CrossMark

REGULAR ARTICLE

Spectroscopic fingerprints of DNA/RNA pyrimidine nucleobases in third-order nonlinear electronic spectra

Angelo Giussani[1] · Javier Segarra-Martí[1] · Artur Nenov[1] · Ivan Rivalta[2] ·
Alessandra Tolomelli[1] · Shaul Mukamel[3] · Marco Garavelli[1,2]

Abstract Accurate ab initio modeling of spectroscopic signals in nonlinear electronic spectra, such as bidimensional (2D) spectra, requires the computation of the electronic transitions induced by the incoming pump/probe pulses, resulting in a challenging calculation of many electronic excited states. A protocol is thus required to evaluate the variations of spectral properties, like transition energies and dipole moments, with the computational level, and to estimate the sensitivity of the spectra to these variations. Such a protocol is presented here within the framework of complete and restricted active space self-consistent field (CASSCF/RASSCF) theory and its second-order perturbation theory extensions (CASPT2/RASPT2). The electronic excited-state manifolds of pyrimidine nucleobases (thymine, uracil, and cytosine) are carefully characterized *in vacuo* employing high-level RAS(0,0|10,8|2,12)//SS-RASPT2 calculations. The results provide a reference data set that can be used for optimizing computational efforts and costs, as required for studying computationally more demanding multichromophoric systems (e.g., di- and oligonucleotides). The spectroscopic signatures of the 2D electronic spectrum of a perfectly stacked uracil–cytosine dimer model are characterized, and experimental setups are proposed that can resolve non-covalent interchromophoric interactions in canonical pyrimidine nucleobase-stacked dimers.

Keywords 2DES · CASSCF/CASPT2 · Nonlinear spectroscopy · Pyrimidines · Photophysics · Quantum chemistry · Electronic excited states · DNA/RNA

✉ Angelo Giussani
angelo.giussani2@unibo.it

✉ Javier Segarra-Martí
javier.segarra@unibo.it

✉ Marco Garavelli
marco.garavelli@unibo.it

1 Dipartimento di Chimica "G. Ciamician", Università di Bologna, Via Selmi 2, 40126 Bologna, Italy

2 Univ Lyon, Ens de Lyon, CNRS, Université Lyon 1, Laboratoire de Chimie UMR 5182, F-69342 Lyon, France

3 Department of Chemistry, University of California, Irvine, CA 92697-2025, USA

1 Introduction

Upon UV excitation, DNA can undergo different relaxation paths which can result in either a non-reactive, radiative (i.e., fluorescence–phosphorescence) [1, 2], or non-radiative (i.e., through conical intersections) [3–7], decays back to the absorbing ground-state structure or to a photochemical process leading to mutagenic and/or carcinogenic interstrand pyrimidine photoproducts (as cyclobutane pyrimidine dimers, pyrimidine–pyrimidone (6–4) adducts, and their Dewar isomers) [8–11]. The understanding and characterization of these phenomena and the factors that determine the fate of the excitation (as, e.g., the nucleobase sequence, the geometry at the time of light irradiation) constitute a fundamental elusive goal, due to the considerable complexity of DNA and the number of different correlated variables affecting its photoresponse [12].

🖄 Springer

The third-order nonlinear pump–probe technique has been extensively employed to study oligo- and poly-nucleotides with single and double strands [1, 2, 4–6, 13–17], providing precious insights on the photophysics and photochemistry of DNA. This technique employs a pair of ultrashort laser pulses to indirectly monitor the dynamics induced by the first (*pump*) pulse simultaneously in all electronic states under the pulse envelope through recording the spectral signatures of each state (i.e., absorption and emission signals) as a function of the time delay (i.e., waiting time) between the pump and the second (*probe*) pulse. The approach has however the drawback since it is unable to resolve concomitant de-excitation paths, due to the impossibility to disentangle concurrent signals coming out from different paths and/or states of a system simultaneously excited by the ultrashort (i.e., few femtoseconds, thus being several thousands cm^{-1} broad) pump pulse. The pump–probe technique is unable to provide unambiguous fingerprints specific for each possible photoinduced decay.

A more advanced third-order nonlinear optical spectroscopy, capable to overcome the pump–probe technique limitations, is the two-dimensional electronic spectroscopy (2DES) [18, 19]. Here, the sample interacts with three pulses: for a fixed time delay between the second and third pulses (t_2), which can be considered as the analogues of the waiting time between the pump and probe radiations in pump–probe spectroscopy, a single experiment is performed by scanning the time delays between the first and second pulse (t_1) and between the third pulse and a local oscillator (t_3), which heterodynes the field emitted by the sample in response to the perturbation by the three incident pulses and facilitates the extraction of the phase and time dependence of the field. Subsequently Fourier transforming the signals along t_1 and t_3, a two-dimensional Ω_1, Ω_3 spectrum is obtained, where all the electronic transitions involved in the three field–matter interactions are recorded with both high spectral and temporal resolution [20–22].

The interpretation of a spectrum strongly relies on the theoretical computation of properties like energies and transitions dipole moments that characterize the system electronic states and the possible electronic transitions among them. In this sense, spectroscopy constitutes a fine example of the interplay between experiment and theory. 2DES is no exception to this rule, being even more dependent on theoretical simulations due to the high number of signals that can be recorded [23–27]. In fact, the use of near-ultraviolet (NUV) and visible (Vis) radiation ideally allows the detection of all the electronic transitions among the electronic states placed up to around 9 eV from the ground state. Depending on the system size and properties, over a hundred excited states may be encountered in this energy window of 0–11 eV, resulting in a large amount of information that can potentially be recorded,

and whose interpretation requires a theoretical characterization of the system excited-state manifold. The latter task poses a considerable computational challenge due to the large number of excited states. The challenge is related to both the required computational effort (memory, CPU time, etc.) and the difficulty to describe a large number of excited states of different nature (e.g., ionic, covalent, doubly excited) on an equal footing. Using complete active space self-consistent field (CASSCF) [28] and complete active space second-order perturbation theory (CASPT2)-related methods [29–31], which is a common choice in photophysical and photochemical studies, the computation of highly accurate results for high-lying states has been documented to require the inclusion of different orbitals outside the valence space in the active space [32], whose number and character are not known a priori. Moreover, one has to face problems as valence-Rydberg mixing [33] and the unbalanced treatment of ionic and covalent states [24, 34], both potentially critical issues when computing many states. Despite the proven ability to provide a correct prediction of photophysical and photochemical data repeatedly shown by the CASPT2//CASSCF protocol for a variety of systems of very diverse size and nature [35–38], the particularly large number of state normally required for simulating 2DES experiments and the necessity of using less computationally expensive treatment, such as the RASPT2//RASSCF methods [39], imposes as a first task in the simulation of a two-dimensional spectrum the determination of reliable results on the excited states of the system under study [40]. The need of such data is reinforced by the almost total absence of both experimental and theoretical information on the high-lying excited-state manifold [41–43], at least for biological relevant molecules as nucleobases-related systems, on which the application of 2DES can bring a significant contribution to their characterization.

Here we provide the first step toward the simulation of two-dimensional electronic spectra of canonical pyrimidine nucleobases-containing systems that is for the reasons explained above, a thoughtful calibration of the electronic excited states that can be populated in 2DES employing NUV and Vis pulses for the thymine, uracil, and cytosine molecules. The obtained spectral signatures here characterized can also be employed in the analysis of widely used pump–probe spectra and other third-order nonlinear spectroscopic techniques. The NUV-pump/Vis- and NUV-pump/NUV-probe 2D electronic spectra of canonical nucleobase chromophores have been simulated (for $t_2 = 0$), highlighting their main features and reporting the principal differences among them. Finally, the spectra arising from dimeric and trimeric ideal aggregates of the molecules studied here are simulated through an "exciton Hamiltonian model."
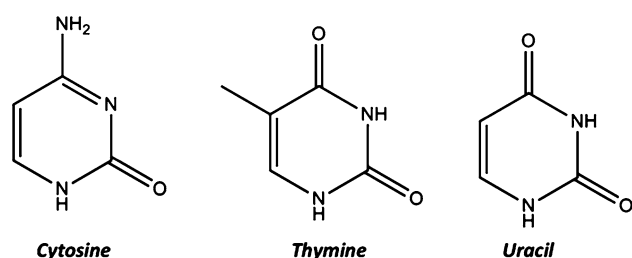
Fig. 1 Chemical structure of the DNA and RNA pyrimidine nucleobases studied in the present work

2 Computational details

A major goal of the present paper is the computation of high-level results on the excited states up to 11 eV for the canonical nucleobases thymine, uracil, and cytosine (see Fig. 1). Such an energy upper limit has been chosen to fully describe the states involved in NUV-pump/Vis-probe and NUV-pump/NUV-probe 2DES experiments. It also allows minimizing the contributions from the higher-energy region where the ionization threshold (~9 eV) might be crossed and where potential contributions from ionized species might arise [44], and their contribution is expected to be rather small above the ionization threshold and gradually increasing over the first few eV based on experimental data on olefins [45]. Our main objective is to obtain a highly accurate estimate for a large number of excited states. To this end we employ large basis sets, providing high flexibility that allows to correctly describe diffuse excited states, being affected more pronouncedly the higher-lying states. Additionally, a large active space beyond the valence space is used, in order to include as much correlation at the CASSCF level as possible in order to improve the reference function subsequently employed for the perturbation treatment at the CASPT2. By doing so, one often encounters problems associated with the appearance of Rydberg-type states that strongly affect the reference function and deteriorate the description of the valence states of interest. To address this, the computations have been performed following a recently presented two-step approach [46]. Firstly, the valence-Rydberg mixing problem, usually present when dealing with high-lying excited states, is treated through the optimization and subsequent deletion of the Rydberg orbitals involved in Rydberg states encountered in the explored energy region [47]. It must be stressed that these contributions and the overall importance of Rydberg-type states have been found to be negligible in solution [48], which is ultimately the media in which DNA/RNA systems are usually measured. In order to describe Rydberg orbitals, the ANO-L basis set [49] contracted to C,N,O[4s,3p,2d]/H[2s1p] employed here has been extended by including a set of uncontracted diffuse 8s 8p and 8d functions in the

center of the charge [47]. Secondly, once deleted the optimized Rydberg orbitals, starting from an active space composed by all the valence π orbitals included in the RAS2 subspace, a systematic increase of the active space is performed expanding the RAS3 subspace by the inclusion of four extra orbitals at each step until reaching convergence in the results. RAS2 and RAS3 refer, as usual, to the restricted active space self-consistent field (RASSCF) active subspaces, i.e., RAS(a,b|c,d|e,f), indicating, respectively, a, the maximum number of holes allowed in the RAS1 subspace; b, the number of orbitals in RAS1; c, the number of electrons included in the RAS2 subspace; d, the number of orbitals in RAS2; e, the maximum number of electron that can be distributed among the RAS3 orbitals; f, the number of orbitals composing the RAS3 subspace [32]. In all cases a maximum number of two electrons were allowed in RAS3. Lone pair n orbitals have been excluded from the active space, in line with previous work on adenine [46], given their relatively dark character (with an exceedingly small oscillator strength) and their blueshifts observed in solution, which will indicate a minor involvement of these states in the range of interest [50–52].

C_s symmetry has been imposed for the three molecules in all the computations here presented. CASSCF ground-state geometries optimizations have been operated including in the active space the most important valence π orbitals (CAS(6,6)). The excited-state wave functions and the transition dipole moments [53] characterizing the different electronic transitions have been computed at the CASSCF and RASSCF levels [54, 55], following the state-averaged (SA) procedure for the optimization of the orbitals. The energies have been obtained using the second-order perturbation theory approach in its single-state variant, SS-CASPT2/RASPT2, which includes the dynamic correlation effects missing in the CASSCF and RASSCF description. Within the CASPT2 calculations, the IPEA shift has been set equal to 0.0 [56], and an imaginary level-shift correction of 0.2 *a.u.* has been used in order to minimize the effects of possible intruder states [57]. The CASPT2 standard zeroth-order Hamiltonian has been employed as originally implemented. The core orbitals have been frozen in the CASPT2 calculations. Such a CASPT2 approach has been validated during the last decades in many different studies on organic molecules, providing a correct prediction, description, and interpretation of the experimental photophysical and photochemical data. The Cholesky decomposition has been used to speed up the calculation of two-electron integrals [58–62]. All quantum–mechanical calculations have been performed with the MOLCAS package [63–65]. Further details on the approach employed are reported in the Supporting Information. Applying this computational protocol, the final outcomes for thymine, uracil, and cytosine have been obtained at the SA-29-RAS(0,0|10,8|2,12)/

SS-RASPT2 level of theory by deleting 29, 23, and 22 optimized Rydberg orbitals, respectively. In the state-averaging procedure, we have included more states than those discussed in the results section (covering a much higher energetic window at the CASSCF level), in order to guarantee that no states with pronounced energetic stabilization upon dynamic correlation corrections were omitted.

The calculated transition dipole moments and energies have been used to simulate two-dimensional electronic spectra employing the Spectron 2.7 program [18], which computes quasi-absorptive 2DES maps by the sum-over-state approach in the dipole approximation [23]. Details on the working equations are presented in the Supporting Information (SI). A static picture is adopted in order to describe the coherence dynamics during time t_1 and t_3, thus assuming that the vibrational dynamics in the system is slower than these timescales. The waiting time t_2 between the pump pulse pair and the probe pulse was set to zero. Using the static approximation, only computations at the Frank–Condon (FC) point are needed for simulate spectra at $t_2 = 0$. Although protocols for computing 2DES that consider spectral diffusion and non-adiabatic effects on an ultrashort timescale have been recently proposed [66, 67], the rather crude static approximation is sufficient for our aim to study the effect of various computational parameters on the shape of the 2D maps and to qualitatively individuate the main signals characterizing the spectra of canonical pyrimidine nucleobases at the Franck–Condon region. Our aim is to have a computational reference of the purely electronic contributions to the excitation energies of all accessible states in order to derive subsequent protocols for their efficient computation. The fine lineshape given by the ultrashort vibrational dynamics as well as by inhomogeneities present in the sample may be addressed at a later stage [66], once the best suited computational protocol for the electronic structure is established. A constant line broadening of 200 cm^{-1} was used throughout. All calculated signals use the all-parallel $xxxx$ pulse polarization configurations and are plotted on a logarithmic scale. Ground-state bleaching (GSB) and stimulated emission (SE) contributions appear as negative (blue) peaks, and excited-state absorptions (ESAs) appear as positive (red) peaks in the 2D spectra.

Bidimensional electronic spectra were generated for simple models of unstacked and stacked uracil–cytosine dimer. The unstacked (and, thus, non-interacting) dimer was generated by combining the separate spectra of both monomers. The coupled dimer was computed using an approximate Frenkel exciton model [68] introducing exciton couplings between the first excited state of uracil and the two lowest lying excited states of cytosine. The couplings were computed in the point dipole approximation [68]. The two bases were assumed to be in a perfectly

stacked configuration with an interbase distance of 4 Å. It was further assumed that the interaction does not affect the energies of the higher-lying states. Two mixed doubly excited states corresponding to the simultaneous excitation of different excitons were added to the manifold of higher-lying states, and a quartic coupling [18] of −1000 cm^{-1} was applied, implying that double excitations on neighboring sites have lower energy compared to the sum of energies of the two single excitations. Further details of the exciton model employed are given in SI (Fig. S7).

3 Results

3.1 Isolated canonical pyrimidine nucleobases

3.1.1 Thymine

Table 1 reports the thymine excited states up to 11 eV based on SA-29-RAS(0,0|10,8|2,12)/SS-RASPT2 computation. Such level of theory is used as the reference since it is in agreement with the benchmarking following the protocol described in Sect. 2, and the SA-29-RAS(0,0|10,8|2,12)/SS-RASPT2 calculation provides converged results with respect to the systematic increase of the RAS3 subspace [32, 46]. The RAS wave functions defining the nature of the computed states are also reported in Table 1. As expected, the lower-states wave functions are mainly composed by one or two configuration state functions, while at higher energies different configurations, among which double excited configuration state functions, contribute to the excited-state wave functions.

A gradual enlargement of the RAS3 subspace is carried out, and a comparison of the electronic structures is drawn utilizing the NUV-pump/Vis-probe (in the 10,000–30,000 cm^{-1} range) and NUV-pump/NUV-probe (in the 30,000–46,000 cm^{-1} range) two-dimensional electronic spectra for $t_2 = 0$ (see Fig. 2). This kind of graphical representation is appealing since the positions of the peaks and their intensities are proportional to the computed vertical excitation energies and transitions dipole moments, respectively. As clearly shown in Fig. 2, both the NUV-pump/Vis-probe and NUV-pump/NUV-probe spectra computed at the SA-29-RAS(0,0|10,8|2,12)/SS-RASPT2 level of theory display almost no changes with respect to the same spectra computed using the SA-29-RAS(0,0|10,8|2,8)/SS-RASPT2 results. The convergence of the SA-29-RAS(0,0|10,8|2,12)/SS-RASPT2 outcomes is further proven in Table S7, where the excitation energies of the computed states with the different active spaces are tabulated (see SI).

Further examination of the NUV-pump/Vis-probe spectra reported in Fig. 2 (lower panel) is possible showing that the Ω_3 values at which the signals appear as

Table 1 Vertical excitation energies (E_{VA}, eV), transition dipole moments (TDM, *a.u.*), leading wave function configurations, and corresponding coefficients for the different thymine excited states calculated at the SA-29-RAS(0,0|10,8|2,12)/SS-RASPT2 level of theory with 29 Rydberg-like deleted orbitals (see Sect. 2)

State	E_{VA} (eV)	TDM GS- > Sn	TDM $\pi\pi^*$- > Sn	Main configuration	Coeff.	Weight
2	5.00	1.18	–	H- > L	−0.86	0.74
3	6.20	0.84	0.65	H-1- > L	0.82	0.67
4	6.51	1.04	0.48	H- > L+1	0.81	0.66
5	7.32	2.23	0.33	H-1- > L+1	−0.81	0.66
6	8.06	1.38	0.55	H-2- > L	−0.65	0.42
				H-3- > L	−0.37	0.14
				H- > L	−0.29	0.08
7	8.42	0.19	0.96	H-3- > L	0.43	0.19
				H- > L+1, L	0.38	0.15
				H-2- > L	−0.37	0.14
				H-1, H- > L	−0.35	0.12
				H- > L+2	0.29	0.08
8	8.95	0.49	0.57	H-2- > L+1	0.58	0.33
				H- > L+2	0.32	0.10
				H- > L+1, L	0.30	0.09
				H-3- > L	−0.24	0.06
9	9.08	0.40	1.16	H- > L	−0.59	0.35
				H- > L+2	−0.38	0.15
				H-3- > L	0.28	0.08
				H-1, H- > L+1, L	0.22	0.05
10	9.45	0.63	0.82	H- > L+2	−0.44	0.20
				H-2- > L+1	0.36	0.13
				H-1, H- > L	−0.32	0.10
				H-1, H- > L+1, L	0.26	0.07
				H-4- > L	−0.25	0.06
				H- > L	0.24	0.06
11	9.75	0.20	0.57	H-1- > L+2	0.47	0.22
				H-1, H- > L+1, L	0.45	0.20
				H-3- > L+1	−0.30	0.09
12	9.81	0.30	0.10	H-3- > L	0.36	0.13
				H-2- > L+1	0.33	0.11
				H-1- > L+2	0.29	0.08
				H-1, H- > L	0.28	0.08
				H-3, H- > L	0.26	0.07
				H-1, H-2– > L	0.24	0.06
13	10.39	0.15	0.54	H-3- > L+1	−0.57	0.33
				H-4- > L	0.39	0.15
14	10.77	0.12	0.78	H-4- > L	−0.44	0.19
				H-3- > L+1	−0.37	0.14
				H-1- > L+2	−0.34	0.11
				H-1, H- > L	0.25	0.06
				H-3- > L	0.23	0.05

well as their intensities are little affected by the level of theory; meanwhile, the position of the trace is constantly moving to higher values of Ω_1, showing the most significant variation passing from the RAS(0,0|10,8|2,4) to the RAS(0,0|10,8|2,8) active space (blueshift equal to 721 cm^{-1}). This reflects a similar increase in the vertical

excitation energies experienced by the states excited by the probe radiation (i.e., states 4, 5, 6, and 7) and the pump reachable state (state 2) that consequently causes the preservation of the energy gap among them, leading to the almost constant position of the peaks along Ω_3. The described behavior could be an indicator of a similar
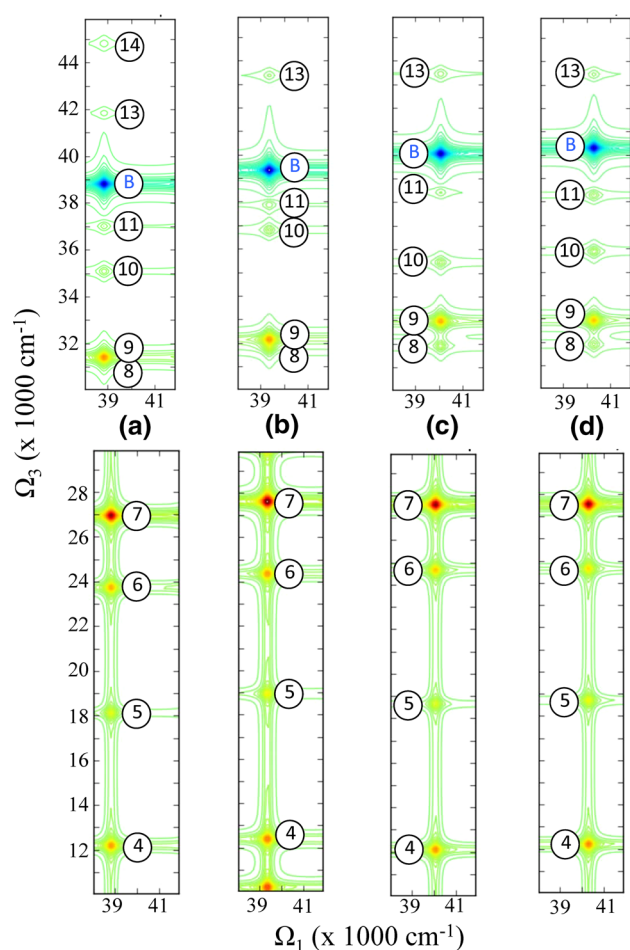
Fig. 2 2D NUV-pump/Vis-probe (*lower panel*) and 2D NUV-pump/NUV-probe (*upper panel*) spectra of thymine in gas phase for different active space sizes: **a** CAS(10,8); **b** RAS(0,0|10,8|2,4); **c** RAS(0,0|10,8|2,8); **d** RAS(0,0|10,8|2,12). Peak numbering follows the excited-state assignment displayed in Table 1

nature of the mentioned states. More pronounced effects are appreciable comparing the NUV-pump/NUV-probe spectra at the different level of theory (see higher panels in Fig. 2). All ESAs exhibit a non-uniform blueshift upon active space enlargement causing the splitting of peaks 8 and 9, which appear to be overlapped in the least correlated cases (Fig. 2a, b). On the other hand, peak 10 undergoes a significant blueshift in the RAS(0,0|10,8|2,4) and then a gradual redshift for the higher-level spectra, resulting in a ~500 cm^{-1} shift from the CAS(10,8) to the RAS(0,0|10,8|2,12).

Using the reference SA-29-RAS(0,0|10,8|2,12)/SS-RASPT2 results, we can conclude that the Franck–Condon thymine NUV-pump/Vis-probe spectrum at $t_2 = 0$ is characterized by four ESA peaks, associated with electronic promotion from state 2 into states 4, 5, 6, and 7, respectively. All those peaks are significantly intense, with signal

7 being the most intense one, and energetically well separated, at least according to the static approximation here assumed. Regarding the analogous NUV-pump/NUV-probe spectrum, five ESA peaks due to excitation of states 8, 9, 10, 11, and 13 plus the expected GSB signal are present on it. Among them, peak 9 and the GSB/SE appear as the most intense features of the 2D map and can be consequently considered as the clearest fingerprints of the Franck–Condon thymine NUV-pump/NUV-probe spectrum. The intense signals resulting from the electron promotions between the initially accessed ππ* and states 6, 7, and 9 appear to be well separated and in a convenient probing window to monitor the specific ππ* excited-state dynamics with transient absorption spectroscopic techniques.

Finally, it is important to compare the obtained results with the available theoretical and experimental data. Previous computations have focused on the lower excited states of the system, providing information on the four lowest ππ* electronic states. CASPT2 results have been presented by the Roos [69–71] and Thiel groups [72, 73]. Szalay et al. studied the excited state of the molecule using a variety of linear response coupled cluster methods, based on CCSD(T) as well as CC3 approximations [41–43]. The corresponding vertical excitation energies are reported in Table S4 (see Supporting Information), comparison with our results shows a reasonable overall agreement, and the present data tend to be slightly blue-shifted when compared with the previous CASPT2 computations while displaying in general lower values with respect to couple cluster results. Those variations can be attributed to: (a) the different levels of theory employed, (b) the much higher number of states here computed, which can significantly affect the outcomes of a CASSCF computation using state-average procedure, and (c) the slightly different geometries used in the cited papers (see Figure S4), which has been recently shown to pronouncedly influence the energetic position of the excited-state manifold by up to 2000 cm^{-1} [46]. Experimentally, between 3 and 4 lower excited states have been described for thymine [74, 75]. Working with the bare molecule *in vacuo*, the most appropriate comparison for our results would be with gas-phase data, although the operated elimination of Rydberg orbitals prevents to compare with gas-phase spectra for regions in which Rydberg states contribute. To the authors' knowledge, only one paper displays experimental values for the vertical excitation energies of states higher than the S_1 in gas phase. In this study, performed by Abouaf et al. employing electronic energy loss spectroscopy [74], three bands associated with ππ* singlet excited-state transitions have been characterized at 4.95, 6.2, and 7.4 eV, respectively. The energetic regions at which those bands appear are in agreement with the vertical excitation energies here computed for states 2, 3, 4, and 5 (i.e., 5.00, 6.20, 6.51, and 7.31). Gustavsson et al. on the

Table 2 Vertical excitation energies (E_{VA}, eV), transition dipole moments (TDM, *a.u.*), leading wave function configurations, and corresponding coefficients for the different uracil excited states calculated at the SA-29-RAS(0,0|10,8|2,12)/SS-RASPT2 level of theory with 23 Rydberg-like deleted orbitals (see Sect. 2)

State	E_{VA} (eV)	TDM GS- > S_n	TDM $\pi\pi$*- > S_n	Main configuration	Coeff.	Weight
2	5.20	1.17	–	H- > L	0.85	0.73
3	6.18	0.54	0.85	H-1- > L	0.81	0.66
4	6.55	1.07	0.43	H- > L + 1	−0.77	0.60
5	7.39	1.99	0.47	H-1- > L + 1	0.77	0.61
6	8.42	0.52	1.03	H- > L + 2	−0.37	0.14
				H-2- > L	−0.41	0.17
				H-3- > L	−0.52	0.27
				H − > L, L + 1	0.34	0.12
7	8.49	1.24	0.15	H-2- > L	−0.52	0.27
				H-3- > L	0.48	0.23
8	8.98	0.14	0.17	H- > L + 2	0.43	0.19
				H-3- > L + 1	0.42	0.18
9	9.35	0.43	0.75	H- > L + 2	0.46	0.21
				H − > L	−0.36	0.13
10	9.55	0.11	0.75	H-4- > L	0.56	0.32
				H-3- > L + 1	−0.41	0.17
				H − > L	−0.43	0.19
11	9.67	0.79	0.64	H-1- > L + 2	−0.60	0.35
12	9.88	0.81	0.75	H-3- > L + 1	0.35	0.12
				H, H-1 − > L, L + 1	−0.33	0.11
13	10.08	0.50	0.27	H-2- > L + 1	0.51	0.26
				H-4- > L	−0.48	0.23
14	10.37	0.13	0.51	H-3- > L + 2	−0.37	0.14
				H-4- > L + 1	0.49	0.24
15	10.39	0.34	0.16	H-2- > L + 1	−0.50	0.25
				H-2- > L	0.33	0.11
				H, H-1 − > L	−0.35	0.12
16	10.59	0.13	0.83	H-3- > L + 2	−0.46	0.21

other hand reported the maximum of the absorption in solution for the first band placing it at 4.7 eV [75], in line with previous theoretical analyses that predict pronounced redshifts in strongly ionic states like the $\pi\pi$* state in thymine [51, 52]. For further comparisons, different experimental results in solutions can be found in previous works by Roos et al. [69–71] and in the references therein.

3.1.2 Uracil

Table 2 summarizes the characterized excited-state reference results up to 11 eV obtained for uracil based on SA-29-RAS(0,0|10,8|2,12)/SS-RASPT2 computations. As shown in Fig. 3 and analogously to thymine, both NUV-pump/Vis-probe and NUV-pump/NUV-probe reference spectra display slight changes with respect to the same spectra computed using the SA-29-RAS(0,0|10,8|2,8)/SS-RASPT2 results. The convergence of the reference computation is again assessed in detail in SI (see Table S8), where the excitation energies of the computed states with

the different active spaces are displayed and compared with respect to previous theoretical studies employing a variety of theoretical techniques, in order to assess the methodology employed (see SI Table S5).

The NUV-pump/Vis-probe spectra reported in Fig. 3 show virtually identical Ω_3 values and intensities associated with the transitions depicted, whereas Ω_1 appears to be more affected by the increase of the active space, showing a gradual blueshift of the trace representing the S_1 excited state (state 2 in Table 2) by going from RAS(0,0|10,8|0,0) to RAS(0,0|10,8|2,12) totaling in a ~2000 cm^{-1} shift. An analogous blueshift occurs for all signals present in the NUV-pump/Vis-probe window (lower panels in Fig. 3), and as a result they remain in identical position along Ω_3 despite being gradually moved to higher energies (states 4, 5, 6, and 7), as it occurs in thymine. This behavior is again attributed to the similar nature of the low-lying excited states accessed by the system and allows explaining their relative stability toward active space enlargement. On the other hand, the NUV-pump/NUV-probe window appears
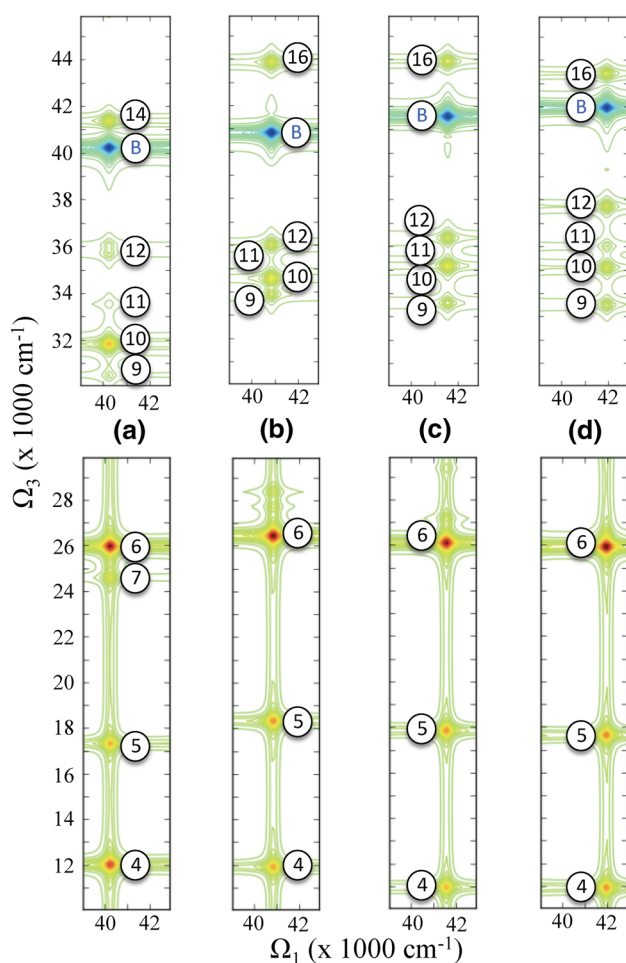
Fig. 3 2D NUV-pump/Vis-probe (*lower panel*) and 2D NUV-pump/NUV-probe (*upper panel*) spectra of uracil in gas phase for different active space sizes: **a** CAS(10,8); **b** RAS(0,0|10,8|2,4); **c** RAS(0,0|10,8|2,8); **d** RAS(0,0|10,8|2,12). Peak numbering follows the excited-state assignment displayed in Table 2

to be more affected in comparison (see higher panels in Fig. 3). A clear dependence in the size of the active space can be readily recognized, by collapsing signals 9 and 10 for the smallest CAS(10,8) space into a degenerated peak while showing a splitting of the two excited states for the more correlated RAS(0,0|10,8|2,4) computations. Both RAS(0,0|10,8|2,4) and RAS(0,0|10,8|2,8) (higher panels in Fig. 3b, c) show degenerated signals (states 11 and 12) at ~36,000 cm^{-1} along Ω_3, which decouple at RAS(0,0|10,8|2,12) yielding a clear ~1500 cm^{-1} splitting between them within the static approach here employed. In general lines, states 9, 10, and 12 blue-shift gradually by increasing the active space, totaling shifts of ~2000, ~3000, and ~1000 cm^{-1} by going from CAS(10,8) to RAS(0,0|10,8|2,12), respectively, while state 14 red-shifts being completely covered by the GSB/SE signal in the more correlated cases. An additional peak emerges at the

high-energy window when the active space is increased beyond the valence orbitals and is attributed to state 16 (see Table 2), which is gradually red-shifted upon active space enlargement being placed at ~43,000 cm^{-1} for the reference spectrum.

Based on the reference SA-29-RAS(0,0|10,8|2,12)/SS-RASPT2 results, we can conclude that the 2DES of uracil at the Franck–Condon region employing NUV-pump/Vis-probe setup at $t_2 = 0$ is characterized by three ESA signals, associated with electronic promotions arising from state 2 and leading to states 4, 5, and 6. All signals appear to be very intense, highlighting the transition to state 6 as the most intense (*cf.* Table 2), and appear to be energetically well separated within the static approach here employed. It is worth noting that signal 6 features strong contributions from the doubly excited HOMO to LUMO (H – > L) configuration, analogously to the aforementioned signal 7 in thymine that is also featured as the most intense transition. Signal 7 appears in the same energetic window as peak 6, and it is obscured due to its exceedingly small transition dipole moment. Regarding the NUV-pump/NUV-probe spectrum, four ESA peaks, due to the excitation of states 9, 10, 12, and 16, are present, as well as the signal associated with the GSB/SE. Among those, peaks 9, 10, 12 and the GSB/SE appear to be among the most intense and well-separated features of the 2D map and can be consequently considered as the clearest fingerprints of the Franck–Condon uracil NUV-pump/NUV-probe spectrum, yielding very similar conclusions to those previously obtained for thymine. Signal 14 appears to be completely covered by the more intense GSB/SE, and its contributions are expected to be negligible. An additional very intense signal appearing in the high-energy window, referred to the transition to state 16, appears to be another fingerprint unique to uracil that may be employed for its unequivocal characterization being the only signal appearing above the GSB/SE.

Our results were then compared to the available theoretical and experimental data. Most theoretical studies have focused on the lower-lying excited-state manifold, providing estimates for the four lowest lying $\pi\pi^*$ electronic excited states. CASPT2 results have been put forth in the literature from Roos et al., employing a similar computational scheme as the one here employed making use of a smaller set of extra diffuse functions [70], whereas Thiel et al. employed more common Dunning-type basis sets [72, 73]. Single-reference linear response coupled cluster estimates have also been reported [76]. All values have been compared to the present findings and can be seen in Table S5 (see SI). The results here obtained tend to be slightly blue-shifted with respect to the previous CASPT2 values reported in the literature, likely due to the large state-averaging procedure employed to cover the whole NUV and Vis energy windows and due to employing large diffuse

extra coefficients in the basis sets, while they tend to be red-shifted compared to high-level CCSD(T) estimates. The usage of different geometries (Figure S5 in the SI) is also expected to profoundly affect the energetic position of the excited-state manifold, as it has been shown elsewhere [46], slightly distancing our estimates from those computed with different geometrical arrangements, like those previously discussed in the literature. Experimentally, three bands have been registered for $\pi\pi^*$ transitions *in vacuo* at 5.1, 6.0, and 6.6 eV, respectively. These are on average placed within a tenth of an eV of the values here obtained for states 2, 3, and 4, at 5.2, 6.18, and 6.55 eV, respectively, being in agreement with the only experimental evidence known to the authors related to high-lying excited states in gas-phase uracil [77]. Recent work by Gustavsson et al. [75] reported the aqueous absorption maximum for the first band, placing it at 4.79 eV, again in agreement as shown above for thymine with what has been here reported given the strong redshift expected due to solvation effects [51]. It is worth noting that the first $n\pi^*$ band, placed at 4.8 eV, has been omitted due to its dark character and therefore its lack of involvement in the 2DES of uracil. The systematic removal of Rydberg-type orbitals in order to improve the description of the valence excited states prevents us to compare our results with high-energy Rydberg studies and their excitation energies. A more in-depth account on the comparison between previous theoretical and experimental evidence is available elsewhere [69–71].

3.1.3 Cytosine

In Table 3, the cytosine excited states up to 11 eV based on SA-29-RAS(0,0|10,8|2,12)/SS-RASPT2 computations are reported. This level of theory provides converged results with respect to the systematic increase of the RAS3 subspace also for cytosine, reaching a point where further active space extension provides negligible changes in the vertical excitation energies (see Table S9 in the SI). Figure 4 shows the NUV-pump/Vis-probe and NUV-pump/NUV-probe reference spectra and how the signals fluctuate by adding correlation in the form of an increase in the active space. These fluctuations, larger than those registered for the other pyrimidine nucleobases studied, may be related to the larger amount of excited states spanned by cytosine within the first 11 eV with respect to the other two pyrimidines. In order to assess the methodology employed, the excitation energies calculated with different active spaces have been compared with respect to previous theoretical studies (Table S6 in the SI) based on several theoretical techniques.

The NUV-pump/Vis-probe spectra of cytosine show similar Ω_3 for the most correlated cases and similar peak intensities, whereas Ω_1 is again heavily affected by the increase of the active space, showing a progressive blueshift of ~2000 cm^{-1} between the most and the least correlated cases (Fig. 4a, d), for the two intense traces found along Ω_1, corresponding to the lowest lying ($\pi\pi^*$, state 2) and second lowest lying and more intense excited state ($\pi\pi^*_{c,2}$, state 3). The first trace, which appears to gather a relatively small intensity in the 2D spectra, comprises analogous fingerprints as those previously described for uracil and thymine, where states 5, 7, and 8 provide monomer-specific fingerprints that could be monitored with an adequate pump–probe setup where their contributions could be enhanced. Nevertheless, the main contributions in cytosine arise from the second trace corresponding to the $\pi\pi^*_{c,2}$ state that will provide appreciable fingerprints while obscuring the weaker contributions arising from the $\pi\pi^*$ state. The second trace along Ω_1 represents the transitions connecting the $\pi\pi^*_{c,2}$ state (state 3, depicting mainly a H-1-> L transition, *cf.* Table 3) to the rest of the high-lying singlet excited-state manifold. This second trace is placed close to the energy regions where uracil and thymine absorb, being crucial for interchromophoric interactions given its energetic position as it will be shown over the next section, and appears to be more strongly affected by the gradual increase of the active space as compared to thymine and uracil. Signals 8, 9, and 10 appear to be accessible from the $\pi\pi^*_{c,2}$ state, the former being relatively intense at the CAS(10,8) level, blue-shifting and reducing its intensity for the more correlated cases. This trend is similar for other signals that are also strongly blue-shifted upon active space enlargement yielding shifts of ~3500 and ~2000 cm^{-1} for states 9 and 10, respectively, by going from CAS(10,8) to RAS(0,0|10,8|2,12). Signal 12 appears to be the most intense contribution, closely followed by peak 10, representing the two main ESAs characterizing the spectrum (*cf.* Table 3). A significant shift can still be observed for the $\pi\pi^*_{c,2}$ state (position of the trace along Ω_1) by going from RAS(0,0|10,8|2,8) to RAS(0,0|10,8|2,12) so its convergence cannot be completely ensured in the present study.

The NUV-pump/NUV-probe window appears to be more strongly affected by the steady increase of the active space than what has been previously showed for thymine and uracil (see higher panel of Fig. 4), yielding not just quantitative but also qualitative changes in the spectra. A clear dependence in the size of the active space can be readily recognized, influencing both energetic position and transition dipole moments describing the electron promotions between the $\pi\pi^*$ and $\pi\pi^*_{c,2}$ states and the rest of the singlet excited-state manifold. The first trace, corresponding to the $\pi\pi^*$ state (state 2, *cf.* Table 3), is again completely obscured by the more intense $\pi\pi^*_{c,2}$ state, contributing to the spectra with its GSB/SE signal. The second trace shows some weak signals at the CAS(10,8) level for states 17, 18, and

Table 3 Vertical excitation energies (E_{VA}, eV), transition dipole moments (TDM, *a.u.*), leading wave function configurations, and corresponding coefficients for the different cytosine excited states calculated at the SA-29-RAS(0,0|10,8|2,12)/SS-RASPT2 level of theory with 22 Rydberg-like deleted orbitals (see Sect. 2)

State	E_{VA} (eV)	TDM GS-> S_n	TDM $\pi\pi^*$-> S_n	TDM $\pi\pi^*_{c,2}$-> S_n	Main configuration	Coeff.	Weight
2	4.49	0.62	–	–	H-> L	−0.89	0.79
3	5.42	1.03	0.83	–	H-1-> L	−0.77	0.60
4	6.31	0.41	0.17	0.87	H-> L + 1	0.79	0.62
					H-> L + 2	−0.41	0.16
5	6.42	2.17	0.42	0.48	H-> L + 2	−0.72	0.52
					H-> L + 3	−0.34	0.12
6	6.86	0.29	0.15	0.10	H-1-> L + 1	−0.74	0.54
					H-1-> L + 2	0.48	0.23
7	6.92	1.20	0.84	0.51	H-1-> L + 2	0.56	0.31
					H-1-> L + 3	0.43	0.18
					H-2-> L	0.36	0.13
8	7.97	1.14	0.74	0.58	H-2-> L	−0.61	0.37
					H-2-> L + 2	0.29	0.09
					H -> L, L + 1	0.23	0.05
9	8.67	1.02	0.23	0.30	H-3-> L	0.62	0.39
					H-1 -> L	0.33	0.11
					H, H-1 -> L	0.26	0.07
10	8.76	0.37	0.04	0.82	H-2-> L + 3	0.40	0.15
					H -> L	−0.39	0.15
11	8.80	0.69	0.42	0.51	H-2-> L + 3	−0.54	0.29
					H-2-> L + 2	0.37	0.14
					H -> L, L + 1	0.26	0.07
12	8.88	0.39	0.14	0.98	H-2-> L + 1	0.36	0.13
					H-3-> L	−0.32	0.10
					H, H-1 -> L	0.24	0.06
13	9.54	0.25	0.76	0.20	H -> L	0.40	0.16
					H-4-> L	−0.35	0.12
					H, H-1 -> L	0.28	0.08
14	10.25	0.21	0.82	0.36	H, H-1 -> L	0.30	0.09
					H -> L, L + 1	0.29	0.08
15	10.33	0.18	0.26	0.14	H -> L, L + 2	0.44	0.19
					H -> L, L + 1	−0.33	0.11
					H, H-1 -> L, L + 1	0.23	0.05
16	10.33	0.17	0.64	0.22	H-3-> L + 2	−0.66	0.44
					H-1 -> L, L + 2	−0.30	0.09
					H-3-> L + 1	0.28	0.08
17	10.36	0.25	0.50	0.19	H, H-1 -> L, L + 1	−0.35	0.12
					H, H-1 -> L, L + 2	0.34	0.12
18	10.44	0.40	0.19	0.25	H-3-> L + 1	0.47	0.22
					H -> L, L + 1	0.29	0.08
					H-3-> L + 2	0.26	0.07
					H-4-> L	−0.23	0.05
19	10.46	0.27	0.61	0.30	H, H-1 -> L, L + 2	−0.47	0.22
					H -> L, L + 2	0.26	0.07

Table 3 continued

State	E_{VA} (eV)	TDM GS-> S_n	TDM $\pi\pi$*-> S_n	TDM $\pi\pi$*$_{c,2}$-> S_n	Main configuration	Coeff.	Weight
20	10.75	0.10	1.18	0.19	H-3 -> L + 1	0.30	0.09
					H -> L, L + 1	0.27	0.07
					H-4- > L	0.25	0.06
					H, H-1 -> L + 1	−0.23	0.05
					H, H-1 -> L	0.23	0.05
21	10.81	0.43	0.54	0.80	H, H-1- > L, L + 2	−0.28	0.08
					H, H-1 -> L, L + 1	−0.27	0.07
					H-4- > L	−0.26	0.07
					H, H-1 -> L	−0.23	0.05
22	11.11	0.17	0.07	0.59	H, H-1 -> L, L + 2	0.51	0.26
					H, H-1 -> L, L + 1	−0.35	0.12
					H -> L, L + 2	0.26	0.07

Fig. 4 2D NUV-pump/Vis-probe (*lower panel*) and 2D NUV-pump/NUV-probe (*upper panel*) spectra of cytosine in gas phase for different active space sizes: **a** CAS(10,8); **b** RAS(0,0|10,8|2,4); **c** RAS(0,0|10,8|2,8); **d** RAS(0,0|10,8|2,12). Peak numbering follows the excited-state assignment displayed in Table 3
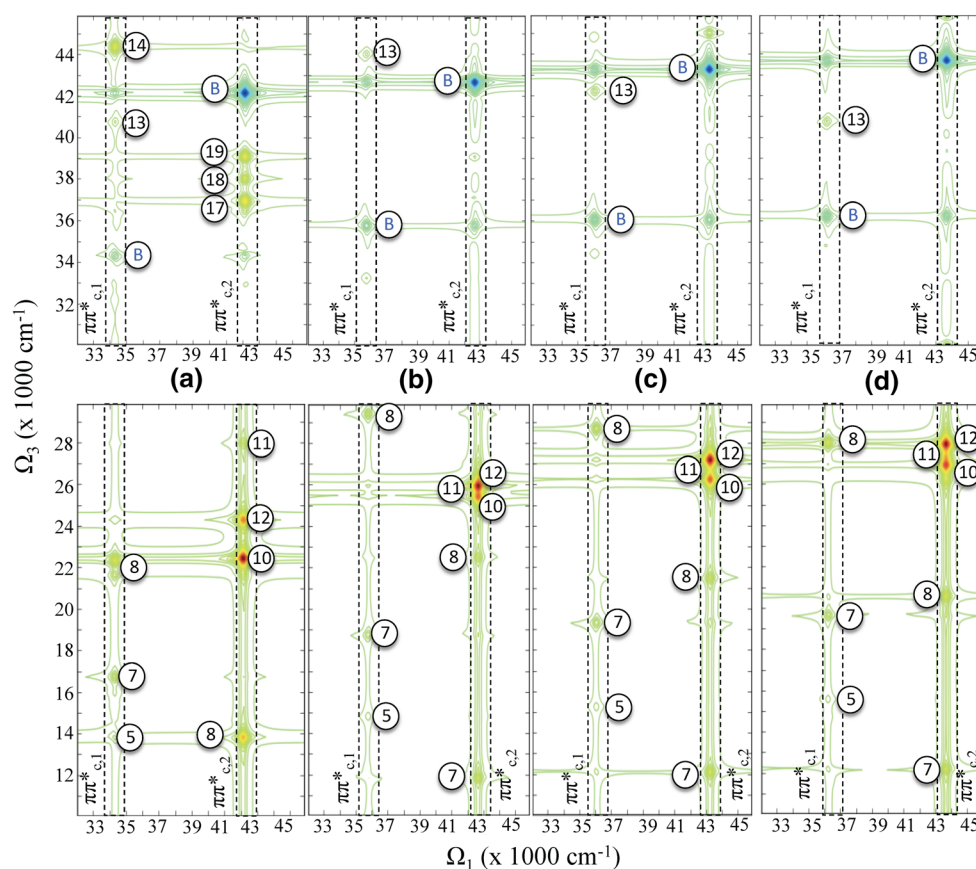
19, but that are completely covered due to the strong intensity shown by the GSB/SE signal upon active space enlargement, showing how the extra correlation added by extending the active space can play a role not only just on the energetics of the system but also on other related properties like its transition dipole moments. For the more correlated cases (Fig. 4b, c, d), only the GSB/SE signal can be discerned, yielding a completely ESA empty spectral window.

Based on the reference SA-29-RAS(0,0|10,8|2,12)/SS-RASPT2 results, we conclude that the 2DES of cytosine at the Franck–Condon region employing a NUV-pump/Vis-probe setup at $t_2 = 0$ is characterized by two bright ESA signals in the second trace along Ω_1, associated with electronic transitions connecting state 3 ($\pi\pi$*$_{c,2}$) to states 10 and 12, the latter being more intense contribution (*cf.* Table 3). The NUV-pump/NUV-probe results reveal a

relatively empty spectrum, showing no intense contributions associated with ESAs and just the GSB/SE signals of the states 2 and 3, especially the latter, are particularly intense. As will be shown over the next section, the absence of strong ESA signals in the trace of $\pi\pi^*_{c,2}$ indicates it as a potential region to probe the mixed excited-state signals when cytosine interacts with uracil, due to their proximity in energy enhancing their interaction.

Previous computations had made use of the CASPT2 to obtain the lowest lying singlet excited states of cytosine [71]. Single-reference linear response coupled cluster methods have also been reported for cytosine [41], all the aforementioned being compared to the present study in Table S6 (see SI). The results here obtained are again slightly blue-shifted with respect to previous CASPT2 computations given the different basis sets and state-averaging conditions employed and slightly red-shifted in comparison with the most correlated linear response CCSD(T) values as it has also been reported for thymine and uracil. Our results have been computed on top of slightly different structures (see Figure S6 in the SI) as compared to other theoretical studies thus accounting for possible strong fluctuations in the relative energetic position of the singlet excited-state manifold, as has been documented elsewhere [46]. Experimentally, four bands have been registered for $\pi\pi^*$ transitions *in vacuo* at 4.6–4.7, 5.2–5.8, 6.1–6.4, and 6.7–7.1 eV [77–79]. These features relate to signal 2 (4.5 eV), being placed within 0.1–0.2 eV to the first experimental band, signal 3 (5.42 eV) predicted to be within the estimated absorption maxima of the second band, signals 4 (6.31 eV) and 5 (6.42 eV) both being within the range of the third band, and signal 7 (6.92 eV) being the most likely candidate for the fourth absorption band registered experimentally, given its vertical excitation energy and transition dipole moment (*cf.* Table 3). Additionally, estimates of vertical excitation in aqueous solution were provided by Clark and co-workers, characterizing six different absorption bands placed at 4.65, 5.33, 5.60, 6.20, 7.69, and 8.06 eV, respectively [80]. By comparing the *in vacuo* and aqueous experimental data, it can readily seen how the second absorption band in gas-phase appears to have two different contributions in solution, the first and third bands being relatively unaffected by solvation, while the fourth and fifth band appears to be blue-shifted, getting close in energy to a sixth band not registered *in vacuo*. Ongoing efforts are being carried out to ascertain the specific states responsible for the diverse absorption bands in solution, which given the complex shifts of the signals cannot be assessed in terms of bathochromic shifts alone. As previously discussed for thymine and uracil, the dark $n\pi^*$ states have been omitted given their exceedingly small oscillator strength associated with their transitions that will prevent them from being directly populated by the pump pulse.

3.1.4 Comparison of the three nucleobases

An overall analysis is made here in order to find the unique features present in each of the pyrimidine nucleobases that may be employed to disentangle their specific contributions in poly- and oligonucleotide single- and double-stranded chains.

We firstly consider the differences between thymine and uracil, the former appearing in DNA, whereas the latter is featured in RNA. Figures 2 and 3 show strong similarities in their 2DES in the different probing windows. The strong resemblance obtained in their respective spectra is in line with the similar chemical structure, where the methyl group present in thymine slightly redshifts the singlet excited manifold with respect to uracil, yielding a displaced analogous picture. Nevertheless, a few qualitative differences can be discerned within the static approach here employed, even though more accurate time-resolved approaches will be required to ascertain the capacity to separate the different fingerprints depending on their specific broadening, which will in turn depend on the excited-state dynamics of the system [66, 67]. These are mainly based on the blueshift of the H- > L transition by the removal of the methyl group in uracil as compared with thymine. Their lowest lying bright excited state is placed roughly in the same region along Ω_1, being shifted by 0.2 eV (~1600 cm^{-1}), in agreement with the experimental shift reported in aqueous solution by Gustavsson et al. [75] of ~1000 cm^{-1}. Consequently, it is found that all states where the H – > L configuration contributes prominently (i.e., states 7, 9 and 10), the corresponding electronic levels are blue-shifted while the other states do not shift substantially. As a consequence, in the visible window, signals 6 and 7, which appear clearly separated in thymine, are overlapped in uracil, while in the NUV window we observe in uracil state mixing between the bright states 9 and 10 and the neighboring states that leads to intensity borrowing. As this effect arises from the mixture of configurations at the CASSCF level, where its magnitude cannot be completely ensured, we cannot exclude the possibility of an artificial mixing. One further fingerprint in the NUV region of uracil is the distinctive ESA signal above the GSB/SE that appears to be less intense thymine.

Cytosine, on the other hand, presents a very different spectrum compared to thymine and uracil given the presence of two well-defined traces along Ω_1. The appearance of these two traces, which presents qualitative different spectra compared to thymine and uracil, yields an easy way to differentiate cytosine from the other pyrimidines. The first trace along Ω_1, corresponding to the lowest lying $\pi\pi^*_{c,1}$ excited state is pronouncedly red-shifted (by ~0.5 eV) with respect to the analogous transitions registered for uracil and thymine. The trace corresponding to this state appears to be weak given its exceedingly small

transition dipole moment compared to any of the other pyrimidines and to the brighter $\pi\pi^*_{c,2}$ state (*cf.* Table 3). The $\pi\pi^*_{c,1}$ state is also placed too far away energetically from the bright $\pi\pi^*$ excited states of thymine and uracil to interact, being the preferred state to monitor cytosine-centered electronic processes given the right combination of pump and probe pulses is adopted. The second trace arising from the population of the $\pi\pi^*_{c,2}$ state (state 3), on the other hand, appears to be relatively bright and free of intense ESA signatures in the NUV region where intermolecular interactions might be probed, given the strong GSB/SE signal featured.

Overall, it can be concluded that the most intense signals in the Vis window for all pyrimidine systems, namely the ESAs featuring states 7, 9, and 10, have large contributions from double excitations involving the HOMO and LUMO orbitals (*cf.* Tables 1, 2, 3). This motif is therefore common to all canonical pyrimidines, including cytosine if a right pulse setup is used to enhance the first trace's intensity and adenine, as previously reported [46]. In cytosine, peak 13 corresponds to an excited state with the largest H- > L contribution, confirming the general trend found for all nucleobases so far. Thus, the H- > L transition is expected to represent the most clear, intense, and unique signature of each DNA base that may be used to monitor the dynamics of its respective $\pi\pi^*$ state in the Vis range. The NUV window, as documented for adenine [46], is strongly affected by the active space size, and thus the importance of this region for tracking the $\pi\pi^*$ excited-state dynamics is mainly based on the consideration that for dimeric and multimeric species the mixed localized excited states, i.e., the fingerprints of interchromophores interaction, are expected to appear just in this high-energy window [25, 26].

3.2 Dimers of canonical pyrimidine nucleobases

Finally, we highlight the spectral signatures due to noncovalent interactions in dimers containing canonical pyrimidine nucleobases. Figure 5 shows the comparison of the bidimensional electronic spectra of an ideal perfectly stacked uracil–cytosine dimer (Fig. 5a), with base separation of 4.0 Å, and of a non-interacting dimer. Some of these fingerprints can be observed in linear absorption spectra or in pump–probe spectra, but only bidimensional spectroscopy is capable of revealing all signatures simultaneously.

The spectrum of the non-interacting dimer can be conceived as composite of the spectra of the monomers (Fig. 5b, c). It is evident that the uracil trace ($\Omega_1 \approx 42,000$ cm^{-1}), associated with the transitions probed out of its lowest $\pi\pi^*$ state ($\pi\pi^*_u$), bears the highest oscillator strength, thereby revealing a set of characteristic ESAs in the range 33,000–39,000 cm^{-1} (peaks 9–12 in Fig. 3). On the other hand, the spectral traces of cytosine ($\Omega_1 \approx 36,000$ cm^{-1}

and $\Omega_1 \approx 43,500$ cm^{-1}), associated with its two lowest $\pi\pi^*$ states ($\pi\pi^*_{c,1}$ and $\pi\pi^*_{c,2}$), show off-diagonal bleach signal at $\Omega_1 \approx 36,000$ cm^{-1}/$\Omega_3 \approx 43,500$ cm^{-1} and $\Omega_1 \approx 43,500$ cm^{-1}/$\Omega_3 \approx 36,000$ cm^{-1}, while being free of intense ESA peaks (Fig. 4).

In the stacked dimer, the excitonic couplings (V) between $\pi\pi^*_u$ and $\pi\pi^*_{c,1}$/$\pi\pi^*_{c,2}$ are calculated to be in the few hundred cm^{-1} range ($V_{\pi\pi^*u/\pi\pi^*c,1} \sim 100$ cm^{-1}, $V_{\pi\pi^*u/\pi\pi^*c,2} \sim 600$ cm^{-1}) when the point dipole approximation is applied. Due to the near orthogonality of the transition dipole vectors μ^c_{01} and μ^u_{01} (Fig. 5a) and the large energetic separation of the corresponding states $\pi\pi^*_{c,1}$ and $\pi\pi^*_u$ the exciton coupling has virtually no effect on the energetic position of the lowest excited state of cytosine. On the contrary, an increase of the gap between $\pi\pi^*_u$ and $\pi\pi^*_{c,2}$ of nearly 400 cm^{-1} is obtained (Fig. 5d, label 1). It should be noted though that the point dipole approximation is known to overestimate the coupling strength for short interchromophore distances [68]. This, together with the peak broadening due to spectral diffusion and the inhomogeneity of the sample, is likely to cover the comparably small shifts.

Furthermore, it is evident that the intensity of the traces associated with $\pi\pi^*_u$ and $\pi\pi^*_{c,2}$ changes as $\pi\pi^*_{c,2}$ borrows oscillator strength from $\pi\pi^*_u$. This scenario is particularly eminent for homodimers [81] where the strong coupling of bright states with identical electronic configuration in each monomer leads to the formation of strongly delocalized dark exciton states in the dimer [82]. In fact the exciton states of the uracil–cytosine dimer should no longer be seen as purely localized on uracil or cytosine, despite the small coupling. This becomes clear, when focusing on the spectral window below the intense GSB signals (Fig. 5e). Due to the exciton delocalization over both chromophores the higher-lying local states of each base become accessible from both exciton states, giving rise to a number of new ESAs along the trace associated with the $\pi\pi^*_{c,2}$ transition in cytosine (Fig. 5e, label 2), which is empty in the isolated or non-interacting cytosine. The emerging signals are correlated with the ones along the uracil trace associated with the $\pi\pi^*_u$ transition, being red-shifted by the same energy which separates $\pi\pi^*_u$ and $\pi\pi^*_{c,2}$. Although weaker in intensity than the diagonal bleach signals, these ESAs can be selectively enhanced by tuning the central frequency of the probe pulse to the region of 35,000 cm^{-1} or by applying polarized pulse sequences [25].

One further characteristic signature of the chromophore coupling is the off-diagonal cross-peaks (Fig. 5d, label 3) along the $\pi\pi^*_u$ and $\pi\pi^*_{c,2}$ traces ($\Omega_1 \approx 42,000$ cm^{-1}/$\Omega_3 \approx 44,000$ cm^{-1} and $\Omega_1 \approx 44,000$ cm^{-1}/$\Omega_3 \approx 42,000$ cm^{-1}), appearing when the pump and the probe pulse involve different transitions in the system. These off-diagonal bleach signals are particularly intriguing from an experimental point of view as they do not
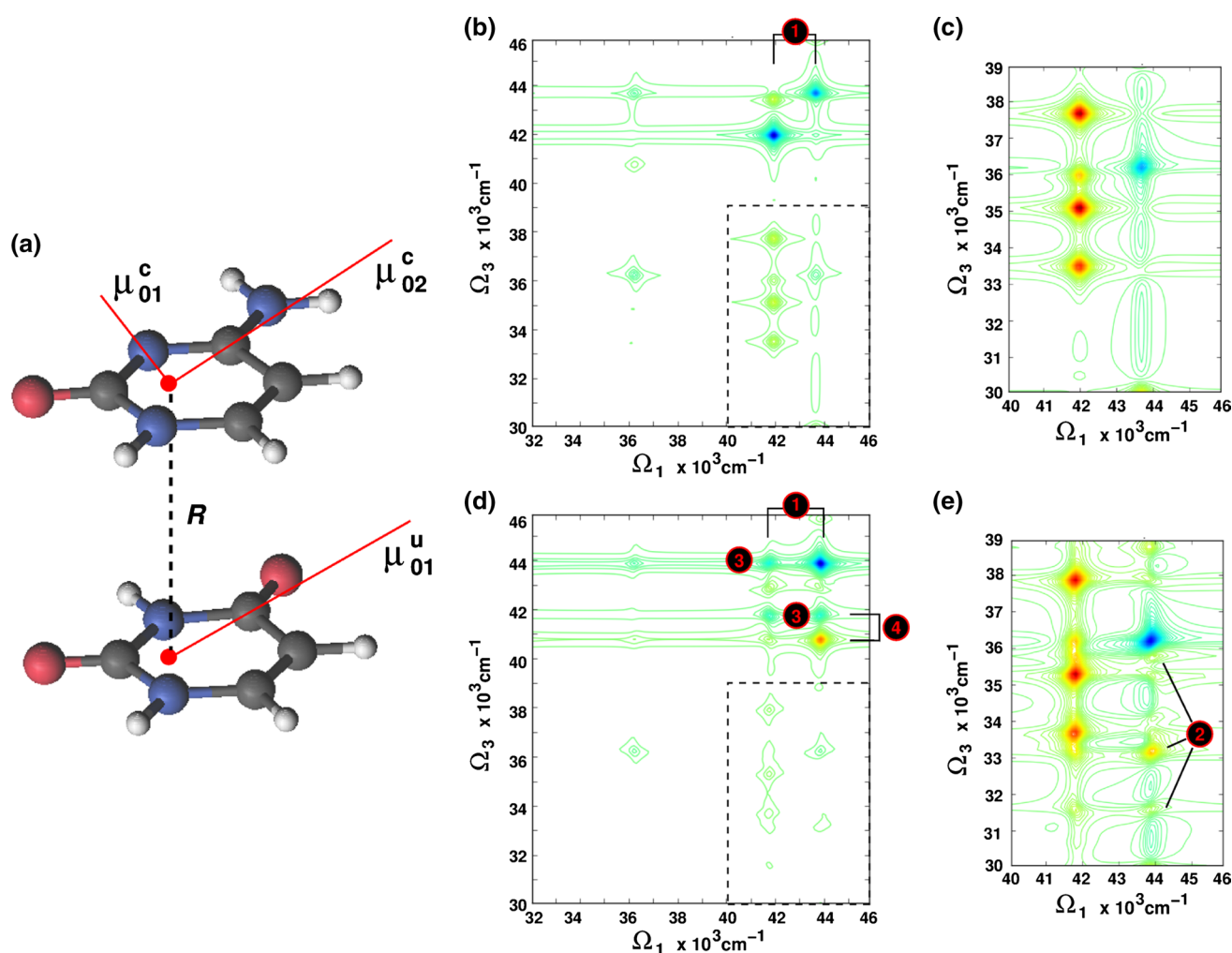
Fig. 5 2D NUV-pump/NUV-probe spectra of the unstacked (**b, c**) and perfectly stacked (**d, e**) uracil–cytosine dimer (**a**) models. A separation **R** of 4 Å was used in the case of the stacked dimer. The direction of the transition dipole vectors is also drawn (*red lines*) originating from the center of mass of each base. The spectral regions encompassed by the *dashed rectangles* in the full spectra (**b, d**) are zoomed-in in the *rightest panels* (**c, e,** respectively)

exhibit large spectral fluctuations (unlike ESA contributions) and can be easily targeted with suitable pump–probe setups ("two-color" setups required). We emphasize that their detection is only possible within the framework of bidimensional electronic spectroscopy, as they remain covered by the more intense diagonal bleach signals in case of pump–probe spectroscopy. In fact, the off-diagonal bleach signal is paired as a general rule to an ESA signal (Fig. 5d, label 4) that is expected to appear when the chromophores interact. The complementary ESA peak in each pair corresponds to a transition from the singly excited state (e.g., $\pi\pi^*_u$) to a mixed doubly excited state where both monomers are excited simultaneously (e.g., $\pi\pi^*_u + \pi\pi^*_{c,2}$). The energetic separation between the bleach and the ESA, known as quartic coupling and corresponding to the Darling–Dennison coupling in infrared spectroscopy [18], is an indicator of the coupling strength [83], as both peaks

should coincide and cancel in a non-interacting system. In the current example we chose a value of -1000 cm^{-1}, which agrees with results from previous ab initio computations of aromatic heterodimers [83], whereas the physical implication is a weak attractive force between the excitons. The cytosine trace proves better suited to visualize the quartic coupling due to being less congested.

Besides the signatures discussed so far, charge transfer transitions can become accessible [25] in stacked oligomers and can be used as markers for the chromophoric interactions, provided they appear in a signal-free region due to their low oscillator strength. The energetic position and signal intensity of charge transfer states are hard to deduce without explicitly computing the electronic structure of the dimer, and therefore, we have not included them in the present exciton model. Furthermore, the higher-lying local states of each monomer will also be affected by the base

stacking, although our experience shows that the effects are rather weak [25, 83].

4 Conclusions

In this work we report a rigorous calibration of the electronic high-lying excited-state manifolds of canonical pyrimidine nucleobases (thymine, uracil, and cytosine) *in vacuo* using accurate multiconfigurational wave function methods. The adopted approach provides accurate estimates of transition energies and dipole moments that have been used for simulating two-dimensional NUV-pump/Vis-probe and NUV-pump/NUV-probe spectra of the nucleobases and for constructing the excitonic Hamiltonian of a perfectly stacked uracil–cytosine dimer model. The nonlinear electronic spectra of thymine and uracil show rich 2D maps, with several ESA signals that represent spectroscopic fingerprints of these nucleobases. Notably, the 2D maps could resolve the tiny energy shifts caused by the small structural difference between thymine and uracil bases. Moreover, in the Vis-probe window, a set of ESA fingerprints can be used to discern between thymine and uracil bases, in analogy to adenine [46], suggesting a probing window for monomer-specific markers. Cytosine 2D spectra differ significantly from the other two pyrimidine nucleobases, due to the presence of two $\pi\pi^*$ absorbing excited states in the NUV window, with one of them involving particularly bright transitions, gathering most of the spectral intensities. The signal trace associated with this cytosine state appears in a similar energetic window as the main $\pi\pi^*$ state of uracil, inducing non-negligible electronic interactions. Interestingly, the 2DES spectrum of cytosine is dominated by the GSB signals and it is lacking ESA signatures unless non-covalent interchromophoric interactions are involved, which makes this nucleobase an ideal marker for interbase interactions. A preliminary study of these interchromophoric electronic couplings, based on a Frenkel excitonic model, was performed on a perfectly stacked uracil–cytosine dimer model. This indicates the presence of spectroscopic fingerprints of non-covalent chromophore interactions that could be detected on the spectral trace of the $\pi\pi^*_{c,2}$ excited state of cytosine, by setting appropriate pump–probe pulse central frequencies, bandwidths, and polarizations.

References

1. Gustavsson T, Improta R, Markovitsi D (2010) DNA/RNA: building blocks of life under UV irradiation. J Phys Chem Lett 1(13):2025–2030. doi:10.1021/jz1004973
2. Vaya I, Gustavsson T (2010) Fluorescence of natural DNA: from the femtosecond to the nanosecond time scales. 10:11834–11835
3. Giussani A, Segarra-Martí J, Roca-Sanjuán D, Merchán M (2015) Excitation of nucleobases from a computational perspective I: reaction paths. In: Barbatti M, Borin AC, Ullrich S (eds) Photoinduced phenomena in nucleic acids I, vol 355. Topics in current chemistry. Springer International Publishing, pp 57–97. doi:10.1007/128_2013_501
4. Crespo-Hernandez CE, Cohen B, Hare PM, Kohler B (2004) Ultrafast excited-state dynamics in nucleic acids. Chem Rev 104(4):1977–2019. doi:10.1021/cr0206770
5. Chen J, Zhang Y, Kohler B (2015) Excited states in DNA strands investigated by ultrafast laser spectroscopy. In: Barbatti M, Borin AC, Ullrich S (eds) Photoinduced phenomena in nucleic acids II, vol 356. Topics in current chemistry. Springer International Publishing, pp 39–87. doi:10.1007/128_2014_570
6. Middleton CT, de La Harpe K, Su C, Law YK, Crespo-Hernandez CE, Kohler B (2009) DNA excited-state dynamics: from single bases to the double helix. In: Annual review of physical chemistry, vol 60. Annual review of physical chemistry. Annual reviews, Palo Alto, pp 217–239. doi:10.1146/annurev.physchem.59.032607.093719
7. Merchán M, González-Luque R, Climent T, Serrano-Andrés L, Rodriguez E, Reguero M, Pelaez D (2006) Unified model for the ultrafast decay of pyrimidine nucleobases. J Phys Chem B 110(51):26471–26476. doi:10.1021/jp066874a
8. Noonan FP, Zaidi MR, Wolnicka-Glubisz A, Anver MR, Bahn J, Wielgus A, Cadet J, Douki T, Mouret S, Tucker MA, Popratiloff A, Merlino G, De Fabo EC (2012) Melanoma induction by ultraviolet A but not ultraviolet B radiation requires melanin pigment. Nat Commun 3:884
9. Cadet J, Mouret S, Ravanat J-L, Douki T (2012) Photoinduced damage to cellular DNA: direct and photosensitized reactions. Photochem Photobiol 88(5):1048–1065. doi:10.1111/j.1751-1097.2012.01200.x
10. Cadet J, Grand A, Douki T (2015) Solar UV radiation-induced DNA bipyrimidine photoproducts: formation and mechanistic insights. In: Barbatti M, Borin AC, Ullrich S (eds) Photoinduced phenomena in nucleic acids II, vol 356. Topics in current chemistry. Springer International Publishing, pp 249–275. doi:10.1007/128_2014_553
11. Giussani A, Serrano-Andrés L, Merchán M, Roca-Sanjuán D, Garavelli M (2013) Photoinduced formation mechanism of the thymine-thymine (6-4) adduct. J Phys Chem B 117(7):1999–2004. doi:10.1021/jp307200g
12. Barbatti M, Borin A, Ullrich S (2014) Photoinduced processes in nucleic acids. In: Topics in current chemistry. Springer, Berlin, pp 1–32. doi:10.1007/128_2014_569
13. Crespo-Hernández CE, Cohen B, Kohler B (2005) Base stacking controls excited-state dynamics in A.T DNA. Nature 436 (7054):1141–1144. doi:10.1038/nature03933

14. Chen J, Kohler B (2014) Base stacking in adenosine dimers revealed by femtosecond transient absorption spectroscopy. J Am Chem Soc 136(17):6362–6372. doi:10.1021/ja501342b

15. Su C, Middleton CT, Kohler B (2012) Base-stacking disorder and excited-state dynamics in single- stranded adenine homo-oligonucleotides

16. Onidas D, Gustavsson T, Lazzarotto E, Markovitsi D (2007) Fluorescence of the DNA double helices (dAdT)n. (dAdT)n studied by femtosecond spectroscopy. Phys Chem Chem Phys 9(37):5143–5148. doi:10.1039/b707882j

17. Markovitsi D, Gustavsson T, Vayá I (2010) Fluorescence of DNA duplexes: from model helices to natural DNA. J Phys Chem Lett 1(22):3271–3276. doi:10.1021/jz101122t

18. Abramavicius D, Palmieri B, Voronine DV, Sanda F, Mukamel S (2009) Chem Rev 109:2350

19. Mukamel S (1995) Principles of nonlinear optical spectroscopy. O.U.P, New York

20. Polli D, Rivalta I, Nenov A, Weingart O, Garavelli M, Cerullo G (2015) Tracking the primary photoconversion events in rhodopsins by ultrafast optical spectroscopy. Photochem Photobiol Sci. doi:10.1039/C4PP00370E

21. C-h Tseng, Sándor P, Kotur M, Weinacht T, Matsika S (2012) Two-dimensional Fourier transform spectroscopy of adenine and uracil using shaped ultrafast laser pulses in the deep UV. J Phys Chem A 116:2654–2661

22. West BA, Moran AM (2012) Two-dimensional electronic spectroscopy in the ultraviolet wavelength range. J Phys Chem Lett 3(18):2575–2581. doi:10.1021/jz301048n

23. Rivalta I, Nenov A, Cerullo G, Mukamel S, Garavelli M (2014) Ab initio simulations of two-dimensional electronic spectra: the SOS//QM/MM approach. Int J Quantum Chem 114(2):85–93. doi:10.1002/qua.24511

24. Nenov A, Rivalta I, Mukamel S, Garavelli M (2014) Bidimensional electronic spectroscopy on indole in gas phase and in water from first principles. Comput Theor Chem 1040–1041:295–303. doi:10.1016/j.comptc.2014.03.031

25. Nenov A, Rivalta I, Cerullo G, Mukamel S, Garavelli M (2014) Disentangling peptide configurations via two-dimensional electronic spectroscopy: ab initio simulations beyond the Frenkel exciton hamiltonian. J Phys Chem Lett 5:767–771

26. Nenov A, Segarra-Martí J, Giussani A, Conti I, Rivalta I, Dumont E, Jaiswal VK, Altavilla SF, Mukamel S, Garavelli M (2015) FD 177: probing deactivation pathways of DNA nucleobases by two-dimensional electronic spectroscopy: first principles simulations. Faraday Discuss 177:345–362. doi:10.1039/C4FD00175C

27. Rivalta I, Nenov A, Weingart O, Cerullo G, Garavelli M, Mukamel S (2014) Modelling time-resolved two-dimensional electronic spectroscopy of the primary photoisomerization event in rhodopsin. J Phys Chem B 118(28):8396–8405. doi:10.1021/jp502538m

28. Roos BO (1987) The complete active space self-consistent field method and its applications in electronic structure calculations. In: Advances in chemical physics. John Wiley & Sons, Inc., New York, pp 399–445. doi:10.1002/9780470142943.ch7

29. Andersson K, Malmqvist PA, Roos BO (1992) 2nd-order perturbation-theory with a complete active space self-consistent field reference function. J Chem Phys 96(2):1218–1226. doi:10.1063/1.462209

30. Roca-Sanjuán D, Aquilante F, Lindh R (2012) Multiconfiguration second-order perturbation theory approach to strong electron correlation in chemistry and photochemistry. Wiley Interdiscip Rev Comput Mol Sci 2(4):585–603. doi:10.1002/wcms.97

31. Roos BO, Andersson K, Fulscher MP, Malmqvist PA, Serrano-Andrés L, Pierloot K, Merchán M (1996) Multiconfigurational perturbation theory: applications in electronic spectroscopy. Adv Chem Phys Vol Xciii 93:219–331. doi:10.1002/9780470141526.ch5

32. Saurí V, Serrano-Andrés L, Shahi ARM, Gagliardi L, Vancoillie S, Pierloot K (2011) Multiconfigurational second-order perturbation theory restricted active space (RASPT2) method for electronic excited states: a benchmark study. J Chem Theory Comput 7(1):153–168. doi:10.1021/ct100478d

33. Rubio M, Serrano-Andrés L, Merchán M (2008) Excited states of the water molecule: analysis of the valence and Rydberg character. J Chem Phys 128 (10). doi:10.1063/1.2837827

34. Segarra-Martí J, Garavelli M, Aquilante F (2015) Multiconfigurational second-order perturbation theory with frozen natural orbitals extended to the treatment of photochemical problems. J Chem Theory Comput 11(8):3772–3784. doi:10.1021/acs.jctc.5b00479

35. Garavelli M (2006) Computational organic photochemistry: strategy. Achiev Perspect Theor Chem Acc 116(1–3):87–105. doi:10.1007/s00214-005-0030-z

36. Giussani A (2014) Toward the understanding of the photophysics and photochemistry of 1-nitronaphthalene under solar radiation: the first theoretical evidence of a photodegradation intramolecular rearrangement mechanism involving the triplet states. J Chem Theory Comput 10(9):3987–3995. doi:10.1021/ct500395f

37. Schapiro I, Ryazantsev MN, Manuel Frutos L, Ferre N, Lindh R, Olivucci M (2011) The ultrafast photoisomerizations of rhodopsin and bathorhodopsin are modulated by bond length alternation and HOOP driven electronic effects. J Am Chem Soc 133(10):3354–3364. doi:10.1021/ja1056196

38. Segarra-Martí J, Coto PB (2014) A theoretical study of the intramolecular charge transfer in 4-(dimethylamino)benzethyne. Phys Chem Chem Phys 16(46):25642–25648. doi:10.1039/C4CP03436H

39. Malmqvist PÅ, Pierloot K, Shahi ARM, Cramer CJ, Gagliardi L (2008) The restricted active space followed by second-order perturbation theory method: theory and application to the study of CuO_2 and Cu_2O_2 systems. J Chem Phys 128(20):204109. doi:10.1063/1.2920188

40. Nenov A, Mukamel S, Garavelli M, Rivalta I (2015) Two-dimensional electronic spectroscopy of benzene, phenol, and their dimer: an efficient first-principles simulation protocol. J Chem Theory Comput 11(8):3755–3771. doi:10.1021/acs.jctc.5b00443

41. Szalay PG, Watson T, Perera A, Lotrich VF, Bartlett RJ (2012) Benchmark studies on the building blocks of DNA. 1. Superiority of coupled cluster methods in describing the excited states of nucleobases in the Franck–Condon region. J Phys Chem A 116(25):6702–6710. doi:10.1021/jp300977a

42. Szalay PG, Watson T, Perera A, Lotrich V, Fogarasi G, Bartlett RJ (2012) Benchmark studies on the building blocks of DNA. 2. Effect of biological environment on the electronic excitation spectrum of nucleobases. J Phys Chem A 116(35):8851–8860. doi:10.1021/jp305130q

43. Szalay PG, Watson T, Perera A, Lotrich V, Bartlett RJ (2013) Benchmark Studies on the Building Blocks of DNA. 3. Watson–Crick and stacked base pairs. J Phys Chem A 117(15):3149–3157. doi:10.1021/jp3100975

44. Roca-Sanjuán D, Rubio M, Merchán M, Serrano-Andrés L (2006) Ab initio determination of the ionization potentials of DNA and RNA nucleobases. J Chem Phys 125(8):084302. doi:10.1063/1.2336217

45. Shida S, Hatano Y (1976) Fragmentation of excited molecules and ions in the radiolysis of hydrocarbons. Int J Radiat Phys Chem 8(1–2):171–185. doi:10.1016/0020-7055(76)90064-4

46. Nenov A, Giussani A, Segarra-martí J, Vishal K, Rivalta I, Cerullo G, Mukamel S (2015) Modeling the high-energy electronic state manifold of adenine: calibration for nonlinear electronic spectroscopy. J Chem Phys 142:212443

47. Lorentzon J, Malmqvist P-Å, Fülscher M, Roos B (1995) A CASPT2 study of the valence and lowest Rydberg electronic states of benzene and phenol. Theor Chim Acta 91(1–2):91–108. doi:10.1007/BF01113865

48. Robin MB (1974) Higher excited states of polyatomic molecules, vol 1. Academic Press Inc, London

49. Widmark PO, Persson BJ, Roos BO (1991) Density matrix averaged atomic natural orbital (ANO) basis sets for correlated molecular wave functions. II. Second row atoms. Theor Chim Acta 79(6):419–432

50. Altavilla SF, Segarra-Martí J, Nenov A, Conti I, Rivalta I, Garavelli M (2015) Deciphering the photochemical mechanisms describing the UV-induced processes occurring in solvated guanine monophosphate. Front Chem 3. doi:10.3389/fchem.2015.00029

51. Nakayama A, Arai G, Yamazaki S, Taketsugu T (2013) Solvent effects on the ultrafast nonradiative deactivation mechanisms of thymine in aqueous solution: excited-state QM/MM molecular dynamics simulations. J Chem Phys 139(21):214304. doi:10.1063/1.4833563

52. Buchner F, Nakayama A, Yamazaki S, Ritze H-H, Lübcke A (2015) Excited-state relaxation of hydrated thymine and thymidine measured by liquid-jet photoelectron spectroscopy: experiment and simulation. J Am Chem Soc 137(8):2931–2938. doi:10.1021/ja511108u

53. Malmqvist P-Å, Roos BO (1989) The CASSCF state interaction method. Chem Phys Lett 155(2):189–194. doi:10.1016/0009-2614(89)85347-3

54. Szalay PG, Müller T, Gidofalvi G, Lischka H, Shepard R (2012) Multiconfiguration self-consistent field and multireference configuration interaction methods and applications. Chem Rev 112(1):108–181. doi:10.1021/cr200137a

55. Malmqvist PA, Rendell A, Roos BO (1990) The restricted active space self-consistent-field method, implemented with a split graph unitary-group approach. J Phys Chem 94(14):5477–5482. doi:10.1021/j100377a011

56. Ghigo G, Roos BO, Malmqvist P-Å (2004) A modified definition of the zeroth-order Hamiltonian in multiconfigurational perturbation theory (CASPT2). Chem Phys Lett 396(1–3):142–149. doi:10.1016/j.cplett.2004.08.032

57. Forsberg N, Malmqvist PA (1997) Multiconfiguration perturbation theory with imaginary level shift. Chem Phys Lett 274(1–3):196–204. doi:10.1016/s0009-2614(97)00669-6

58. Bostrom J, Delcey MG, Aquilante F, Serrano-Andres L, Pedersen TB, Lindh R (2010) Calibration of Cholesky auxiliary basis sets for multiconfigurational perturbation theory calculations of excitation energies. J Chem Theory Comput 6(3):747–754. doi:10.1021/ct900612k

59. Aquilante F, Malmqvist P-Å, Pedersen TB, Ghosh A, Roos BO (2008) Cholesky decomposition-based multiconfiguration second-order perturbation theory (CD-CASPT2): application to the spin-state energetics of CoIII(diiminato)(NPh). J Chem Theory Comput 4(5):694–702. doi:10.1021/ct700263h

60. Pedersen T, Aquilante F, Lindh R (2009) Density fitting with auxiliary basis sets from Cholesky decompositions. Theoret Chem Acc 124(1–2):1–10. doi:10.1007/s00214-009-0608-y

61. Aquilante F, Pedersen TB, Lindh R (2007) Low-cost evaluation of the exchange Fock matrix from Cholesky and density fitting representations of the electron repulsion integrals. J Chem Phys 126(19):194106. doi:10.1063/1.2736701

62. Aquilante F, Lindh R, Bondo Pedersen T (2007) Unbiased auxiliary basis sets for accurate two-electron integral approximations. J Chem Phys 127(11):114107. doi:10.1063/1.2777146

63. Aquilante F, De Vico L, Ferré N, Ghigo G, Malmqvist PA, Neogrady P, Pedersen TB, Pitonak M, Reiher M, Roos BO, Serrano-Andrés L, Urban M, Veryazov V, Lindh R (2010) Software news and update MOLCAS 7: the next generation. J Comput Chem 31(1):224–247. doi:10.1002/jcc.21318

64. Aquilante F, Pedersen TB, Veryazov V, Lindh R (2013) MOLCAS—a software for multiconfigurational quantum chemistry calculations. Wiley Interdiscip Rev Comput Mol Sci 3(2):143–149. doi:10.1002/wcms.1117

65. Aquilante F, Autschbach J, Carlson R, Chibotaru L, Delcey MG, De Vico L, Fernández Galvan I, Ferré N, Frutos LM, Gagliardi L, Garavelli M, Giussani A, Hoyer C, Li Manni G, Lischka H, Ma D, Malmqvist PA, Müller T, Nenov A, Olivucci M, Pedersen TB, Peng D, Plasser F, Pritchard B, Reiher M, Rivalta I, Schapiro I, Segarra-Martí J, Stenrup M, Truhlar DG, Ungur L, Valentini A, Vancoillie S, Veryazov V, Vysotskiy V, Weingart O, Zapata F, Lindh R (2015) Molcas 8: new capabilities for multiconfigurational quantum chemical calculations across the periodic table. J Comput Chem. doi:10.1002/jcc.24221.10.1002/jcc.24221

66. Nenov A, Giussani A, Fingerhut BP, Rivalta I, Dumont E, Mukamel S, Garavelli M (2015) Spectral lineshape in nonlinear electronic spectroscopy †. Phys Chem Chem Phys 17:30925–30936. doi:10.1039/C5CP01167A

67. Tempelaar R, van der Vegte CP, Knoester J, Jansen TLC (2013) Surface hopping modeling of two-dimensional spectra. J Chem Phys 138(16):164106. doi:10.1063/1.4801519

68. May V, Kühn O (2011) Charge and energy transfer dynamics in molecular systems. Wiley VCH, Germany

69. Fülscher MP, Serrano-Andrés L, Roos BO (1997) A theoretical study of the electronic spectra of adenine and guanine. J Am Chem Soc 119(26):6168–6176. doi:10.1021/ja964426i

70. Lorentzon J, Fuelscher MP, Roos BO (1995) Theoretical study of the electronic spectra of uracil and thymine. J Am Chem Soc 117(36):9265–9273. doi:10.1021/ja00141a019

71. Fülscher MP, Roos BO (1995) Theoretical study of the electronic spectrum of cytosine. J Am Chem Soc 117(7):2089–2095. doi:10.1021/ja00112a024

72. Schreiber M, Silva MR, Sauer SPA, Thiel W (2008) Benchmarks for electronically excited states: CASPT2, CC2, CCSD, and CC3. J Chem Phys 128 (13). doi:10.1063/1.2889385

73. Silva MR, Schreiber M, Sauer SPA, Thiel W (2008) Benchmarks for electronically excited states: time-dependent density functional theory and density functional theory based multireference configuration interaction. J Chem Phys 129 (10). doi:10.1063/1.2973541

74. Abouaf R, Pommier J, Dunet H (2003) Electronic and vibrational excitation in gas phase thymine and 5-bromouracil by electron impact. Chem Phys Lett 381(3–4):486–494. doi:10.1016/j.cplett.2003.09.121

75. Gustavsson T, Banyasz A, Lazzarotto E, Markovitsi D, Scalmani G, Frisch MJ, Barone V, Improta R (2006) Singlet excited-state behavior of uracil and thymine in aqueous solution: a combined experimental and computational study of 11 uracil derivatives. J Am Chem Soc 128(2):607–619. doi:10.1021/ja056181s

76. Kannar D, Szalay PG (2014) Benchmarking coupled cluster methods on singlet excited states of nucleobases. J Mol Model 20:2503

77. Clark LB, Peschel GG, Tinoco I (1965) Vapor spectra and heats of vaporization of some purine and pyrimidine bases1. J Phys Chem 69(10):3615–3618. doi:10.1021/j100894a063

78. Voet D, Gratzer WB, Cox RA, Doty P (1963) Absorption spectra of nucleotides, polynucleotides, and nucleic acids in the far ultraviolet. Biopolymers 1(3):193–208. doi:10.1002/bip.360010302

79. Yamada T, Fukutome H (1968) Vacuum ultraviolet absorption spectra of sublimed films of nucleic acid bases. Biopolymers 6(1):43–54. doi:10.1002/bip.1968.360060104

80. Zaloudek F, Novros JS, Clark LB (1985) The electronic spectrum of cytosine. J Am Chem Soc 107:7344–7351

81. Kozak CR, Kistler KA, Lu Z, Matsika S (2010) Excited-state energies and electronic couplings of DNA base dimers. J Phys Chem B 114(4):1674–1683. doi:10.1021/jp9072697

82. Olaso-González G, Merchán M, Serrano-Andrés L (2009) The role of adenine excimers in the photophysics of oligonucleotides. J Am Chem Soc 131(12):4368–4377. doi:10.1021/ja808280j

83. Nenov A, Beccara S, Rivalta I, Cerullo G, Mukamel S, Garavelli M (2014) Tracking conformational dynamics of polypeptides by nonlinear electronic spectroscopy of aromatic residues: a first-principles simulation study. Chem Phys Chem 15(15):3282–3290. doi:10.1002/cphc.201402374