Characterising 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*†

†Instituto de Ciencia Molecular, Universitat de Valencia, P.O. Box 22085, ES-46071 Valencia, Spain
‡Dipartimento di Chimica Industriale “Toso Montanari”, Università di Bologna, Viale del Risorgimento 4, I-40136 Bologna, Italy
¶Department of Chemistry − BMC, Uppsala University, Uppsala, Sweden
§ENSL, CNRS, Laboratoire de Chimie UMR 5182, 46 Allée d’Italie, 69364 Lyon France
∥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

E-mail: ignacio.fernandez@kemi.uu.se; roland.lindh@kemi.uu.se; javier.segarra@uv.es

Abstract

We characterise 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 towards their description of photoinduced phenomena. Our results show 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 towards 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 only relevant 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 highlight instead the importance of correctly describing the electronic states involved at these crossing points. Our work shows 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.\textsuperscript{1–4} Their photostability is strongly related to the ability possessed by DNA bases to dissipate the excess energy gained upon absorption in a non-radiative manner at ultrafast timescales, 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.\textsuperscript{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,\textsuperscript{7–10} by choosing the most suitable (photostable) compounds and thus aiding in its photo-protective design, further securing efficient DNA replication.

The existence of conical intersections is nowadays widely recognised and their importance often highlighted in modern photochemistry,\textsuperscript{11–13} their direct characterisation being still elusive by experimental means\textsuperscript{14} even though indirect fingerprints have been recorded.\textsuperscript{15} Given the importance of CIs for rationalizing the outcome of a given photochemical reaction,\textsuperscript{16–18} recent efforts have been put towards efficiently facilitating their characterisation in realistic systems. This includes extending the available protocols to hybrid quantum mechanics/molecular mechanics (QM/MM) schemes\textsuperscript{19,20} hence including solvent effects, or locating CIs with alternative penalty functions\textsuperscript{21} and updated branching plane techniques\textsuperscript{22} that do not require the evaluation of expensive non-adiabatic 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\textsuperscript{23,24} and later on within states of the same symmetry.\textsuperscript{25,26} Different efficient methodologies have been devised over the years for CI optimisation: these range from linear-\textsuperscript{27} to quadratic-based\textsuperscript{28} projection techniques to the use of nudged elastic band method\textsuperscript{29} being mostly combined with multiconfigurational techniques that possess first-order non-adiabatic coupling formulations and implementations.\textsuperscript{30–34} Characterising the potential energy surfaces around conical intersections is important as their shape or topography is believed to influence excited state reactivity,\textsuperscript{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 Fig. 1). Despite nucleobase CIs being reported in the literature
over the years, this work represents to our knowledge the first systematic analysis of the active space dependence on the characterisation of such critical points, and provides also an overview of the potential impact shown by the diffuseness of the basis set employed in the optimisation procedure.

Active space selection 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 totalling over 350 optimised CIs), we find 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, it was observed 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.

Interestingly, comparing optimised structures for a given conical intersection revealed
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 modelling and understanding photoinduced phenomena in DNA and RNA nucleobases.

The manuscript is organised 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 featuring in the diverse DNA/RNA building blocks. A discussion follows from the similarities that emerge amongst the different nucleobases, followed by a conclusions section that summarises the present findings.

2 Computational Details

All computations were carried out with the OpenMOLCAS package. Different complete active space self-consistent field (CASSCF) 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 Figs. S4 and S5). Tables S1-S5 in
the SI contains the specific orbitals included in each of the conical intersections for each of the nucleobases studied. 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)\textsuperscript{54,55} were used in their double-\(\zeta\) (VDZP) and triple-\(\zeta\) (VTZP) polarised contractions. The atomic compact Cholesky decomposition (acCD)\textsuperscript{56} was used throughout to speed-up the two-electron integrals\textsuperscript{57–59} as well as for computing analytic CASSCF gradients\textsuperscript{60} and non-adiabatic couplings.\textsuperscript{34} Conical intersections were characterised using the method introduced by Fdez. Galván et al.\textsuperscript{34} and available in OpenMolcas.\textsuperscript{50,51}

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

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

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

where \(\sigma\) refers to the pitch, \(\Delta_{gh}\) to the asymmetry, and \(\theta_s\) to the tilt heading.\textsuperscript{34} Depending on the value taken by these two parameters (\(\mathcal{P}\) and \(\mathcal{B}\)), we have the following classification system:

\[
\mathcal{P} \begin{cases} > 1 \text{ Sloped} \\ < 1 \text{ Peaked} \end{cases} \quad \text{and} \quad \mathcal{B} \begin{cases} > 1 \text{ Single-path} \\ < 1 \text{ Bifurcating} \end{cases}
\]

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 the following figure.

![Figure 2: Classification of the conical intersections in terms of \( P \) and \( B \) parameters with a three-dimensional representation of how are the two potential energy surfaces in the branching space.](image)

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 analysing 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 analysing 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 behaviours 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 neighbouring 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)\textsuperscript{6,10,45,61} to rationalise the rich photophysical landscape observed experimentally in DNA pyrimidine-based monomers, their characterisation 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\textsuperscript{62}) have also been reported in the literature for these systems, but they are beyond the scope of the present manuscript.\textsuperscript{10,40,44,63}

($^1\pi\pi^*/S_0$)$_{CI}$ is perhaps the best known intersection in DNA/RNA pyrimidine nucleobases and entails a pronounced “ethylene-like”\textsuperscript{64–66} twisting leading to an out-of-plane motion in C5 position\textsuperscript{10,39,41–43,45,48,61,67–70,70–77} shared across all pyrimidine-based nucleobases\textsuperscript{10,78} and that is associated to their fastest component of their decay.\textsuperscript{67,79} 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 in-sets in Figure 3.

Upon characterising 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 (Fig. 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 $B$ slightly above 1. Uracil (Fig. 3b), on the other hand, favours a peaked and single-path topography with the exception of the smaller (6,5) and (4,3) spaces, even if these values approach $B \approx 1$. Thymine (Fig. 3c) displays a much more complex behaviour as it features intersections classed in almost every single quadrant (except sloped and bifurcating) depending on the active space used: the
Figure 3: $P$ and $B$ parameters of $(1\pi \pi^*/S_0)_{CI}$ using multiple different active spaces (see Computational Details) for a) cytosine, b) uracil and c) thymine. The largest (14,10) reference active space in each sub-panel is labelled in the figure, with its associated symbols being marked with a black dot at their centre. A picture with the superimposed geometries of all optimised conical intersections are provided as in-sets, where structures with “borderline” classifications (marked with squares in the plot) are highlighted in a different colour.

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 Fig. 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 optimised 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 $\pi$, $\pi^*$ and lone
pair occupied $n_{O/N}$ orbitals, uracil and thymine (Fig. 3b and 2c) show a consistent description of this intersection as peaked and single-path, whereas cytosine (Fig. 3a) appears to be peaked and bifurcating.

Despite the vast changes observed in topography, structurally all optimised $(^1\pi\pi^*/S_0)_{CI}$ conical intersections are very similar. Root mean square deviation (RMSD) analyses provided in the SI (Figures S6-S8) 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.

---

Figure 4: $\mathcal{P}$ and $\mathcal{B}$ parameters of $(^1n_{O}\pi^*/^1\pi\pi^*)_{CI}$ using multiple different active spaces (see Computational Details) for a) cytosine, b) uracil and c) thymine. The largest (14,10) reference active space in each sub-panel is labelled in the figure, with its associated symbols being marked with a black dot at their centre. A picture with the superimposed geometries of all optimised conical intersections are provided as in-sets, where structures with “borderline” classifications (marked with squares in the plot) are highlighted in a different colour.

We look next at $(^1n_{O}\pi^*/^1\pi\pi^*)_{CI}$ that facilitates population transfer to the optically dark
$^1n_O\pi^*$ state and that is present in all pyrimidine nucleobases.$^{80-84}$

In cytosine (Fig. 4a) we predict a sloped and single-path character for all different active spaces, with (4,3) approaching borderline values of both $P \approx 1$ and $B \approx 1$ 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, with both (10,8) and (8,6) displaying a sloped and bifurcating topography nearing values of $B \approx 1$, and with two more active spaces featuring within the sloped and single-path quadrant. Thymine (Fig. 4c), like cytosine (Fig. 4a), shows a sloped and single-path topography with the exception of the (4,3) calculation that has nevertheless values approaching $B \approx 1$, particularly for its triple-ζ basis set result.

Looking at the outliers from Fig. 4 we can see in this case some clear differences in their structure, particularly for thymine (cf. Fig. 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 optimised structures. Cytosine (Fig. 4a) and uracil (Fig. 4b) also feature pronounced differences observable by the naked eye in the in-sets in Fig. 4, even if their RMSD analyses are not fully conclusive (see SI Figures S9-S11): 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, which are of the order of $\sim 0.5$ Å and therefore larger than those observed in the $(^1\pi\pi^*/S_0)_{CI}$ yet still relatively small, may moreover be also understood as the structures pucker either upwards or downwards, the varying active spaces leading to either of these analogous (but slightly different) geometries in the gas phase. Optimisations of both upwards and downwards 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 upwards and downwards (or endo and exo)$^{85}$ conformations result in markedly different
energies due to steric hindrance and/or other intermolecular interactions.

Figure 5: $P$ and $B$ parameters of the $^{1}n_{O\pi^{*}/S_{0}}$ CI using multiple different active spaces (see Computational Details) for a) cytosine, b) uracil and c) thymine. The largest (14,10) reference active space in each sub-panel is labelled in the figure, with its associated symbols being marked with a black dot at their centre. A picture with the superimposed geometries of all optimised conical intersections are provided as in-sets, where structures with "borderline" classifications (marked with squares in the plot) are highlighted in a different colour.

The last intersection shared by all pyrimidines is $^{1}n_{O\pi^{*}/\pi\pi^{*}}$ CI, which is shown in Figure 5. Cytosine (Fig. 5a) displays a well-defined sloped and single-path topography for all active spaces tested. Uracil (Fig. 5b) also features a sloped and single-path topography for its most correlated (14,10) calculation, a sloped and bifurcating character for the (8,7) active space with values approaching $B \approx 1$, while (8,6), (6,5) and (4,3) calculations feature a peaked and bifurcating topography. Thymine (Fig. 5c), like cytosine, shows a sloped and single-path character for all active spaces studied, with the largest deviation coming from the (4,4) calculation which is however still well-defined within its quadrant.

Similarities arise between $^{1}n_{O\pi^{*}/\pi\pi^{*}}$ CI and $^{1}n_{O\pi^{*}/S_{0}}$ 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 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 (upwards or downwards) 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 (Fig. 5a), H-N$_3$-C$_6$-H in uracil (Fig. 4b) and O-C$_2$-C$_5$-C in thymine (Fig. 5c), and where the changes in these dihedrals correlate with the increases in RMSD observed with respect to the (14,10) reference structures (see Figs S12-S14).

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-$\zeta$ basis set. 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 $P$. $(^1n_N\pi^*/^1\pi\pi^*)_{CI}$, presented in Fig. 6b, favours a peaked and bifurcating topography for the (14,10) reference calculation. It has three outliers, two of these being 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 structure, $(^1n_N\pi^*/S_0)_{CI}$ features a very pronounced NH$_2$ out-of-plane motion (see Fig. 6a) that is well captured by all active spaces tested, leading to almost identical geometries. This is reminiscent of the result reported above in Fig. 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 H-N$_1$-C$_4$-N dihedral angle, which shows the largest difference for the (4,3) calculation with an associated $\sim$0.6 Å RMSD, the rest of the structures being almost identical to the reference (14,10) simulation (See Fig S15).

$(^1n_N\pi^*/^1\pi\pi^*)_{CI}$, on the other hand, features a much more planar structure (Fig. 6b) with the exception of the (4,3) outlier that results in a heavily puckered structure. This pronounced change may correlate with the (4,3) calculation for the $(^1n_O\pi^*/^1\pi\pi^*)_{CI}$ in thymine (Fig. 4c), where the smallest active space also leads to an overly puckered structure. This puckering is also related to the H-N$_1$-C$_4$-N dihedral angle mentioned above and depicted in Figure S16 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 $\sim$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 which strongly depend on the size of the active space. Structurally, less pronounced changes are observed, which suggests these differences in topography to be rooted on the ability (or inability) of the different active spaces to appropriately represent the partaking 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.

Due to their larger conjugated frame, purine-based nuclebases feature two low-lying $^1\pi\pi^*$ states that are relevant to the photophysics, often named $L_a(^1\pi\pi^*)$ and $L_b(^1\pi\pi^*)$ adapted from Platt’s notation,\(^86\) as well as a number of lone pair $n_{N/O}$ states. This leads to $(L_a(^1\pi\pi^*)/S_0)_{CI}$, $(L_a(^1\pi\pi^*)/L_b(^1\pi\pi^*))_{CI}$ and $(L_b(^1\pi\pi^*)/^1n_N\pi^*)_{CI}$ that are common to both bases, as well as $(^1n_N\pi^*/S_0)_{CI}$ and $(L_a(^1\pi\pi^*)/^1n_N\pi^*)_{CI}$ for adenine and $(L_a(^1\pi\pi^*)/^1n_O\pi^*)_{CI}$ for guanine.

We refrain from studying $(^1n_O\pi^*/S_0)_{CI}$ and $(L_a(^1\pi\pi^*)/^1n_O\pi^*)_{CI}$ for guanine as they could not be successfully optimised. References in the literature for $(^1n_O\pi^*/S_0)_{CI}$ in guanine\(^47\) suggest this intersection to feature a pronounced ring-opening component that cannot be properly described with the active spaces used here and its therefore considered beyond the scope of the present work.

The first intersection to analyse is $(L_a(^1\pi\pi^*)/S_0)_{CI}$, which is equivalent to the $(^1\pi\pi^*/S_0)_{CI}$ in pyrimidines, and that is believed to be the main responsible behind the ultrafast decay of purine-base DNA nucleobases.\(^45-47,87-96\) This crossing is also characterised by an out-of-plane motion that resembles, like in pyrimidyne-based nucleobases, the pyramidalised intersection in ethylene.\(^64\)
Figure 7: \( \mathcal{P} \) and \( \mathcal{B} \) parameters of \( (L_a^1\pi\pi^*/S_0)_{CI} \) for a) guanine and b) adenine and c) for \( (1n_N\pi^*/S_0)_{CI} \) in adenine using multiple different active spaces (see Computational Details). The largest \((18,13)\) reference active space in each sub-panel is labelled in the figure, with its associated symbols being marked with a black dot at their centre. A picture with the superimposed geometries of all optimised conical intersections are provided as in-sets, where structures with “borderline” classifications (marked with squares in the plot) are highlighted in a different colour.

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

Fig. 7a shows the different intersection topographies obtained with varying active spaces for the \( (L_a^1\pi\pi^*/S_0)_{CI} \) 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 the more correlated \((20,14)\) and \((16,12)\) active spaces are classed as sloped and single-path, with smaller active spaces (with the exception of \((6,4)\), dark red symbols placed at the top left corner in the in-set) being placed within the peaked quadrant, even if they still feature values very close to \( \mathcal{P} \approx 1 \). We observe \((14,11)\) displays different topography for double-\( \zeta \) and triple-\( \zeta \) 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 \( (L_a(1 \pi \pi^*)/S_0)_{CI} \) optimised structures in guanine are very similar, as shown by the superimposed geometries displayed as an in-set 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 Fig S17), confirming all optimised structures are essentially equivalent. The small RMS deviations do correlate to an extent with differences registered across geometries in the H-C₅-C₂-N dihedral angle, which 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 \( (L_a(1 \pi \pi^*)/S_0)_{CI} \) across multiple active spaces in adenine are shown in Figure 7b. Similar to guanine, we observe a clustering of most active spaces 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 single-path: interestingly, larger differences are observed due to basis set size in this case, the double-ζ estimates of the more correlated spaces (14,11) and (12,10) being placed within the bifurcating quadrant.

Structurally, all geometries are almost identical and display a negligible \( \sim 0.1 \) Å with respect to the most correlated (18,13) geometry (see Fig S18). The very small differences across structures can be mostly associated to changes in the \( N_1-C_6-N-H \) dihedral angle, which approaches \( -20^\circ \) for the more correlated (18,13) calculation and \( \sim 5^\circ \) for the least correlated (2,2). Interestingly, in this case we observe how the \( N_1-C_6-N-H \) dihedral angle diverges for double-ζ and triple-ζ calculations when considering small active spaces, particularly (2,2), where there is a significant \( \sim 30^\circ \) difference between them.

Adenine also displays a well-known \( (^1n_N \pi^*/S_0)_{CI} \) featuring a pronounced out-of-plane \( \text{NH}_2 \) motion\(^{45,46,87} \), which is depicted in Figure 7c, together with the different CI topographies obtained upon varying active space size. This particular intersection shows a wide spread of topographies, similar to what was found in cytosine (Fig. 6a), featuring intersections across the single-path sloped and peaked quadrants with the most correlated (18,13) calculation.
belonging to the sloped and single-path type. Most resulting active space optimisations are placed within the sloped and single-path quadrant, with (6,5) and (10,8) appearing within the peaked and sloped regions and with (10,9) and (4,4) being classed as peaked and single-path.

As shown in the in-set of Figure 7c, the superimposed structures show very small differences across the optimised conical intersections, with associated RMSD values with respect to the (18,13) structure \(\sim 0.2\) Å. The largest differences are observed for the (10,9) and (4,4) calculations, which also display a very different \(C_5-C_6-N-H\) dihedral angle converging towards \(-40^\circ\) (as opposed to the \(\sim 20^\circ\) obtained for all other cases, see Figure S19 in the SI, and that also happen to be the two cases featuring a different intersection topography.

Figure 8: \(\mathcal{P}\) and \(\mathcal{B}\) parameters of \((L_a(1\pi^\ast)/L_b(1\pi^\ast))_C\) using multiple different active spaces (see Computational Details) for guanine (a) and adenine (b). The largest reference active space, (18,13) for guanine and (16,12) for adenine, is labelled in each sub-panel in the figure, with its associated symbols being marked with a black pattern at their centre. A picture with the superimposed geometries of all optimised conical intersections are provided as insets, with those structures “borderline” classifications (marked with squares in the plot) are highlighted in a different colour.

The \((L_a(1\pi^\ast)/L_b(1\pi^\ast))_C\) conical intersection for both purine nucleobases can be found in figure 8. In guanine (Figure 8a), all conical intersections with different active spaces converge to the same topography, peaked and bifurcating, with almost all cases clustered at values of less than 0.5 for both, \(\mathcal{P}\) and \(\mathcal{B}\). In addition, it can be observed in the overlapping structures that in all the cases we have almost the same geometry. The RMSD values obtained with respect to the most correlated (18,13) calculation are lower than 0.4 Å (see Figure S20) which reinforces the fact that all optimized geometries are practically the same, among which
the most significant change is the one occurring in the $N_1-C_2-N-H$ dihedral angle varying from $\sim -80^\circ$ for the active space (18,13) to $\sim -30^\circ$ in the case of the lower active space, (4,4).

The situation is somewhat different in adenine (Figure 8b). We note that most cases are classified as peaked and bifurcating with the (16,12) on the boundary with the peaked and single-path quadrant, with the exception of (8,6), (8,7) and (4,3) which are classified as peak and single path. Here the main differences between the superimposed geometries are due to the dihedral angle $H-C_2-C_8-H$. Interestingly, the RMS deviations shown in Figure S21, are almost zero for the active spaces classified as peaked and bifurcating and slightly higher for those with a different classification with the highest value for (4,3) that is the one with the higher $B$ value and therefore the one with the most different topography to (16,12).

The last conical intersections studied are depicted in the figure 9. The first one is the $(L_a(1\pi\pi^*)/1n_N\pi^*)_{CI}$ of adenine where it can be seen that in the most cases, the topography converged is peaked and bifurcating. The most correlated active space is (16,12) and it can be found at the frontier of being of peaked and bifurcating character. In addition to (16,12), the active spaces (8,7) and (8,8) are at the limit of the classification too, with a very interesting situation for (8,7) with a different topography depending on the basis set used in the optimization. When using a triple-$\zeta$ basis set, the resulting topography is peaked and single-path, whereas with a double-$\zeta$ basis set, the conical intersection is considered peaked and bifurcating. In the case of (8,8) and the smaller (4,3) active space, the differences observed between the double and triple-$\zeta$ basis sets are minimal having both at the same quadrant. By analyzing the superimposed geometries and obtaining the RMS deviations of the geometries respect to the (16,12), it is possible to see how similar they all are with the exception of (8,7) and (4,3) that present the larger RMSD (see Figure S22 in the SI) ($\sim 0.3\text{Å}$) values what reinforced the fact that they are the most different ones.

The conical intersection of adenine’s $(L_b(1\pi\pi^*)/1n_N\pi^*)_{CI}$ can be observed in sub-panel b of figure 9. As shown in the graph, this conical intersection appears to be relatively insensitive
Figure 9: $P$ and $B$ parameters of $(L_a(1\pi\pi^*)/1n_N\pi^*)_{CI}$ (a) and $(L_b(1\pi\pi^*)/1n_N\pi^*)_{CI}$ (b) for adenine and $(L_a(1\pi\pi^*)/1n_O\pi^*)_{CI}$ (c) and $(L_b(1\pi\pi^*)/1n_N\pi^*)_{CI}$ (d) for guanine using multiple different active spaces (see Computational Details). The largest reference active space, (18,13) for guanine and (16,12) for adenine, is labelled in each sub-panel in the figure, with its associated symbols being marked with a black pattern at their centre. A picture with the superimposed geometries of all optimised conical intersections are provided as in-sets, with those structures “borderline” classifications (marked with squares in the plot) are highlighted in a different colour.

to correlation, with all cases classified as peaked and bifurcating except for (6,6), which has a slightly higher value of $B$. This small deviation changes its classification to single-path, but its value of $P$ remains below one, indicating that the peaked classification is not affected. With the structural analysis carried out by comparing the superimposed geometries with the RMSD values and the change in the dihedral angle $C_5$-$C_6$-$N$-$H$ (Figure S23 in the SI), one can checked the similarities between all the optimized geometries having the larger values, which are still quite small ($\sim 0.1\text{Å}$), the ones optimized with the (6,6) and (6,5) active spaces and those who are more in the frontier between the two quadrants with relevance here.

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 9c), almost all cases
were classified as peaked and bifurcating, including the one optimized with the larger active space \((18,13)\). Notably, the differences between double and triple-\(\zeta\) were greater for this conical intersection than for the others examined. For example, the active space \((6,5)\) was classified as peaked and bifurcating with double-\(\zeta\), but as peaked and single-path with triple-\(\zeta\). Likewise, the active space \((10,8)\) was classified as peaked and single-path with double-\(\zeta\), but as sloped and single-path with triple-\(\zeta\). These cases fall at the frontier between three quadrants, and thus small changes in the value of \(P\) or \(B\) can have a significant impact on the topography of the conical intersection. The RMS deviations observed for this conical intersection were less than 0.2 Å (Figure S24 in the SI) and, therefore, these small deviations suggest that structural changes are not a significant factor in explaining the different topographies observed above.

In the \((L_b(1\pi\pi^*))/n_N\pi^*)_{CI}\) conical intersection of guanine (Figure 9d), all the active spaces led to the same result of a peaked and bifurcating topography. Similar to the previous conical intersection, there were differences between the results obtained with double and triple-\(\zeta\), but these differences did not affect the classification of the different active spaces since none of the values of \(P\) or \(B\) were at the boundaries of the quadrants. This suggests that this conical intersection is relatively independent of the correlation included in the calculation, as was in the case of \((L_b(1\pi\pi^*))/n_N\pi^*)_{CI}\) for adenine (Figure 9b). Both nucleobases have similar results observed in the \(P\) vs \(B\) representations with all the cases in the same quadrant (with the exception of \((6,6)\) in adenine), as well as in the RMSD figures of the SI (Figures S23 and S25), where no correlation could be found between the structural changes and the topographies observed.

Based on the data presented, it is evident that purine nucleobases also exhibit changes in the conical intersections as the active space is varied, but the degree of change is not as significant as observed in pyrimidine nucleobases. Conical intersections connecting excited states, regardless of their nature, with ground states are more influenced by the amount of correlation included in the calculation, with many active spaces at the boundaries of quad-
rants. However, the small structural changes observed here, similar to those in pyrimidines, again suggest that these are not a primary factor in classifying conical intersections. Instead, the problem seems to be more related to the correct description of the electronic states involved.

4 Discussion

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

In structural terms, RMSD analyses show changes across geometries are very small, \(~0.2\,\text{Å}\) for most cases, with \(0.6-0.7\,\text{Å}\) for the largest deviations registered for \((^1n_O/^1\pi\pi^*)_{CI}\) in thymine (see Figure S11 in the SI) and \((^1n_O\pi^* / S_0)_{CI}\) and \((^1n_N/^1\pi\pi^*)_{CI}\) in cytosine (see Figures S12 and S16 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-\(\zeta\) (red circles) and triple-\(\zeta\) (green triangles) 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 \((L_a(^1\pi\pi^*)/^1n_N\pi^*)_{CI}\) for adenine and \((L_a(^1\pi\pi^*)/^1n_O\pi^*)_{CI}\) for guanine (Figure 9a and c 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 towards electron correlation beyond those orbitals directly included in the active space (i.e. due to strong/static correlation), and due to their localised 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, 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, but this does not necessarily mean 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, or by defining variants of these schemes aimed at selecting orbitals at energy degeneracy points, but which are beyond the scope of the present work.

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. Figure 10a shows $P$ and $B$ values for $(1\pi\pi^*/S_0)_{CI}$ and $(L_a(1\pi\pi^*)/S_0)_{CI}$, which are expected to be the main deactivation funnels responsible for the fastest decay component in DNA/RNA nucleobases. As can be seen, no clear trends emerge when comparing the different nucleobases whilst 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 the specific conical intersection topography, or rather the topographies obtained with the most correlated CASSCF simulations used, do not seem to converge to a shared intersection type accounting for a unified or common ultrafast decay mechanism in
DNA/RNA nucleobase monomers.\textsuperscript{10,78}

Conical intersection topography has been suggested as a potential key shared aspect regulating reactivity: work by Robb and co-workers\textsuperscript{103–105} point at sloped single-path intersections as potential promoters of photostability, as the ground state gradient in these types of intersections was shown to point towards the regeneration of the reactants. Other authors have also highlighted an active role of conical intersection topography in excited state decay\textsuperscript{13,106} of a wide range of molecular species,\textsuperscript{17,107–110} being believed to be a crucial component controlling their reactivity. Our discussion here focusses solely on intersection topographies as we have neglected their accessibility,\textsuperscript{18} i.e. the presence of potential energy barriers, given CASSCF energies are known to be not quantitative.

We check next whether there are other connections to be made in terms of conical intersection topography beyond those outlined above for (\(1\pi\pi^*/S_0\))\textsubscript{CI}, focussing on (\(1n_{O/N}/1\pi\pi^*\))\textsubscript{CI}. This intersection controls the population of dark \(1n_{O/N}\pi^*\) 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 of the DNA/RNA nucleobase monomers: Figure 10b shows \(P\) and \(B\) parameters for all \(n_{O/N}\pi^*\) and \(\pi\pi^*\) intersections, where we find almost all intersections display a peaked and bifurcating character that may contribute to the efficient \(1\pi\pi^* \rightarrow 1 n_{O/N}\pi^*\) population transfer observed.
in these species. This includes \((1_{N\pi^*}^1{\pi\pi^*})_{CI}\) in cytosine, \((1_{O\pi^*}^1{\pi\pi^*})_{CI}\) in uracil, both \((L_a(1{\pi\pi^*})/1_{N\pi^*})_{CI}\) and \((L_b(1{\pi\pi^*})/1_{N\pi^*})_{CI}\) in adenine, and both \((L_a(1{\pi\pi^*})/1_{O\pi^*})_{CI}\) and \((L_b(1{\pi\pi^*})/1_{N\pi^*})_{CI}\) in guanine. The only intersections which depict different topography when considering their most correlated calculation is \((1_{O\pi^*}^1{\pi\pi^*})_{CI}\) for both cytosine and thymine, which features a sloped and single-path character.

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

On the other hand, it is worth noting that the similarities across the different DNA/RNA monomers for the \((1_{O/N\pi^*}^1{\pi\pi^*})_{CI}\) intersections, as well as the dissimilarities found for the \((1{\pi\pi^*}/S_0)_{CI}\) topographies, may be due to differential correlation effects. These are expected to be less significant when \(1_{O\pi^*}\) (strongly covalent) states are involved, 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 using their largest (14,10) active space and a double-\(\zeta\) 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, but a more thorough report will be presented in the near future covering this and other important aspects related to the effects of electron correlation at the intersection seam in DNA/RNA species.
5 Conclusions

In this work we have, for the first time to our knowledge, analysed 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- and higher-order descriptions of the intersection seam, resorting to diabatic instead of adiabatic schemes, 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 optimised 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 at 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 optimised. 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 sizeable impact on the topography of conical intersections in purine nucleobases: we find basis set size can impact intersections at quadrant boundaries, with $(L_a(^1\pi\pi^*/^1n_N\pi^*)_{CI}$ in adenine 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 at 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 conical intersections are known to facilitate and even control photochemical reactivity,\textsuperscript{13,17,18,107} our results show 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 analysing 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 non-adiabatic events.

**Acknowledgement**

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 thank 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 Education, University and Research (MIUR) through the PRIN program (project PSI-MOVIE, prot. 2020HTSXMA) and the use of computing resources of the “Pôle Scientifique de Modélisation Numérique” (PSMN) at the École Normale Supérieure de Lyon, France.

**Supporting Information Available**

All Cartesian xyz coordinate files for all optimised (minimum energy) conical intersections in this work are readily available through an open-access repository. Figures with molecular orbitals included in the different active spaces, as well as their labelling, 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.

**References**


(38) Worth, G. A.; Cederbaum, L. S. Beyond Born–Oppenheimer: Molecular Dynamics


(54) Widmark, P.-O.; Malqvist, P.-A.; Roos, B. O. Density matrix averaged atomic natural


(62) Matsika, S.; Yarkony, D. R. Accidental conical intersections of three states of the


(71) Kistler, K. A.; Matsika, S. Photophysical pathways of cytosine in aqueous solution. 


(83) 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.


(100) Kaufold, B. W.; Chintala, N.; Pandeya, P.; Dong, S. S. Automated Active Space Selection with Dipole Moments. *Journal of Chemical Theory and Computation* 2023, 0, null, PMID: 37040135.


(103) Boggio-Pasqua, M.; Robb, M. A.; Bearpark, M. J. Photostability via a Sloped Conical


(124) Segarra-Martí, J.; Nouri, S. M.; Bearpark, M. J. Modelling Photoionisations in Tau-


Graphical TOC Entry