

Cite this: DOI: 10.1039/xxxxxxxxxx

# Ultrafast and Radiationless Electronic Excited State Decay of Uracil and Thymine Cations: Computing The Effects of Dynamic Electron Correlation<sup>†</sup>

Javier Segarra-Martí,<sup>\*a</sup> Thierry Tran,<sup>a</sup> and Michael J. Bearpark<sup>\*a</sup>

Received Date

Accepted Date

DOI: 10.1039/xxxxxxxxxx

www.rsc.org/journalname

In this article we characterise the radiationless decay of the first few electronic excited states of the cations of DNA/RNA nucleobases uracil and thymine, including the effects of dynamic electron correlation on energies and geometries (optimised with XMS-CASPT2). In both systems, we find that one state of  ${}^2n_O^+$  and another two of  ${}^2\pi^+$  character can be populated following photoionisation, and their different minima and interstate crossings are located. We find strong similarities between uracil and thymine cations: with accessible conical intersections suggesting that depopulation of their electronic excited states takes place on ultrafast timescales in both systems, suggesting that they are photostable in agreement with previous theoretical (uracil<sup>+</sup>) evidence. We find that dynamic electron correlation separates the energy levels of the "3-state" conical intersection ( $D_2/D_1/D_0$ )<sub>CI</sub> previously located with CASSCF for uracil<sup>+</sup>, which will therefore have a different geometry and higher energy. Simulating the electronic and vibrational absorptions allows us to characterise spectral fingerprints that could be used to monitor these cation photo-processes experimentally.

## Introduction

Understanding the effect of different types of UV-light on DNA/RNA aggregates has been the target of many scientific studies over the past two decades, owing to the outstanding photo-physical properties displayed by their chromophoric species, the nucleobases, which dissipate the excess energy gained upon light irradiation in an ultrafast and harmless manner.<sup>1–5</sup> This particular feature has been linked to the impressive photostability shown by the genomic material, where the vast majority of light absorbed is converted harmlessly into thermal energy and dumped into the cellular environment.

Most studies in the literature have focused on the UV-B (280–315 nm) and low-UV-C (100–280 nm) wavelengths, as these match the first intense absorption band registered in neutral DNA/RNA nucleobases<sup>1,2</sup> and are particularly convenient as they can be probed with the third harmonic of femtosecond Ti:Sapphire laser pulses (often tuned at ~267 nm), which are

commonly used to initiate the photo-reactions in pump-probe experimental set-ups.

DNA/RNA photo-ionisation dynamics has been less explored to date due to the high energies (< 150 nm, i.e. strong UV-C) required to remove electrons from the nucleobases (although this is reduced upon N1 substitution when moving to nucleosides and nucleotides, the nucleobase derivatives actually present in the DNA double helix chain).<sup>6–9</sup> This has for many years prevented in-depth studies, but recent developments in the generation of high-energy few-cycle UV-C pulses<sup>10,11</sup> and the advent of high-energy free electron laser sources<sup>12</sup> have opened up the possibility of studying time-resolved photoionisation processes in such small organic compounds and their use in pump-probe set-ups to monitor photo-reactions.

Photoionisation has also recently been detected at lower energies due to intermolecular interactions within the DNA chain leading to charge separation, which effectively generates cationic and anionic species from photo-excitation,<sup>13,14</sup> thus making their ground and excited state decays of interest in the overall photo-initiated mechanisms of DNA.

The purpose of this article is to characterise the ultrafast radiationless decay of excited state cationic uracil and thymine with dynamically correlated electronic structure methods, and to provide spectral fingerprints that may allow monitoring these photo-processes experimentally in the future.

Several studies in the literature have focused on the accurate

<sup>a</sup> Department of Chemistry, Molecular Sciences Research Hub, Imperial College London, White City Campus, 80 Wood Lane, W12 0BZ, London, UK. E-mail: j.segarramarti@imperial.ac.uk; m.bearpark@imperial.ac.uk

<sup>†</sup> Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/b000000x/

<sup>‡</sup> Additional footnotes to the title and authors can be included e.g. 'Present address:' or 'These authors contributed equally to this work' as above using the symbols: ‡, §, and ¶. Please place the appropriate symbol next to the author's name and include a \footnotetext entry in the the correct place in the list.

characterisation of the ionisation potentials (both vertical and adiabatic) of DNA nucleobases,<sup>15</sup> which are strongly modulated due to solvation,<sup>6–8</sup> while less work has been done in the characterisation of the doublet excited state reactivity. Matsika and co-workers have analysed theoretically the role of the different excited cationic states taking part in the photoionisation of uracil from both static<sup>16–19</sup> and dynamic<sup>20,21</sup> standpoints, including also their subsequent fragmentation.<sup>22</sup> These studies shown that, upon photoionisation, a majority of the population will reach the more favourable  $D_2$  ( ${}^2\pi_{HOMO-1}^+$ ) state, which will then rapidly decay to  $D_1$  ( ${}^2n_O^+$ ) and then reach  $D_0$  ( ${}^2\pi_{HOMO}^+$ ) within 100 fs.

In this contribution we have characterised the different key structures (ground and excited state minima and the conical intersections connecting them) that define the excited state decay of cationic uracil and thymine with state-of-the-art dynamically correlated XMS-CASPT2 energy gradients and couplings. We have then compared these novel results with less computationally demanding CASSCF estimates.<sup>23</sup> Doing this, we find that dynamic electron correlation mainly affects the "3-state" conical intersection previously characterised for uracil<sup>+</sup>,<sup>19</sup> which becomes more clearly a 2-state intersection, as that found for thymine<sup>+</sup> here.

Excited state absorption simulations on top of the characterised cationic  $D_0$  and  $D_1$  minima show similar signals in the 200–400 nm range for both states. These are hard to disentangle from one another, whereas the associated infrared (IR) signals display significant differences between the  $D_0$  and  $D_1$  states related to the intense carbonyl stretching bands predicted for the latter at 1500–1600  $\text{cm}^{-1}$ . This is within the probe windows currently being employed experimentally, which could be used to monitor these photo-ionisation events.

## Computational details

Most computations reported here were carried out with the OpenMOLCAS software package.<sup>24</sup> A large atomic natural orbital basis set (ANO-L) was used throughout in its valence double- $\zeta$  polarised contraction.<sup>25,26</sup> The active space for both uracil and thymine comprises the full  $\pi$  valence occupied and virtual space plus the two  $n_O$  occupied lone-pair orbitals to account for  ${}^2n_O^+$  states, totalling 14 electrons in 10 orbitals for the neutral and 13 electrons in 10 orbitals for the cationic species.

CASSCF wave functions were averaged over four doublet states and were subsequently used for single-point CASPT2 energy corrections.<sup>27–29</sup> An imaginary level shift of 0.2 a.u. was employed in the perturbative step to avoid the presence of intruder states,<sup>30</sup> and IPEA shifts<sup>31</sup> of 0.0 and 0.25 a.u. were tested as this correction has been shown to improve the description of cationic open-shell states in these systems.<sup>15</sup>

CASPT2 computations were performed in its single-state,<sup>27,28</sup> multistate (MS),<sup>32</sup> and extended multistate (XMS)<sup>33</sup> variants to benchmark the effect of the zeroth order Hamiltonian on the cationic manifold. All results here are however reported at the XMS-CASPT2 level of theory unless explicitly stated otherwise, as this was previously found to treat covalent and ionic excited states in a balanced manner.<sup>33</sup>

The resolution of identity based on the Cholesky decomposition was used throughout to speed up the calculation of the electron

repulsion integrals,<sup>34–36</sup> and was used for both energy evaluations<sup>37</sup> as well as in calculating CASSCF analytical gradients<sup>38,39</sup> and non-adiabatic couplings.<sup>40</sup> Second-order nuclear derivatives were computed numerically employing the aforementioned gradients.<sup>41</sup>

CASSCF conical intersections were characterised with the method of Fdez Galván *et al.*<sup>40</sup> Given the current lack of tools to explicitly search for 3-state crossings at the CASPT2 level of theory, we relied on a  $D_2/D_0$  crossing optimisation in order to provide us with an estimate of this crossing, which features a relatively small energy gap between  $D_2/D_0$  and  $D_1$ , which geometrical should be closer to the actual 3-state degeneracy.

The characterised cationic ground ( $D_0$ ) and excited ( $D_1$ ) state minima, as well as the different low-lying conical intersections (CIs) were also optimised at the CASPT2 level of theory to establish the role of dynamic electron correlation on the geometrical parameters in these cations, as it has been shown to significantly impact the singlet manifold.<sup>42–45</sup> CASPT2 minima and CI optimisations (using the projection method of Bearpark *et al.*<sup>46</sup>) were also carried out with analytical gradients<sup>47–49</sup> and couplings<sup>50</sup> as implemented in BAGEL<sup>51</sup> to ensure that those obtained using numerical differentiation are correct. Linear interpolations in internal coordinates were performed between the key structures optimised (minima and conical intersections) in order to rule out the presence of potential energy barriers.

Optimisations are only considered at the XMS-CASPT2 level of theory here, because: i) SS-CASPT2 is known to provide the wrong dimensionality of the branching space in conical intersections,<sup>52</sup> which are central to the present study; ii) MS-CASPT2 requires the enlargement of the active space to balance the dynamic electron correlation included to states of different character around conical intersections, i.e. crossings between ionic and covalent states, as reported by Serrano-Andrés *et al.*,<sup>53</sup> and this is unfeasible for optimisations; and iii) XMS-CASPT2 has been recently shown to provide smoother potential energy surface profiles around conical intersections due to treating crossings between ionic and covalent states more accurately than SS- and MS-CASPT2 formulations.<sup>54,55</sup>

Additional computations averaging over the lowest-lying 30 electronic doublet states were carried out on top of the different minima to evaluate the excited state absorption signals.<sup>56,57</sup> Our simulations of these spectral signals assume that excited state absorptions of the individual  ${}^2n_O^+$  and  ${}^2\pi^+$  states are dominated by the electronic structure at their corresponding minima,<sup>58,59</sup> thus neglecting the time-evolution of the system and its lineshape as their simulation requires more sophisticated approaches that are out of the scope of the present study.<sup>60–63</sup> The CAS state interaction method<sup>64</sup> was employed to evaluate transition dipole moments and oscillator strengths and the energies were corrected with the standard (single-state) CASPT2 formulation with an IPEA shift of 0.0.

The transitions were broadened with Gaussian functions having full width at half maximum of 0.3 eV, as used in similar organic systems,<sup>65</sup> while the IR signals were broadened with Gaussian functions having full width at half maximum of 10  $\text{cm}^{-1}$ .<sup>66</sup> This is an approximation: we have not computed rovibrational

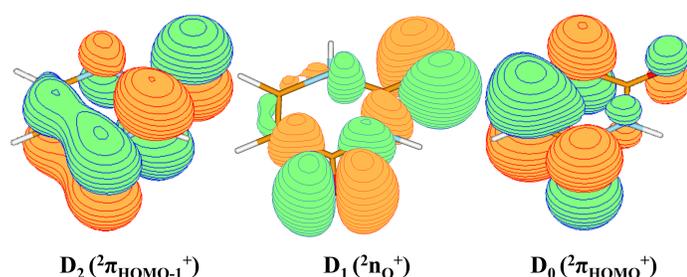
broadenings explicitly. Using the same approach as here, our previous studies on related systems (Cyt and 5mCyt)<sup>66</sup> where experimental evidence is also available reproduce experiment reasonably well in terms of UV/Vis excited state absorptions and IR signals, which we expect should also translate to the very similar Ura and Thy that are studied here. The short lifetime of the accessed excited states means there will be extra broadenings that we cannot reliably estimate.

Excited state absorptions and IR signals were broadened as implemented in Gabedit<sup>67</sup> and orbital visualisation was performed with Molden.<sup>68</sup>

## Results and discussion

### Uracil<sup>+</sup> energies

We start by analysing and comparing the results obtained for the two well-defined experimental magnitudes often employed to benchmark the cationic manifold computationally, that is, the vertical and adiabatic ionisation potentials.



**Fig. 1** The singly occupied orbitals of the three electronic excited states of uracil cation studied here.

Averaging values from Table 1 for each state in turn, the different formulations of the CASPT2 method predict the first three vertical ionisation potentials to be placed at 9.53, 10.00 and 10.59 eV, while the first adiabatic ionisation potential is placed at 9.35 eV, in agreement with the available experimental evidence<sup>69–72</sup>, as well as previous theoretical estimates.<sup>6,15,74</sup> As has been reported by Roca-Sanjuán *et al.*,<sup>15</sup> the introduction of the IPEA shift appears to blue-shift the estimates obtained at the CASPT2 level, bringing them closer to the experimental values. This effect, which amounts to  $\sim 0.2$  eV for all CASPT2 zeroth-order Hamiltonians, is however overestimated when employing it together with multistate CASPT2 variants (either MS or XMS), as these already feature ionisation energies within the experimental range with IPEA=0.0 a.u. XMS-CASPT2 is shown to be blue-shifted with respect to MS and single-state formulations, a feature previously reported for the singlet manifold.<sup>55,75</sup>

In contrast, CASSCF places the first three vertical and first adiabatic ionisation potentials at 8.58, 9.36, 9.47 and 8.18 eV, respectively (Table 1), which deviate from the experimental evidence by  $\sim 1$  eV. Nevertheless, CASSCF still provides acceptable energies when considering solely the cationic manifold and the differences among the diverse cationic states. Concretely, CASSCF predicts a difference of 0.78 eV between the  $D_0$  and  $D_1$  states and 0.11 eV between  $D_1$  and  $D_2$  in the Franck-Condon (FC) region, compared to the averaged values over all CASPT2 formulations of  $\sim 0.48$

and 0.59 eV, respectively, while preserving the correct ordering of the states. These values also agree qualitatively with those obtained by Assmann *et al.*<sup>20</sup> at the Equation Of Motion Ionisation Potential Coupled Cluster Singles and Doubles (EOM-IP-CCSD) level of theory. Concretely, EOM-IP-CCSD reports  $D_0 - D_1$  and  $D_1 - D_2$  energy gaps of 0.69 and 0.36 eV whereas averaged CASPT2 values yield differences of 0.48 and 0.59 eV, respectively. These differences point towards a red-shift of the  $D_1(n_O^+)$  state at the CASPT2 level, which is however quantitatively comparable with the experimental evidence available for the ionisation potential of this state.

Similar differences of  $\sim 0.3$  eV are observed between CASSCF and the different CASPT2 methods for the  $D_0 - D_1$  energy gaps estimated at the  $D_0$  minimum (1.35 vs 1.01 eV), while the  $D_2 - D_1$  gap is almost quantitatively recovered (0.69 vs 0.78 eV). Albeit significant, these differences are much smaller than those observed in the singlet manifold in DNA/RNA nucleobases<sup>43</sup> where state ordering with CASSCF often differs from the one obtained with correlated methods, an issue that impacts the simulation of their photo-processes.

Overall, differences within the cationic manifold due to the inclusion of dynamic correlation are on a quantitative level, which supports the use of CASSCF as a cost-effective option for qualitatively characterising photo-ionisation events in uracil.

### Uracil<sup>+</sup> geometries

We start by looking at the geometric distortions undergone in the cationic manifold upon photo-ionisation with respect to the starting singlet FC ( $S_0$ ) equilibrium geometry at the XMS-CASPT2 level of theory (Fig. 2, distances in red). The  $D_0$  ( $2\pi_{HOMO}^+$ ) minimum shows N1-C2 and C5-C6 bond elongations and a C6-N1 bond shortening of 0.05 Å, which reflects the singly occupied molecular orbital describing this cationic state (see Fig. 1). The  $D_1$  ( $2n_O^+$ ) minimum, on the other hand, displays its most noticeable distortions around the C4-O carbonyl and surrounding atoms where this cationic state localises, with a lengthening for C2-N3 of 0.04 and of 0.05 Å for the C4-O bond, and a shortening of 0.05 Å for N3-C4.

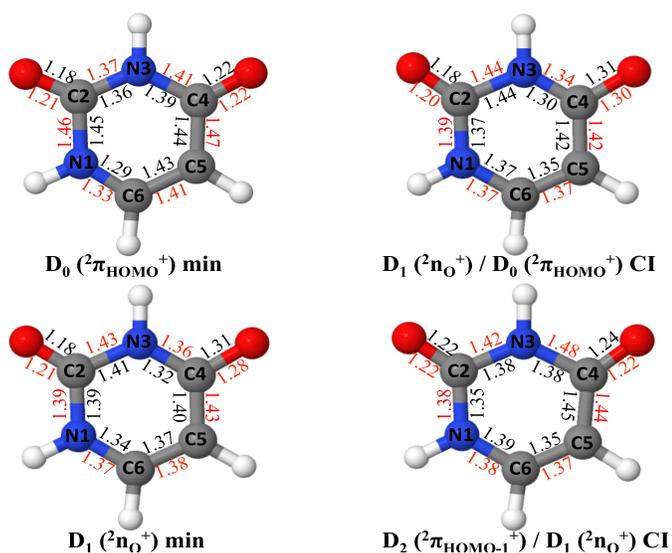
The  $D_2/D_1$  CI is characterised by a large N3-C4 bond elongation of 0.07 Å with respect to the FC reference. The  $D_1/D_0$  CI, on the other hand, displays more pronounced distortions: bond elongations of 0.05 Å along C2-N3 and 0.07 Å along C4-O and a shortening of N3-C4 by 0.07 Å.

We look next at the optimised geometries at the CASSCF level (Fig. 2, distances in black) to determine the geometry changes induced by including dynamic electron correlation. Noticeable differences in the  $D_0$  minimum are observed in the C6-N1 bond length, which is 0.04 Å shorter than with XMS-CASPT2, while C4-C5 is 0.03 Å shorter. The  $D_1$  minimum presents bond length shortenings of 0.04 Å for N3-C4, 0.03 Å for C4-C5 and 0.03 Å for C6-N1, and an elongation of 0.03 Å for C4-O.

The  $D_2/D_1$  CI shows the most marked difference among all structures studied: the N3-C4 distance is shortened by 0.1 Å on going from XMS-CASPT2 to CASSCF, which is also followed by shortenings of 0.04 and 0.03 Å for the C2-N3 and N1-C2 bonds,

**Table 1** Gas-phase vertical and adiabatic ionisation potentials of uracil computed at the CASSCF level of theory and with a range of zeroth-order CASPT2 Hamiltonians, and their comparison with the available experimental data where appropriate. Both  $S_0$  and  $D_0$  minima are computed at the CASSCF level of theory. All energies are given in eV.

	IPEA shift	Vertical Ionisation Potentials ( $S_0$ min)			Ionisation potentials ( $D_0$ min)		
		${}^2\pi_{HOMO}^+$	${}^2n_O^+$	${}^2\pi_{HOMO-1}^+$	${}^2\pi_{HOMO}^+$	${}^2n_O^+$	${}^2\pi_{HOMO-1}^+$
CASSCF	-	8.58	9.36	9.47	8.18	9.52	10.22
CASPT2	0.0	9.38	9.82	10.41	9.22	10.26	10.92
	0.25	9.56	10.04	10.61	9.37	10.41	11.17
MS-CASPT2	0.0	9.43	9.86	10.46	9.26	10.31	10.98
	0.25	9.61	10.08	10.65	9.41	10.44	11.21
XMS-CASPT2	0.0	9.51	10.01	10.62	9.36	10.27	11.16
	0.25	9.66	10.21	10.78	9.47	10.43	11.36
Experiment		9.45 - 9.60 <sup>69-72</sup>	10.02 - 10.13 <sup>69-72</sup>	10.51 - 10.55 <sup>69-72</sup>	9.20 - 9.32 <sup>73</sup>	-	-



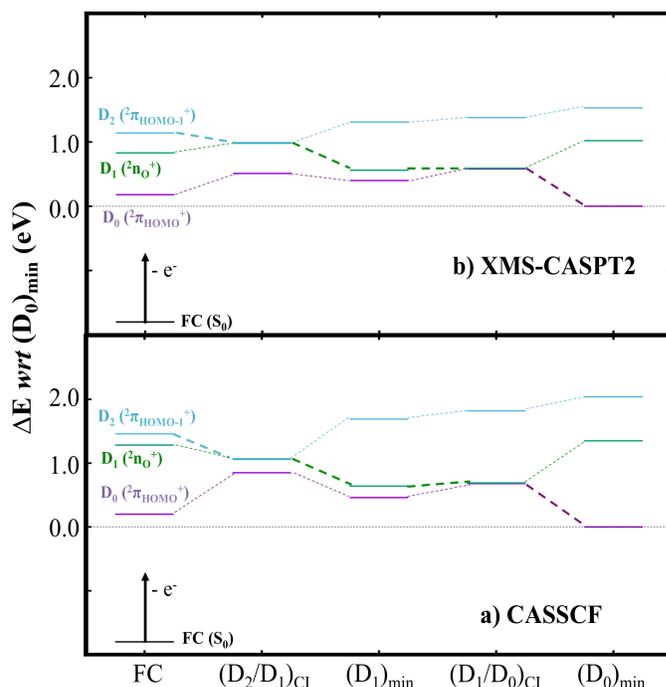
**Fig. 2** Minima (min) and conical intersections (CI) characterised for the lowest-lying cationic states of uracil. All bond lengths are given in Å. CASSCF values in black; CASPT2 values in red.

respectively. This conical intersection, however, presents an effective 3-state degeneracy at the CASSCF level as has been previously reported by Matsika<sup>19</sup> while it features a 2-state crossing at the XMS-CASPT2 level. The  $(D_1/D_0)_{CI}$  again shows bond length shortenings of 0.02, 0.04 and 0.02 Å for N1-C2, N3-C4 and C5-C6, respectively.

Apart from the  $D_2/D_1$  CI, CASSCF geometries are shown to be qualitatively similar to those at the XMS-CASPT2 level and thus present a cost-effective alternative approach for their characterisation.

### Uracil<sup>+</sup> evolution and decay

Fig. 3 shows the sequence of geometries encountered following photo-ionisation with energies given with respect to  $(D_0)_{min}$ . We have focused on two different approaches here: a) CASSCF and b) XMS-CASPT2, which both give the same state ordering.



**Fig. 3** Potential energy surfaces of cationic uracil computed at the a) CASSCF and b) XMS-CASPT2 levels of theory. All energies are given in eV with respect to  $(D_0)_{min}$ . Thick lines represent the evolution of the excited state population assuming initial activation of the  $D_2$  state.

Following previous studies,<sup>20</sup> we assume direct population into the  $D_2({}^2\pi_{HOMO-1}^+)$  state. An appealing aspect of photoionisation is that standard selection rules such as those commonly used for singlet-to-singlet transitions are not applicable and thus all sorts of ionisations are allowed and can be tuned by the energy of the ionising pulse employed in the experiment. We analyse photoionisation events upon  $D_2$  population as it crosses the  $(D_2/D_1)$  intersection region, which features the largest differences due to the inclusion of dynamic correlation.

Initial access to  $D_2$  undergoes an ultrafast relaxation to the  $(D_2/D_1)_{CI}$ : there is no separate  $D_2$  minimum. At the CASSCF level,  $D_0$  has almost the same energy at this geometry ( $\sim 0.2$  eV off the  $D_1$  and  $D_2$  states), leading effectively to a 3-state CI.<sup>19,20</sup> Interestingly, gradient difference and non-adiabatic coupling vectors show marked distortions in bond lengths while preserving

ring planarity. On the other hand, XMS-CASPT2 presents a 2-state ( $D_2/D_1$ )<sub>CI</sub> where the  $D_0$  state is placed 0.5 eV lower in energy. Attempts were made to locate the 3-state CI at the XMS-CASPT2 level by effectively optimising the crossing between the  $D_2(\pi_{HOMO-1}^+)$  and  $D_0(\pi_{HOMO}^+)$  states, which features a closer lying (but not degenerate)  $D_1(n_O^+)$  state. This crossing is however placed at higher energies than the  $D_2$  state in the FC region and thus the involvement of a 3-state crossing in the ensuing photoionisation events appears to be ruled out.

Notable differences arise also when comparing the energy of the  $D_2(\pi_{HOMO-1}^+)$  state at the FC and ( $D_2/D_1$ )<sub>CI</sub> geometries, showing a 0.46 eV stabilisation at CASSCF and 0.15 eV at the XMS-CASPT2 level, in accordance with the pronounced differences previously described for their respective characterised geometries. This energy difference implies a more significant relaxation at the CASSCF level, which dynamically would be associated with larger kinetic energies when reaching the crossing.

We then assume a  $D_2 \rightarrow D_1$  population transfer, which upon further relaxation leads to a well-defined excited state minimum ( $D_1$ )<sub>min</sub> that has not been characterised in previous studies.<sup>19</sup> This minimum is however placed at energies close to the ( $D_1/D_0$ )<sub>CI</sub> and displays negligible barriers to this at both CASSCF and XMS-CASPT2 (0.05 vs 0.02 eV, respectively) enabling ultrafast decay to the cationic ( $D_0$ )<sub>min</sub> ground state.<sup>20</sup>

From this we conclude that dynamic correlation impacts particularly the effective 3-state CI and thus its direct involvement in the ultrafast excited state decay to the ground cationic state. It is worth noting that using the SS-CASPT2//CASSCF protocol - i.e. without further geometry optimisation - leads to pronounced differential correlation effects at ( $D_2/D_1$ )<sub>CI</sub>, giving the wrong state ordering and altering the nature of the states taking part in the crossing, thus discouraging its use.

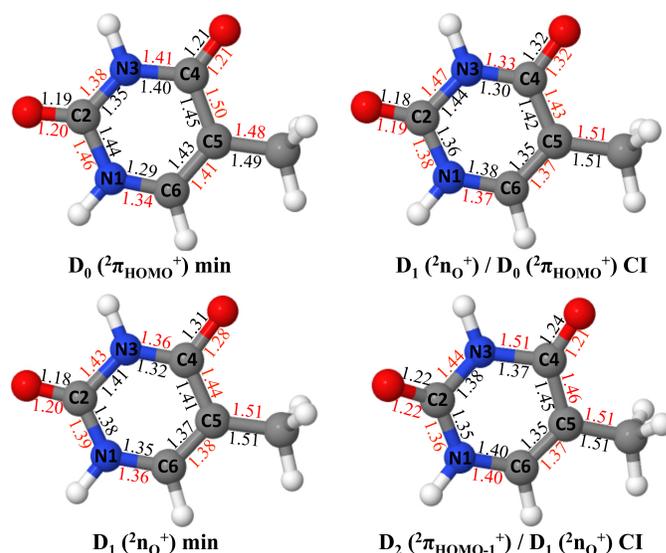
Overall we predict a significant decrease in the probability of direct  $D_2 \rightarrow D_0$  deactivation when moving from CASSCF to XMS-CASPT2 potential energy surfaces, which increases the  $D_1 \rightarrow D_0$  decay. Given the close-lying energetic position of ( $D_1$ )<sub>min</sub> and ( $D_1/D_0$ )<sub>CI</sub> (negligible barrier to reach the CI) we expect the process to still occur on ultrafast (sub-ps) timescales.<sup>20</sup>

### Thymine<sup>+</sup> energies

We consider the vertical and adiabatic ionisation potentials of thymine<sup>+</sup> next, where we particularly focus on the differences introduced due to 5-methylation by comparing the results with those previously obtained for uracil<sup>+</sup>. Due to their structural similarities, the lowest-lying cationic states of thymine are equivalent to those of uracil with singly occupied molecular orbitals as reported in Fig. 1.

Averaged CASPT2 estimates of the first three vertical and first adiabatic ionisation potential are 9.22, 9.93, 10.51 and 9.05 eV, respectively, which are placed roughly within a tenth of an eV from the experimental evidence (see Table 2).<sup>69–72</sup>

We observe a 0.2 ( $D_2$ ) to 0.3 ( $D_0$ ) eV red-shift on vertical and adiabatic ionisation potentials of  $2\pi^+$  character related to 5-methylation, i.e. moving from uracil<sup>+</sup> to thymine<sup>+</sup>. On the other hand, this effect appears to be negligible for  $2n_O^+$  states, which



**Fig. 4** Minima (min) and conical intersections (CI) characterised for the lowest-lying cationic states of thymine. All bond lengths are given in Å. CASSCF values in black; CASPT2 values in red.

leads to smaller  $D_2 - D_1$  gaps in thymine<sup>+</sup> compared to uracil<sup>+</sup>. We reproduce the trend registered experimentally whereby 5-methylation induces a stabilisation on the  $2\pi^+$  cationic manifold.<sup>69–72</sup>

Within the cationic manifold, CASSCF predicts 1 eV  $D_0 - D_1$  and 0.08 eV  $D_1 - D_2$  energy gaps at the FC region, in contrast with the XMS-CASPT2 estimates of 0.71 and 0.57 eV, respectively, leading to differences of  $\sim 0.3$  eV for the former and  $\sim 0.5$  eV for the latter. The red-shift on  $D_2$  and the negligible effect on  $D_1$  upon methylation leads to a near degeneracy between these states, which is however broken when including dynamic correlation.

Summarising, CASSCF estimates qualitatively reproduce XMS-CASPT2 values in the FC region with the exception of the  $D_2 - D_1$  energy gap, which leads to a near-degeneracy in the former and a half an eV gap in the latter.

### Thymine<sup>+</sup> geometries

We next compare thymine cationic structures with those of its FC region, starting with those computed at the XMS-CASPT2 level (see Fig. 4, distances in red).

The  $D_0$  minimum shows a marked 0.07 and 0.05 Å bond lengthening in N1-C2 and C5-C6, respectively, and a shortening of 0.04 Å for the C6-N1 bond. The  $D_1$  minimum shows its main distortions to be localised in the C4-O carbonyl group and surroundings, displaying N3-C4 and C4-C5 bond shortenings of 0.04 and 0.03 Å, respectively, and an elongation of C4-O of 0.05 Å. These changes mirror those found in uracil<sup>+</sup>.

The  $D_2/D_1$  CI shows bond length differences similar to those previously outlined for uracil<sup>+</sup>. However, this crossing features out-of-plane distortions with puckered carbonyl groups as opposed to the planar geometry found for uracil<sup>+</sup>. The  $D_1/D_0$  CI also presents large differences with respect to the FC reference, comparable to those in uracil<sup>+</sup> but the structure remains planar.

**Table 2** Gas-phase vertical and adiabatic ionisation potentials of thymine computed at the CASSCF level of theory and with a range of zeroth-order CASPT2 Hamiltonians, and their comparison with the available experimental data where appropriate. Both  $S_0$  and  $D_0$  minima are computed at the CASSCF level of theory. All energies are given in eV.

	Vertical Ionisation Potentials ( $S_0$ min)				Ionisation potentials (@ $D_0$ min)		
	IPEA shift	${}^2\pi_{HOMO}^+$	${}^2n_O^+$	${}^2\pi_{HOMO-1}^+$	${}^2\pi_{HOMO}^+$	${}^2n_O^+$	${}^2\pi_{HOMO-1}^+$
CASSCF	-	8.27	9.27	9.35	7.88	9.55	10.09
CASPT2	0.0	9.00	9.67	10.25	8.84	10.18	10.77
	0.25	9.20	9.91	10.46	9.01	10.35	11.01
MS-CASPT2	0.0	9.05	9.70	10.31	8.90	10.23	10.82
	0.25	9.24	9.94	10.50	9.05	10.38	11.05
XMS-CASPT2	0.0	9.33	10.08	10.67	9.18	10.41	11.20
	0.25	9.50	10.30	10.84	9.31	10.61	11.41
Experiment		9.02 - 9.20 <sup>69-72</sup>	9.95 - 10.05 <sup>69-72</sup>	10.39 - 10.44 <sup>69-72</sup>	8.80 - 8.87 <sup>73</sup>	-	-

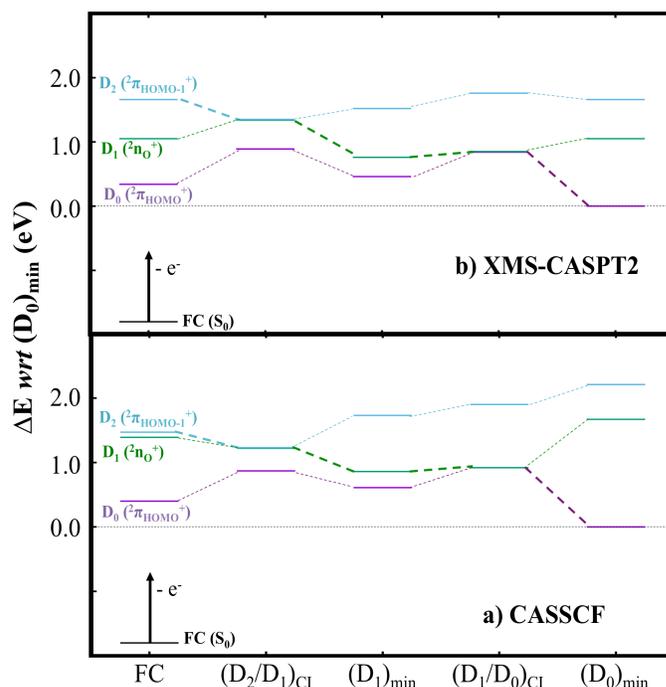
CASSCF geometries (Fig. 4, distances in black) show systematic bond shortenings compared to XMS-CASPT2, in agreement with what was previously found for uracil<sup>+</sup>. The  $D_2/D_1$  CI displays the largest differences, highlighted by the 0.14 Å shortening of N3-C4, which is the same coordinate where the most drastic differences in uracil<sup>+</sup> were located. These differences go beyond bond length distances, featuring a planar structure for CASSCF in contrast to the puckered geometry obtained with XMS-CASPT2. The  $D_1/D_0$  CI presents small differences between the CASSCF and XMS-CASPT2 results, the main ones being bond length shortenings of 0.03 Å for the C2-N3 and N3-C4 bonds.

Aside from  $D_2/D_1$  CI, CASSCF geometries appear to qualitatively reproduce those obtained for XMS-CASPT2, which is in line with what has been obtained previously for uracil<sup>+</sup>.

### Thymine<sup>+</sup> evolution and decay

We assume the photo-ionisation of thymine<sup>+</sup> to occur in the FC equilibrium region and to initially populate the  $D_2({}^2\pi_{HOMO-1}^+)$  state (Fig. 5), as uracil<sup>+</sup>.

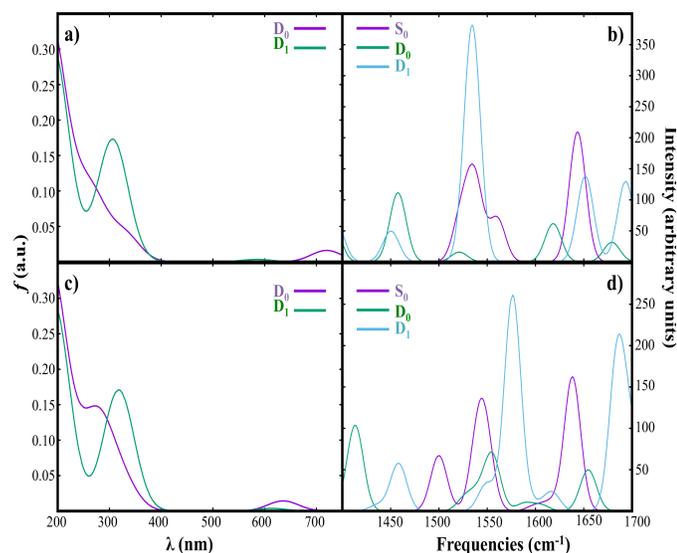
Upon initial promotion to  $D_2({}^2\pi_{HOMO-1}^+)$ , a swift relaxation is expected to lead the excited state population to the  $(D_2/D_1)_{CI}$ . Importantly, despite featuring large geometrical differences between the CASSCF and XMS-CASPT2 optimised structures highlighted above, CASSCF does not predict an effective "3-state" CI as for uracil<sup>+</sup> and presents a 2-state crossing with a large  $D_1 - D_0$  energy gap of 0.36 eV, in line with the 0.46 eV obtained at the XMS-CASPT2 level.



**Fig. 5** Potential energy surfaces of cationic thymine computed at the a) CASSCF and b) XMS-CASPT2 levels of theory. All energies are given in eV with respect to  $(D_0)_{min}$ . Thick lines represent the evolution of the excited state population assuming initial activation of the  $D_2$  state.

We thus predict population to be funnelled to the  $D_1({}^2n_O^+)$  state and to proceed relaxing towards the  $D_1$  minimum. Both computational methods support the same mechanism whereby population will move to reach the  $(D_1/D_0)_{CI}$  by surmounting very small potential energy barriers (0.06 eV at CASSCF; 0.09 eV at XMS-CASPT2). Upon reaching this crossing, the population is expected to be funnelled down to the cationic ground state and relax towards its minimum.

Dynamic electron correlation affects the initial separation of the  $D_1$  and  $D_2$  states, which are almost degenerate at the CASSCF level but display a sizeable energy gap with respect XMS-CASPT2. We expect this to affect the lifetime of the  $D_2$  state. Overall, the effect of 5-methylation on energies and geometries is shown to be more marked when accounting for dynamic correlation but CASSCF estimates are still expected to be cost-effective alterna-



**Fig. 6** Electronic (excited state; panels a and c) and vibrational (IR; panels b and d) absorption signals of cationic uracil (top panels) and thymine (bottom panels) in their ground ( $D_0$ ) and excited state ( $D_1$ ) minima. The vibrational signals at  $S_0$  are also added for the sake of comparison.

tives to yield qualitative results.

Interestingly, and despite dynamic electron correlation leading to more pronounced differences, its impact in the overall photoionisation mechanism seems to be less than that described for uracil<sup>+</sup> above. 5-methylation lifts the effective 3-state degeneracy at the CASSCF level in the  $(D_2/D_1)_{CI}$ , reconciling CASSCF and XMS-CASPT2 estimates. Differential correlation effects arise also for this geometry, showing a qualitatively different state ordering at the SS-CASPT2//CASSCF level and thus discouraging the use of this protocol for cationic DNA/RNA species.

The large geometric differences among the different levels of theory are shown to not alter the shape of the potential energy surface and lead to qualitatively similar results (Fig. 5). We thus expect the main differences between CASSCF and XMS-CASPT2 in thymine<sup>+</sup> to be related to the  $D_2$  lifetime, which we predict will be much shorter at the former due to the energy degeneracy featured with  $D_1$  in the FC region.

Overall, we expect thymine to follow very similar dynamics at CASSCF and XMS-CASPT2 levels of theory, predicting ultrafast  $D_2 \rightarrow D_1$  and subsequent  $D_1 \rightarrow D_0$  deactivations. This is in contrast with uracil<sup>+</sup>, where the appearance of the effective "3-state" CI at the CASSCF level suggests an efficient  $D_2 \rightarrow D_0$  decay channel<sup>20</sup>, which we expect to be hampered upon inclusion of dynamic electron correlation as shown above.

### Electronic and IR absorption signals

We simulate next potential spectral fingerprints that may allow monitoring of these photo-initiated events. Fig. 6 shows both electronic excited state absorptions and IR originating from the two well-defined cationic ground  $D_0$  ( ${}^2\pi_{HOMO}^+$ ) and excited  $D_1$  ( ${}^2n_O^+$ ) minima. Standard UV/Vis probe ranges (200-750 nm) are considered for excited state absorptions computed at the CASPT2 level and encompassing a range of dipole allowed bright

${}^2\pi^* \leftarrow {}^2\pi^+$  and  ${}^2n_O^* \leftarrow {}^2n_O^+$  transitions, whereas a 1400-1700  $\text{cm}^{-1}$  probe window as often used in the transient experiments<sup>76,77</sup> is analysed for the IR, which employs frequencies computed at the CASSCF level of theory.

Top panels (a and b) in Fig. 6 display the different UV/Vis and IR signals predicted for uracil<sup>+</sup>. Excited state absorptions in the UV/Vis window refer to those often recorded in pump-probe set-ups, which in this case display overlapping signals for  $D_0$  and  $D_1$  states in the high energy UV (200-400 nm) region while showing a small absorption contribution unique to  $D_0$  in the low-energy Vis ( $\sim 700$  nm) window. IR signals appear to be more easily disentangled, with a large peak at  $\sim 1525$   $\text{cm}^{-1}$  representing a carbonyl stretching that relates to the carbonyl bond lengthening observed in the  $D_1$  ( ${}^2n_O^+$ ) minimum while a prominent peak at  $\sim 1450$  appears to be the more distinctive fingerprint for  $D_0$  ( ${}^2\pi_{HOMO}^+$ ). It is worth noting that the large peak  $\sim 1525$   $\text{cm}^{-1}$  in  $D_1$  largely overlaps with the carbonyl signal of the FC region, which may make its characterisation difficult at early times.

The UV/Vis and IR spectra for thymine<sup>+</sup> are given in the lower panels (c and d) of Fig. 6. As can be seen, they are very similar to those previously introduced for uracil<sup>+</sup> and present analogous fingerprints, which are however slightly blue-shifted due to 5-methylation. The main UV/Vis intense signals (200-400 nm) appear relatively unchanged whereas the low-energy absorption signal ( $\sim 650$  nm) is slightly blue-shifted compared to uracil<sup>+</sup>, while the IR spectra features the prominent  $D_1$  carbonyl peaks slightly shifted at  $\sim 1575$   $\text{cm}^{-1}$  and the  $D_0$  signals at  $\sim 1425$   $\text{cm}^{-1}$ , which prevents the overlap with the signals arising from the FC ( $S_0$ ).

Overall, we expect the low-energy Vis window to provide fingerprints that will allow monitoring the  $D_0$  state, whereas the carbonyl stretching at  $\sim 1550$   $\text{cm}^{-1}$  appears to be the best signal to follow the  $D_1$  decay.

## Conclusions

In this article we thoroughly analyse the role played by dynamic electron correlation in the description of the cationic state manifold in uracil and thymine in terms of energies and geometries, the latter being considered at the XMS-CASPT2 level for the first time.

We find that energies are sensitive to dynamic electron correlation, a range of CASPT2 formulations yielding results in line with the experimental observables (vertical and adiabatic ionisation potentials). CASSCF energies are shown to be qualitatively correct when considering solely the cationic manifold in comparison with the XMS-CASPT2 reference, but appear to be red-shifted with respect to the available experimental evidence. 5-methylation (thymine) introduces red-shifts in the cationic manifold of  ${}^2\pi^+$  character with respect to uracil while leaving the  ${}^2n_O^+$  states unaffected, which is reproduced at both CASSCF and XMS-CASPT2 levels of theory, leading to smaller cationic energy gaps.

The largest differences observed between the two systems studied are related to the effective "3-state" intersection  $(D_2/D_1/D_0)_{CI}$  previously characterised at the CASSCF level for uracil<sup>+</sup> by Matsika,<sup>19,20</sup> which is not featured in thymine<sup>+</sup>. Inclusion of dynamic electron correlation leads to a well separated  $D_2/D_1$  CI

with a large energy gap with the ground state ( $\sim 0.5$  eV for both systems). A real 3-state crossing at the CASPT2 level will lie at a different geometry and higher energy.

Using XMS-CASPT2 potential energy surfaces, we show how, upon accessing initially the ( ${}^2\pi_{HOMO-1}^+$ ) state, the excited state population will reach the ( $D_2/D_1$ ) CI and relax to a well-defined  $D_1$  minimum, which we report here for the first time. Negligible potential energy barriers between this minimum and the ( $D_1/D_0$ ) CI funnel down the population to the cationic ground state where it will relax to the  $D_0$  minimum. CASSCF results are shown to provide analogous qualitative results for thymine<sup>+</sup> but feature the effective "3-state" degeneracy at the ( $D_2/D_1$ ) CI for uracil<sup>+</sup> that has been suggested to enable a direct  $D_2 \rightarrow D_0$  decay.<sup>20</sup>

Dynamic electron correlation therefore impacts the photoionisation events studied: in uracil<sup>+</sup> it disfavours any direct  $D_2 \rightarrow D_0$  decay channel and increases the sequential  $D_2 \rightarrow D_1$  and  $D_1 \rightarrow D_0$  deactivation by disrupting the "3-state" intersection. In thymine<sup>+</sup> it will increase the initial lifetime of the initially accessed  $D_2$  state, which appears degenerate with  $D_1$  at the CASSCF level in the FC region. Our results for these systems show large differential correlation effects that lead to a wrong state ordering when employing the popular CASPT2//CASSCF protocol, discouraging its use for cationic DNA/RNA species.

Simulations of the UV/Vis and IR transient absorption signals on top of the characterised  $D_0$  ( ${}^2\pi_{HOMO}^+$ ) and  $D_1$  ( ${}^2n_O^+$ ) minima reveal strong similarities between uracil<sup>+</sup> and thymine<sup>+</sup>. In the UV/Vis window, weak signals in the 600-700 nm range associated with the  $D_0$  state are expected to be the main fingerprints of this state, while the IR probe window presents large carbonyl stretching fingerprints at  $\sim 1550$  unique to  $D_1$  that could allow monitoring of the photoionisation decays of both uracil and thymine.

## Acknowledgments

J.S.-M. acknowledges support from the European Commission through the Marie Curie actions (*AttoDNA*, FP8-MSCA-IF, grant no 747662). We wish to thank Dr. João Malhado and Prof. Mike Robb for useful discussions. We dedicate this work to Prof. Mike Ashfold on his 65<sup>th</sup> birthday, in recognition of his outstanding research in physical chemistry and chemical physics, and his support for computational chemistry.

## Notes and references

- 1 C. E. Crespo-Hernández, B. Cohen, P. M. Hare and B. Kohler, *Chem. Rev.*, 2004, **104**, 1977–2020.
- 2 C. T. Middleton, K. de La Harpe, C. Su, Y. K. Law, C. E. Crespo-Hernández and B. Kohler, *Annual Review of Physical Chemistry*, 2009, **60**, 217–239.
- 3 A. Giussani, J. Segarra-Martí, D. Roca-Sanjuán and M. Merchán, *Top. Curr. Chem.*, 2015, **355**, 57–97.
- 4 R. Improta, F. Santoro and L. Blancafort, *Chemical Reviews*, 2016, **116**, 3540–3593.
- 5 W. J. Schreier, P. Gilch and W. Zinth, *Annual Review of Physical Chemistry*, 2015, **66**, 497–519.
- 6 P. Slaviček, B. Winter, M. Faubel, S. E. Bradforth and P. Jungwirth, *Journal of the American Chemical Society*, 2009, **131**, 6460–6467.
- 7 E. Pluhařová, P. Jungwirth, S. E. Bradforth and P. Slaviček, *The Journal of Physical Chemistry B*, 2011, **115**, 1294–1305.
- 8 E. Cauët, M. Valiev and J. H. Weare, *The Journal of Physical Chemistry B*, 2010, **114**, 5886–5894.
- 9 M. Rubio, D. Roca-Sanjuán, M. Merchán and L. Serrano-Andrés, *The Journal of Physical Chemistry B*, 2006, **110**, 10234–10235.
- 10 S. L. Horton, Y. Liu, P. Chakraborty, S. Matsika and T. Weinacht, *The Journal of Chemical Physics*, 2017, **146**, 064306.
- 11 S. L. Horton, Y. Liu, P. Chakraborty, P. Marquetand, T. Rozgonyi, S. Matsika and T. Weinacht, *Phys. Rev. A*, 2018, **98**, 053416.
- 12 T. J. A. Wolf, F. Holzmeier, I. Wagner, N. Berrah, C. Bostedt, J. Bozek, P. Bucksbaum, R. Coffee, J. Cryan, J. Farrell, R. Feifel, T. J. Martinez, B. McFarland, M. Mucke, S. Nandi, F. Tarantelli, I. Fischer and M. Gühr, *Applied Sciences*, 2017, **7**, 681.
- 13 D. B. Bucher, B. M. Pilles, T. Carell and W. Zinth, *Proceedings of the National Academy of Sciences*, 2014, **111**, 4369–4374.
- 14 A. Banyasz, T. Ketola, L. Martínez-Fernández, R. Improta and D. Markovitsi, *Faraday Discuss.*, 2018, **207**, 181–197.
- 15 D. Roca-Sanjuán, M. Rubio, M. Merchán and L. Serrano-Andrés, *The Journal of Chemical Physics*, 2006, **125**, 084302.
- 16 M. Kotur, T. C. Weinacht, C. Zhou and S. Matsika, *Phys. Rev. X*, 2011, **1**, 021010.
- 17 M. Kotur, T. Weinacht, C. Zhou and S. Matsika, *IEEE Journal of Selected Topics in Quantum Electronics*, 2012, **18**, 187–194.
- 18 S. Matsika, M. Spanner, M. Kotur and T. C. Weinacht, *The Journal of Physical Chemistry A*, 2013, **117**, 12796–12801.
- 19 S. Matsika, *Chemical Physics*, 2008, **349**, 356–362.
- 20 M. Assmann, T. Weinacht and S. Matsika, *The Journal of Chemical Physics*, 2016, **144**, 034301.
- 21 M. Assmann, H. Köppel and S. Matsika, *The Journal of Physical Chemistry A*, 2015, **119**, 866–875.
- 22 C. Zhou, S. Matsika, M. Kotur and T. C. Weinacht, *The Journal of Physical Chemistry A*, 2012, **116**, 9217–9227.
- 23 K. F. Hall, M. Boggio-Pasqua, M. J. Bearpark and M. A. Robb, *The Journal of Physical Chemistry A*, 2006, **110**, 13591–13599.
- 24 F. Aquilante, J. Autschbach, R. K. Carlson, L. F. Chibotaru, M. G. Delcey, L. De Vico, I. Fdez. Galván, N. Ferré, L. M. Frutos, L. Gagliardi, M. Garavelli, A. Giussani, C. E. Hoyer, G. Li Manni, H. Lischka, D. Ma, P. Å. Malmqvist, T. Müller, A. Nenov, M. Olivucci, T. B. Pedersen, D. Peng, F. Plasser, B. Pritchard, M. Reiher, I. Rivalta, I. Schapiro, J. Segarra-Martí, M. Stenrup, D. G. Truhlar, L. Ungur, A. Valentini, S. Vancoillie, V. Veryazov, V. P. Vysotskiy, O. Weingart, F. Zapata and R. Lindh, *Journal of Computational Chemistry*, 2016, **37**, 506–541.
- 25 P.-O. Widmark, P.-Å. Malmqvist and B. O. Roos, *Theor. Chim. Acta*, 1990, **77**, 291.
- 26 P.-O. Widmark, B. J. Persson and B. O. Roos, *Theor. Chim. Acta*, 1991, **79**, 419.

- 27 K. Andersson, P. A. Malmqvist, B. O. Roos, A. J. Sadlej and K. Wolinski, *J. Phys. Chem.*, 1990, **94**, 5483–5488.
- 28 K. Andersson, P.-Å. Malmqvist and B. O. Roos, *J. Chem. Phys.*, 1992, **96**, 1218–1226.
- 29 D. Roca-Sanjuán, F. Aquilante and R. Lindh, *WIREs Comput. Mol. Sci.*, 2012, **2**, 585–603.
- 30 N. Forsberg and P.-Å. Malmqvist, *Chem. Phys. Lett.*, 1997, **274**, 196 – 204.
- 31 G. Ghigo, B. O. Roos and P.-Å. Malmqvist, *Chem. Phys. Lett.*, 2004, **396**, 142 – 149.
- 32 J. Finley, P.-Å. Malmqvist, B. O. Roos and L. Serrano-Andrés, *Chem. Phys. Lett.*, 1998, **288**, 299 – 306.
- 33 A. A. Granovsky, *J. Chem. Phys.*, 2011, **134**, 214113.
- 34 F. Aquilante, R. Lindh and T. Bondo Pedersen, *J. Chem. Phys.*, 2007, **127**, 114107.
- 35 F. Aquilante, T. B. Pedersen and R. Lindh, *J. Chem. Phys.*, 2007, **126**, 194106.
- 36 F. Aquilante, T. B. Pedersen, R. Lindh, B. O. Roos, A. Sánchez de Merás and H. Koch, *J. Chem. Phys.*, 2008, **129**, 024113.
- 37 F. Aquilante, P.-Å. Malmqvist, T. B. Pedersen, A. Ghosh and B. O. Roos, *J. Chem. Theory Comput.*, 2008, **4**, 694–702.
- 38 M. G. Delcey, L. Freitag, T. B. Pedersen, F. Aquilante, R. Lindh and L. González, *J. Chem. Phys.*, 2014, **140**, 174103.
- 39 M. G. Delcey, T. B. Pedersen, F. Aquilante and R. Lindh, *J. Chem. Phys.*, 2015, **143**, 044110.
- 40 I. Fdez. Galván, M. G. Delcey, T. B. Pedersen, F. Aquilante and R. Lindh, *Journal of Chemical Theory and Computation*, 2016, **12**, 3636–3653.
- 41 A. Bernhardsson, R. Lindh, J. Olsen and M. Fulscher, *Molecular Physics*, 1999, **96**, 617–628.
- 42 J. Segarra-Martí, M. Garavelli and F. Aquilante, *J. Chem. Theory Comput.*, 2015, **11**, 3772–3784.
- 43 J. Segarra-Martí, A. Francés-Monerris, D. Roca-Sanjuán and M. Merchán, *Molecules*, 2016, **21**, 1666.
- 44 A. J. Pepino, J. Segarra-Martí, A. Nenov, R. Improta and M. Garavelli, *The Journal of Physical Chemistry Letters*, 2017, **8**, 1777–1783.
- 45 A. J. Pepino, J. Segarra-Martí, A. Nenov, I. Rivalta, R. Improta and M. Garavelli, *Phys. Chem. Chem. Phys.*, 2018, **20**, 6877–6890.
- 46 M. J. Bearpark, M. A. Robb and H. B. Schlegel, *Chemical Physics Letters*, 1994, **223**, 269 – 274.
- 47 T. Shiozaki, W. Győrffy, P. Celani and H.-J. Werner, *The Journal of Chemical Physics*, 2011, **135**, 081106.
- 48 M. K. MacLeod and T. Shiozaki, *The Journal of Chemical Physics*, 2015, **142**, 051103.
- 49 B. Vlaisavljevich and T. Shiozaki, *Journal of Chemical Theory and Computation*, 2016, **12**, 3781–3787.
- 50 J. W. Park and T. Shiozaki, *Journal of Chemical Theory and Computation*, 2017, **13**, 2561–2570.
- 51 T. Shiozaki, *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2018, **8**, e1331.
- 52 B. G. Levine, C. Ko, J. Quenneville and T. J. Martínez, *Molecular Physics*, 2006, **104**, 1039–1051.
- 53 L. Serrano-Andrés, M. Merchán and R. Lindh, *J. Chem. Phys.*, 2005, **122**, 104107.
- 54 T. Shiozaki, C. Woywod and H.-J. Werner, *Phys. Chem. Chem. Phys.*, 2013, **15**, 262–269.
- 55 S. Sen and I. Schapiro, *Molecular Physics*, 2018, **116**, 2571–2582.
- 56 A. Giussani, J. Segarra-Martí, A. Nenov, I. Rivalta, A. Tolomelli, S. Mukamel and M. Garavelli, *Theoretical Chemistry Accounts*, 2016, **135**, 121.
- 57 J. Segarra-Martí, A. J. Pepino, A. Nenov, S. Mukamel, M. Garavelli and I. Rivalta, *Theoretical Chemistry Accounts*, 2018, **137**, 47.
- 58 Q. Li, A. Giussani, J. Segarra-Martí, A. Nenov, I. Rivalta, A. A. Voityuk, S. Mukamel, D. Roca-Sanjuán, M. Garavelli and L. Blancafort, *Chemistry – A European Journal*, 2016, **22**, 7497–7507.
- 59 J. Segarra-Martí, S. Mukamel, M. Garavelli, A. Nenov and I. Rivalta, *Top. Curr. Chem.*, 2018, **376**, 24.
- 60 A. Nenov, A. Giussani, B. P. Fingerhut, I. Rivalta, E. Dumont, S. Mukamel and M. Garavelli, *Phys. Chem. Chem. Phys.*, 2015, **17**, 30925–30936.
- 61 A. Nenov, R. Borrego-Varillas, A. Oriana, L. Ganzer, F. Segatta, I. Conti, J. Segarra-Martí, J. Omachi, M. Dapor, S. Taioli, C. Manzoni, S. Mukamel, G. Cerullo and M. Garavelli, *The Journal of Physical Chemistry Letters*, 2018, **9**, 1534–1541.
- 62 M. Richter and B. P. Fingerhut, *Journal of Chemical Theory and Computation*, 2016, **12**, 3284–3294.
- 63 M. Kowalewski, B. P. Fingerhut, K. E. Dorfman, K. Bennett and S. Mukamel, *Chemical Reviews*, 2017, **117**, 12165–12226.
- 64 P.-Å. Malmqvist and B. O. Roos, *Chem. Phys. Lett.*, 1989, **155**, 189–194.
- 65 J. Segarra-Martí, E. Zvereva, M. Marazzi, J. Brazard, E. Dumont, X. Assfeld, S. Haacke, M. Garavelli, A. Monari, J. Léonard and I. Rivalta, *Journal of Chemical Theory and Computation*, 2018, **14**, 2570–2585.
- 66 L. Martínez-Fernández, A. J. Pepino, J. Segarra-Martí, J. Jovaišaitė, I. Vaya, A. Nenov, D. Markovitsi, T. Gustavsson, A. Banyasz, M. Garavelli and R. Improta, *Journal of the American Chemical Society*, 2017, **139**, 7780–7791.
- 67 A.-R. Allouche, *Journal of Computational Chemistry*, 2011, **32**, 174–182.
- 68 G. Schaftenaar and J. Noordik, *Journal of Computer-Aided Molecular Design*, 2000, **14**, 123–134.
- 69 D. Dougherty, K. Wittel, J. Meeks and S. P. McGlynn, *Journal of the American Chemical Society*, 1976, **98**, 3815–3820.
- 70 G. Lauer, W. Schäfer and A. Schweig, *Tetrahedron Letters*, 1975, **16**, 3939 – 3942.
- 71 S. Urano, X. Yang and P. R. LeBreton, *Journal of Molecular Structure*, 1989, **214**, 315 – 328.
- 72 K. D. Fulfer, D. Hardy, A. A. Aguilar and E. D. Poliakoff, *The Journal of Chemical Physics*, 2015, **142**, 224310.
- 73 V. Orlov, A. Smirnov and Y. Varshavsky, *Tetrahedron Letters*, 1976, **17**, 4377 – 4378.

- 74 D. M. Close, *The Journal of Physical Chemistry A*, 2004, **108**, 10376–10379.
- 75 C. Wiebeler, V. Borin, A. V. Sanchez de Araújo, I. Schapiro and A. C. Borin, *Photochemistry and Photobiology*, 2017, **93**, 888–902.
- 76 P. M. Keane, M. Wojdyla, G. W. Doorley, G. W. Watson, I. P. Clark, G. M. Greetham, A. W. Parker, M. Towrie, J. M. Kelly and S. J. Quinn, *Journal of the American Chemical Society*, 2011, **133**, 4212–4215.
- 77 K. de La Harpe, F. R. Kohl, Y. Zhang and B. Kohler, *The Journal of Physical Chemistry A*, 2018, **122**, 2437–2444.