-dimensional representation of how are the two potential energy surfaces in the branching space.

3. RESULTS AND DISCUSSION

The results are divided in two sections: one for pyrimidine-based (uracil, thymine, and cytosine) and one for purine-based (adenine and guanine) DNA/RNA nucleobases and their (photochemically) most relevant low-energy conical intersections. We start by analyzing the common trends across the different pyrimidine-based systems upon active space change, as they are often grouped together due to their similar photochemistry, and then move onto analyzing how strong correlation affects purine-based systems that display larger reactivity and structural differences. We end by discussing the diverse trends that emerge from comparison across different molecular structures and their resulting conical intersection topographies and attempt to draw some structure−function relationships based on this to compare against the different photochemical behaviors recorded experimentally in the literature for DNA/RNA nucleobases.

3.1. Pyrimidines. The pyrimidine nucleobases uracil, thymine, and cytosine have a range of low-energy accessible conical intersections, which are depicted in Figures 3−6. These CIs connect the initially accessed $^1\pi\pi^*$ with neighboring dark states of $^1n_O\pi^*$ (as well as $^1n_N\pi^*$ for cytosine) and the ground state. This leads to $(^1\pi\pi^*/S_0)_{CI}$, $(^1\pi\pi^*/^1n_O\pi^*)_{CI}$, and $(^1n_O\pi^*/S_0)_{CI}$ (as well as $(^1n_N\pi^*/S_0)_{CI}$ and $(^1\pi\pi^*/^1n_N\pi^*)_{CI}$ for cytosine). All these different intersections have been invoked (to different degrees)[6,10,45,62] to rationalize the rich photophysical landscape observed experimentally in DNA pyrimidine-based monomers, their characterization being paramount to fully understand UV-induced DNA photochemistry.

Besides these, it should be mentioned that a number of 3-state conical intersections (i.e., where three different singlet states are degenerate[63]) have also been reported in the literature for these systems, but they are beyond the scope of the present manuscript.[10,40,44,64]

$(^1\pi\pi^*/S_0)_{CI}$ is perhaps the best known intersection in DNA/RNA pyrimidine nucleobases and entails a pronounced "ethylene-like"[65−67] twisting leading to an out-of-plane motion in the C5 position[10,39,41−43,45,48,62,68−78] shared across all pyrimidine-based nucleobases[10,79] and that is associated with their fastest component of their decay.[68,80] The main distortions shown by this structure are a pronounced C5−C6 elongation followed by an out-of-plane motion, which are qualitatively shown as insets in Figure 3.

Upon characterizing this intersection with a wide range of active spaces, and thus including varying amounts of electron correlation, we observe different trends for the different pyrimidine nucleobases: cytosine (Figure 3a) shows a consistent peaked and bifurcating character throughout the different active spaces (and basis sets) tested with the exception of (8,6) that appears classed as peaked and single-path with a value of $\mathcal{B}$ slightly above 1. Uracil (Figure 3b), on the other hand, favors a peaked and single-path topography with the exception of (6,5) and (4,3) spaces, even if these values approach $\mathcal{B} \approx 1$. Thymine (Figure 3c) displays a much more complex behavior as it features intersections classed in almost every single quadrant (except sloped and bifurcating) depending on the active space used: the largest and presumably most accurate (14,10) space predicts this intersection to be peaked and single-path, while the second most correlated calculation (referred to the (12,9) active space, marked with a brown square in Figure 3c) is classed as peaked and bifurcating but very close to the limits imposed by the classification, as is also (8,7) but bordering in this case from the sloped and single-path quadrant. The remaining optimized structures with other active spaces cluster up at values referred to sloped and single-path topographies.

Significant changes in the $(^1\pi\pi^*/S_0)_{CI}$ conical intersection topography emerge when varying the correlation included in the model, with no clear trends discernible, which highlights how electron correlation strongly impacts the properties of these important structures. Interestingly, changes are also prominent when considered across pyrimidine nucleobase species: assuming the most correlated (14,10) calculation, which includes all valence π, π^*, and lone pair occupied $n_{O/N}$ orbitals, uracil and thymine (Figure 3b,c) show a consistent

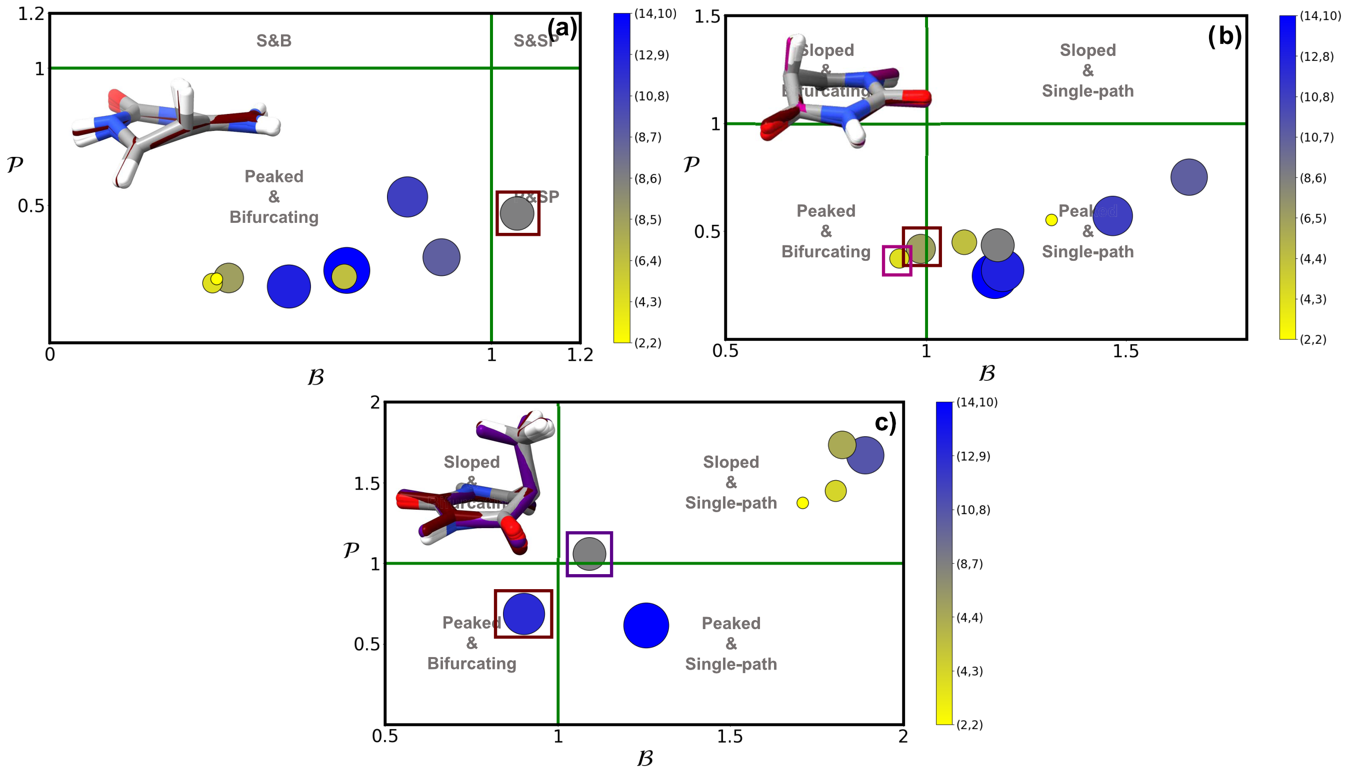


Figure 3. $\mathcal{P}$ and $\mathcal{B}$ parameters of $(^1\pi\pi^*/S_0)_{CI}$ using multiple different active spaces (see Section 2) for (a) cytosine, (b) uracil and (c) thymine. Active space size is denoted by both marker size and the contour gradient color provided in the right-hand side of each panel. A picture with the superimposed geometries of all optimized conical intersections are provided as insets, with the colored structures representing the outlier intersections marked with a square. An analogous picture with the results obtained using a triple-ζ basis set can be found in the SI.

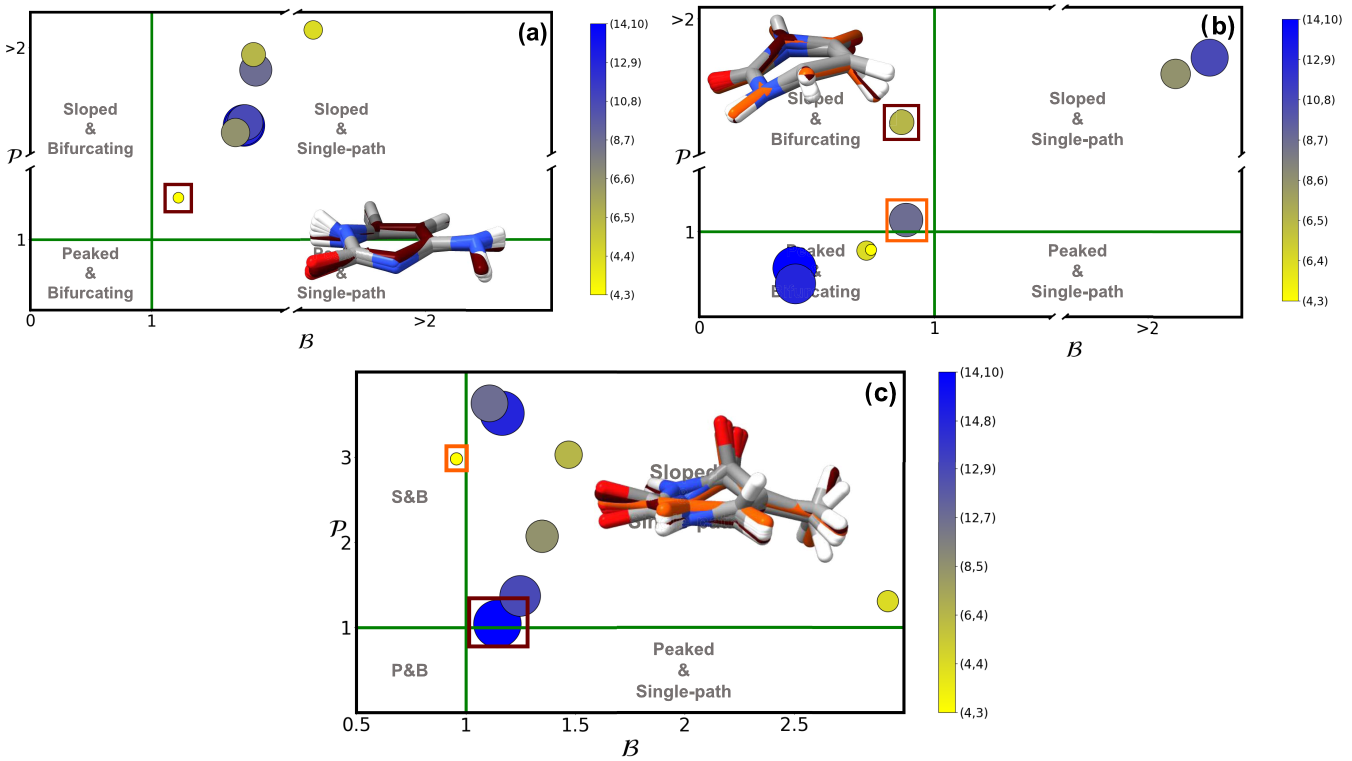


Figure 4. $\mathcal{P}$ and $\mathcal{B}$ parameters of $(^1n_O\pi^*/^1\pi\pi^*)_{CI}$ using multiple different active spaces (see Section 2) for (a) cytosine, (b) uracil, and (c) thymine. Active space size is denoted by both marker size and the contour gradient color provided in the right-hand side of each panel. A picture with the superimposed geometries of all optimized conical intersections are provided as insets, with the colored structures representing the outlier intersections marked with a square. An analogous picture with the results obtained using a triple-ζ basis set can be found in the SI.

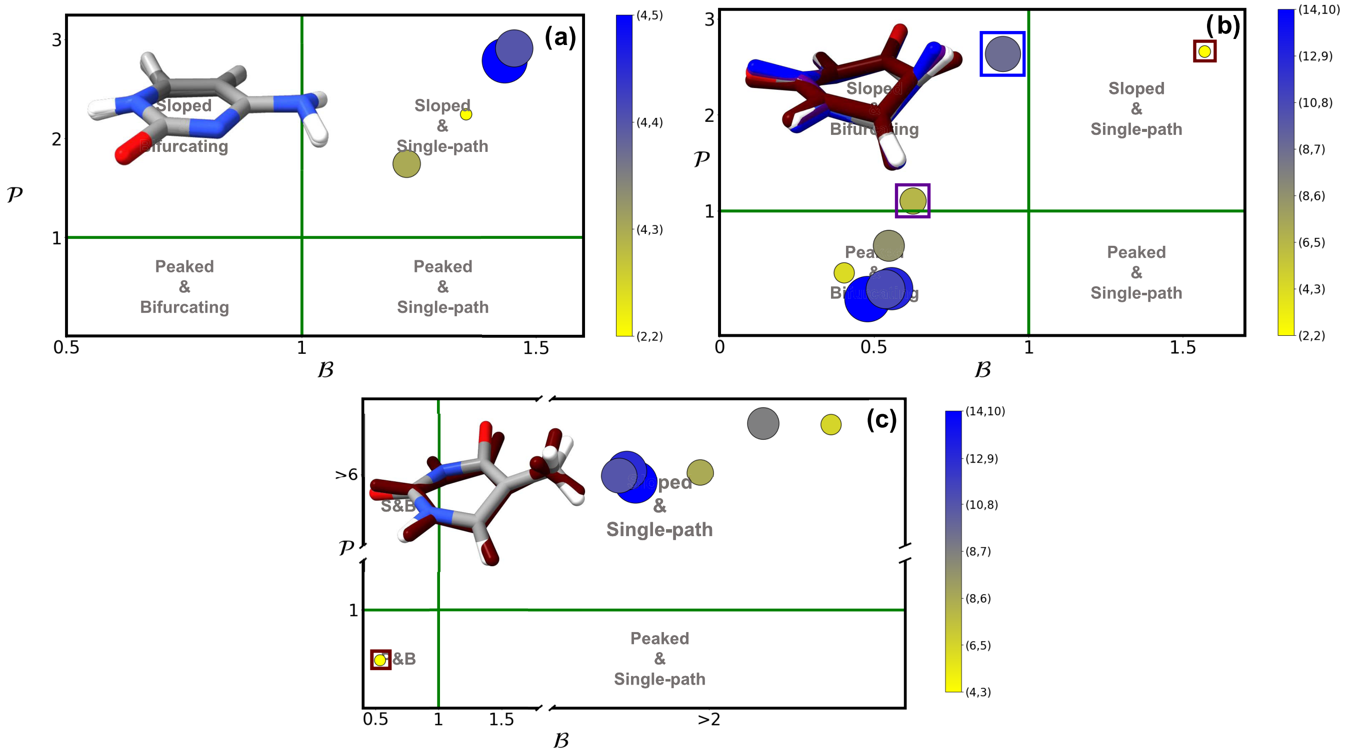


Figure 5. $\mathcal{P}$ and $\mathcal{B}$ parameters of the $({}^1\mathrm{n_O}\pi^*/\mathrm{S_0})_{\mathrm{CI}}$ using multiple different active spaces (see Section 2) for (a) cytosine, (b) uracil, and (c) thymine. Active space size is denoted by both marker size and the contour gradient color provided in the right-hand side of each panel. A picture with the superimposed geometries of all optimized conical intersections are provided as insets, with the colored structures representing the outlier intersections marked with a square. An analogous picture with the results obtained using a triple-ζ basis set can be found in the SI.

description of this intersection as peaked and single-path, whereas cytosine (Figure 3a) appears to be peaked and bifurcating.

Despite the vast changes observed in topography, structurally all optimized $({}^1\pi\pi^*/\mathrm{S_0})_{\mathrm{CI}}$ conical intersections are very similar. Root-mean-square deviation (RMSD) analyses provided in the SI (Figures S13−S15) show very small values, with deviations below 0.2 Å, which are not sufficient to appropriately discriminate even the outlier cases described above in terms of intersection topography. Interestingly, these small RMS deviations do correlate with noticeable changes in a dihedral angle showcasing out-of-plane motions for each of the pyrimidine nucleobases: H-N$_1$-C$_4$-O for uracil and thymine and H-N$_1$-C$_4$-N in cytosine.

We look next at $({}^1\mathrm{n_O}\pi^*/{}^1\pi\pi^*)_{\mathrm{CI}}$ that facilitates population transfer to the optically dark ${}^1\mathrm{n_O}\pi^*$ state and that is present in all pyrimidine nucleobases.[81−85]

In cytosine (Figure 4a), we predict a sloped and single-path character for all different active spaces, with (4,3) approaching borderline values of both $\mathcal{P} \approx 1$ and $\mathcal{B} \approx 1$ but still remaining within the same quadrant as the other calculations. Uracil, on the other hand, displays the largest deviations: the most correlated (14,10) calculation points at a peaked and bifurcating character, while (10,8) and (8,6) display a sloped and single-path topography with values of $\mathcal{B} \approx 1$ and (6,5) and (8,7) classified as sloped and bifurcating with values of $\mathcal{P}$ close to 1 in the case of the latter. Thymine (Figure 4c), like cytosine (Figure 4a), shows a sloped and single-path topography with the exception of the (4,3) calculation that has nevertheless values approaching $\mathcal{B} \approx 1$, particularly for its triple-ζ (see SI Figure S7) basis set result.

Looking at the outliers from Figure 4, we can see in this case some clear differences in their structure, particularly for thymine (cf. Figure 4c), where the (4,3) active space puckers much more heavily around the O$=$C$_2$-N$_3$-C$_4$ $=$ O frame and is clearly discernible from all other optimized structures. Cytosine (Figure 4a) and uracil (Figure 4b) also feature pronounced differences observable by the naked eye in the insets in Figure 4, even if their RMSD analyses are not fully conclusive (see SI Figures S16−S18): we observe how the structures displaying the largest RMSD values (with respect to the most precise (14,10) calculations) correlate with pronounced changes in the H-N$_3$-C$_6$-H dihedral angle for uracil and thymine and with the H-C$_5$-C$_2$-O angle for cytosine. These RMS deviations are of the order of ∼0.5 Å and are therefore larger than those observed in the $({}^1\pi\pi^*/\mathrm{S_0})_{\mathrm{CI}}$ while being relatively small. They may also be understood as the structures pucker either upward or downward, the varying active spaces leading to either of these analogous (but slightly different) geometries in the gas phase. Optimizations of both upward and downward intersections for each case were attempted, but they always converged to either one of the out-of-plane motions for the different active space studies: we expect this to change when considering nucleobases embedded in complex realistic environments, such as a double-helix DNA structure, where upward and downward (or endo and exo)[86] conformations result in markedly different energies due to steric hindrance and/or other intermolecular interactions.

The last intersection shared by all pyrimidines is $({}^1\mathrm{n_O}\pi^*/\mathrm{S_0})_{\mathrm{CI}}$, which is shown in Figure 5. Cytosine (Figure 5a) displays a well-defined sloped and single-path topography for all active spaces optimized. In this particular case, we encountered a problem when converging larger active spaces

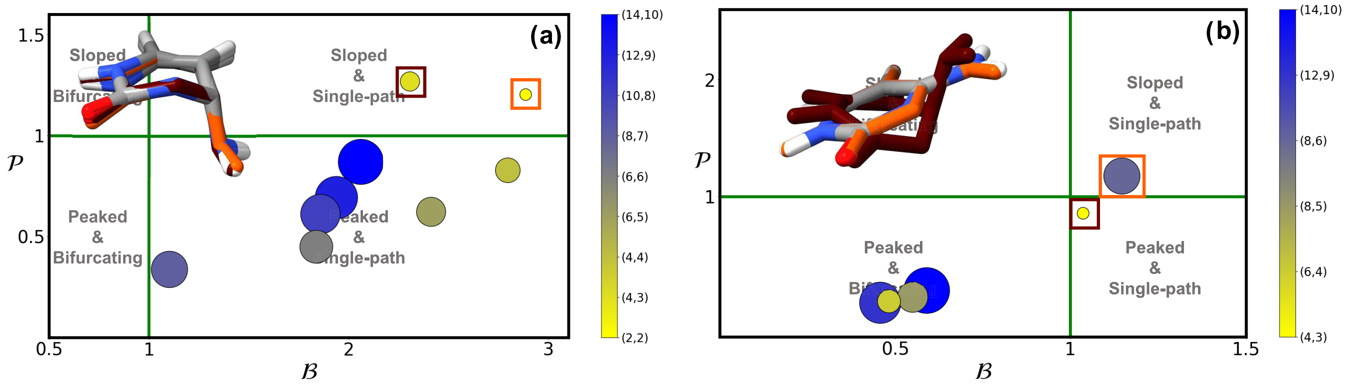


Figure 6. $\mathcal{P}$ and $\mathcal{B}$ parameters of the (a) $(^1n_N\pi^*/S_0)_{CI}$ and (b) $(^1n_N\pi^*/^1\pi\pi^*)_{CI}$ of cytosine. Active space size is denoted by both marker size and the contour gradient color provided in the right-hand side of each panel. Pictures with the superimposed geometries of all optimized conical intersections are provided as insets, with the colored structures representing the outlier intersections marked with a square. An analogous picture with the results obtained using a triple-ζ basis set can be found in the SI.

for cytosine: by including more π-type orbitals (both occupied and virtual) optimization leads to $(^1n_O\pi^*/^1\pi\pi^*)_{CI}$ instead. It is worth noting, however, that an almost 3-state degeneracy is observed in all cases reported, which is in line with previous work that assigns this to be a 3-state conical intersection.[44] Uracil (Figure 5b) has a peaked and bifurcating character in most of the conical intersections, with the exceptions of (6,5) and (8,7) classified as sloped and bifurcating and the smaller active space used, (2,2), in the sloped and single-path quadrant. Thymine (Figure 5c), like cytosine, shows a sloped and single-path character for all active spaces studied, with the largest deviation coming from the (4,3) calculation, which is placed in the peaked and bifurcating quadrant.

Similarities arise between $(^1n_O\pi^*/\pi\pi^*)_{CI}$ and $(^1n_O\pi^*/S_0)_{CI}$: they are both consistent upon active space change for cytosine and thymine but appear to be very sensitive to changes in the active space of uracil. This suggests that there is a significant correlation between active space size and the ability to adequately describe the $^1n_O\pi^*$ excited state in uracil that is not present in cytosine or thymine, despite the latter featuring almost the same molecular structure.

Structurally, $(^1n_O\pi^*/S_0)_{CI}$ features similar differences to those observed above for $(^1n_O\pi^*/^1\pi\pi^*)_{CI}$: the intersection triggers a ring puckering that can go either of two ways (upward or downward) depending on the specific active space employed. This leads to larger deviations that are readily observed by the out-of-plane motions at H-N_1-C_4-N in cytosine (Figure 5a), H-N_3-C_6-H in uracil (Figure 5b), and O-C_2-C_5-C in thymine (Figure 5c) and where the changes in these dihedrals correlate with the increases in RMSD observed with respect to the (14,10) reference structures (see Figures S19−S21).

The last two conical intersections, $(^1n_N\pi^*/S_0)_{CI}$ and $(^1n_N\pi^*/^1\pi\pi^*)_{CI}$, involve a dark $^1n_N\pi^*$ state and are therefore present only in cytosine.

Figure 6a shows the different estimates obtained for $(^1n_N\pi^*/S_0)_{CI}$, which is classed as peaked and single-path but whose reference (14,10) calculation lies almost at the frontier ($\mathcal{P} \approx 1$) of being of sloped character, particularly when using a triple-ζ basis set (see Figure S9). The rest of the calculations remain within the same quadrant with the exception of the smaller (4,3) and (2,2) active spaces, which appear as sloped and single-path by having overestimated values of $\mathcal{P}$. $(^1n_N\pi^*/^1\pi\pi^*)_{CI}$, presented in Figure 6b, favors a peaked and bifurcating topography for the (14,10) reference calculation. It has two outliers, the (4,3) and (8,6), which lie at the boundaries of the classification within the peaked and sloped single-path quadrants, respectively.

In terms of the structure, $(^1n_N\pi^*/S_0)_{CI}$ features a very pronounced NH_2 out-of-plane motion (Figure 6a) that is well captured by all active spaces tested, leading to almost identical geometries. This is reminiscent of the result reported above in Figure 3 for the $(^1\pi\pi^*/S_0)_{CI}$ ethylene-like intersection in all pyrimidine-based systems, where a pronounced out-of-plane motion leads to almost identical structures, which however present largely different CI topographies. The main differences in structure originate from the O-C_2-C_5-H dihedral angle, which shows minimal changes in all the cases with RMSD values around ∼0.2 Å (see Figure S22).

$(^1n_N\pi^*/^1\pi\pi^*)_{CI}$, on the other hand, features a much more planar structure (Figure 6b) with the exception of the (4,3) outlier, which results in a heavily puckered structure. This puckering is related to the H-N_1-C_4-N dihedral angle mentioned above and depicted in Figure S23 in the SI, which shows a concomitant increase of this dihedral angle with the RMSD computed against the (14,10) reference structure, reaching a value of ∼0.7 Å for the most distorted (4,3) geometry. From this, we may conclude that $(^1n\pi^*/^1\pi\pi^*)_{CI}$ intersections do appear to converge to different results in terms of both CI topographies and geometries when computed with minimal active spaces.

Overall, we observe rather pronounced changes in conical intersection topography when varying the active space in DNA/RNA pyrimidine nucleobase monomers: with the exception of those intersections featuring $^1n_O\pi^*$ states in cytosine and thymine, the rest display a vast array of conical intersection topographies that strongly depend on the size of the active space. Structurally, less-pronounced changes are observed, which suggests these differences in topography to be rooted in the ability (or inability) of the different active spaces to appropriately represent the participating electronic states, with the $^1n_O\pi^*$ state in uracil being the most challenging case.

3.2. Purines.

We turn our attention next to the purine-based canonical nucleobases guanine and adenine. They feature two fused 6- and 5-member cyclic moieties that result in larger molecular frames that correlate more active molecular orbitals and thus leads to more low-lying relevant electronic excited states.

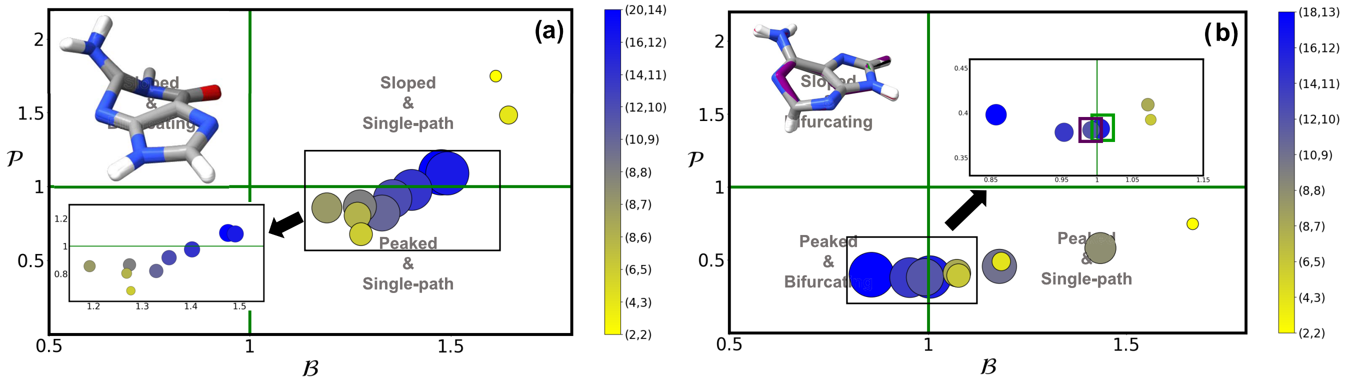


Figure 7. $\mathcal{P}$ and $\mathcal{B}$ parameters of $(\mathrm{L_a}(^1\pi\pi^*)/\mathrm{S_0})_{\mathrm{CI}}$ for (a) guanine and (b) adenine using multiple different active spaces (see Section 2). Active space size is denoted by both marker size and the contour gradient color provided in the right-hand side of each panel. Pictures with the superimposed geometries of all optimized conical intersections are provided as insets, with the colored structures representing the outlier intersections marked with a square. An analogous picture with the results obtained using a triple-ζ basis set can be found in the SI.

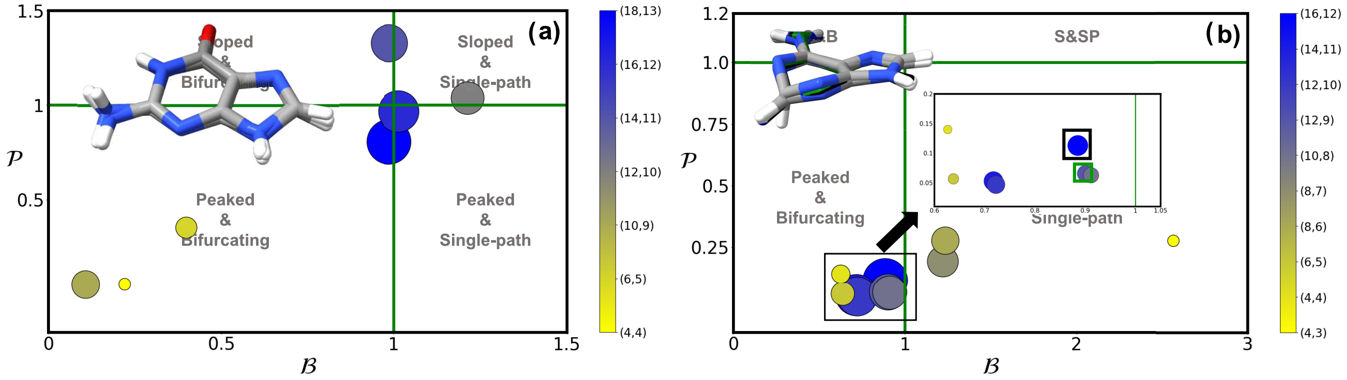


Figure 8. $\mathcal{P}$ and $\mathcal{B}$ parameters of $(\mathrm{L_a}(^1\pi\pi^*)/\mathrm{L_b}(^1\pi\pi^*))_{\mathrm{CI}}$ using multiple different active spaces (see Section 2) for guanine (a) and adenine (b). Active space size is denoted by both marker size and the contour gradient color provided in the right-hand side of each panel. Pictures with the superimposed geometries of all optimized conical intersections are provided as insets.

Due to their larger conjugated frame, purine-based nucleobases feature two low-lying $^1\pi\pi^*$ states that are relevant to the photophysics, often named $\mathrm{L_a}(^1\pi\pi^*)$ and $\mathrm{L_b}(^1\pi\pi^*)$ adapted from Platt's notation,[87] as well as a number of lone pair $\mathrm{n_{N/O}}$ states. This leads to $(\mathrm{L_a}(^1\pi\pi^*)/\mathrm{S_0})_{\mathrm{CI}}$, $(\mathrm{L_a}(^1\pi\pi^*)/\mathrm{L_b}(^1\pi\pi^*))_{\mathrm{CI}}$, and $(\mathrm{L_b}(^1\pi\pi^*)/^1\mathrm{n_N}\pi^*)_{\mathrm{CI}}$ that are common to both bases, as well as $(^1\mathrm{n_N}\pi^*/\mathrm{S_0})_{\mathrm{CI}}$ and $(\mathrm{L_a}(^1\pi\pi^*)/^1\mathrm{n_N}\pi^*)_{\mathrm{CI}}$ for adenine and $(\mathrm{L_a}(^1\pi\pi^*)/^1\mathrm{n_O}\pi^*)_{\mathrm{CI}}$ and $(^1\mathrm{n_O}\pi^*/\mathrm{S_0})_{\mathrm{CI}}$ for guanine.

We refrain from including the last four conical intersections mentioned above as they could not be successfully optimized. References in the literature for $(^1\mathrm{n_O}\pi^*/\mathrm{S_0})_{\mathrm{CI}}$ in guanine[47] and for $(^1\mathrm{n_N}\pi^*/\mathrm{S_0})_{\mathrm{CI}}$ in adenine[88] suggest they feature a pronounced ring-opening component that may not be properly described with the active spaces used here and are therefore considered beyond the scope of the present work. $(\mathrm{L_a}(^1\pi\pi^*)/^1\mathrm{n_N}\pi^*)_{\mathrm{CI}}$ for adenine and $(\mathrm{L_a}(^1\pi\pi^*)/^1\mathrm{n_O}\pi^*)_{\mathrm{CI}}$ for guanine, on the other hand, have been excluded due to most active spaces converging to their respective $(\mathrm{L_a}(^1\pi\pi^*)/\mathrm{L_b}(^1\pi\pi^*))_{\mathrm{CI}}$ intersections instead, making their systematic study unfeasible.

The first intersection to analyze is $(\mathrm{L_a}(^1\pi\pi^*)/\mathrm{S_0})_{\mathrm{CI}}$, which is equivalent to the $(^1\pi\pi^*/\mathrm{S_0})_{\mathrm{CI}}$ in pyrimidines, and that is believed to be the main responsible behind the ultrafast decay of purine-based DNA nucleobases.[45−47,88−97] This crossing is also characterized by an out-of-plane motion that resembles, like in pyrimidine-based nucleobases, the pyramidalized intersection in ethylene.[65]

Both guanine and adenine present $(\mathrm{L_a}(^1\pi\pi^*)/\mathrm{S_0})_{\mathrm{CI}}$, the different CI topographies along active space change being depicted in panels a and b of Figure 7, respectively.

Figure 7a shows the different intersection topographies obtained with varying active spaces for the $(\mathrm{L_a}(^1\pi\pi^*)/\mathrm{S_0})_{\mathrm{CI}}$ in guanine, most of them clustering up along the dividing line between peaked and sloped quadrants within the single-path character. A zoom-in on this region shows that the more correlated (20,14) and (16,12) active spaces are classed as sloped and single-path, with smaller active spaces being placed within the peaked quadrant, even if they still feature values very close to $\mathcal{P} \approx 1$. We observe that (14,11) displays different topographies for double-ζ and triple-ζ (see Figure S10) basis sets: the values are however very close to the dividing value of $\mathcal{P} = 1$, and any differences due to basis set size are therefore expected to be minor.

All $(\mathrm{L_a}(^1\pi\pi^*)/\mathrm{S_0})_{\mathrm{CI}}$ optimized structures in guanine are very similar, as shown by the superimposed geometries displayed as an inset in Figure 7a. Analysis of the RMSD comparing each structure to the most correlated (20,14) active space yields differences in the order of less than 0.1 Å for the largest (and least correlated) cases (see Figure S24), confirming that all optimized structures are essentially equivalent. The small RMS deviations do correlate to an extent with differences registered across geometries in the H-$\mathrm{C_8}$-$\mathrm{C_2}$-N dihedral angle, which

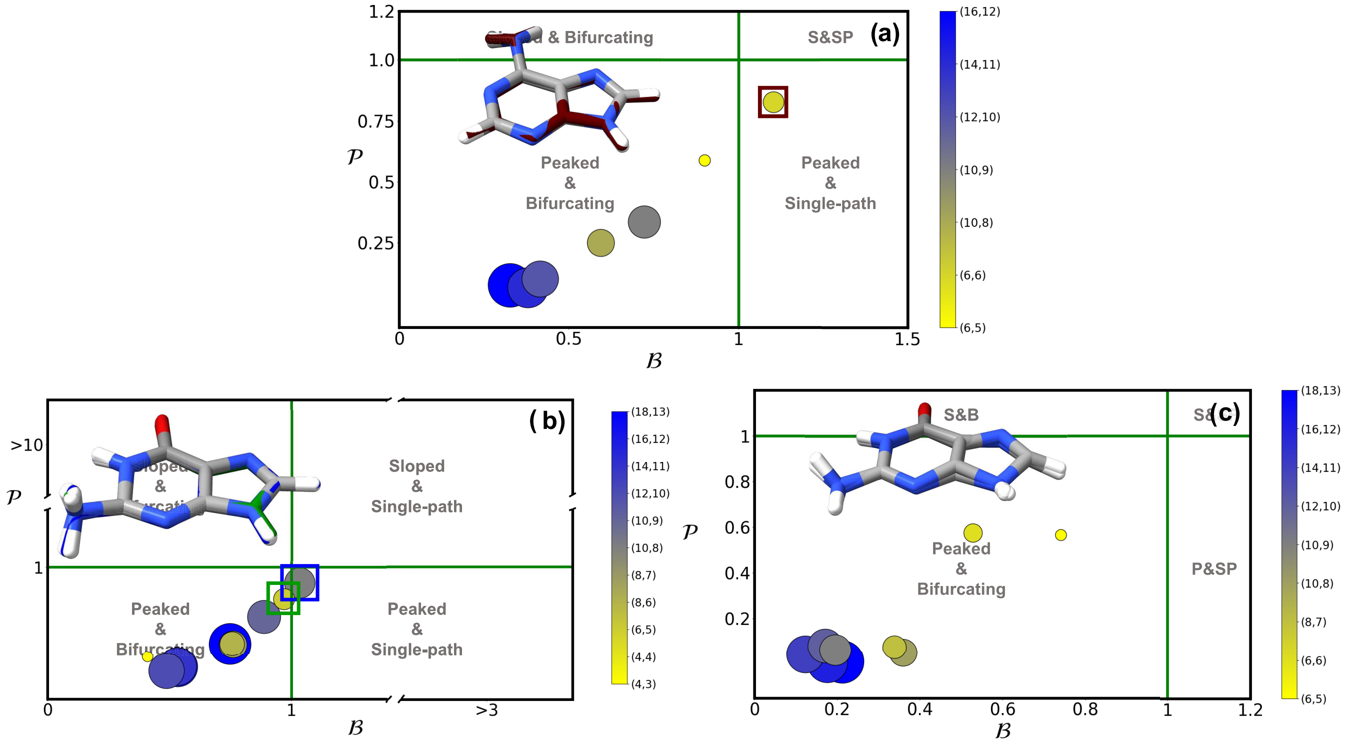


Figure 9. $\mathcal{P}$ and $\mathcal{B}$ parameters of $(\mathrm{L_b}(^1\pi\pi^*)/^1\mathrm{n_N}\pi^*)_{\mathrm{CI}}$ (a) for adenine and $(\mathrm{L_a}(^1\pi\pi^*)/^1\mathrm{n_O}\pi^*)_{\mathrm{CI}}$ (b) and $(\mathrm{L_b}(^1\pi\pi^*)/^1\mathrm{n_N}\pi^*)_{\mathrm{CI}}$ (c) for guanine using multiple different active spaces (see Section 2). Active space size is denoted by both marker size and the contour gradient color provided in the right-hand side of each panel. A picture with the superimposed geometries of all optimized conical intersections is provided as an inset, with the colored structures representing the outlier intersection marked with a purple square.

decreases concomitantly with the active space size, leading to more planar structures when using a less-correlated (2,2) active space.

The different CI topographies for $(\mathrm{L_a}(^1\pi\pi^*)/\mathrm{S_0})_{\mathrm{CI}}$ across multiple active spaces in adenine are shown in Figure 7b. Similar to guanine, we observe a clustering of the most active space topographies along a dividing line, in this case between peaked single-path and peaked bifurcating quadrants. A zoom-in in this region shows the most accurate (18,13) active space being classed as bifurcating, while most of the rest are in the limit between the two quadrants: interestingly, larger differences are observed due to basis set size in this case, changing from the peaked and bifurcating quadrant with double-ζ to peaked and single-path with triple-ζ (see Figure S10).

Structurally, all geometries are almost identical and display a negligible ~0.1 Å RMS difference with respect to the most correlated (18,13) geometry (see Figure S25). The very small differences across structures can be mostly associated with changes in the $\mathrm{N_1}$-$\mathrm{C_6}$—N-H dihedral angle, which approaches −20° for the more correlated (18,13) calculation and ~5° for the least correlated (2,2). Interestingly, in this case, we observe how the $\mathrm{N_1}$-$\mathrm{C_6}$—N-H dihedral angle diverges for double-ζ and triple-ζ calculations when considering small active spaces, particularly (2,2), where there is a significant ~30° difference between them.

Figure 8 reports results for the $(\mathrm{L_a}(^1\pi\pi^*)/\mathrm{L_b}(^1\pi\pi^*))_{\mathrm{CI}}$ conical intersection for both purine nucleobases. In guanine (Figure 8a), the largest active spaces are spread across all four quadrants. The most correlated (18,13) is classified as peaked and bifurcating, while (16,12) is peaked and single-path, the difference being due to a slightly higher value of $\mathcal{B}$. We find that (14,11) lies at the boundary between sloped and bifurcating and sloped and single-path, similar to (12,10) with a value of $\mathcal{P}$ close to 1. The remaining active spaces all lie within the peaked and bifurcating quadrants, in line with the (18,13) reference. In this particular case, more substantial differences are observed due to increasing basis set size: as shown in SI Figure S11, all active spaces are classified as peaked and bifurcating.

In addition, by looking at the superimposed structures provided as an inset, we observe that they all yield essentially the same geometry. This is further verified by the RMSD values with respect to the most correlated (18,13) calculation, which are less than 0.4 Å in all cases (see Figure S26 in the SI). The most significant change is that occurring in the $\mathrm{N_1}$-$\mathrm{C_2}$—N-H dihedral angle, which varies from ~−80° for (18,13) to ~−30° for the smallest (4,4) active space.

We observe different trends for adenine (Figure 8b). Most cases are classified as peaked and bifurcating, with (16,12) being placed at the boundary with the peaked and single-path quadrant. (8,6), (8,7), and (4,3) are, on the other hand, classified as peak and single-path. The largest structural differences are found for the H-$\mathrm{C_2}$-$\mathrm{C_8}$-H dihedral angle. RMSDs shown in SI Figure S27 are almost zero across active spaces classified as peaked and bifurcating and slightly higher for those with a different classification, (4,3) having the largest value that is in line with its larger $\mathcal{B}$ and thus different classification.

The last conical intersections studied are depicted in Figure 9. The first one is the $(\mathrm{L_b}(^1\pi\pi^*)/^1\mathrm{n_N}\pi^*)_{\mathrm{CI}}$ of adenine, and it can be observed in subpanel a of Figure 9. This conical intersection appears to be relatively insensitive to (strong) correlation, with all cases classified as peaked and bifurcating except for (6,6), which has a slightly higher value of $\mathcal{B}$. This

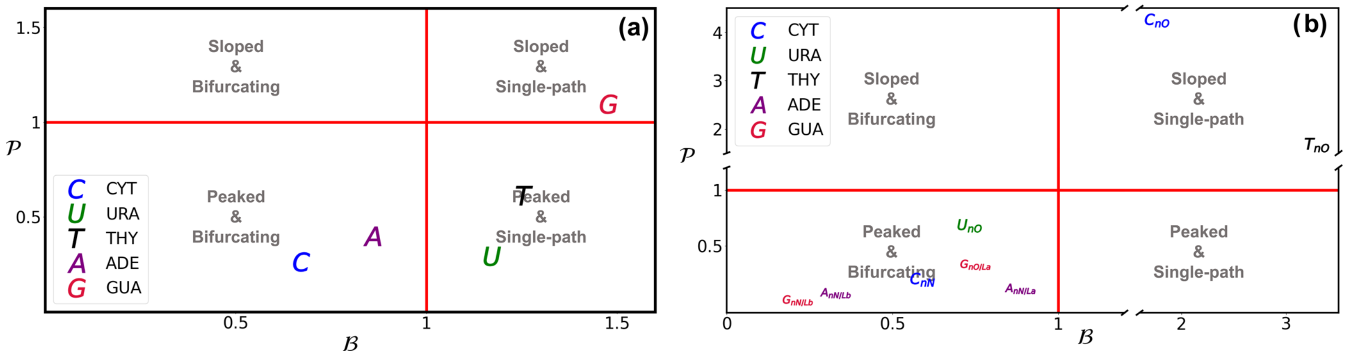


Figure 10. $\mathcal{P}$ and $\mathcal{B}$ parameters of photochemically relevant conical intersections $(^1\pi\pi^*/S_0)_{CI}$ and $(L_a(^1\pi\pi^*)/S_0)_{CI}$ (a) and the different $(^1n_{O/N}/^1\pi\pi^*)_{CI}$ (b) for all DNA/RNA nucleobases, obtained using the largest active space feasible for each system, except for guanine where (18,13) was used in subpanel (b).

small deviation changes its classification to single-path, but its value of $\mathcal{P}$ remains below one, indicating that it is still classified as peaked. RMSD structural analyses and C5-C6−N-H dihedral changes (SI Figure S28) show very small differences, the largest (0.1 Å) being for (6,6) and (6,5) active spaces, which are those appearing at quadrant boundaries in Figure 9a marked by blue and green squares, respectively.

The $(L_a(^1\pi\pi^*)/^1n_O\pi^*)_{CI}$ and $(L_b(^1\pi\pi^*)/^1n_N\pi^*)_{CI}$ of guanine represent the last two conical intersections of interest. In the case of $(L_a(^1\pi\pi^*)/^1n_O\pi^*)_{CI}$ (Figure 9b), almost all cases were classified as peaked and bifurcating, including the most correlated (18,13). The differences between double and triple-ζ (Figure S12) basis sets were the largest observed across all systems and conical intersections. For example, (6,5) appears as peaked and bifurcating with double-ζ and as peaked and single-path with triple-ζ. Likewise, (10,8) is classed as peaked and single-path with double-ζ and sloped and single-path with triple-ζ. These cases fall at the frontier between three quadrants, and thus small changes in $\mathcal{P}$ or $\mathcal{B}$ can have a significant impact on topography. The RMSDs for this conical intersection are less than 0.2 Å (Figure S29 in the SI), which suggests very small structural changes despite the vast changes in topography noted above.

In the $(L_b(^1\pi\pi^*)/^1n_N\pi^*)_{CI}$ conical intersection of guanine (Figure 9c), all active spaces lead to a peaked and bifurcating topography. At a difference from the previous intersection, small changes arise between double- and triple-ζ (Figure S12) results, which do not affect their classification. Both adenine (Figure 9a) and guanine (Figure 9b) have similar $(L_b(^1\pi\pi^*)/^1n_N\pi^*)_{CI}$ results: with the exception of (6,6) in adenine, all other active spaces appear within the peaked and bifurcating quadrant. This is also observed in the RMSD figures of the SI (Figures S28 and S30), where no correlation could be found between any significant structural changes and the topographies observed.

Summarizing, purine nucleobases also exhibit significant changes in conical intersection topography upon active space change, but to a minor extent to what was observed above for pyrimidines. Conical intersections connecting excited states (regardless of their nature) with the ground state are more influenced by the amount of correlation included in the calculation, with many active spaces at boundary quadrants. The small structural changes observed, similar to those also seen in pyrimidines, again suggest that changes in topography correlate better with the (better or worse) description of the electronic states involved.

4. DISCUSSION

A number of conclusions can be drawn upon comparison of the different systems studied.

In structural terms, RMSD analyses show that changes across geometries are very small, ∼0.2 Å for most cases, with 0.6−0.7 Å for the largest deviations registered for $(^1n_O\pi^*/^1\pi\pi^*)_{CI}$ in thymine (see Figure S18 in SI) and $(^1n_O\pi^*/S_0)_{CI}$ and $(^1n_N\pi^*/^1\pi\pi^*)_{CI}$ in cytosine (see Figures S19 and S23 in the SI). The resulting structures (obtained with different active spaces) are therefore analogous, and this associates the vast changes in conical intersection topography to the varying electron correlation introduced in the different models describing the partaking electronic states by means of changing (increasing) the correlating orbitals.

Another aspect worth highlighting and perhaps expected is the small contribution of basis set size to the calculation outcomes: both double-ζ and triple-ζ (results available in the SI) contractions yield very similar results across all nucleobases and conical intersection types, with some larger deviations being observed for purine-based adenine and guanine. In some particular cases, the topography of conical intersections can be affected by borderline values, leading to different results depending on the basis set used, as was found for example in $(L_a(^1\pi\pi^*)/L_b(^1\pi\pi^*))_{CI}$ and $(L_a(^1\pi\pi^*)/^1n_O\pi^*)_{CI}$ for guanine (Figures 8a and 9b respectively), as well as $(L_a(^1\pi\pi^*)/S_0)_{CI}$ for both guanine and adenine in Figure 7.

As CASSCF does not add contributions toward electron correlation beyond those orbitals directly included in the active space (i.e., due to strong/static correlation), and due to their localized nature, expanding the basis set has a relatively small effect that is likely more discernible in purine-based species due to their larger molecular frames (and larger amount of correlating orbitals included therein). We expect a (potentially) more prominent role of basis set size when including also dynamic electron correlation,[98] which we plan to look at in future work.

A brief note should be included regarding active space selection: here we used a systematic approach based on natural orbital occupation numbers (a more in-depth discussion on this based on the natural orbital occupation numbers obtained for each optimized intersection is available in the SI), but this does not necessarily mean that a balanced amount of correlation is included in the model in all cases, particularly when referring to the balanced description of the different electronic states partaking in a given conical intersection. Some of the pronounced deviations observed may therefore arise due to inadequate active space selection, which could be potentially improved by either using automated schemes for active space selection such as those available in the literature[99−103] or defining variants of these schemes aimed at selecting orbitals at energy degeneracy points, but which are beyond the scope of the present work. This systematic procedure of active space enlargement does aim to progressively increase the amount of static correlation included in the model. It is worth noting, however, that by including more and more molecular orbitals, a degree of dynamic electron correlation is also being retained, particularly for the largest active spaces studied, hence making the changes observed not solely due to static but also partly (even if to a very small degree) due to dynamic electron correlation.

Perhaps a more interesting and pertinent comparison within a chemical perspective is that across molecular systems: DNA/RNA nucleobases are expected to display similar excited-state decays following the recorded experimental evidence in the literature.[2,4,6] Figure 10a shows $\mathcal{P}$ and $\mathcal{B}$ values for $(^1\pi\pi^*/\mathrm{S_0})_{\mathrm{CI}}$ and $(\mathrm{L_a}(^1\pi\pi^*)/\mathrm{S_0})_{\mathrm{CI}}$, which are expected to be the main deactivation funnels responsible for the fastest decay component in DNA/RNA nucleobases.[80] As can be seen, no clear trends emerge when comparing the different nucleobases while employing their most correlated (and accurate) active spaces tested in this work: uracil and thymine feature a peaked and single-path character and are consistent with one another, as it would be expected from simple methylation; cytosine, on the other hand, showcases a peaked and bifurcating topography; adenine appears to resemble cytosine in displaying a peaked and bifurcating character even when not having the carbonyl group, whereas guanine, despite its structural resemblance to cytosine, shows a sloped and single-path topography.

This suggests that the specific conical intersection topography, or rather the topographies obtained with the most correlated CASSCF simulations used, does not seem to converge to a shared intersection type accounting for a unified or common ultrafast decay mechanism in DNA/RNA nucleobase monomers.[10,79]

Conical intersection topography has been suggested as a potential key shared aspect regulating reactivity: work by Robb and co-workers[104−106] points at sloped single-path intersections as potential promoters of photostability, as the ground-state gradient in these types of intersections was shown to point toward the regeneration of the reactants. Other authors have also highlighted an active role of conical intersection topography in excited-state decay[13,107] of a wide range of molecular species,[17,108−111] being believed to be a crucial component controlling their reactivity. An important aspect arising from this study is defining the smallest amount of static correlation required to qualitatively reproduce the different conical intersection topographies in DNA/RNA nucleobases, the specific active spaces reported in Table 1. The missing dynamic electron correlation can be added to the model in more cost-effective manners (i.e., through perturbation theory as in CASPT2), the size of the active space being the main bottleneck in conical intersection characterization, as well as in relatable simulations like nonadiabatic molecular dynamics. Moreover, working with smaller active spaces allows tackling larger systems, i.e., DNA/RNA dimers and multimers, while still accounting for the essential (competing) monomer-like features observed experimentally for these species.[6] Our discussion here focuses solely on intersection topographies as we have neglected their accessibility,[18] i.e., the presence of potential energy barriers, given that CASSCF energies are known to be not quantitative.

Table 1. Minimum Active Space Found to Correctly Describe (i.e., in Agreement to the Largest Space Reference) All Different Conical Intersections for Each Nucleobase

	active space
cytosine	(12,9)
uracil	(14,10)
thymine	(14,10)
guanine	(16,12)
adenine	(12,10)

Our CASSCF results (summarized in Table 1) suggest that cytosine requires at least (12,9) to account for all different conical intersections in agreement with their (14,10) reference. For uracil and thymine, we find that the full (14,10) is required as no other space matches the reference for all different conical intersections; we however find that (12,9) reasonably reproduces all but one of the intersections and may be a cost-effective solution. For guanine, we find at least (16,12) to be required to reproduce $(\mathrm{L_a}(^1\pi\pi^*)/\mathrm{S_0})_{\mathrm{CI}}$, which forces the need to use this large active space: this may be due to most active spaces clustering along the $\mathcal{P} = 1$ boundary (see Figure 7a). For adenine, we find that a smaller (12,10) suffices to represent all different conical intersections with respect to their (16,12) reference.

We check next whether there are other connections to be made in terms of conical intersection topography beyond those outlined above for $(^1\pi\pi^*/\mathrm{S_0})_{\mathrm{CI}}$, focusing on $(^1\mathrm{n_{O/N}}/^1\pi\pi^*)_{\mathrm{CI}}$. This intersection controls the population of dark $^1\mathrm{n_{O/N}}\pi^*$ states in DNA/RNA nucleobases and is therefore crucial to understand their overall photochemistry.

In this case, we observe a shared conical intersection topography across most DNA/RNA nucleobase monomers: Figure 10b shows $\mathcal{P}$ and $\mathcal{B}$ parameters for all $\mathrm{n_{O/N}}\pi^*$ and $\pi\pi^*$ intersections, where we find that almost all intersections display a peaked and bifurcating character that may contribute to the efficient $^1\pi\pi^* \rightarrow {^1\mathrm{n_{O/N}}\pi^*}$ population transfer observed in these species.[112−114] This includes $(^1\mathrm{n_N}\pi^*/^1\pi\pi^*)_{\mathrm{CI}}$ in cytosine, $(^1\mathrm{n_O}\pi^*/^1\pi\pi^*)_{\mathrm{CI}}$ in uracil, $(\mathrm{L_b}(^1\pi\pi^*)/^1\mathrm{n_N}\pi^*)_{\mathrm{CI}}$ in adenine, and both $(\mathrm{L_a}(^1\pi\pi^*)/^1\mathrm{n_O}\pi^*)_{\mathrm{CI}}$ and $(\mathrm{L_b}(^1\pi\pi^*)/^1\mathrm{n_N}\pi^*)_{\mathrm{CI}}$ in guanine. The only intersection that depicts different topographies when considering the most correlated calculation is $(^1\mathrm{n_O}\pi^*/^1\pi\pi^*)_{\mathrm{CI}}$ for both cytosine and thymine (values of $\mathcal{P}$ and $\mathcal{B}$ very close to 1), which features a sloped and single-path character.

Our calculations therefore show a more convergent picture across nucleobases for $(^1\mathrm{n_{O/N}}/^1\pi\pi^*)_{\mathrm{CI}}$ conical intersection topographies than what was found above for $(^1\pi\pi^*/\mathrm{S_0})_{\mathrm{CI}}$. Interestingly, the sole exceptions are cytosine and thymine, which have however been proposed to feature 3-state conical intersections between $\mathrm{S_0}$, $^1\pi\pi^*$ and $^1\mathrm{n_O}\pi^*$ states[10,64,115] and that feature distinct (i.e., longer-lived) excited-state dynamics[62,77,80,113−119] compared to those registered for uracil.[80,120]

On the other hand, it is worth noting that the similarities across the different DNA/RNA monomers for the $(^1\mathrm{n_{O/N}}/^1\pi\pi^*)_{\mathrm{CI}}$ intersections, as well as the dissimilarities found for the $(^1\pi\pi^*/\mathrm{S_0})_{\mathrm{CI}}$ topographies, may be due to differential correlation effects.[68,69] These are expected to be less significant when $^1\mathrm{n_O}\pi^*$ (strongly covalent) states are involved[121] and could therefore point at the lack of dynamic electron correlation as the potential missing ingredient to find a shared intersection topography across the different DNA/RNA monomers for their main $(^1\pi\pi^*/S_0)_{CI}$ ultrafast decay channel in our model.

Interestingly, preliminary calculations with dynamically correlated XMS-CASPT2 surfaces[122−124] using their largest active space and a double-ζ basis set show converging results across DNA/RNA pyrimidine nucleobases for the "ethylene-like" $(^1\pi\pi^*/S_0)_{CI}$ topographies leading to a peaked and bifurcating shape. Purine nucleobases (adenine and guanine), on the other hand, converge to a sloped and single-path topography instead for their equivalent $(L_a(^1\pi\pi^*)/S_0)_{CI}$. This suggests that pyrimidine and purine DNA/RNA nucleobases have different conical intersection topographies underpinning their ultrafast decay channel, despite featuring similar lifetimes.[125] A more thorough report will be presented soon covering this and other important aspects related to the effects of dynamic electron correlation at the intersection seam in DNA/RNA species.[98,126]

5. CONCLUSIONS

In this work, we have, for the first time to our knowledge, analyzed how conical intersection topographies change upon modifying the amount of strong (or static) electron correlation included in the model using CASSCF calculations with varying active space sizes in DNA/RNA canonical nucleobases.

Overall, we observe very large differences in the resulting conical intersection topography due to active space change, which are hard to ascribe to obvious trends in the way strong correlation is included in the model. These large discrepancies, which do not systematically converge upon active space increase, may partly be due to the inadequacy of the conical intersection classification scheme, which appears to be not robust enough to account for the subtle changes introduced due to differences in static correlation. Schemes employing second-[28] and higher-order descriptions of the intersection seam, resorting to diabatic instead of adiabatic schemes,[127] or exploring farther regions of the potential energy surfaces may be required to provide an unambiguous classification of the CI topographies in DNA/RNA bases due to changes in active space size.

Interestingly, changing active spaces within the same DNA/RNA nucleobase and intersection type largely results in almost identical optimized geometries with very different topographies. This is the case of the well-known ethylene-like ring-puckering intersection in pyrimidine-based bases, as well as its analogous $(L_a(^1\pi\pi^*)/S_0)_{CI}$ in purine-based bases, which produce a wide range of conical intersection topographies while leading to analogous structures. The most correlated level of theory points to different intersection topographies for the distinct DNA/RNA nucleobase monomers, even though they are expected to underpin analogous ultrafast decay channels.

Conical intersections featuring $^1n_O\pi^*$ states, on the other hand, appear to feature more marked structural differences due to out-of-plane motions occurring in two distinct yet analogous directions. More obvious correlations are observed in this case, where despite featuring larger structural differences, we observe a more consistent conical intersection classification across active spaces and where CI topography outliers correlate with the most different structures optimized. Interestingly, intersections between $^1\pi\pi^*$ and $^1n\pi^*$ states display the same topography across the different nucleobases, with the exception of cytosine and thymine, which are reported to feature 3-state crossings connecting $(^1n_O/\pi\pi^*)_{CI}$ with S_0.

While basis set size is found to be negligible in most cases, including all in pyrimidine derivatives, it can have a sizable impact on the topography of conical intersections in purine nucleobases: we find that basis set size can impact intersections at quadrant boundaries, with $(L_a(^1\pi\pi^*)/L_b(^1\pi\pi^*))_{CI}$ and $(L_a(^1\pi\pi^*)/^1n_O\pi^*)_{CI}$ in guanine as examples. Additionally, the classification of $(L_a(^1\pi\pi^*)/S_0)_{CI}$ for both guanine and adenine is affected by basis set size: this points to larger basis set dependencies with increasing molecular size.

Our findings highlight the vast changes introduced in conical intersection topography upon varying active space size. Given that conical intersections are known to facilitate and even control photochemical reactivity,[13,17,18,108] our results show that their specific shape may depend very strongly on the amount of static electron correlation included in the model, much more so than previously thought. Ongoing work is analyzing how the inclusion of dynamic electron correlation further affects intersection topographies, aiming to uncover potential biases in their description, which may have undesired effects in the simulation of nonadiabatic events.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.3c00577.

> All Cartesian xyz coordinate files for all optimized (minimum energy) conical intersections in this work are readily available through an open-access repository.[128] Figures and tables with molecular orbitals included in the different active spaces, as well as their labeling, and root-mean-square deviation figures together with a plot displaying the most relevant dihedral angle for each of the different conical intersections (with each active space) and DNA/RNA nucleobases (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Ignacio Fdez. Galván − *Department of Chemistry − BMC, Uppsala University, SE-75123 Uppsala, Sweden;* orcid.org/0000-0002-0684-7689; Email: ignacio.fernandez@kemi.uu.se

Roland Lindh − *Department of Chemistry − BMC, Uppsala University, SE-75123 Uppsala, Sweden;* orcid.org/0000-0001-7567-8295; Email: roland.lindh@kemi.uu.se

Javier Segarra-Martí − *Instituto de Ciencia Molecular, Universitat de Valencia, ES-46071 Valencia, Spain;* orcid.org/0000-0002-2076-3406; Email: javier.segarra@uv.es

Authors

Juliana Cuéllar-Zuquin − *Instituto de Ciencia Molecular, Universitat de Valencia, ES-46071 Valencia, Spain;* orcid.org/0000-0002-0127-579X

Ana Julieta Pepino − *Dipartimento di Chimica Industriale "Toso Montanari", Università di Bologna, I-40136 Bologna, Italy*

Ivan Rivalta − *Dipartimento di Chimica Industriale "Toso Montanari", Università di Bologna, I-40136 Bologna, Italy;*

ENSL, CNRS, Laboratoire de Chimie UMR 5182, 69364 Lyon, France; orcid.org/0000-0002-1208-602X

Francesco Aquilante − *Theory and Simulation of Materials (THEOS), and National Centre for Computational Design and Discovery of Novel Materials (MARVEL), École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland;* orcid.org/0000-0003-4422-3938

Marco Garavelli − *Dipartimento di Chimica Industriale "Toso Montanari", Università di Bologna, I-40136 Bologna, Italy;* orcid.org/0000-0002-0796-289X

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jctc.3c00577

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The project that gave rise to these results received the support of a fellowship from "La Caixa" Foundation (ID 100010434) and from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 847648, fellowship code "LCF/BQ/PI20/11760022". We gratefully acknowledge computing resources from *Tirant III* and technical support provided by the *Servei d'Informatica de la Universitat de Valencia*. We also wish to acknowledge grant RYC2021-034191-I funded by MCIN/AEI/10.13039/501100011033 and by "ESF Investing in your future" or by "European Union NextGenerationEU/PRTR". R.L. and I.F.G acknowledge the Swedish Research Council (VR, Grant No. 2020-03182) for funding. I.R. gratefully acknowledge the support of Italian Ministry of University and Research (MUR) through the PRIN 2020 program (project PSI-MOVIE, prot. 2020HTSXMA, CUP J33C21000040006) and the use of computing resources of the "Pôle Scientifique de Modélisation Numérique" (PSMN) at the École Normale Supérieure de Lyon, France.

■ REFERENCES

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

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

(3) 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. *Annu. Rev. Phys. Chem.* **2009**, *60*, 217−239.

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

(5) Domcke, W.; Yarkony, D. R. Role of Conical Intersections in Molecular Spectroscopy and Photoinduced Chemical Dynamics. *Annu. Rev. Phys. Chem.* **2012**, *63*, 325−352.

(6) 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.

(7) Beckstead, A. A.; Zhang, Y.; de Vries, M. S.; Kohler, B. Life in the light: nucleic acid photoproperties as a legacy of chemical evolution. *Phys. Chem. Chem. Phys.* **2016**, *18*, 24228−24238.

(8) Boldissar, S.; de Vries, M. S. How nature covers its bases. *Phys. Chem. Chem. Phys.* **2018**, *20*, 9701−9716.

(9) 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* **2009**, *10*, 21−32.

(10) 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.

(11) Yarkony, D. R. Diabolical conical intersections. *Rev. Mod. Phys.* **1996**, *68*, 985−1013.

(12) Yarkony, D. R. Nonadiabatic Quantum Chemistry−Past and Future. *Chem. Rev.* **2012**, *112*, 481−498.

(13) Malhado, J. P.; Bearpark, M. J.; Hynes, J. T. Non-adiabatic dynamics close to conical intersections and the surface hopping perspective. *Front. Chem.* **2014**, *2*, No. 97.

(14) Kowalewski, M.; Bennett, K.; Dorfman, K. E.; Mukamel, S. Catching Conical Intersections in the Act: Monitoring Transient Electronic Coherences by Attosecond Stimulated X-Ray Raman Signals. *Phys. Rev. Lett.* **2015**, *115*, 193003.

(15) Polli, D.; Altoè, P.; Weingart, O.; Spillane, K. M.; Manzoni, C.; Brida, D.; Tomasello, G.; Orlandi, G.; Kukura, P.; Mathies, R. A.; Garavelli, M.; Cerullo, G. Conical intersection dynamics of the primary photoisomerization event in vision. *Nature* **2010**, *467*, 440−443.

(16) Zhu, X.; Yarkony, D. R. Non-adiabaticity: the importance of conical intersections. *Mol. Phys.* **2016**, *114*, 1983−2013.

(17) Ben-Nun, M.; Molnar, F.; Schulten, K.; Martínez, T. J. The role of intersection topography in bond selectivity of cis-trans photoisomerization. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 1769−1773.

(18) Schuurman, M. S.; Stolow, A. Dynamics at Conical Intersections. *Annu. Rev. Phys. Chem.* **2018**, *69*, 427−450.

(19) Bearpark, M. J.; Larkin, S. M.; Vreven, T. Searching for Conical Intersections of Potential Energy Surfaces with the ONIOM Method: Application to Previtamin D. *J. Phys. Chem. A* **2008**, *112*, 7286−7295.

(20) Weingart, O.; Nenov, A.; Altoè, P.; Rivalta, I.; Segarra-Martí, J.; Dokukina, I.; Garavelli, M. COBRAMM 2.0 − A software interface for tailoring molecular electronic structure calculations and running nanoscale (QM/MM) simulations. *J. Mol. Model.* **2018**, *24*, No. 271.

(21) Levine, B. G.; Coe, J. D.; Martínez, T. J. Optimizing Conical Intersections without Derivative Coupling Vectors: Application to Multistate Multireference Second-Order Perturbation Theory (MS-CASPT2). *J. Phys. Chem. B* **2008**, *112*, 405−413.

(22) Maeda, S.; Ohno, K.; Morokuma, K. Updated Branching Plane for Finding Conical Intersections without Coupling Derivative Vectors. *J. Chem. Theory Comput.* **2010**, *6*, 1538−1545.

(23) Herzberg, G.; Longuet-Higgins, H. C. Intersection of potential energy surfaces in polyatomic molecules. *Discuss. Faraday Soc.* **1963**, *35*, 77−82.

(24) Teller, E. The Crossing of Potential Surfaces. *J. Phys. Chem. A* **1937**, *41*, 109−116.

(25) Manaa, M. R.; Yarkony, D. R. On the intersection of two potential energy surfaces of the same symmetry. Systematic characterization using a Lagrange multiplier constrained procedure. *J. Chem. Phys.* **1993**, *99*, 5251−5256.

(26) Atchity, G. J.; Xantheas, S. S.; Ruedenberg, K. Potential energy surfaces near intersections. *J. Chem. Phys.* **1991**, *95*, 1862−1876.

(27) Bearpark, M. J.; Robb, M. A.; Schlegel, H. B. A direct method for the location of the lowest energy point on a potential surface crossing. *Chem. Phys. Lett.* **1994**, *223*, 269−274.

(28) Sicilia, F.; Bearpark, M. J.; Blancafort, L.; Robb, M. A. An analytical second-order description of the S0/S1 intersection seam: fulvene revisited. *Theor. Chem. Acc.* **2007**, *118*, 241−251.

(29) Mori, T.; Martínez, T. J. Exploring the Conical Intersection Seam: The Seam Space Nudged Elastic Band Method. *J. Chem. Theory Comput.* **2013**, *9*, 1155−1163.

(30) Bak, K. L.; Jørgensen, P.; Jensen, H. J.; Aa; Olsen, J.; Helgaker, T. First-order nonadiabatic coupling matrix elements from multiconfigurational self-consistent-field response theory. *J. Chem. Phys.* **1992**, *97*, 7573−7584.

(31) Snyder, J. W.; Fales, B. S.; Hohenstein, E. G.; Levine, B. G.; Martínez, T. J. A direct-compatible formulation of the coupled perturbed complete active space self-consistent field equations on graphical processing units. *J. Chem. Phys.* **2017**, *146*, No. 174113.

(32) Dudley, T. J.; Olson, R. M.; Schmidt, M. W.; Gordon, M. S. Parallel coupled perturbed CASSCF equations and analytic CASSCF second derivatives. *J. Comput. Chem.* **2006**, *27*, 352−362.

(33) Lischka, H.; Dallos, M.; Szalay, P. G.; Yarkony, D. R.; Shepard, R. Analytic evaluation of nonadiabatic coupling terms at the MR-CI level. I. Formalism. *J. Chem. Phys.* **2004**, *120*, 7322−7329.

(34) Fdez. Galván, I.; Delcey, M. G.; Pedersen, T. B.; Aquilante, F.; Lindh, R. Analytical State-Average Complete-Active-Space Self-Consistent Field Nonadiabatic Coupling Vectors: Implementation with Density-Fitted Two-Electron Integrals and Application to Conical Intersections. *J. Chem. Theory Comput.* **2016**, *12*, 3636−3653.

(35) Virshup, A. M.; Chen, J.; Martínez, T. J. Nonlinear dimensionality reduction for nonadiabatic dynamics: The influence of conical intersection topography on population transfer rates. *J. Chem. Phys.* **2012**, *137*, No. 22A519.

(36) Young, R. A.; Yarkony, D. R. A novel conical intersection topography and its consequences: The 1, 2 ^{2}A conical intersection seam of the vinoxy radical. *J. Chem. Phys.* **2005**, *123*, 084315.

(37) Robb, M. A.; Bernardi, F.; Olivucci, M. Conical intersections as a mechanistic feature of organic photochemistry. *Pure Appl. Chem.* **1995**, *67*, 783−789.

(38) Worth, G. A.; Cederbaum, L. S. Beyond Born-Oppenheimer: Molecular Dynamics Through a Conical Intersection. *Annu. Rev. Phys. Chem.* **2004**, *55*, 127−158.

(39) Ismail, N.; Blancafort, L.; Olivucci, M.; Kohler, B.; Robb, M. A. Ultrafast Decay of Electronically Excited Singlet Cytosine via a π,π^* to n_O,π^* State Switch. *J. Am. Chem. Soc.* **2002**, *124*, 6818−6819.

(40) Blancafort, L.; Robb, M. A. Key Role of a Threefold State Crossing in the Ultrafast Decay of Electronically Excited Cytosine. *J. Phys. Chem. A* **2004**, *108*, 10609−10614.

(41) Asturiol, D.; Lasorne, B.; Robb, M. A.; Blancafort, L. Photophysics of the π,π^* and n,π^* States of Thymine: MS-CASPT2 minimum-Energy Paths and CASSCF on-the-Fly Dynamics. *J. Phys. Chem. A* **2009**, *113*, 10211−10218.

(42) Perun, S.; Sobolewski, A. L.; Domcke, W. Conical Intersections in Thymine. *J. Phys. Chem. A* **2006**, *110*, 13238−13244.

(43) Matsika, S. Radiationless Decay of Excited States of Uracil through Conical Intersections. *J. Phys. Chem. A* **2004**, *108*, 7584−7590.

(44) Matsika, S. Three-State Conical Intersections in Nucleic Acid Bases. *J. Phys. Chem. A* **2005**, *109*, 7538−7545.

(45) Barbatti, M.; Aquino, A. J. A.; Szymczak, J. J.; Nachtigallová, D.; Hobza, P.; Lischka, H. Relaxation mechanisms of UV-photoexcited DNA and RNA nucleobases. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 21453−21458.

(46) Serrano-Andrés, L.; Merchán, M.; Borin, A. C. Adenine and 2-aminopurine: Paradigms of modern theoretical photochemistry. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 8691−8696.

(47) Serrano-Andrés, L.; Merchán, M.; Borin, A. C. A Three-State Model for the Photophysics of Guanine. *J. Am. Chem. Soc.* **2008**, *130*, 2473−2484.

(48) 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.

(49) Veryazov, V.; Malmqvist, PÅ.; Roos, B. O. How to select active space for multiconfigurational quantum chemistry? *Int. J. Quantum Chem.* **2011**, *111*, 3329−3338.

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

(51) Aquilante, F.; Autschbach, J.; Baiardi, A.; Battaglia, S.; Borin, V. A.; Chibotaru, L. F.; Conti, I.; De Vico, L.; Delcey, M.; Fdez. Galván, I.; Ferré, N.; Freitag, L.; Garavelli, M.; Gong, X.; Knecht, S.; Larsson, E. D.; Lindh, R.; Lundberg, M.; Malmqvist, PÅ.; Nenov, A.; Norell, J.; Odelius, M.; Olivucci, M.; Pedersen, T. B.; Pedraza-González, L.; Phung, Q. M.; Pierloot, K.; Reiher, M.; Schapiro, I.; Segarra-Martí, J.; Segatta, F.; Seijo, L.; Sen, S.; Sergentu, D.-C.; Stein, C. J.; Ungur, L.; Vacher, M.; Valentini, A.; Veryazov, V. Modern quantum chemistry with [Open]Molcas. *J. Chem. Phys.* **2020**, *152*, No. 214117.

(52) Li Manni, G.; Fdez. Galván, I.; Alavi, A.; Aleotti, F.; Aquilante, F.; Autschbach, J.; Avagliano, D.; Baiardi, A.; Bao, J. J.; Battaglia, S.; Birnoschi, L.; Blanco-González, A.; Bokarev, S. I.; Broer, R.; Cacciari, R.; Calio, P. B.; Carlson, R. K.; Carvalho Couto, R.; Cerdán, L.; Chibotaru, L. F.; Chilton, N. F.; Church, J. R.; Conti, I.; Coriani, S.; Cuéllar-Zuquin, J.; Daoud, R. E.; Dattani, N.; Decleva, P.; de Graaf, C.; Delcey, M. G.; De Vico, L.; Dobrautz, W.; Dong, S. S.; Feng, R.; Ferré, N.; Filatov Gulak, M.; Gagliardi, L.; Garavelli, M.; González, L.; Guan, Y.; Guo, M.; Hennefarth, M. R.; Hermes, M. R.; Hoyer, C. E.; Huix-Rotllant, M.; Jaiswal, V. K.; Kaiser, A.; Kaliakin, D. S.; Khamesian, M.; King, D. S.; Kochetov, V.; Krośnicki, M.; Kumaar, A. A.; Larsson, E. D.; Lehtola, S.; Lepetit, M.-B.; Lischka, H.; López Ríos, P.; Lundberg, M.; Ma, D.; Mai, S.; Marquetand, P.; Merritt, I. C. D.; Montorsi, F.; Mörchen, M.; Nenov, A.; Nguyen, V. H. A.; Nishimoto, Y.; Oakley, M. S.; Olivucci, M.; Oppel, M.; Padula, D.; Pandharkar, R.; Phung, Q. M.; Plasser, F.; Raggi, G.; Rebolini, E.; Reiher, M.; Rivalta, I.; Roca-Sanjuán, D.; Romig, T.; Safari, A. A.; Sánchez-Mansilla, A.; Sand, A. M.; Schapiro, I.; Scott, T. R.; Segarra-Martí, J.; Segatta, F.; Sergentu, D.-C.; Sharma, P.; Shepard, R.; Shu, Y.; Staab, J. K.; Straatsma, T. P.; Sørensen, L. K.; Tenorio, B. N. C.; Truhlar, D. G.; Ungur, L.; Vacher, M.; Veryazov, V.; Voß, T. A.; Weser, O.; Wu, D.; Yang, X.; Yarkony, D.; Zhou, C.; Zobel, J. P.; Lindh, R. The OpenMolcas Web: A Community-Driven Approach to Advancing Computational Chemistry. *J. Chem. Theory Comput.* **2023**, *19* (20), 6933−6991.

(53) Roos, B. O. *Advances in Chemical Physics*; John Wiley & Sons, Ltd, 1987; pp 399−445.

(54) Fdez. Galván, I.; Brakestad, A.; Vacher, M. Role of conical intersection seam topography in the chemiexcitation of 1,2-dioxetanes. *Phys. Chem. Chem. Phys.* **2022**, *24*, 1638−1653.

(55) Widmark, P.-O.; Malqvist, P.-A.; Roos, B. O. Density matrix averaged atomic natural orbital (ANO) basis sets for correlated molecular wave functions. I. First row atoms. *Theor. Chim. Acta* **1990**, *77*, 291−306.

(56) Widmark, P.-O.; Persson, B. J.; Roos, B. O. Density matrix averaged atomic natural orbital (ANO) basis sets for correlated molecular wave functions. II. Second row atoms. *Theor. Chim. Acta* **1991**, *79*, 419−432.

(57) Aquilante, F.; Gagliardi, L.; Pedersen, T. B.; Lindh, R. Atomic Cholesky decompositions: A route to unbiased auxiliary basis sets for density fitting approximation with tunable accuracy and efficiency. *J. Chem. Phys.* **2009**, *130*, No. 154107.

(58) Aquilante, F.; Lindh, R.; Pedersen, T. B. Unbiased auxiliary basis sets for accurate two-electron integral approximations. *J. Chem. Phys.* **2007**, *127*, No. 114107.

(59) Aquilante, F.; Pedersen, T. B.; Lindh, R. Low-cost evaluation of the exchange Fock matrix from Cholesky and density fitting representations of the electron repulsion integrals. *J. Chem. Phys.* **2007**, *126*, No. 194106.

(60) Pedersen, T. B.; Aquilante, F.; Lindh, R. Density fitting with auxiliary basis sets from Cholesky decompositions. *Theor. Chem. Acc.* **2009**, *124*, 1−10.

(61) Delcey, M. G.; Pedersen, T. B.; Aquilante, F.; Lindh, R. Analytical gradients of the state-average complete active space self-consistent field method with density fitting. *J. Chem. Phys.* **2015**, *143*, No. 044110.

(62) 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.

(63) Matsika, S.; Yarkony, D. R. Accidental conical intersections of three states of the same symmetry. I. Location and relevance. *J. Chem. Phys.* **2002**, *117*, 6907−6910.

(64) Kistler, K. A.; Matsika, S. Three-state conical intersections in cytosine and pyrimidinone bases. *J. Chem. Phys.* **2008**, *128*, No. 215102.

(65) Ben-Nun, M.; Martínez, T. J. Photodynamics of ethylene: ab initio studies of conical intersections. *Chem. Phys.* **2000**, *259*, 237−248.

(66) Serrano-Pérez, J. J.; Bearpark, M. J.; Robb, M. A. The extended S1/S0 conical intersection seam for the photochemical 2 + 2 cycloaddition of two ethylene molecules. *Mol. Phys.* **2012**, *110*, 2493−2501.

(67) Karafiloglou, P.; Malrieu, J.-P. Origin of the conical intersection between the singlet ionic excited surfaces of twisted ethylene. *Theor. Chim. Acta* **1985**, *67*, 275−286.

(68) Merchán, M.; Serrano-Andrés, L. Ultrafast Internal Conversion of Excited Cytosine via the Lowest $\pi\pi^*$ Electronic Singlet State. *J. Am. Chem. Soc.* **2003**, *125*, 8108−8109.

(69) 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*, No. 1666.

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

(71) 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*, 2650−2661.

(72) Kistler, K. A.; Matsika, S. Photophysical pathways of cytosine in aqueous solution. *Phys. Chem. Chem. Phys.* **2010**, *12*, 5024−5031.

(73) Fingerhut, B. P.; Dorfman, K. E.; Mukamel, S. Probing the Conical Intersection Dynamics of the RNA Base Uracil by UV-Pump Stimulated-Raman-Probe Signals; Ab Initio Simulations. *J. Chem. Theory Comput.* **2014**, *10*, 1172−1188.

(74) Aziz, S. G.; Elroby, S. K.; Alyoubi, A.; Hilal, R. Exploring the Conical Intersection Seam in Cytosine: A DFT and CASSCF Study. *Procedia Comput. Sci.* **2014**, *29*, 1384−1391. 2014 International Conference on Computational Science.

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

(76) Welborn, V. V.; Voorhis, T. V. Non-radiative deactivation of cytosine derivatives at elevated temperature. *Mol. Phys.* **2018**, *116*, 2591−2598.

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

(78) Hudock, H. R.; Levine, B. G.; Thompson, A. L.; Satzger, H.; Townsend, D.; Gador, N.; Ullrich, S.; Stolow, A.; Martínez, T. J. Ab Initio Molecular Dynamics and Time-Resolved Photoelectron Spectroscopy of Electronically Excited Uracil and Thymine. *J. Phys. Chem. A* **2007**, *111*, 8500−8508.

(79) Merchán, M.; González-Luque, R.; Climent, T.; Serrano-Andrés, L.; Rodríguez, E.; Reguero, M.; Peláez, D. Unified Model for the Ultrafast Decay of Pyrimidine Nucleobases. *J. Phys. Chem. B* **2006**, *110*, 26471−26476.

(80) Borrego-Varillas, R.; Nenov, A.; Kabaciński, P.; Conti, I.; Ganzer, L.; Oriana, A.; Jaiswal, V. K.; Delfino, I.; Weingart, O.; Manzoni, C.; Rivalta, I.; Garavelli, M.; Cerullo, G. Tracking excited state decay mechanisms of pyrimidine nucleosides in real time. *Nat. Commun.* **2021**, *12*, No. 7285.

(81) Asturiol, D.; Lasorne, B.; Worth, G. A.; Robb, M. A.; Blancafort, L. Exploring the sloped-to-peaked S2/S1 seam of intersection of thymine with electronic structure and direct quantum dynamics calculations. *Phys. Chem. Chem. Phys.* **2010**, *12*, 4949−4958.

(82) Green, J. A.; Jouybari, M. Y.; Aranda, D.; Improta, R.; Santoro, F. Nonadiabatic Absorption Spectra and Ultrafast Dynamics of DNA and RNA Photoexcited Nucleobases. *Molecules* **2021**, *26*, No. 1743.

(83) 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 $\pi\pi^*/n\pi^*$ Decay of Thymine in Water as a Test Case. *J. Chem. Theory Comput.* **2018**, *14*, 820−832.

(84) Yaghoubi Jouybari, M.; Liu, Y.; Improta, R.; Santoro, F. Quantum dynamics of the $\pi\pi^*/n\pi^*$ decay of the epigenetic nucleobase 1,5-dimethyl-cytosine in the gas phase. *Phys. Chem. Chem. Phys.* **2020**, *22*, 26525−26535.

(85) 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.

(86) 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. *Chem. - Eur. J.* **2016**, *22*, 7497−7507.

(87) Platt, J. R. Classification of Spectra of Cata-Condensed Hydrocarbons. *J. Chem. Phys.* **1949**, *17*, 484−495.

(88) Serrano-Andrés, L.; Merchán, M.; Borin, A. C. A Three-State Model for the Photophysics of Adenine. *Chem. - Eur. J.* **2006**, *12*, 6559−6571.

(89) Conti, I.; Garavelli, M.; Orlandi, G. Deciphering Low Energy Deactivation Channels in Adenine. *J. Am. Chem. Soc.* **2009**, *131*, 16108−16118.

(90) Conti, I.; Altoè, P.; Stenta, M.; Garavelli, M.; Orlandi, G. Adenine deactivation in DNA resolved at the CASPT2//CASSCF/AMBER level. *Phys. Chem. Chem. Phys.* **2010**, *12*, 5016−5023.

(91) Altavilla, S. F.; Segarra-Martí, J.; Nenov, A.; Conti, I.; Rivalta, I.; Garavelli, M. Deciphering the photochemical mechanisms describing the UV-induced processes occurring in solvated guanine monophosphate. *Front. Chem.* **2015**, *3*, No. 29.

(92) Heggen, B.; Lan, Z.; Thiel, W. Nonadiabatic decay dynamics of 9H-guanine in aqueous solution. *Phys. Chem. Chem. Phys.* **2012**, *14*, 8137−8146.

(93) Lan, Z.; Fabiano, E.; Thiel, W. Photoinduced Nonadiabatic Dynamics of 9H-Guanine. *ChemPhysChem* **2009**, *10*, 1225−1229.

(94) Lu, Y.; Lan, Z.; Thiel, W. Monomeric adenine decay dynamics influenced by the DNA environment. *J. Comput. Chem.* **2012**, *33*, 1225−1235.

(95) Fabiano, E.; Thiel, W. Nonradiative Deexcitation Dynamics of 9H-Adenine: An OM2 Surface Hopping Study. *J. Phys. Chem. A* **2008**, *112*, 6859−6863.

(96) Barbatti, M.; Lischka, H. Nonadiabatic Deactivation of 9H-Adenine: A Comprehensive Picture Based on Mixed Quantum-Classical Dynamics. *J. Am. Chem. Soc.* **2008**, *130*, 6831−6839.

(97) Barbatti, M.; Szymczak, J. J.; Aquino, A. J. A.; Nachtigallová, D.; Lischka, H. The decay mechanism of photoexcited guanine − A nonadiabatic dynamics study. *J. Chem. Phys.* **2011**, *134*, No. 014304.

(98) Segarra-Martí, J.; Tran, T.; Bearpark, M. J. 3-Methylation alters excited state decay in photoionised uracil. *Phys. Chem. Chem. Phys.* **2022**, *24*, 27038−27046.

(99) Stein, C. J.; Reiher, M. Automated Selection of Active Orbital Spaces. *J. Chem. Theory Comput.* **2016**, *12*, 1760−1771.

(100) Stein, C. J.; Reiher, M. autoCAS: A Program for Fully Automated Multiconfigurational Calculations. *J. Comput. Chem.* **2019**, *40*, 2216−2226.

(101) Kaufold, B. W.; Chintala, N.; Pandeya, P.; Dong, S. S. Automated Active Space Selection with Dipole Moments. *J. Chem. Theory Comput.* **2023**, *19*, 2469−2483.

(102) Jeong, W.; Stoneburner, S. J.; King, D.; Li, R.; Walker, A.; Lindh, R.; Gagliardi, L. Automation of Active Space Selection for

Multireference Methods via Machine Learning on Chemical Bond Dissociation. *J. Chem. Theory Comput.* **2020**, *16*, 2389−2399.

(103) Bensberg, M.; Reiher, M. Corresponding Active Orbital Spaces along Chemical Reaction Paths. *J. Phys. Chem. Lett.* **2023**, *14*, 2112−2118.

(104) Boggio-Pasqua, M.; Robb, M. A.; Bearpark, M. J. Photostability via a Sloped Conical Intersection: A CASSCF and RASSCF Study of Pyracylene. *J. Phys. Chem. A* **2005**, *109*, 8849−8856.

(105) Hall, K. F.; Boggio-Pasqua, M.; Bearpark, M. J.; Robb, M. A. Photostability Via Sloped Conical Intersections: A Computational Study of the Excited States of the Naphthalene Radical Cation. *J. Phys. Chem. A* **2006**, *110*, 13591−13599.

(106) Tokmachev, A. M.; Boggio-Pasqua, M.; Bearpark, M. J.; Robb, M. A. Photostability via Sloped Conical Intersections: A Computational Study of the Pyrene Radical Cation. *J. Phys. Chem. A* **2008**, *112*, 10881−10886.

(107) Farfan, C. A.; Turner, D. B. A systematic model study quantifying how conical intersection topography modulates photochemical reactions. *Phys. Chem. Chem. Phys.* **2020**, *22*, 20265−20283.

(108) Mignolet, B.; Curchod, B. F. E.; Martínez, T. J. Rich Athermal Ground-State Chemistry Triggered by Dynamics through a Conical Intersection. *Angew. Chem., Int. Ed.* **2016**, *55*, 14993−14996.

(109) Tuna, D.; Sobolewski, A. L.; Domcke, W. Conical-Intersection Topographies Suggest That Ribose Exhibits Enhanced UV Photostability. *J. Phys. Chem. B* **2016**, *120*, 10729−10735.

(110) Blancafort, L. Photochemistry and Photophysics at Extended Seams of Conical Intersection. *ChemPhysChem* **2014**, *15*, 3166−3181.

(111) Neville, S. P.; Wang, Y.; Boguslavskiy, A. E.; Stolow, A.; Schuurman, M. S. Substituent effects on dynamics at conical intersections: Allene and methyl allenes. *J. Chem. Phys.* **2016**, *144*, No. 014305.

(112) Picconi, D.; Barone, V.; Lami, A.; Santoro, F.; Improta, R. The Interplay between $\pi\pi^*/n\pi^*$ Excited States in Gas-Phase Thymine: A Quantum Dynamical Study. *ChemPhysChem* **2011**, *12*, 1957−1968.

(113) 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.; Martìnez, 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 $\pi\pi^*$/nπ^* internal conversion in organic chromophores via K-edge resonant absorption. *Nat. Commun.* **2017**, *8*, No. 29.

(114) Picconi, D.; Avila Ferrer, F. J.; Improta, R.; Lami, A.; Santoro, F. Quantum-classical effective-modes dynamics of the $\pi\pi^* \rightarrow n\pi^*$ decay in 9H-adenine. A quadratic vibronic coupling model. *Faraday Discuss.* **2013**, *163*, 223−242.

(115) 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. *J. Am. Chem. Soc.* **2015**, *137*, 2931−2938.

(116) 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*, 19045−19057.

(117) Wolf, T. J. A.; Paul, A. C.; Folkestad, S. D.; Myhre, R. H.; Cryan, J. P.; Berrah, N.; Bucksbaum, P. H.; Coriani, S.; Coslovich, G.; Feifel, R.; Martinez, T. J.; Moeller, S. P.; Mucke, M.; Obaid, R.; Plekan, O.; Squibb, R. J.; Koch, H.; Gühr, M. Transient resonant Auger-Meitner spectra of photoexcited thymine. *Faraday Discuss.* **2021**, *228*, 555−570.

(118) Pilles, B. M.; Maerz, B.; Chen, J.; Bucher, D. B.; Gilch, P.; Kohler, B.; Zinth, W.; Fingerhut, B. P.; Schreier, W. J. Decay Pathways of Thymine Revisited. *J. Phys. Chem. A* **2018**, *122*, 4819−4828.

(119) Stojanović, L.; Bai, S.; Nagesh, J.; Izmaylov, A. F.; Crespo-Otero, R.; Lischka, H.; Barbatti, M. New Insights into the State Trapping of UV-Excited Thymine. *Molecules* **2016**, *21*, No. 1603.

(120) Brister, M. M.; Crespo-Hernández, C. E. Excited-State Dynamics in the RNA Nucleotide Uridine 5′-Monophosphate Investigated Using Femtosecond Broadband Transient Absorption Spectroscopy. *J. Phys. Chem. Lett.* **2019**, *10*, 2156−2161.

(121) Segarra-Martí, J.; Garavelli, M.; Aquilante, F. Multiconfigurational Second-Order Perturbation Theory with Frozen Natural Orbitals Extended to the Treatment of Photochemical Problems. *J. Chem. Theory Comput.* **2015**, *11*, 3772−3784.

(122) Nishimoto, Y.; Battaglia, S.; Lindh, R. Analytic First-Order Derivatives of (X)MS, XDW, and RMS Variants of the CASPT2 and RASPT2Methods. *J. Chem. Theory Comput.* **2022**, *18*, 4269−4281.

(123) Park, J. W. Single-State Single-Reference and Multistate Multireference Zeroth-Order Hamiltonians in MS-CASPT2 and Conical Intersections. *J. Chem. Theory Comput.* **2019**, *15*, 3960−3973.

(124) Park, J. W.; Shiozaki, T. Analytical Derivative Coupling for Multistate CASPT2 Theory. *J. Chem. Theory Comput.* **2017**, *13*, 2561−2570.

(125) 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*, No. 074316.

(126) Segarra-Martí, J.; Nouri, S. M.; Bearpark, M. J. Modelling Photoionisations in Tautomeric DNA Nucleobase Derivatives 7H-Adenine and 7H-Guanine: Ultrafast Decay and Photostability. *Photochem* **2021**, *1*, 287−301.

(127) Fdez. Galván, I.; Lindh, R. Smooth Things Come in Threes: A Diabatic Surrogate Model for Conical Intersection Optimization. *J. Chem. Theory Comput.* **2023**, *19*, 3418−3427.

(128) Cuéllar-Zuquin, J.; Pepino, A. J.; Galván, I. F.; Rivalta, I.; Aquilante, F.; Garavelli, M.; Lindh, R.; Segarra-Martí, J. Characterising conical intersections in DNA/RNA nucleobases with multiconfigurational wave functions of varying active space size. Supplementary Information. 2023; https://zenodo.org/record/8348402.