Title

Towards Accurate Simulation of Two-Dimensional Electronic Spectroscopy

Journal

Topics in Current Chemistry, 376(3)

ISSN

2365-0869

Authors

Segarra-Martí, Javier
Mukamel, Shaul
Garavelli, Marco
et al.

Publication Date

2018-06-01

DOI

10.1007/s41061-018-0201-8

Peer reviewed

Top Curr Chem (Z) (2018) 376:24
https://doi.org/10.1007/s41061-018-0201-8

REVIEW

Towards Accurate Simulation of Two-Dimensional Electronic Spectroscopy

Javier Segarra-Martí[1] · Shaul Mukamel[2] · Marco Garavelli[3] · Artur Nenov[3] · Ivan Rivalta[1]

Received: 21 December 2017 / Accepted: 24 April 2018 / Published online: 1 June 2018

Abstract We introduce the basic concepts of two-dimensional electronic spectroscopy (2DES) and a general theoretical framework adopted to calculate, from first principles, the nonlinear response of multi-chromophoric systems in realistic environments. Specifically, we focus on UV-active chromophores representing the building blocks of biological systems, from proteins to nucleic acids, describing our progress in developing computational tools and protocols for accurate simulation of their 2DUV spectra. The roadmap for accurate 2DUV spectroscopy simulations is illustrated starting with benchmarking of the excited-state manifold of the chromophoric units in a vacuum, which can be used for building exciton Hamiltonians for large-scale applications or as a reference for first-principles simulations with reduced computational cost, enabling treatment of minimal (still realistic) multi-chromophoric model systems. By adopting a static approximation that neglects dynamic processes such as spectral diffusion and population transfer, we show how 2DUV is able to characterize the ground-state conformational space of dinucleosides and small peptides comprising dimeric chromophoric units (in their native environment) by tracking inter-chromophoric electronic couplings. Recovering the excited-state coherent vibrational dynamics and population transfers,

This article is part of the Topical Collection "Multidimensional Time-Resolved Spectroscopy"; edited by Tiago Buckup, Jeremie Leonard.

✉ Marco Garavelli

✉ Artur Nenov

✉ Ivan Rivalta
ivan.rivalta@ens-lyon.fr

[1] Université de Lyon, École Normale Supérieure de Lyon, CNRS, Université Claude Bernard Lyon 1, Laboratoire de Chimie UMR 5182, 69342 Lyon, France

[2] Department of Chemistry, University of California, Irvine, CA 92697-2025, USA

[3] Dipartimento di Chimica Industriale, Università degli Studi di Bologna, Viale del Risorgimento 4, 40136 Bologna, Italy

we observe a remarkable agreement between the predicted 2DUV spectra of the pyrene molecule and the experimental results. These results further led to theoretical studies of the excited-state dynamics in a solvated dinucleoside system, showing that spectroscopic fingerprints of long-lived excited-state minima along the complex photoinduced decay pathways of DNA/RNA model systems can be simulated at a reasonable computational cost. Our results exemplify the impact of accurate simulation of 2DES spectra in revealing complex physicochemical properties of fundamental biological systems and should trigger further theoretical developments as well as new experiments.

Keywords Nonlinear electronic spectroscopy · Theoretical simulations · Wavefunction methods · QM/MM computations · DNA/RNA nucleobases · Aromatic amino acids

1 Introduction

Two-dimensional (2D) optical spectroscopy based on multipulse laser sequences originated as an extension of the 2D nuclear magnetic resonance (2DNMR) technique [1] to the optical regime [2]. 2DNMR had an impressive impact in several fields, with first handover to the infrared (IR) regime (2DIR) now a well-established method often employed in the characterization of the structure and dynamics of complex molecular systems by directly mapping their vibrational couplings [3]. Mapping of electronic couplings by 2D electronic spectroscopy (2DES) [4] has become feasible thanks to advances [5, 6] in ultrafast optical techniques that allow the targeting of electronic transitions in the visible (Vis) range. 2DES in the Vis range (2DVis) has become increasingly popular over the last decade, showcasing its great potential by deciphering energy transfer processes in photosynthesis [7–10]. The desirable extension to the UV domain (2DUV), however, where many fundamental biomolecules display strong absorption bands, has been slow thus far, impeded by technical difficulties associated with the use of UV laser pulses [11]. Recent progress [12–18] in attaining interferometric phase stability and sufficient laser bandwidth has enabled access to the first examples of experimental 2DUV spectra [19–22]. Simulation studies have demonstrated that 2DUV of aromatic residues in the near UV (NUV) and the backbone in the far UV can effectively probe protein secondary structure [23, 24]. Among the most recent developments in multidimensional spectroscopy related to 2DES, it is important to highlight the advances in electronic-vibrational spectroscopy [25] and 2D Stark spectroscopy for characterizing dark charge-transfer states [26], along with the impressive potential of developing techniques in the X-ray regime [27].

The signals recorded in 2DES correlation plots refer to the third-order nonlinear response of the sample, and contain a wealth of information on the excited-state manifold of the chromophoric units and its photo-induced evolution in time, which are strongly related due to intra- and inter-chromophoric electronic couplings and coherence/decoherence effects in chemical and biophysical systems [28].

The electronic transitions involving excited-state absorptions, especially in 2DUV spectra, would encompass high-energy electronic levels whose nature is

hitherto unexplored. The complexity of the information contained in 2DES maps and the involvement of high-energy electronic states calls for the implementation of computational tools based on a combination of ab initio electronic structure and nonlinear response formalism. Accurate simulation of 2DUV electronic spectroscopy may, in fact, lead to both interpretation and prediction of nonlinear spectra, enhancing the potential application of this technique in various fields. Here, we outline a route towards accurate simulation of 2DUV spectra of fundamental biological systems. After describing the general 2DES technique and the target UV-active bio-chromophores in this section, we present a basic background in Sect. 2, aimed at providing the theoretical foundations of our simulation protocols and showing the approximations adopted. In Sect. 3, the main results of our developments and applications are illustrated for both nucleic acid and protein model systems, showing the potential for 2DUV in characterizing the ground-state (GS) conformational space and excited-state dynamics in these target systems. Finally, an outlook summary and perspective are provided in Sect. 4.

2 The 2DES Technique

Two-dimensional electronic spectroscopy [29] is based on a sequence of three ultrashort laser pulses interacting with the sample, generating a third-order nonlinear polarization and emitting signal fields in phase-matched directions (Fig. 1). The major advantage of 2DES is the higher spectral resolution relative to 1D-PP time-resolved techniques, with the nonlinear signal containing information on system dynamics and electronic couplings spread over two frequency axes (pump and probe, Ω_1 and Ω_3, respectively; see Fig. 1b). The spectral resolution in two dimensions enables accurate characterization of inhomogeneous and homogeneous broadening processes, with the ability to separate these two contributions (not possible with 1D techniques), providing information about solvent reorganization timescales as well as enabling the detection of signals associated with coupling between electronic excitations and charge/energy transfer processes.

The heterodyne-detected three-pulse photon echo (3PPE) non-collinear scheme [5, 30] is the experimental setup that can fully resolve (in amplitude and phase) the third-order nonlinear response, by collecting a four-wave mixing signal in a background-free direction, which is heterodyned by a local oscillator (LO); see Fig. 1a. The signal field emitted can be detected in the so-called rephasing K_{I} ($k_{\mathrm{LO}} = -k_1 + k_2 + k_3$) and non-rephasing K_{II} ($k_{\mathrm{LO}} = +k_1 - k_2 + k_3$) phase-matching directions, as a function of three controlled excitation-pulse time delays (t_1, t_2 and t_3). Instead of 3PPE, the partially collinear pump–probe (PCPP) geometry [21, 31, 32] can be adopted by using a pair of collinear pump pulses (k_1, k_2) non-collinearly combined with a probe pulse (k_3), providing a nonlinear response signal which is heterodyned by the probe pulse itself (self-heterodyning). The PCPP experiment is disadvantaged by the strong background signal, and it intrinsically provides the combined $K_{\mathrm{I}} + K_{\mathrm{II}}$ "quasi-absorptive" response, with loss of information on specific rephasing and non-rephasing signals, although removal of slowly decaying dispersive contributions due to phase cancellation of the concurring K_{I} and K_{II} signals might help in resolving weak signals.

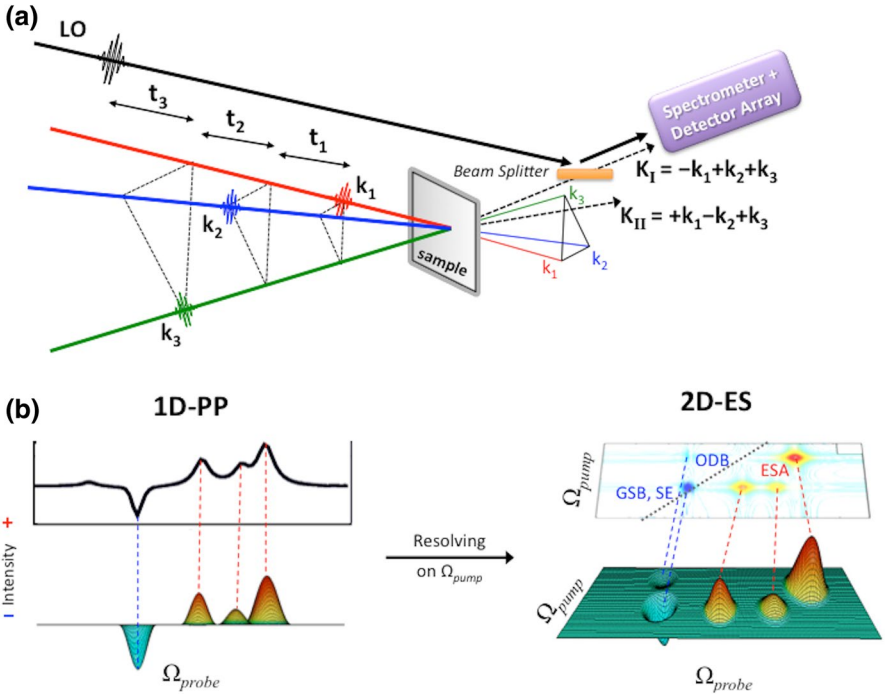

Fig. 1 Schematic representation of **a** heterodyne-detected three-pulse photon echo (3PPE) 2DES experiment (LO: local oscillator) showing both rephasing K_I ($k_{LO} = -k_1+k_2+k_3$) and non-rephasing K_{II} ($k_{LO} = +k_1 -k_2+k_3$) phase-matching directions, and **b** 1D pump–probe (1D-PP) and analogous 2D electronic spectra (2D-ES) at a fixed waiting time t_2, showing how the different shapes and signs of the nonlinear signals are properly resolved in the 2DES maps

For a given "population time" t_2 (also known as "waiting time", and analogous to the time delay between the pump and the probe pulse in 1D pump–probe [1D-PP] spectroscopy), a 2D spectrum (or map) is recorded as a function of "excitation frequency" Ω_1 (Ω_{pump}) and "detection frequency" Ω_3 (Ω_{probe}) by Fourier transformation with respect to the t_1 and t_3 time intervals. Thus, the nonlinear spectral signatures of the GS are obtained by setting the time t_2 to zero, while the state-specific response along excited-state dynamics can be monitored for $t_2 > 0$.

As illustrated in Fig. 1b, the induced nonlinear polarization measured in a 2DES map at any certain population time is richer than that of a 1D-PP spectrum, as by adding the spectral resolution along Ω_{pump}, it is possible to disentangle the contributions arising from different pump frequencies and the line broadening along both pump and probe frequencies. In general, a 2DES spectrum at $t_2 = 0$ contains diagonal peaks that correspond to ground-to-excited-state absorptions (ground-state bleaching, GSB) and stimulated emissions (SE), symmetric off-diagonal signals associated with coupling between ground- and excited-state electronic excitations (off-diagonal bleaching, ODB) and asymmetric off-diagonal

peaks (with the opposite sign relative to diagonal peaks) due to excited-state absorptions (ESAs) to the high-lying states. For $t_2 > 0$, the s dynamics shape the 2DES spectra, with SE and ESA peaks exhibiting at short waiting times, fluctuations along Ω_{probe}, an indirect probe of the coherent vibrational dynamics along the photoactive state, and at longer waiting times, Stoke shifts associated with internal energy redistribution and dissipation in the environment, as well as intensity decay, a function of the finite excited-state lifetimes. For additional details on 2DES experimental techniques and map features, we recommend the review by Maiuri and Brazard in this series [33].

2.1 2DES in the Ultraviolet: A New Light in Photobiology

The emergence of novel UV pulse technologies [14] has enabled the use of pump and/or probe laser pulses in the UV window. These efforts may have a particular impact in monitoring biologically relevant processes [34]. Indeed, cyclic aromatic groups are often the UV-active chromophores that play a major role in photoinduced events in biological systems [35]. UV light, in fact, is absorbed by an incredibly large portion of biomolecules, ranging from proteins to nucleic acids. As shown in Fig. 2, adsorption by both protein and nucleic acid backbones occurs exclusively at high energies, i.e. above ca. 5 eV (i.e. at wavelengths below ca. 250 nm). Aromatic rings in proteins and DNA/RNA, including tryptophan (Trp), phenylalanine (Phe), tyrosine (Tyr) and histidine (His) amino acid side chains and adenine (Ade), guanine (Gua), cytosine (Cyt), thymine (Thy) and uracil (Ura) nucleobases, feature their most intense absorptions in the same spectral window as their backbones. However, all these aromatic units (except histidine) have distinctive absorption bands associated with $\pi \rightarrow \pi^*$ transitions at lower energies (i.e. around 4.4–5.0 eV, ca. 250–280 nm), thus providing a UV range where their excitations are clearly separated from those of the backbone. Notably, when considering even lower transition energies in the NUV range around 300–400 nm (i.e. 3.1–4.1 eV), quite important UV-active biomolecules are found, including essential cofactors such as plastoquinones, flavins and reduced nicotinamide adenine dinucleotide (NADH), to name a few. This evidence suggests that the low-energy UV window represents a promising option for monitoring many photoinduced biological phenomena involving aromatic units in proteins or DNA/RNA and eventually coenzymes.

Thus, since 2DES experiments may provide unique information on electronic coupling and related electron/energy transfer processes (see Sect. 1.1), applications of this nonlinear spectroscopic technique in the UV range could shed new light on a vast number of biological processes, from various enzymatic redox catalysis to photosynthesis, and could be used either to obtain structural information (in the GS) or to track photoinduced excited-state events. Notably, given its high spectral and temporal resolution, 2DUV spectroscopy can be used to disentangle complex decay processes such as in DNA/RNA polynucleotides, where multiple chromophores exhibit various competitive deactivation pathways, with timescales from sub-100 fs to nanoseconds. In this context, developing computational protocols for accurate simulation of 2DUV spectroscopy is

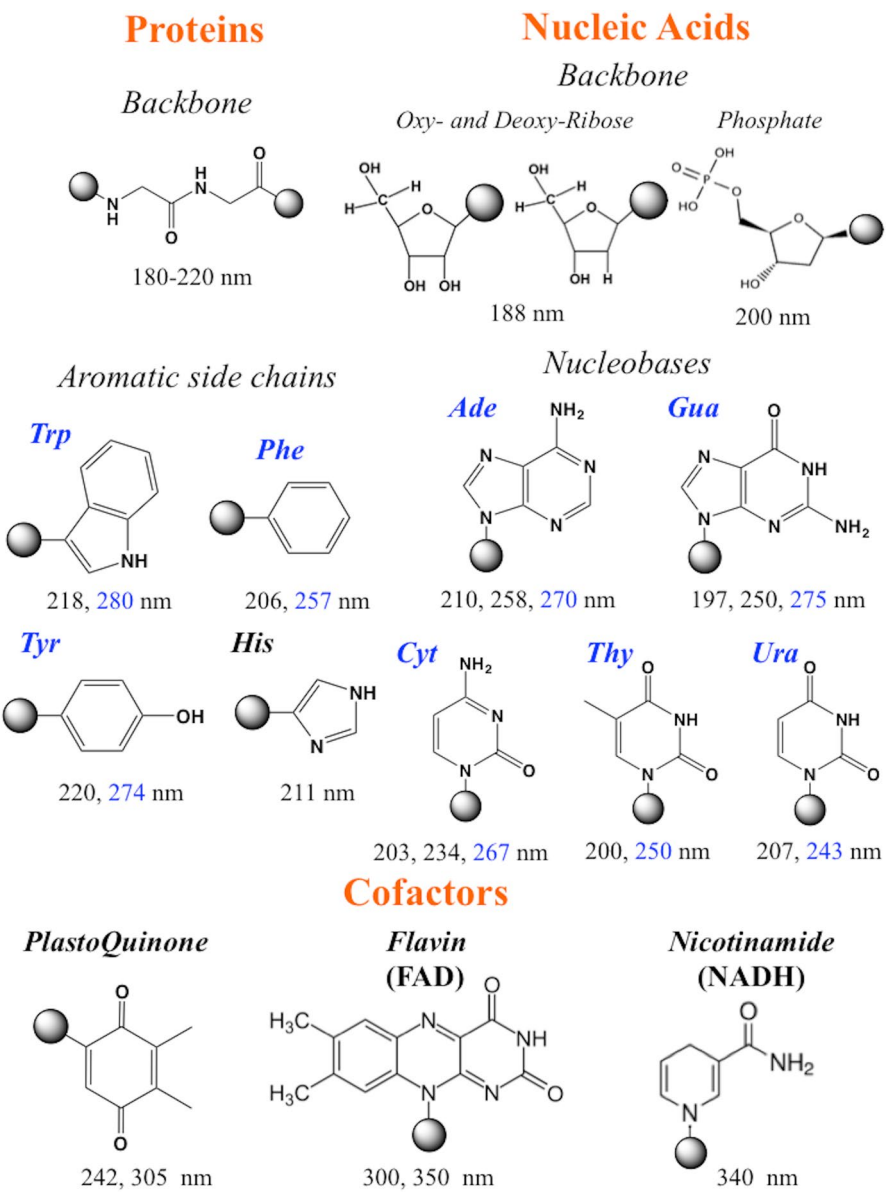

Fig. 2 UV-absorbing bio-chromophores in proteins and nucleic acids, together with associated absorption wavelengths. States involved in absorption out of the spectral region of backbones are highlighted in blue for the UV-chromophores reported in this review. Their respective backbone units, absorbing in the high-energy UV window, and important protein cofactors are also depicted for the sake of completeness

highly desirable, as together with current advances in experimental techniques, it would ultimately lead to a comprehensive understanding of these processes.

In the following sections, after presenting the basic theoretical background for 2DES and the approximation frameworks adopted in our studies, we review our recent efforts in developing computational tools for simulating 2DUV spectra of two major biological macromolecules, proteins and nucleic acids, focusing on the chromophoric units that provide absorption distinct from their backbones.

3 Theoretical Background

3.1 Density Matrix Formalism and 2DES Response Functions

A unified theory for nonlinear optical spectroscopy [2] has been developed using density matrix formalism to represent the state of matter and its evolution in the Liouville space in order to determine the nonlinear response generated by the light–matter interactions. As mentioned above, the source of the signal field recorded in 2DES experiments is the induced nonlinear polarization, which is defined as the expectation value of the dipole operator $\hat{\mu}$

$$P^{(n)}(t) = \left\langle \hat{\mu}\hat{\rho}^{(n)}(t) \right\rangle = \mathrm{Tr}\left[\hat{\mu}\hat{\rho}^{(n)}(t)\right] \tag{1}$$

where $\hat{\rho}(t)$ is the density matrix that describes the quantum state of an ensemble of optically active sites, i.e. chromophores (or more generally, molecules)

$$\hat{\rho} = \sum_{\mathrm{ab}} \rho_{\mathrm{ab}} |a\rangle\langle b| \tag{2}$$

where the sum runs over the electronic states of the chromophores, with density matrix diagonal elements representing populations and off-diagonal elements, coherences. The dipole operator for this ensemble is analogously defined as

$$\hat{\mu} = \sum_{\mathrm{ab}} \mu_{\mathrm{ab}} |a\rangle\langle b| \tag{3}$$

, and μ_{ab} is the transition dipole between states a and b.

The time evolution of the density matrix (Eq. 2) satisfies the Liouville-von Neumann equation,

$$i\hbar \frac{\mathrm{d}\hat{\rho}}{\mathrm{dt}} = [\hat{H}, \hat{\rho}] \tag{4}$$

where $\hat{H}$ is the system Hamiltonian. In the absence of an external field ($\hat{H} = \hat{H}_0$), the free evolution of an unperturbed density matrix, $\hat{\rho}^{(0)}(t)$, is conveniently described using the retarded Green's function (forward propagator) *G(t)*; see detailed description in the Appendix.

The inclusion of an external optical electric field (provided it is weak) is achieved using a perturbative approach, where the time-dependent external field perturbation$\hat{H}'(t)$, defined as

$$\hat{H}'(t) = -\hat{\mu} \cdot \mathrm{E(t)} \tag{5}$$

, is included in the system Hamiltonian, i.e. $\hat{H} = \hat{H}_0 + \hat{H}'(t)$. The density matrix is expanded as

$$\hat{\rho}(t) = \hat{\rho}^{(0)}(t) + \hat{\rho}^{(1)}(t) + \cdots + \hat{\rho}^{(n)}(t) + \cdots = \hat{\rho}^{(0)}(t) + \sum_{n=1}^{\infty} \hat{\rho}^{(n)}(t), \tag{6}$$

where the superscript denotes the *n*th-order expansion. In this perturbative scheme, the Liouville-von Neumann equation can be solved by sorting in powers of $\hat{\rho}$, followed by an iterative integration of the resulting equations (see Appendix for the explicit form of the third-order density matrix).

Writing out the perturbation (Eq. 5) and reformulating the density matrix as a function of the time intervals t_i between the pulses (with $t_1 = \tau_2 - \tau_1, t_2 = \tau_3 - \tau_2, t_3 = t - \tau_3$), the *n*th-order polarization $P^{(n)}(t) = \left\langle \hat{\mu}\hat{\rho}^{(n)}(t) \right\rangle = \mathrm{Tr}\left[\hat{\mu}\hat{\rho}^{(n)}(t)\right]$ becomes third-order,

$$P^{(3)}(t) = \mathrm{Tr}[\hat{\mu}\hat{\rho}^{(3)}(t)] = \int_0^\infty \mathrm{d}t_3 \int_0^\infty \mathrm{d}t_2 \int_0^\infty \mathrm{d}t_1 R^{(3)}(t_3, t_2, t_1) \times E(t - t_3)E(t - t_3 - t_2)E(t - t_3 t_2 - t_1) \tag{7}$$

where the system-specific (third-order) response $R^{(3)}(t_1,t_2,t_3)$

$$R^{(3)}(t_1, t_2, t_3) = \left(\frac{i}{\hbar}\right)^3 \mathrm{Tr}[\hat{\mu}G(t_3)[\hat{\mu}G(t_2)[\hat{\mu}G(t_1)[\hat{\mu}, \rho(0)]]]] \tag{8}$$

is formally separated from the incident electric fields. Equation (7) is the general equation for the computation of the nonlinear signals in 2D optical spectroscopy. Recording in a given phase-matched direction translates into the selective detection of subgroups of contributions in the Liouville space (eight in total), referred to as Liouville pathways, following directly from Eq. 8. For example, detection in the rephasing (K_1) phase-matching direction (see Fig. 1a) selects three contributions to the response

$$R^{(3)}_{k_1}(t_1, t_2, t_3) = \sum_{i=\mathrm{GSB,ESA,SE}} R^{(3)}_{k_1,i}(t_1, t_2, t_3) \tag{9}$$

that are associated with different physical processes occurring in the system during the interaction with the external fields (see GSB, SE and ESA in Sect. 1.1) that can be represented by Feynman diagrams.

Typically, simulations of ultrafast spectroscopy assume temporally well-separated ultrashort laser pulses and work in the so-called impulsive limit (i.e. the limit in which the pulse duration is shorter than a single vibrational oscillation period). In that case, the third-order polarization $P^{(3)}(t_1,t_2,t_3)$ (Eq. 7) becomes equivalent to the nonlinear response of the system $R^{(3)}(t_1,t_2,t_3)$ (Eq. 8). This approximation simplifies

the calculations, as undesired effects due to the temporal overlap of the pulses do not need to be considered. Furthermore, the obscuring of any coherent dynamics, which occurs on the same timescale as the duration of the pulses, is avoided, and the highest possible temporal resolution is achieved. The impulsive limit has been adopted throughout this review.

Simulation protocols for nonlinear optical spectroscopy deal with solving Eq. (7) or, more generally, Eq. 8 when working with realistic pulse shapes. This requires computing the interactions of the system with three time-dependent electric fields as well as its field-free evolution between them. In principle, Eq. 7 could be solved without any approximation, given knowledge of the multidimensional potentials associated with each electronic state involved in the evolution of the density matrix, enabling propagation of populations and coherences. In this case, non-adiabatic couplings and transition dipole moments along the potential energy surfaces (PES) would be required to allow population transfer and interaction with the field, respectively. It would then be possible to treat electronic and nuclear degrees of freedom fully quantum-mechanically by quantum dynamics, thereby resolving coherent oscillatory dynamics, line broadening (due to finite excited-state lifetimes) and vibrational progressions due to wave packet decoherence and recoherence. Furthermore, arbitrary pulses could be straightforwardly incorporated into the simulations, providing an opportunity to study the effects of the pulse's central frequency, bandwidth and polarization [36, 37].

The drawback to quantum dynamics simulations is the high cost of pre-computing the PES, a task that quickly becomes infeasible with an increasing number of degrees of freedom. As a workaround, simulations are often run on parameterized two- or three-dimensional potentials that approximate the exact PES through analytical expressions, with parameters chosen to fit the energetics of a few representative points along the PES. This approach was recently applied to simulate 2DES in Rhodopsin [38]. Methods such as multi-configuration time-dependent Hartree have facilitated simulations on multidimensional surfaces [39].

Another approach for solving Eqs. (7) and (8) involves the stochastic modeling of bath fluctuations in a semi-classical fashion (stochastic Liouville equations, SLE) [40, 41]. At the heart of SLE is the explicit inclusion of collective bath modes in the system's Hamiltonian. The Hamiltonian is formally time-independent, but depends on some classical time-dependent stochastic variables $\sigma(t)$, which represent the bath. The response function is calculated by averaging over an ensemble of realizations, propagated through trajectory-based molecular dynamics simulations, where at every subsequent time-step it is evaluated as

$$R^{(3)}(t_1,t_2,t_3) = \left(\frac{i}{\hbar}\right)^3 \mathrm{Tr}[\hat{\mu}G_{\sigma(t)}(t_3+t_2+t_1,t_2+t_1)[\hat{\mu}G_{\sigma(t)}(t_2+t_1,t_1)[\hat{\mu}G_{\sigma(t)}(t_1,0)[\hat{\mu},\rho(0)]]]] \tag{10}$$

where $G_{\sigma(t)}(\tau_i,\tau_j)$ is the Green's function solution of the field-free evolution between two time steps

$$\frac{\mathrm{d}\hat{\rho}_{\sigma(t)}}{\mathrm{dt}} = \frac{i}{\hbar}\left[\hat{H}_{0;\sigma(t)},\rho_{\sigma(t)}\right] \tag{11}$$

Since bath dynamics is non-Markovian, solving Eq. 10 requires knowledge of the past evolution of the system. Bulk-induced dephasing effects are automatically included via ensemble averaging. Trajectory-based approaches face the challenge of the computational cost associated with the propagation of a swarm of trajectories (particularly those in the excited state) and the absence of a well-defined density matrix responsible for the quantum feedback on the classical bath during a coherence state evolution. Various trajectory-based implementations that address these issues have been documented in recent years [42–45].

For systems with a classical bath following Gaussian statistics and linear system–bath coupling, Eqs. (8) can be solved using the second-order cumulant expansion, i.e. the cumulant expansion of Gaussian fluctuations (CGF) [46]. This method makes it possible to calculate the shapes of electronic transition bands coupled to a bath (for fluctuations with arbitrary timescales) using the formalism of line shape functions, $g_{ij}(t)$; see Appendix for details. Within the CGF framework, the population transfer can be accounted for phenomenologically according to the Lindblad equation (see Appendix), with secular approximation to the Green's function enabling partitioning of the nonlinear response (Eqs. 9, 10) into population and coherence contributions. For instance, considering the manifold of ground (g) and excited (e, f) states, and just the ESAs detected in the rephasing phase-matching direction (with the expressions for GSB and SE given in Ref. [36]), the population ($e' = e$ during delay time t_2) contributions are stated as

$$R^{(3)}_{k_1\mathrm{ESA},i} = +i \sum_{e',e,f} \mu^2_{fe'} \mu^2_{ge} G_{e'e',ee}(t_2) \times e^{-i(\varepsilon_f - \varepsilon_{e'})t_3 + (\varepsilon_e - \varepsilon_g)t_1 + \varphi^{\mathrm{ESA},i}_{fe'e}(t_1, t_1+t_2, t_1+t_2+t_3, 0)} \tag{12}$$

where $G_{e'e',ee}(t)$ is the Green's function controlling the population transport (see Appendix for additional details), while the coherence ($e' \neq e$ during delay time t_2) contributions are instead stated as

$$R^{(3)}_{k_1\mathrm{ESA},ii} = +i \sum_{e',e\,f}^{e' \neq e} \sum_{f} \mu_{fe'} \mu_{fe} \mu_{ge'} \mu_{ge} e^{-i(\varepsilon_f - \varepsilon_e)t_3} \times e^{-i(\varepsilon_{e'} - \varepsilon_e)t_2 + (\varepsilon_e - \varepsilon_g)t_1 + \varphi^{\mathrm{ESA},ii}_{fe'e}(t_1, t_1+t_2, t_1+t_2+t_3, 0)} \tag{13}$$

where ε_a, with $a \in \{g, e, f\}$, are the energies of the eigenstates of the system's Hamiltonian H_0, their energy differences ($\varepsilon_a - \varepsilon_b$) being the transition energies (TEs), μ_{ab} are the associated transition dipole moments (TDMs), and $\varphi_{e'fe}(\tau_4, \tau_3, \tau_2, \tau_1)$ is the phase function that describes spectral diffusion, thus translating the vibrational structure of the evolving electronic states into a series of oscillating diagonal and off-diagonal peaks. The mathematical formulation of the phase functions therefore depends on the level of sophistication adopted for describing the system vibrational dynamics [2, 47]. In any case, the phase function is based on the line shape functions, and two examples of phase functions built from line shape functions will be given in Sects. 3.3 and 3.4 (Eqs. 16 and 19, respectively).

3.2 Coupling Accurate Electronic Structure Computations to 2DES

Two-dimensional electronic spectroscopy targets the manifold of (localized and delocalized) excited states of the chromophoric units in the sample. A thorough characterization of the electronic structure of the chromophoric systems is thus required to properly interpret 2D electronic spectra, which generally represents a great challenge even for the most advanced computational techniques, despite a plethora of quantum mechanics (QM) methods currently available. In particular, the QM treatment of isolated large chromophoric moieties (usually absorbing in the Vis) or rather small UV-chromophores that still involve high-energy electronic levels and thus encompass a large number of excited states in the target manifolds, might already push the limits of what is currently computationally feasible. Treating multiple interacting chromophores, i.e. molecular aggregates, at the ab initio level then quickly becomes prohibitive with increasing number or size of chromophoric units. At the same time, as shown in Eqs. 12 and 13, TEs and TDMs are the fundamental ingredients for simulating the third-order nonlinear response recorded in 2DES maps, and even if largely approximate, their estimations cannot be circumvented. A common strategy for coping with these limitations is to adopt Frenkel exciton models that make simulations of 2DES spectra of realistic model systems computationally feasible [36, 48]. Exciton modeling works well for vibrational excited-state absorptions in 2DIR spectroscopy (namely overtone absorptions), since anharmonicity for overtones of a chromophore (local overtones) or between coupled chromophores can be reliably computed perturbatively (without explicit knowledge of the high-lying energy level manifold). In general, such a perturbative treatment is less effective for electronic transitions, since local excited states behave quite differently with respect to local overtones, showing large anharmonic couplings that cannot be described perturbatively. Still, when dealing with coupled electronic excitations between interacting chromophores, the TEs and TDMs of each isolated chromophore can be initially calculated at the QM level (usually in the gas phase and for a given geometry) and then employed as parameters to build the exciton Hamiltonian. The electronic couplings in multi-chromophoric systems can then be estimated, while neglecting electron exchange between chromophores. Clearly, the limited description provided by Frenkel exciton modeling has computational advantages that become evident only when just a few excited states (i.e. energy levels) are considered in the model Hamiltonian (i.e. few QM computations are initially performed). This implies that many of the excited states related to (potential) ESA signals in 2DES maps are often neglected in most conventional spectroscopy simulation protocols. This assumption generally holds if ESA signals are expected to fall outside the spectral window of interest. However, broadband transient absorption [49] and, more recently, also 2D [16, 50, 51] spectra show that this assumption breaks down regularly, especially in UV-active bio-chromophores. In fact, as we will show in the following sections, ESAs are ubiquitous, system-dependent, state-specific spectroscopic fingerprints that, for instance, make it possible to selectively study the excited-state dynamics of different decay channels. In cases such as charge-transfer states, ESAs might represent the only spectroscopic signature of a given state, being of particular interest for studying its photophysics.

The ability to accurately predict ESA signals would then open the door for simulating broadband 2D electronic spectroscopy [37, 52–61] as well as double quantum coherences spectroscopy [53, 62, 63].

As will be illustrated throughout Sect. 3, obtaining the electronic structure of biologically relevant chromophores from first principles, at a level of accuracy needed for comparison against experimental data, is far from trivial. A protocol for computing reliable TEs and TDMs of isolated bio-chromophores (or dimeric aggregates) is presented in Sect. 3.1, which comes with additional benefits: (a) the results can be used to expand currently available exciton models beyond the lowest transitions; (b) the results can be used to benchmark low-cost approaches that can then be applied for larger-scale computations; (c) the results essentially make simulations independent from the experiment, providing them with predictive power and allowing us to envision problem-driven experimental setups. To this aim, we take a step back in the theoretical descriptions given in Sect. 2.2 and start from a drastic approximation, i.e. treating the chromophoric system uncoupled from the bath of vibrations and thus as a closed quantum system having only electronic degrees of freedom. Within this framework the electronic states become eigenstates of the system's Hamiltonian, and their dynamics are reduced to a $e^{-i\varepsilon_f t}$ phase factor (herein referred to as a "static picture").

This approximation implies setting the phase functions $\varphi = 0$ and neglecting population transport $G(t)$ in Eqs. 12 and 13, which then simplify to

$$R^{(3)}_{k_1\mathrm{ESA},i} = +i \sum_{e,f} \mu^2_{fe}\mu^2_{ge} e^{-i(\varepsilon_f - \varepsilon_e - i\gamma_{fe})t_3 + i(\varepsilon_e - \varepsilon_g - i\gamma_{eg})t_1} \tag{14}$$

and

$$R^{(3)}_{k_1\mathrm{ESA},ii} = +i \sum_{e',e}^{e'\neq e} \sum_{f} \mu_{fe'}\mu_{fe}\mu_{ge'}\mu_{ge} e^{-i(\varepsilon_f - \varepsilon_e - i\gamma_{fe})t_3} e^{-i(\varepsilon_{e'} - \varepsilon_e - i\gamma_{e'e})t_2 + i(\varepsilon_e - \varepsilon_g i\gamma_{eg})t_1} \tag{15}$$

All dephasing processes are thus condensed into the phenomenological dephasing constants γ_{ab}, which induce homogeneous broadening of all $a \rightarrow b$ electronic transitions. Equations 14 and 15 allow us to simulate the nonlinear responses for systems with hundreds of excited states, assuming all the corresponding TEs ($\varepsilon_b - \varepsilon_a$) and TDMs ($\mu_{ab}$), a protocol known as the sum-over-states (SOS). Figure 3 shows in greater detail, for the specific case of an ESA signal recorded in the rephasing K_1 phase-matching direction (Fig. 3a), how the nonlinear response is computed through Eq. 14. The final relation obtained for the third-order nonlinear response is accomplished by looking at the various contributions adding up in time, following the evolution of the system density matrix elements as described by the corresponding (ESA) Feynman diagram (Fig. 3b) and by making explicit the terms contributing to the density matrix evolutions during each light–matter interaction (Fig. 3c) during a 3PPE experiment (see Fig. 1a).

Recently [52], we combined this quasi-static (as all dynamic effects have been neglected) protocol with a hybrid scheme combining a QM electronic structure with a molecular mechanics (MM) treatment of environmental effects (i.e. the SOS//QM/MM

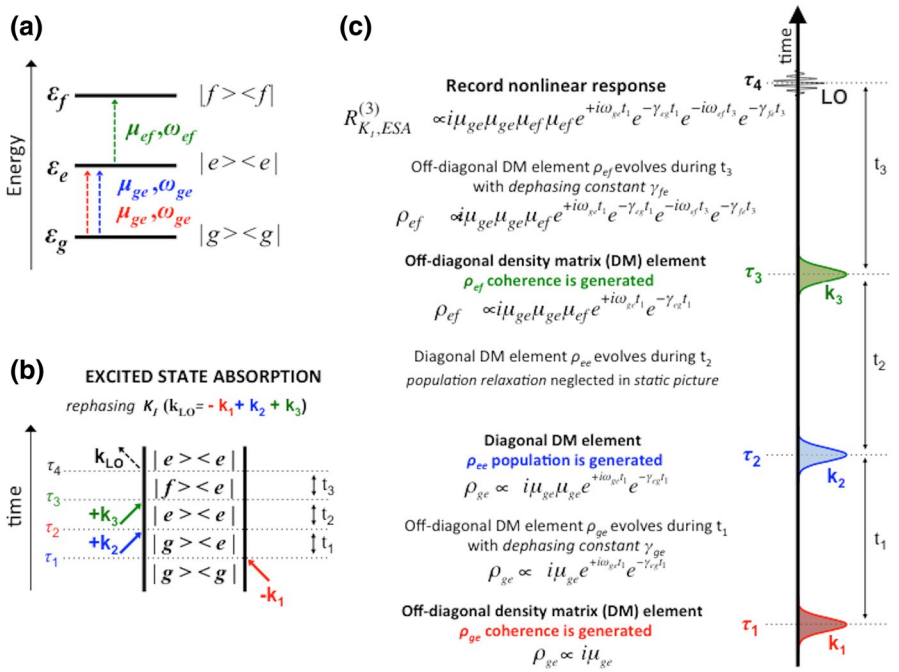

Fig. 3 **a** A three-level system, showing transition dipole moments (μ) and frequencies (ω) related to the ground (g) and excited (e, f) states and their energies ($\varepsilon_{g,e,f}$). **b** Feynman diagram of an ESA signal recorded in the rephasing K_I phase-matching direction. **c** Schematic representation of a 3PPE experiment (see also Fig. 1a), including explicit terms contributing to the density matrix evolutions during each light–matter interaction and the final relation for the third-order nonlinear response (see Eq. 14)

protocol) in order to generate 2DES spectra for several biological key players in their native environment. Some applications of the SOS//QM/MM protocol to bio-chromophores are showcased in the Results section, mainly evidencing how a number of bright excited states in the UV–Vis spectral range often represent the major contribution to 2DES maps and could embody specific spectroscopic fingerprints for tracking ground- or excited-state dynamics. Accurate characterization of the excited-state manifold by advanced electronic structure computations thus represents a pillar supporting the reliable simulation of 2D electronic spectra.

4 Results and Discussion

The basic theoretical background for understanding how to compute nonlinear electronic spectra has been illustrated in Sect. 2, showing how a complete 2DES simulation technique would need to combine high-precision electronic structure

computations and non-adiabatic quantum dynamics simulations to resolve the elaborate nonlinear response equations. On one hand, the practical use of quantum dynamics makes it possible to accurately model spectral dynamics while relying only on approximate electronic potentials, capturing just the essence of the electronic structure complexity, which is reflected in the absence of ESA contributions. On the other hand, the electronic structure complexity can be straightforwardly incorporated within a static approach, recovering ESA contributions. However, as dynamics are neglected, this approach is not able to provide realistic spectral line shapes. Trajectory-based approaches offer a compromise by permitting us to adjust the level of precision applied to molecular dynamics and PES sampling. In the following section, we begin our discussion by utilizing the static approach to outline a protocol for the accurate computation of transition energies and dipole moments (Sect. 3.1), as well as for benchmarking low-cost methods. Sections 3.2 and 3.3 demonstrate how the latter can be employed to study GS conformational dynamics of oligopeptides and DNA nucleobase dimers within the framework of the SOS//QM/MM protocol. We then go beyond the static approximation and re-introduce dynamic features in Eqs. 14 and 15 within the CGF framework for the study of excited states. Specifically, coherent intramolecular vibrational dynamics is discussed in Sect. 3.3, while Sect. 3.4 covers population transfer.

4.1 Benchmarking the Excited-State Manifolds

Accurate prediction of the basic spectroscopic parameters that define the space of electronic transitions (i.e. TEs and TDMs) accessible by 2DES experiments is the first computational challenge. The excited-state manifolds of UV-active chromophores in the 2DUV energy range (from ca. 3.5 to 11 eV) comprise various types of electronic transitions, generally including single and double excitations with local or (inter- and intramolecular) charge-transfer character, and thus involving covalent as well as ionic (valence bond-like) excited states. Wave function approaches introduced in the 1980s, based on the combination of CASSCF multi-configurational wave functions [64] and PT2 perturbative energy corrections [65] (CASSCF//CASPT2), currently represent the most widely used methodology for handling such a variety of excited states on equal footing [66], generally providing good quantitative predictions of TEs and TDMs, with expected error of around 0.2 eV (ca. 1600 cm^{-1}). Application of the CASSCF//CASPT2 methodology to larger and larger molecules has become possible through many developments over the years, including the implementation of the restricted active space self-consistent field (RASSCF) methodology [67], efficient approximations for two-electron integral estimates [68] and large-scale parallelization [69], to cite some of those most widely used in the results reported here. The single-state PT2 treatment (hereafter "PT2") has likely been preferred to the more expensive (and in most cases infeasible) multistate PT2 approach. However, the accuracy of CASSCF//PT2 predictions strongly depends on the choice of active spaces and basis sets, two parameters that involve a critical increase in computational cost. Moreover, the number of excited states present in

the energy ranges of 2DUV experiments increases dramatically with the size of the (multi-)chromophoric system under consideration, requiring the use of large state-averaging procedures to converge the multi-configurational wave functions, with a consequent further increase in computational cost. To cope with the computational feasibility issues of multi-configurational treatments, we have generally adopted the RASSCF methodology, as its flexibility makes it possible either to improve the electronic wave function of the largest feasible CASSCF computation, usually involving the full-valence π-orbital space, or to reduce the overall computational cost by limiting the CAS treatment to a relatively small active space (RAS2).

Indeed, with the RASSCF approach, extra virtual orbitals can be included in the active space where a restricted number of electrons are allowed (RAS3 space), increasing the number of electronic excitations and expanding the multi-configurational character of the electronic wave functions relative to what is already accounted for in a full-valence RAS2 space. This type of RAS scheme can be employed to benchmark $S_0 \rightarrow S_N$ vertical excitations of chromophore monomeric units. To this end, we have generally attempted to effect a direct comparison between vertical excitations in the Franck–Condon (FC) region, i.e. at the GS equilibrium geometries, and experimentally recorded cross-sections in a vacuum, when available. The choice of the gas phase reflects the need for comparison in the absence of environmental effects that would both complicate the computational modeling strategy and reduce the experimental spectral resolution (as line broadening is larger in solution), and is aimed at assessing the prediction of the overall distribution of the absorption bands in the whole UV range, keeping in mind that vibronic effects [70] are not accounted for in this study.

Benzene and phenol molecules are the chromophores related to the Phe and Tyr amino acid aromatic side chains (see Fig. 2). As shown in Fig. 4, CASSCF//PT2 computations of the $S_0 \rightarrow S_N$ vertical excitations (in the FC region) indicate that the full-valence active spaces of benzene and phenol, i.e. CAS(6,6) and CAS(8,7), respectively, are not sufficiently accurate when compared with the values of experimental cross-sections in a vacuum [58]. Instead, increasing the active space by allowing single and double excitations to additional extravalence RAS3 orbitals enables an appropriate description of the dynamic σ–π polarization and significantly improves accuracy, reaching convergence with eight extra orbitals and quantitative agreement with experiments. It is worth mentioning that these computations have been performed applying a computational recipe to eliminate Rydberg contamination in the valence states after adding a set of diffuse and uncontracted basis functions in the center of charge of the molecules (ANO-L-aug) [71, 72], which has been further used to simulate dimers of amino acid side chains and other chromophores.

While the $S_0 \rightarrow S_N$ excitations do not cover all possible excitations involved in a 2DUV experiment, because the $S_N \rightarrow S_M$ excitations would strongly contribute as ESA signals in the 2D spectra, the energy range considered in this benchmark study involves quite large TEs (with energies up to 11 eV from the GS). In fact, considering a one-color 2DUV experiment with pump and probe frequencies centered at the $S_0 \rightarrow S_1$ transition energy (4.5–4.7 eV), the S_M excited states considered here already comprise those that could be involved as ESA signals (i.e. the $S_1 \rightarrow S_M$ transitions, with $M < 9$) and are properly described at the RASSCF//PT2

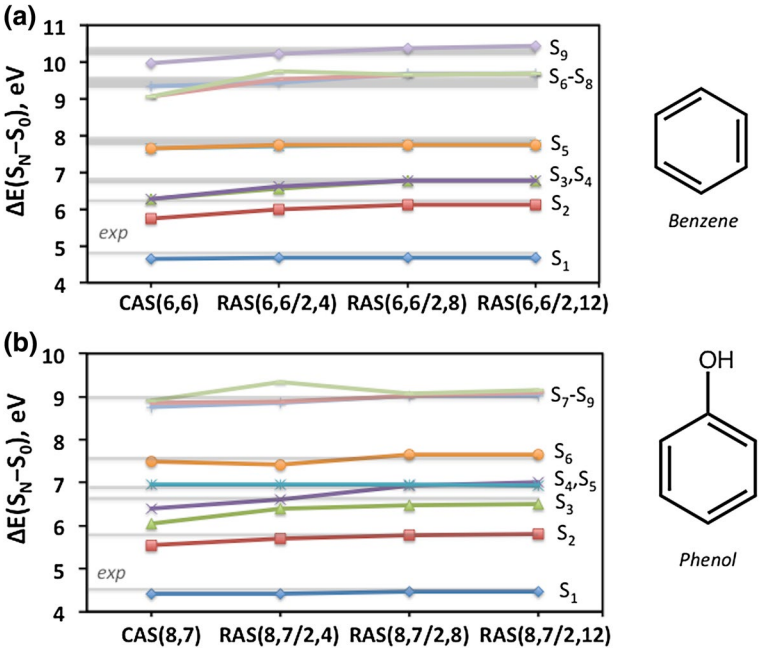

Fig. 4 CASSCF//PT2 and RASSCF//PT2 vertical $S_0 \rightarrow S_N$ excitation energies (in eV) of benzene and phenol monomers in a vacuum obtained using the large ANO-L(432,21)-aug basis set at different levels of theory. Minimal active space CAS(X,Y) results are compared to restricted active space (RAS2/RAS3) calculations with an increasing number of virtual orbitals (4, 8 and 12) and experimental values (gray bars, indicating ranges of experimental absorption maxima recorded). Reproduced from data reported in Ref. [58]

level. This outcome suggests that theoretical predictions of vertical excitations in the UV can be reliably obtained by extended RAS schemes. Notably, convergence was shown by TDMs as well as TEs, thus providing robust theoretical and reference RAS schemes for employment in further studies.

When extending the benchmark study of benzene and phenol monomers to the remaining aromatic protein chromophores, i.e. the indole moiety of the Trp amino acid, and to the canonical nucleobases of DNA and RNA, we found a lack of experimental data. In fact, the gas-phase experimental cross-sections of indole, Ade, Gua, Cyt, Thy and Ura are limited to energies up to 6.4 [73], 7.7 [57, 74], 4.4 [75], 7.1 [75–77], 7.4 [78] and 6.6 eV [75], respectively. Therefore, after a preliminary comparison of $S_0 \rightarrow S_N$ vertical excitations against the available cross-section values, the benchmark study of these chromophores was conducted monitoring the $S_N \rightarrow S_M$ excitations, giving rise to ESA signals that mainly contribute to the 2D electronic spectra. In particular, the S_N states are chosen among the lowest-lying and brightest (i.e. with largest $S_0 \rightarrow S_N$ TDM) excited states of each chromophore, as these states are generally the targets of experimental 2DUV spectra. This approach provides a direct visualization of the influence of the level

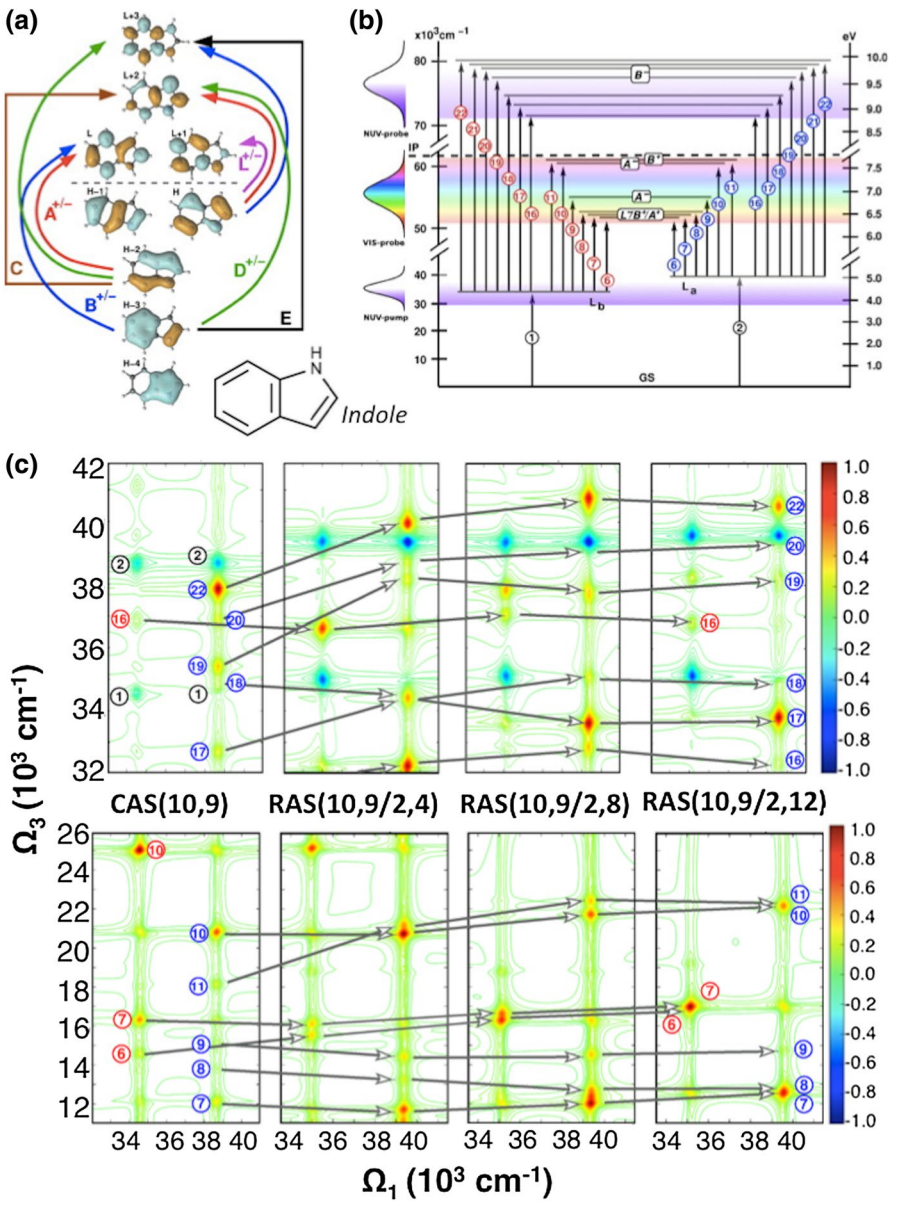

Fig. 5 Benchmark computations on indole: **a** molecular orbitals in the CAS(10,9) active space displaying a labeling scheme for **b** the various types of transitions associated with ESAs in the Vis and UV probing (Ω_3 frequency) windows, as arising from the primary $\pi\pi^*$ (L_b and L_a) states (in the pump frequency range $\Omega_1 = 34{,}000\text{-}40{,}000$ cm^{-1}). **c** 2D electronic spectra simulated at the CAS(10,9), RAS(10,9|2,4), RAS(10,9|2,6) and RAS(10,9|2,12) levels of theory, using the ANO-L(432,21)-aug basis set. Note that in this RAS nomenclature, the empty RAS1 subspaces are omitted for simplicity. Negative (blue) signals are related to bleaching, while positive (red) peaks arise from $S_{1,2} \rightarrow S_M$ ESAs. Reproduced from data reported in Ref. [53]

of theory adopted on the resulting 2D map, which is shaped by both TE and TDM computed values.

Figure 5 summarizes the benchmark study of indole [53] performed with a large basis set [ANO-L(432,21)-aug], indicating (as in benzene and phenol) that going beyond a CASSCF treatment with full-π valence active space (by introducing at least eight extravalence orbitals in the RAS3 subspace) is required in order to converge the energies and dipole moments of the electronic transitions involved and, thus, the overall aspect of the 2D spectra (see Fig. 5c). Both one-color (UV-pump and UV-probe, 2DUV–UV) and two-color (UV-pump and Vis-probe, 2DUV–Vis) 2D spectra are reported in Fig. 5c, showing a broad set of ESA and bleaching signals that vary in position and intensity as a function of the level of theory. Overall, in the 2DUV spectra of indole, two signal traces related to the first two ($\pi\pi^*$) excited states, labeled L_b and L_a in Platt notation [79], are monitored, as they are quite close in energy, and their transitions from the GS have similar TDMs. In particular, 2DUV–UV spectra (see Fig. 5) feature off-diagonal bleaching signals (ODB, see Fig. 1) related to couplings between L_b and L_a states. It is worth mentioning that these signals are characteristic 2DUV fingerprints of excited states localized on the same molecule. They can be seen as strongly coupled "chromophores", where the energy of their "bi-exciton" (i.e. the state given by the sum of the two electronic transitions) is significantly shifted from the sum of the single-exciton energies due to the large anharmonicity constant, which cannot be computed perturbatively with exciton Hamiltonians.

Here, we consider the 2D map converged with respect to the levels of theory when the intensities of all the peaks are maintained and their positions do not vary more than 1600 cm^{-1}, i.e. the expected error in TE for this methodology (0.2 eV). It is worth mentioning that such type of benchmark study requires a full assignment of the 2D peaks in order to assess the variations in the 2D maps according to the theoretical level employed. This work is particularly tedious, as it requires labeling of all relevant transitions (Fig. 5b) according to the molecular orbitals (MO) involved in the electronic excitations.

Figure 6 shows the benchmark computations for the Ura, Thy and Cyt pyrimidine nucleobases and the effect of the larger active space on the 2DUV–UV spectra [59]. All these canonical nucleobases feature the same π-orbital valence AS, i.e. CAS(10,8), while the RAS3 spaces involving 4, 8 and 12 virtual orbitals and the large ANO-L(432,21)-aug basis set have been employed, as for the benchmarks of amino acid side chains (Figs. 4, 6, 5). While some ESA signals show more AS-dependent fluctuations than others, the overall the 2DUV–UV spectra converge with the addition of eight extravalence orbitals. Cyt is particularly interesting, as the first excited state is very bright but is not the brightest low-lying $\pi\pi^*$ state (as in the other pyrimidine bases), giving rise to two signal traces along two relatively close-lying pump frequencies. This occurrence offers the opportunity to monitor the off-diagonal GSB signals related to the two low-lying $\pi\pi^*$ states (as in indole and purine bases) and the performances of the various levels of theory in producing reliable 2DUV–UV maps.

Adenine and guanine are the canonical purine nucleobases, which feature larger aromatic heterocyclic structure than pyrimidine bases, with the pyrimidine ring

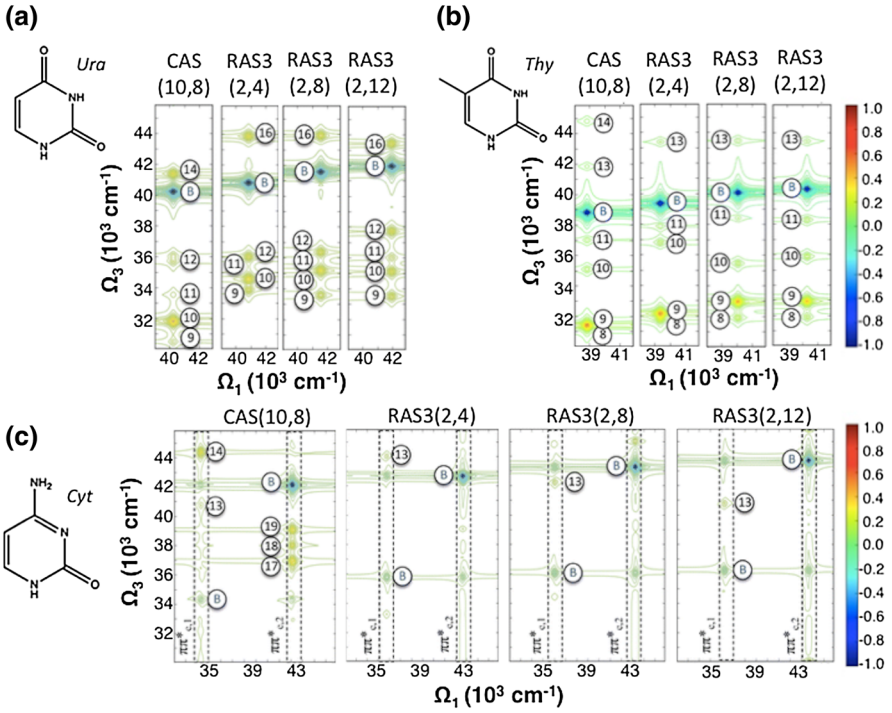

Fig. 6 Benchmark computations on pyrimidine nucleobases **a** uracil, **b** thymine and **c** cytosine, showing the 2DES maps in the UV-pump/UV-probe window (2DUV–UV) as computed at the CAS(10,8), RAS(10,8|2,4), RAS(10,8|2,8) and RAS(10,8|2,12) levels of theory and using the ANO-L(432,21)-aug basis set. Note that only the (variable) RAS3 spaces are indicated in the picture, for simplicity. Reproduced from data reported in Ref. [59]

fused to an imidazole ring. Adenine has just one more π-orbital in the valence active space than indole (MOs depicted in Fig. 5a), i.e. CAS(12,10), and similar MOs that give rise to analogous electronic transitions (Figs. 5, 7). As with indole, adenine features two close-lying $\pi\pi^*$ excited states, but in the purine base the GS → L_a transition has higher TDM than GS → L_b. As shown in Fig. 7b, the benchmark of 2DUV–UV maps demonstrates remarkable changes when the active space is increased with respect to the full valence space, with positions of the Ω_1 traces more sensible to the level of theory than in indole, while convergence is still reached when eight extravalence orbitals are added in a RAS(12,10|2,8) scheme.

A comparison of the computed $S_0 \rightarrow S_N$ excitation energies with the available experimental cross-sections indicated unexpected discrepancies in the case of adenine [57] relative to other aromatic systems. Therefore, for this nucleobase, the benchmark study was extended to investigate the role of the geometry optimization methodology used to define the GS equilibrium structure. In particular, the lack of dynamic correlation at the CASSCF level introduces significant shortening of double bonds in the Ade heterocycle that are significantly elongated when optimizing the geometry at density functional theory (DFT) or second-order

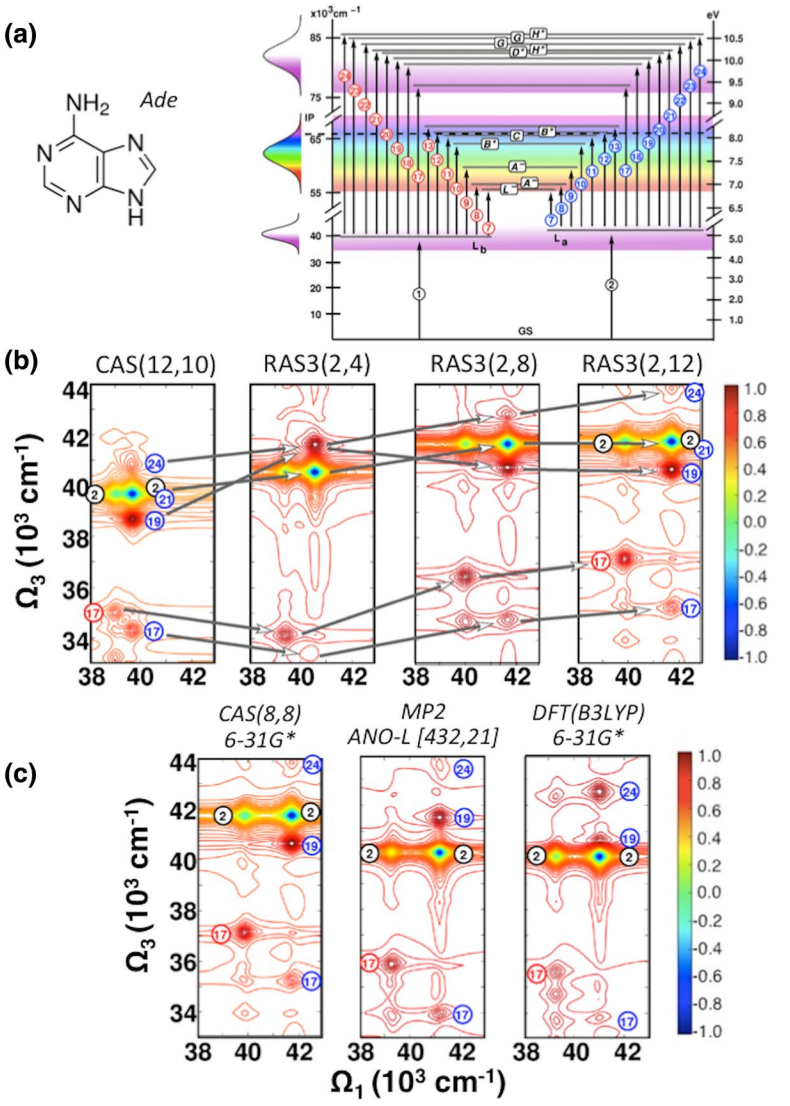

Fig. 7 Benchmark computations of the adenine excited-state manifold (**a**), showing the 2DES maps in the UV-pump/UV-probe window (2DUV–UV) as computed (**b**) at the CAS(12,10), RAS(12,10|2,4), RAS(12,10|2,8) and RAS(12,10|2,12) levels of theory and using the ANO-L(432,21)-aug basis set, or (**c**) at the RAS(12,10|2,12) level with the ANO-L(432,21)-aug basis set while using the GS geometry optimized at the CAS(8,8)/6-31G*, MP2/ANO-L(432,21)-aug and DFT(B3LYP)/6-31G* levels. Note that only the (variable) RAS3 spaces are indicated in the picture, for simplicity. Reproduced from data reported in Ref. [57]

Møller–Plesset perturbation theory (MP2) levels. As shown in Fig. 7c, indeed, the choice of geometry optimization method may have a significant effect on the simulations of 2D electronic spectra, and it should be considered as another factor that could affect comparisons with experimental data.

Thus far, we have reviewed the benchmark studies of the aromatic protein chromophores and the canonical nucleobases of DNA and RNA (with the exception of Gua, which is a work in progress), aimed at obtaining converged results among the highest (still computationally feasible) levels of multi-configurational treatment for each chromophore so that reference theoretical data are made available. These reference transition energies and dipole moments could indeed be used both as ab initio parameters in exciton Hamiltonians and as target values for benchmarking computationally cheaper approaches, in order to tackle 2DUV spectra simulation of larger (multi)chromophoric systems and/or several of their structural conformations. To this end, we have used computational recipes to account for the σ-π polarization effects when dealing with reduced active spaces [71, 72] (as necessary for increased system size). In the benchmark study reported above, we observed that reduced active space schemes (i.e. lacking virtual orbitals with higher orbital momenta) can dramatically underestimate transition energies, as they overestimate the dynamic correlation of ionic (in valence bond terms) states. In the search for cost-effective protocols to counteract or at least damp this effect, we came up with two semi-empirical procedures. On the one hand we observed that by deleting a number of virtual extravalence π^*-orbitals with higher angular momentum in the perturbation treatment, the number of dynamic correlations is reduced, producing a state-dependent blue-shift of the excitation energies, thus minimizing the mean deviation from the reference values in the whole excited-state manifold. On the other hand, we realized that real and imaginary shift parameters (originally introduced to cure intruder state problems) [80, 81] invoke a non-uniform decrease in the correlation contribution, much more pronounced for ionic than for covalent states. This makes their use well suited for our purposes, even if we need to resort to larger values than suggested in the literature (a detailed argumentation can be found in Ref. [82]). However, the choice of shift parameter is chromophore-dependent, and its application, while simpler with respect to orbital removal procedure, is limited to systems where transition energies of different chromophores can be corrected with a similar shift parameter, such as homodimeric systems. We must stress that the application of either of the two protocols makes sense only in the framework of semi-empirical parameterization against a reliable reference data set, and their use is otherwise discouraged.

As an example case study, here we report a benchmark for benzene and phenol and their dimer in the gas phase and for their corresponding amino acid side chains (Phe and Tyr, respectively) in a model tetrapeptide in water solution [58].

As shown in Fig. 8, the reference TEs computed for benzene and phenol (see Fig. 4), i.e. those at RAS(6,6|2,12) and RAS(8,7|2,12) levels with ANO-L(432,21)-aug basis set, respectively, can be used as target values for determining computationally cheaper approaches that could accurately simulate 2DES spectroscopy of related multi-chromophoric systems. In fact, reference computations with large RAS schemes are unaffordable even for the smallest multimeric system comprising benzene and phenol monomers, i.e. the benzene–phenol dimer. The first step towards reducing the computational costs is thus to develop a recipe that at least allows the use of minimal (still full-valence) active spaces (*m*AS), i.e. CAS(6,6) and CAS(8,7) for benzene and phenol, respectively, and relatively small basis sets, without compromising TEs and TDMs estimates. As shown in Figs. 4 and 8, the full-valence *m*AS in combination with the large basis set,

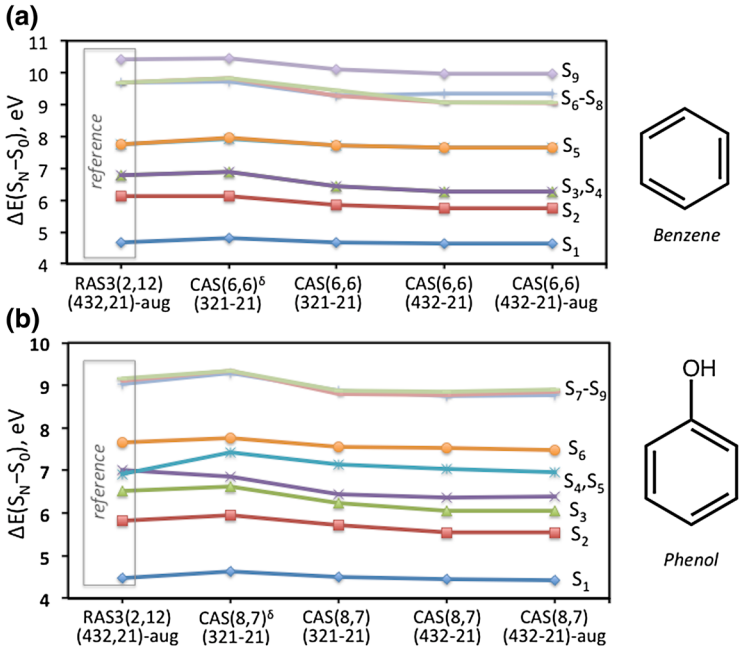

Fig. 8 CASSCF//PT2 and RASSCF//PT2 vertical $S_0 \rightarrow S_N$ excitation energies (in eV) of benzene and phenol monomers in a vacuum obtained at various levels of theory. Reference data are those computed at the RAS3(2,12)//PT2/ANO-L(432,21)-aug levels and are compared to cheaper computational approaches, including minimal (full-valence) active spaces (*m*AS), i.e. CAS(6,6) and CAS(8,7) for benzene and phenol, respectively, with various ANO-L basis sets and refined *m*AS (r-*m*AS)$^{\delta}$ with the smallest ANO-L(321-21) basis set. Reproduced from data reported in Ref. [58]

i.e. ANO-L(432,21)-aug, results in a large underestimation of several excited-state energies (i.e. those with ionic character), and this discrepancy is essentially due to $\sigma-\pi$ polarization effects associated with the smaller active space size, with basis set size having a minor effect. Deleting a number of virtual extravalence π^*-orbitals with higher angular momentum (e.g. six for benzene and seven for phenol) in the perturbation treatment reduces the dynamic correlations, significantly improving the performances of the *m*AS (Fig. 8). This active space refinement procedure allows us to obtain reasonably accurate excited-state energies of benzene and phenol monomers in the gas phase. Thus such type of refined-*m*AS (r-*m*AS)$^{\delta}$, i.e. CAS(6,6)$^{\delta}$ and CAS(8,7)$^{\delta}$ for benzene and phenol, respectively, could be employed in conjunction with the ANO-L(321,21) basis set, yielding a significant reduction in computational cost and allowing us to move forward to the study of small multimeric systems.

Extending the refined-*m*AS approach from the benzene and phenol monomers to their dimer has proven to work well for a non-interacting aggregate in the gas phase, a model system with distant (~ 11 Å) monomers that allows direct comparison with reference calculations and experimental data on monomers (Fig. 9) [58]. The full-valence *m*AS for the benzene–phenol dimer is the CAS(14,13) space, i.e. the sum of the two *m*AS of the monomers, which already comprises a large number of configuration state functions (CSFs), i.e. 736'164 CSFs. Simulation of the 2DUV–UV

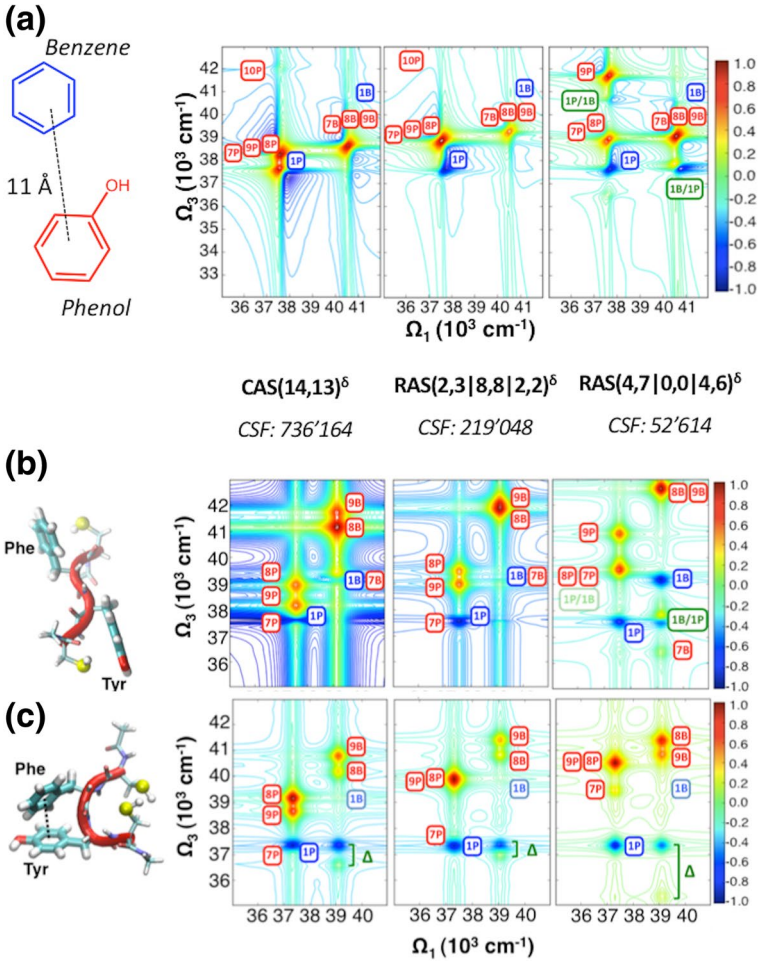

Fig. 9 Comparison of simulated one-color 2DUV–UV spectra of benzene–phenol dimers with different active spaces, including CAS(14,13)$^{\delta}$ and RAS(2,3|8,8|2,2)$^{\delta}$ and RAS(4,7|0,0|4,6)$^{\delta}$ schemes and using the ANO-L(321,21) basis set. The numbers of CSFs are also indicated for each level of theory. Spectra are shown for **a** a non-interacting dimer in the gas phase, and **b** non-interacting and **c** stacked dimers in the water-solvated CFYC tetrapeptide model. The signal labels (*n*B, *n*P) refer to benzene (B) or phenol (P) Sn excited states, while cross-peaks arising from "artificial" quartic coupling (Δ) are indicated as 1B/1P and 1P/1B. Reproduced from data reported in Ref. [58]

spectrum of the non-interacting dimer with the refined-*m*AS CAS(14,13)$^{\delta}$ provides a 2D map that is consistent with the sum of the 2D maps of the monomers computed at the CAS(6,6)$^{\delta}$ and CAS(8,7)$^{\delta}$ levels [58]. This preliminary check indicated that the CAS(14,13)$^{\delta}$ 2DUV–UV spectrum of the non-interacting dimer, in turn, could be used as a reference spectrum, to further reduce the computational cost. The CAS(14,13)$^{\delta}$ computation of the dimer, indeed, almost reaches the limit of a reasonable computational time with current computer capabilities, and more efficient protocols would be beneficial. In this context, we have adopted the RASSCF

scheme to reduce the computational cost. In fact, a small (or even empty) RAS2 space could be adopted, while including the remaining valence orbitals in the RAS1/RAS3 active spaces, where a restricted number of holes are allowed in the RAS1 space and the corresponding electrons are promoted to the RAS3 orbitals. This type of RAS1|RAS2|RAS3 scheme has been tested to minimize the active space and the computational cost for simulation of the non-interacting benzene and phenol dimer. Thus, here, we apply the RASSCF protocol to reduce the active space (and consequently the computational cost) by limiting the number of excitations, opposite to what is done for the reference computations, where the number of excitations has been increased for improving the accuracy.

Figure 9 shows the comparison of simulated one-color 2DUV–UV spectra of benzene–phenol dimers obtained with different refined active spaces, including the refined-*m*AS, i.e. CAS(14,13)$^{\delta}$, and the computationally cheaper RAS(2,3|8,8|2,2)$^{\delta}$ and RAS(4,7|0,0|4,6)$^{\delta}$ schemes, while using the relatively small ANO-L(321,21) basis set. As compared to the dimer refined-*m*AS, the RAS spaces involve much less CSF, with the RAS(2,3|8,8|2,2) and the RAS(4,7|0,0|4,6) spaces corresponding to 219'048 and 52'641 CSF, respectively. As shown in Fig. 9, the former RAS scheme provides a reliable one-color spectrum when compared to that of the CAS(14,13)$^{\delta}$, while the latter cannot not properly describe some of the states located in the UV probing region (e.g. phenol signals related to the $S_0 \rightarrow S_{7-9}$ transitions, namely 7P, 8P and 9P in Fig. 9a), as well as mixed doubly excited states that give rise to off-diagonal signals (1B/1P or 1P/1B in Fig. 9a). In fact, the electronic structure of the dimer involve several of these mixed double excitations, given by the collective one-electron excitations on both monomers and appearing at the energy sum of the localized single excitations. The electronic coupling between two interacting monomers, denoted as "quartic" coupling (Δ) [2, 83], shifts the energies of mixed states relative to the sum of the corresponding single excitations, yielding off-diagonal cross-peaks in the 2DUV–UV maps. Thus, quartic coupling and corresponding off-diagonal 2D signals are expected to be absent (or negligible) in the non-interacting benzene–phenol dimer. This is correctly observed in the CAS(14,13)$^{\delta}$ and RAS(2,3|8,8|2,2)$^{\delta}$ spectra, but it is broken in the RAS(4,7|0,0|4,6)$^{\delta}$ simulations, indicating that too great a reduction of the active space degrades the description of the excited-state manifold and the corresponding 2D maps.

In order to validate the benchmark study on the gas-phase non-interaction and to establish an efficient protocol for 2DUV spectra simulations of realistic systems, the cysteine-phenylalanine-tyrosine-cysteine (CFYC) tetrapeptide (Fig. 9b, c), solvated in water solution, was considered as a model protein system [54]. As will be shown in the next section, the folding/unfolding dynamics of this tetrapeptide involve configurations with interacting and non-interacting benzene and phenol chromophores, occurring as aromatic side chains of the phenylalanine (Phe, F) and tyrosine (Tyr, Y) amino acid residues, respectively. In Fig. 9, we report two representative structures of these configurations and their corresponding 2DUV–UV spectra, simulated with our SOS//QM/MM approach [52]. Notably, the comparison of various levels of theory shown for the non-interacting dimer in the gas phase (Fig. 9a) is completely preserved for the non-interacting aromatic side chains in the realistic solvated tetrapeptide model, with the computed RAS(2,3|8,8|2,2)$^{\delta}$ spectrum showing the same accuracy as that of the CAS(14,13)$^{\delta}$, while a cheaper RAS scheme yields artificial quartic coupling. When considering a

configuration with interacting aromatic side chain in a stacked conformation (Fig. 9c), an analogous trend is found. In this case, a sizeable electronic quartic coupling is expected, and it is found equally at CAS(14,13)$^{\delta}$ and RAS(2,3|8,8|2,2)$^{\delta}$ levels, while the energy shift (Δ) is significantly overestimated at the RAS(4,7|0,0|4,6)$^{\delta}$ level.

Here, a successful benchmark study aimed at identifying efficient multi-configurational/multi-reference approaches in terms of the computational cost/accuracy ratio has been demonstrated in detail for benzene and phenol chromophores involved in protein systems. Analogous studies have been performed for indole [82] and adenine [57] chromophores, while the remaining nucleobases are currently under investigation. Overall, we believe that such theoretical benchmark studies will provide a full set of parameters and computational recipes that will enable the construction of excitonic model Hamiltonians for application in large systems (such as full proteins or large DNA/RNA sequences) and simulating 2DUV spectra of dimeric and small oligomeric systems under realistic conditions with unprecedented accuracy.

In the next sections, we will illustrate how accurate simulation of 2DUV spectroscopy could be used as a powerful tool to investigate physicochemical properties of biological systems.

4.2 2DES for Tracking GS Conformational Dynamics

In Sect. 3.1, we illustrated how 2DUV spectroscopy holds the potential to resolve electronic couplings associated with the interaction between UV chromophores in a model protein system containing two aromatic side chains. In particular, the energy shift (Δ) of mixed doubly excited states due to quartic coupling clearly affects the 2DUV–UV spectrum of the solvated CFYC tetrapeptide, enabling discrimination between peptide conformations with interacting and non-interacting chromophores. The CFYC peptide was chosen for modeling protein folding/unfolding dynamics because its terminal cysteine residues can form a disulfide bond that holds the tetrapeptide in a cyclic *closed* conformation while, if this bond breaks, CFYC will naturally unfold in an *open* conformation (Fig. 10). QM/MM geometry optimization of the closed CFYC, solvated in water solution, yields a structure in which the two aromatic side chains are rather close to other, in a so-called T-stacked conformation [52]. On the other hand, the open CFYC conformation is associated with unstacked (i.e. non-interacting) chromophoric units (Fig. 10a).

As shown in Fig. 10b, the excited-state manifolds of both closed T-stacked and open unstacked QM/MM optimized structures have been fully characterized at the CAS(14,13)$^{\delta}$ level [37, 52] showing how they are strongly differentiated by the chromophoric electronic couplings, which have a twofold effect: (i) red-shift of the mixed states by quartic coupling, and (ii) strong red-shift of charge-transfer (CT) states (from Phe to Tyr and vice versa) with a concomitant increase in the TDMs associated with the corresponding $S_{1\text{-}2} \rightarrow$CT transitions. The former would cause off-diagonal (negative) signals to appear in the 2DUV–UV maps (see Fig. 9c), while the latter would show up exclusively at lower energies, i.e. in 2DUV–Vis spectra. Figure 10c shows the SOS//QM/MM computed 2DUV–Vis spectra for closed T-stacked and open unstacked conformations. In these spectra, few (positive) ESA

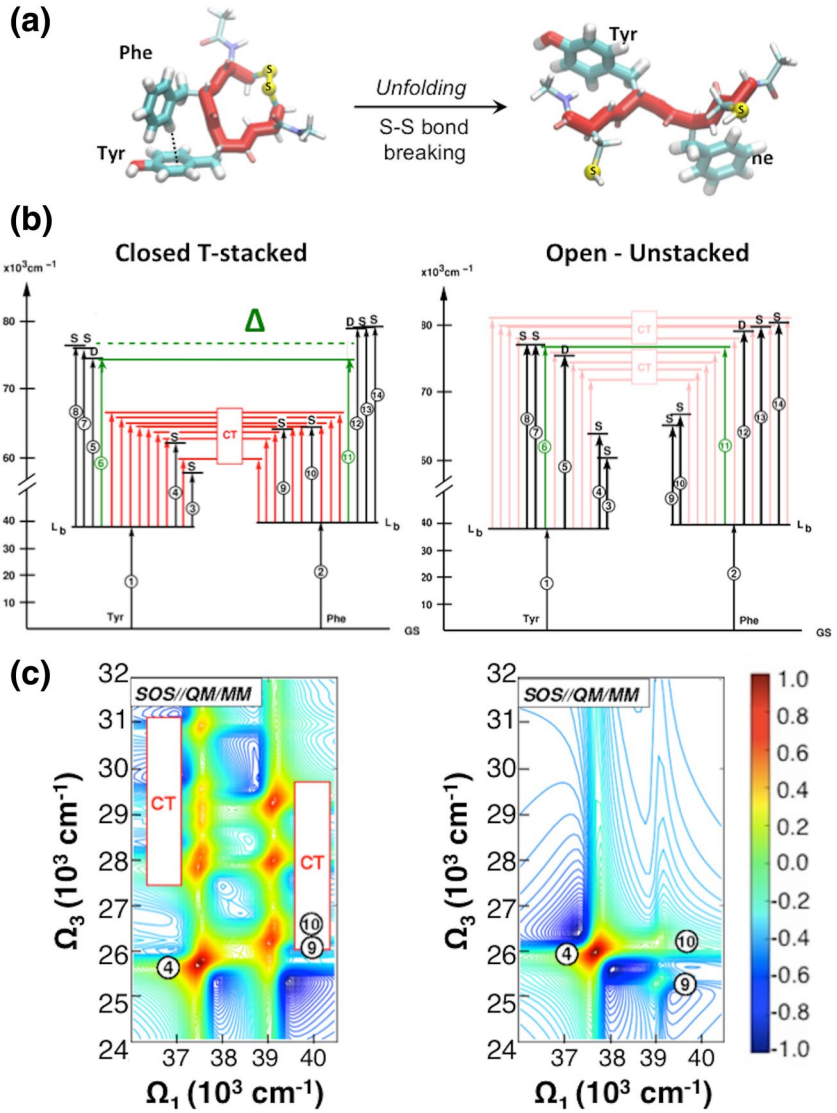

Fig. 10 **a** Optimized QM/MM GS geometries of the solvated CFYC tetrapeptide in the closed T-stacked and (upon disulfide bond breakage) open unstacked conformations. **b** Excited-state manifolds of the two CFYC conformations computed at the CAS(14,13)$^{\delta}$//PT2/ANO-L(432-2) level, with local (single and double) excitations in black, mixed double excitations in green, and bright or dark CT states in red or light red, respectively. **c** Simulated two-color 2DUV–Vis spectra of both CFYC conformations, with ESAs labels as defined in the scheme of panel b. Reproduced from data reported in Ref. [37]

signals arising from the $S_{1-2} \rightarrow S_M$ transitions (with S_1 and S_2 being the first $\pi\pi^*$ states of Tyr and Phe, respectively, namely L_b in Platt notation) are present for non-interacting side chains, while they are accompanied by a large number of ESA

signals in the T-stacked conformation, due to bright $S_{1-2} \rightarrow$CT transitions. It is worth noting that the appearance of these important ESA signals in the 2D maps would be completely missed in standard exciton modeling, with SOS//QM/MM computations of chromophore aggregates here showing their relevance.

In another example, we benchmarked low-cost methods for indole [82]. The first two ($\pi\pi^*$) excited states (L_b and L_a) of indole (see Fig. 5) have essentially opposite character: L_b is covalent and apolar, with low oscillator strength, while L_a is ionic, polar and bright. Thus, indole represents a more intricate challenge compared to benzene and phenol, which possess a single (covalent) state absorbing in the NUV. Understandably, additional compromises had to be made in the calibration procedure. Nevertheless, a RAS(4,5|0,0|4,4) level of theory followed by a single-state RASPT2 energy correction utilizing an imaginary level shift parameter of 0.5 a.u. (Figure 11a) shows reasonable agreement with the reference data obtained with RAS(10,9|2,12) level of theory (Figs. 5c, 11a) for states below 8 eV, with significantly reduced computational effort. Employing the cost-efficient protocols developed for computing the electronic structures of indole and phenol, we applied the SOS//QM/MM scheme to the Trp-cage peptide, a common protein model for studying protein folding/unfolding, which contains Trp and Tyr side chains. As for the Phe- and Tyr-containing CFYC tetrapeptide (Fig. 10), two-color 2DUV–Vis spectroscopy can be used to detect ESA signatures of chromophore–chromophore interactions. We provide an indirect way to detect Tyr and Phe, whose absorption bands remain hidden under the (more) intense envelope of Trp in linear absorption experiments. A signal-free probing window (between 18,000 and 24,000 cm^{-1}) along the Ω_1 trace of Trp seems to facilitate this undertaking. Specifically, weak CT signals are resolved in this spectral window in a folded Trp-cage conformation with the two chromophores at a distance of about 5 Å. This outcome indicates that an experimentally detected enhancement of the ESA signal in Tyr- and Trp-containing peptides, in the (otherwise) signal-free Vis spectral window between 18,000 and 24,000 cm^{-1}, should be regarded as a clear signature of chromophore–chromophore proximity of 5 Å or below.

From the above examples, we can thus conclude that 2DUV spectroscopy can be used to distinguish between folded and unfolded structures of protein models with interacting and non-interacting UV-chromophores, respectively. The question then arises as to whether, given the high temporal (femtosecond) resolution of 2DUV experiments, this technique could also be used to monitor folding/unfolding dynamics of protein models and, more generally, whether it represents a useful tool for tracking GS conformational dynamics. We have exploited the SOS//QM/MM computational recipes described above to tackle these questions by simulating the 2DUV spectra along the unfolding dynamics of the CFYC tetrapeptide [54]. Figure 12 shows the structural changes relative to the inter-chromophore distance (d) and angle (between vectors normal to the aromatic planes, α) observed during a ratchet-and-pawl biased molecular dynamics (rMD) simulation [84], indicating how, starting from a closed T-stacked conformation (with $d = 4.5$–6 Å and $\alpha > 60°$), the CFYC peptide unfolds ($d > 6.5$ Å), passing through twisted offset stacked conformations (with $d = 5$–6.5 Å and $\alpha = 40$–$60°$). Notably, given the restricted flexibility of the peptide backbone, the two aromatic side chains are never observed in a parallel

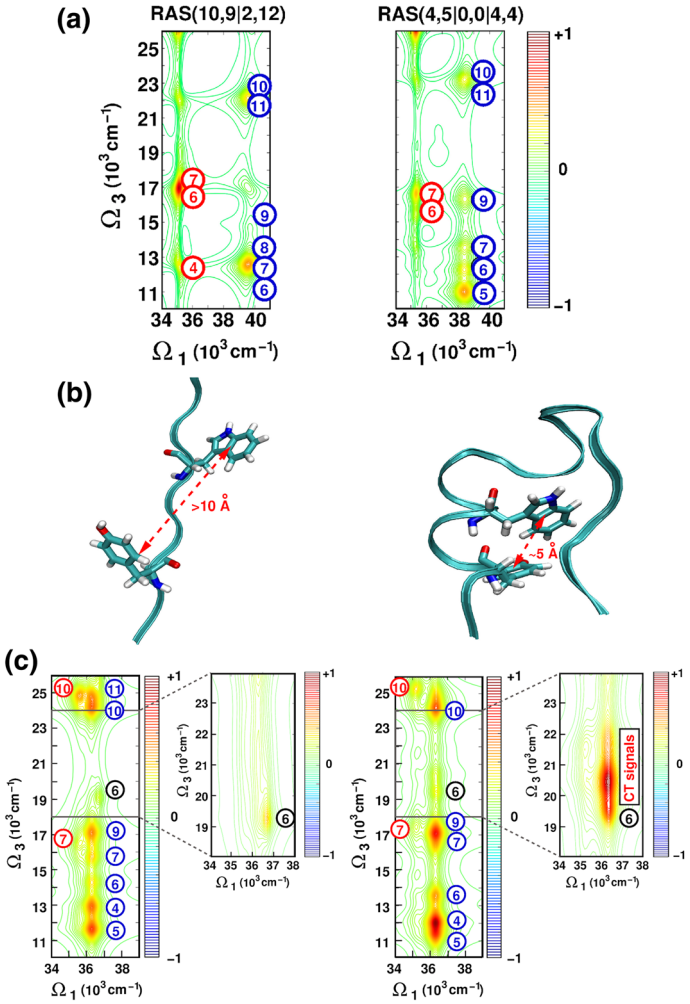

Fig. 11 **a** Comparison of simulated 2DUV–Vis spectra of indole in the gas phase, including the reference RAS(10,9|2,12) (left panel) and the low-cost RAS(4,9|0,0|4,7) level (right panel). Structures (**b**) and 2DUV–Vis spectra (**c**) of unfolded (left panel) and folded (right panel) Trp-cage from representative snapshots. ESAs associated with the indole side chain of Trp (L_b in red, L_a in blue) or with the phenol side chain of Tyr (black) are labeled. Probing of the spectral region around $\Omega_3 = 18{,}000$–$24{,}000$ cm^{-1} is zoomed in to highlight the weak signatures of charge-transfer states emerging in the stacked conformation. Labeling in (**a**) and (**c**) according to Fig. 5b. Reproduced from data reported in Ref. [82]

(namely π-) stacking conformation (i.e. with $d < 6$ Å and $\alpha < 30°$). The 2DUV–UV spectra computed for selected snapshots along the rMD simulation indicate a good correlation between the inter-chromophore distance and the quartic electronic coupling, with smaller d accompanied by larger shifts Δ (see Fig. 12b), which is less sensitive to the relative orientation of the chromophores, with T-stacked or twisted offset conformations possessing large quartic couplings. Significant correlations are also observed in the 2DUV–Vis spectra [54], where the energetic positions

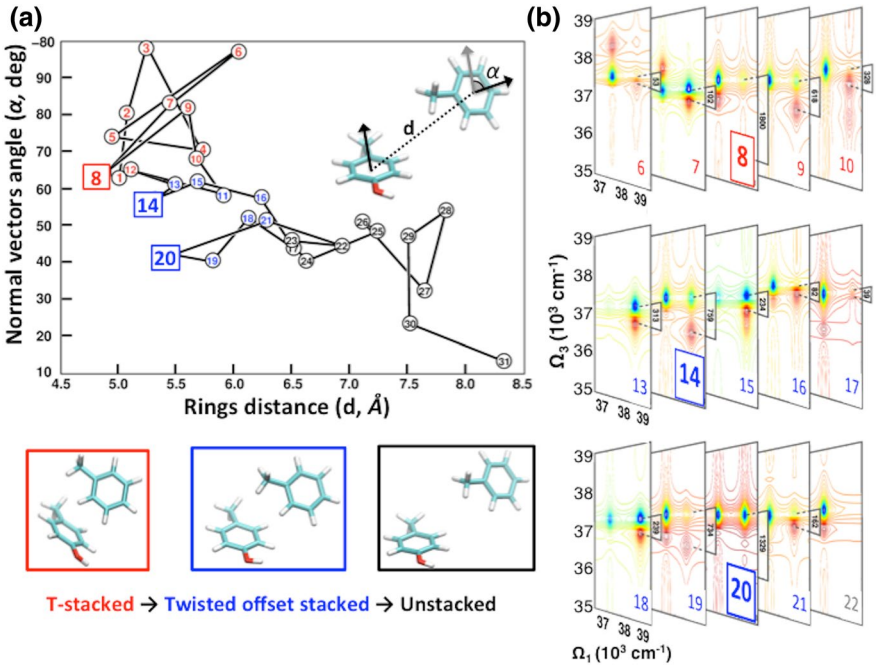

Fig. 12 **a** Inter-chromophore distances d (in Å) and angles α (measured between vectors normal to the aromatic planes, in degrees) for selected snapshots along the biased rMD simulation of the CYFC tetrapeptide unfolding dynamics. Color code according to the three-step unfolding/unstacking process (T-shaped in red, twisted offset stacked in blue, and unstacked in black), with representative structural arrangements of the UV chromophores depicted within boxes. **b** Simulated one-color 2DUV–UV spectra of selected structures, with snapshots having the shortest inter-chromophore distances highlighted. Reproduced from data reported in Ref. [54]

and brightness of the CT states are affected by inter-chromophore distances, their π-orbital overlaps and environmental effects explicitly accounted for in our SOS// QM/MM approach.

Considering the benchmark study of the protein aromatic chromophores described in the previous section, which provides converged results that are consistent with available experimental data, we can conclude that our predictions for two-dimensional electronic spectra strongly support the use of the 2DUV technique as an analytical tool for monitoring GS conformational dynamics of protein systems containing aromatic UV-chromophores.

We expect that the adoption of 2DUV will be equally well suited for tracking of GS conformational changes in nucleic acids. Dinucleoside monophosphates, i.e. dimers of DNA/RNA nucleobases, represent excellent model systems for exploring this hypothesis from our theoretical perspective. In fact, they are realistic biomolecules that are treatable by our QM approaches [57, 60] and assume various stacking conformations in water solution [85], currently representing model systems for studying the complex excited-state dynamics of nucleic acids

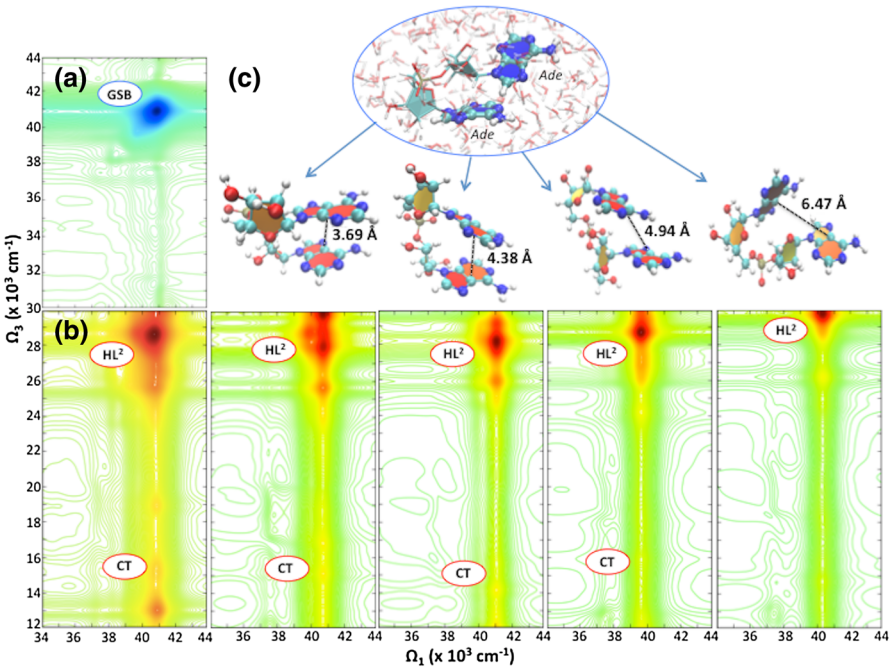

Fig. 13 Simulated 2DUV–UV (**a**) and 2DUV–Vis (**b**) spectra of the water-solvated ApA dinucleoside monophosphate, considering the conformational space sampled with unbiased MD in the GS. Panel (**c**) shows the molecular structure of the ApA system and a few selected representative conformations displaying the different intermolecular interactions possible. Reproduced from data reported in Ref. [61]

[86]. As an example case study, we very recently [61] performed 2DUV spectra simulations of the adenine–adenine monophosphate dinucleoside (ApA) in water solution. Figure 13 shows the computed 2DUV–UV and 2DUV–Vis for a set of structures extracted from an unbiased MD simulation of the solvated ApA in the GS, featuring various intermolecular arrangements that are known to deeply affect the photophysical properties of the adenine moieties [87]. Computations were performed using the SOS//QM/MM approach and adopting the RAS(4,12|0,0|4,6) active space, i.e. the cheapest (but still reliable) multi-configurational scheme calibrated in our benchmark study of adenine monomer and homodimer [57]. By adopting this efficient scheme we could achieve computing at a reasonable cost the excited-state manifold of multiple MD snapshots. Still, the electronic structures computations must be computed for selected (most representative) conformations among those thermally accessible in the ApA conformational space. The presence of explicit solvent molecules in the QM/MM treatment and the inclusion of multiple conformations (Fig. 13c) give rise to inhomogeneous broadening of the 2D signals. In particular, the GSB signals found in the 2DUV–UV spectrum of ApA (Fig. 13a) feature the characteristic broadening along the diagonal due to interactions with the environment and between the chromophores. In fact, the couplings among adenine chromophores

introduce energy splitting in the main signal trace along Ω_1, i.e. that associated with the fundamental $S_0 \rightarrow S_{3-4}$ transitions (with S_3 and S_4 being the lowest $\pi\pi^*$ bright state of the two adenines, namely ^1L_a), which results in an effective additional broadening of the GSB signals. Analogous broadening along Ω_1 is observed in the 2DUV–Vis spectrum (Fig. 13b) for the dominant ESA signals arising from excitations to the adenine doubly excited state (referred to as HL^2), lying at Ω_3 around 28,000–30,000 cm^{-1}.

The solvated ApA dinucleoside is a highly flexible system, featuring conformations ranging from T-shaped to quasi-planar π-stacked, to completely unstacked adenines along the MD trajectory. Cluster analysis has been used to group molecular conformations and select those that are most representative, a few of which are depicted in Fig. 13c. Generally, short inter-chromophore distance in the 3.5–4.5 Å range (measured as C5–C5 distance) is associated with rotated π-stacked structures that yield substantial broadening in 2D maps along Ω_1 (i.e. ^1L_a trace splitting), while also featuring CT signals in the low-energy Vis probing window (at Ω_3 around 16,000 cm^{-1}). At inter-chromophore distances of around 5 Å, the ApA structures can be found in T-shaped conformations, usually with the amino group of one adenine pointing towards the 5-ring of the opposite adenine moiety. Generally, these T-stacked conformations can be differentiated from the others because the presence of ESA signals from CT states is not accompanied by broadening along Ω_1 in this case. As expected, unstacked conformations behave as non-interacting dimeric systems, featuring dark CT states (thus not contributing to the 2DUV–Vis maps) and ESA signals with negligible Ω_1 broadening.

Overall, these results demonstrate how 2DUV spectroscopy is theoretically able to characterize the conformational space of a dinucleoside monophosphate by exploiting the two-dimensional spectral resolution, representing a powerful alternative or complement to standard pump–probe experiments for elucidating the role of structural arrangements on nucleic acid photophysics and photochemistry. However, the spectral line shapes reported in the examples above are not realistic, as they account for only partial contributions to the spectral line broadening. Various effects (coupling to nuclear degrees of freedom, coupling to environment, finite excited-state lifetimes) are the source of dephasing-induced broadening. In real 2D, electronic spectra broadening could obscure the spectral fingerprints predicted by ab initio simulations. Therefore, inclusion of dynamic effects is indispensable to move towards accurate 2DES simulation, and this is the focus of the next two sections.

4.3 Excited-State Coherent Vibrational Dynamics Resolved by 2DES

Real quantum systems are not closed. Upon interaction with the incident electric field, the electron density of the molecular system rearranges instantaneously in a new discrete quantum state. The changed electron density exerts a force on the nuclei, which are set in motion, and vibrational dynamics is initiated. As the nuclei are much heavier than electrons, their vibrations have vibrational periods from 10 fs (fast hydrogen stretch vibrations) to a few hundred femtoseconds (for wagging,

twisting and other slow vibrational modes). Nonlinear spectroscopy with ultrashort sub-10-fs pulses (an order or magnitude shorter than most of the vibrational periods) are capable of detecting this vibrational dynamic, which manifests in coherent oscillations of the spectral signatures. These coherences provide insight into the photoactive vibrational modes of the system and help in understanding its reactivity.

We begin this section by re-introducing a linear coupling of the electronic degrees of freedom to a Gaussian bath, while still neglecting population transfer. Regarding the description of ESAs, this corresponds to adding the phase function $\varphi_{e'fe}(\tau_4, \tau_3, \tau_2, \tau_1)$ to Eqs. (14) and (15), where it is worth noting the absence of population transfer $e' = e$ in Eq. 14. The phase function acquires the following complex form:

$$\begin{aligned}\varphi_{e'fe}(\tau_4, \tau_3, \tau_2, \tau_1) = &-g_{e'e'}(\tau_{43}) - g_{ff}(\tau_{32}) - g_{ee}(\tau_{21}) \\ &- g_{e'f}(\tau_{42}) - g_{e'f}(\tau_{43}) - g_{e'f}(\tau_{32}) \\ &- g_{e'e}(\tau_{41}) - g_{e'e}(\tau_{42}) - g_{e'e}(\tau_{31}) - g_{e'e}(\tau_{32}) \\ &- g_{fe}(\tau_{31}) - g_{fe}(\tau_{32}) - g_{fe}(\tau_{21})\end{aligned} \tag{16}$$

where $g_{ij}(\tau_{ij})$ are line shape functions (see Appendix) and $\tau_{ij} = \tau_i - \tau_j$. Equation 16 depends in a non-trivial way on the three delay times t_1, t_2 and t_3 between the pulses, and it captures coherences which survive during the entire duration of the multipulse experiment. This model can describe bath fluctuations of arbitrary timescales. In the following, we show an example were we focus on the spectral signatures of the strong coupling to a bath of discrete high-frequency intramolecular vibrational modes. The line shape function for this model is formulated on the basis of the multidimensional uncoupled displaced harmonic oscillator (DHO) [2]:

$$g_{ij}^{\mathrm{DHO}}(t) = \sum_k \frac{\omega_k \tilde{\mathrm{d}}_{ik} \tilde{\mathrm{d}}_{jk}}{2} \left[\coth\left(\frac{\omega_k}{2k_B T}\right)(1 - \cos(\omega_k t)) + i \sin(\omega_k t) \right] \tag{17}$$

where the one-dimensional potential for the ith state along each normal mode k is fully characterized by two parameters, frequency ω_k and relative displacement $\tilde{\mathrm{d}}_{ik}$. In the simulations, a composite line shape function is used, constructed by combining the DHO line shape function with the line shape function of the semi-classical Brownian oscillator (see Appendix), which describes the coupling to a continuum of low-frequency modes (of the environment), inducing decoherence on the timescale of a few tens of femtoseconds and giving rise to the homogeneous broadening of the peaks. Note that the omission of a mechanism for population transfer decay implies infinite excited-state lifetimes. However, in reality, excited states have finite lifetimes. When ultrafast (< 100 fs) decay channels are present, the dephasing due to population transfer may induce additional signal broadening, which in the present framework is treated phenomenologically.

Pyrene (see Fig. 14a) is a polycyclic aromatic hydrocarbon that has attracted attention for its prominent photophysical properties such as its remarkably long

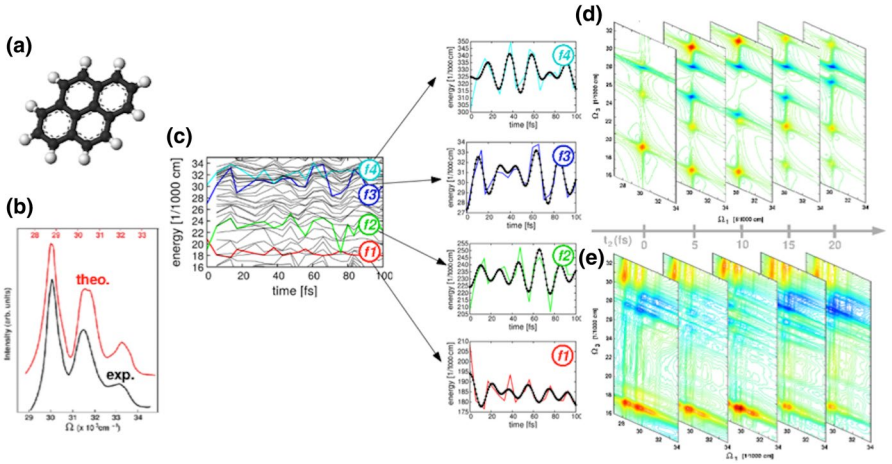

Fig. 14 **a** Structure of pyrene; **b** comparison of recorded (black) and simulated (red) linear absorption spectrum; **c** state density up to 34,000 cm^{-1}, highlighting the energy profiles of the bright states *f*–f_4, fingerprints of the L_a state, with superimposed fits according to Eq. 18 (black dotted lines); (**d**, **e**) simulated quasi-absorptive 2D electronic spectra of pyrene for waiting times t_2 in the range 0–20 fs obtained through pumping at the frequency of the L_a transition and supercontinuum probing in the UV–Vis region within the "quasi"-static picture (**d**) and with bath fluctuations included (**e**). Color code: GSB (blue), ESA (red). Reproduced from data reported in Ref. [93]

fluorescence lifetime and high fluorescence quantum yields [88]. It is used to probe solvent polarity [89] and constitutes the basis of commercial dyes, fluorescence probes for electro-chemiluminescence and many larger photoactive molecules [90–92]. Pyrene features well-separated absorption bands (see Fig. 14b) with clear FC progressions in the NUV (lowest bright state L_a at 320 nm) and in the deep UV (second bright state at 280 nm), making it an excellent model for assessing novel spectroscopic techniques. Pyrene has been the subject of some of the first documented broadband 2DES spectra in the NUV [16] and deep UV [18], therefore representing a perfect target for benchmarking our theoretical protocols against 2D experimental data [93].

Here, we review 2DES simulations which resolve the coherent spectral dynamics initiated immediately after photoexcitation of the first bright excited state (L_a) of pyrene, exhibiting a vibrational progression with a fundamental (most intense) transition and two overtones between 28,500 cm^{-1} (~350 nm) and 33,500 cm^{-1} (~300 nm) (see Fig. 14b). In particular, we aim at reproducing the spectra for waiting times of 0–20 fs in a broadband probe window ranging from 16,000 cm^{-1} (~600 nm) to 32,000 cm^{-1} (~300 nm). These simulations require knowledge of (i) the electronic structure of pyrene up to 65,000 cm^{-1} (ca. 8 eV), thus covering not only the L_a state (labeled *e*) but also the manifold of higher-lying states (labeled *f*) that fall within the envelope of the probe pulse (i.e. 16,000–32,000 cm^{-1} above the L_a state); (ii) the FC active modes of the L_a state, source of the vibrational dynamics appearing as coherences in the 2D spectra; and (iii) the response of the states from the *f*-manifold to the vibrational dynamics in the L_a state. ESA bands are an indirect probe of the coherent vibrational dynamics in the photoactive state, featuring high

sensitivity by oscillations of signals intensities and/or spectral positions as a function of the delay time.

To obtain the required information listed above, 60 excited states were computed simultaneously at the RASSCF(4,8|0,0|4,8)//PT2 level of theory, where the RAS comprise all valence π-orbitals of the molecule, allowing up to quadruple excitations only in the RAS1/RAS3 spaces. Coherent vibrational dynamics were described with the DHO model. The frequency (ω_k) and displacement ($\tilde{d}_{ik}$) parameters used to construct the line shape functions g_{ij} (Eq. 17) were extracted from a 100-fs mixed quantum–classical molecular dynamics simulation treating electrons quantum-mechanically by solving the time-dependent Schrödinger equation, while applying classical Newtonian dynamics to the nuclei. The dynamics were initiated in the L_a state of pyrene at the FC point without initial kinetic energy. The choice of a reference Hamiltonian acting on the classical bath during the coherence propagation, as is the case for the delay times t_1 and t_3, is not unique [94], as the *bra* and the *ket* sides of the density matrix are subject to different electronic potentials. We chose to propagate the nuclear degrees of freedom as though interacting with the L_a state. The frequency and displacement parameters themselves were obtained by fitting the electronic gaps $E_e(t)$–$E_g(t)$ (giving rise to GSB and SE signals) and $E_f(t)$–$E_e(t)$ (giving rise to ESAs), computed at the aforementioned RASSCF(4,8|0,0|4,8)//PT2 level along the quantum–classical dynamics, to the analytical expression for the classical time-dependent fluctuation of the $L_a - S_n$ energy gap (n belonging to a state from either the GS, g, or the f-manifold)

$$E_h(t) - E_e(t) = -\sum_k \omega_k^2(\tilde{d}_{ek} - \tilde{d}_{hk})\tilde{d}_{ek}\cos(\omega_k t) + \sum_k \frac{\omega_k^2(\tilde{d}_{ek} - \tilde{d}_{hk})^2}{2} + (\omega_{hg} + \omega_{eg}) \tag{18}$$

with $h \in \{g,f\}$, which made it possible to extract the mass-weighted displacement coefficients d_{ek} and d_{fk}, as well as the electronic transition energies ω_{eg} and ω_{fg} (the parameters d_{gk} and ω_{gg} describing the GS per definition, are zero). In the adopted DHO framework, the spectral dynamics of the ESA are a function of the dynamics in the photoactive state (namely S_3 for pyrene), i.e. of the relative displacement of the higher-lying excited-state PES relative to the PES of the photoactive state, and can induce positive or negative frequency correlations [95]. By definition, it is restricted to the normal modes describing the molecular dynamics in the e-manifold (i.e. if d_{ek} is zero, then so is d_{fk}). This is an implication of the missing state-specific modes. We emphasize that the extraction of the energy fluctuations of states from the f-manifold must be performed in a diabatic representation (i.e. by following the excited-state wave function rather than the adiabatic root) in order to stay in the framework of the uncoupled harmonic oscillators, as due to the high state density in the Vis and NUV, PES of higher-lying states intersect, thus rendering the resulting adiabatic potentials highly anharmonic (see Fig. 14c). To achieve this goal, we selected at the FC point the states from the manifold of higher-lying states f within the probed spectral window and with significant oscillator strength out of the L_a state (i.e. the reference diabatic states, shown in red, green, blue and cyan in Fig. 14c), and tracked the temporal evolution of the wave function associated with

each reference state along the dynamics by searching for the adiabatic state with the greatest overlap with the reference. The energy profiles extracted for the four brightest transitions are depicted in Fig. 14c, showing clear oscillatory dynamics properly captured by the fitting procedure (black dotted lines). Note that the energy state f_1 (red) oscillates out of phase with respect to f_2, f_3 and f_4.

Figure 14d, e compares simulated 2DES spectra between 0 fs and 20 fs simulated in the "quasi"-static picture (Fig. 14d) and with bath fluctuations included (Fig. 14e). The dynamic evolution of the spectra in the static picture is approximated by assuming fast bath fluctuations during t_1 and t_3, inducing homogenous Lorentzian broadening of the signals. Within this approximation, the 2DES at a waiting time $t_2 \neq 0$ can be computed with the SOS protocol (note that with a single state in the *e*-manifold, the coherence term Eq. 14 vanishes), setting the pump pulse to interact with the *e*-manifold of the FC point, while setting the probe pulse to interact with the *f*-manifold of the snapshot reached along the trajectory at t_2. It is apparent that the "quasi"-static picture reproduces qualitatively the main spectral features: (a) the GSB around 30,000 cm^{-1} (shown in blue); (b) the oscillatory dynamics of the SE between 30,000 cm^{-1} and 25,000 cm^{-1}, with a period of ca. 20 fs (shown in blue); (c) the aforementioned three ESA contributions (shown in red) around 19,000 cm^{-1} (f_1), 24,000 cm^{-1} (f_2) and 32,000 cm^{-1} (f_4); and (d) the time-dependent fluctuation of the intensities, a consequence of the coordinate dependence of the TDM. However, the static approach fails to correctly describe the spectral line shapes. This becomes evident when slow bath fluctuations are taken into account (Fig. 14e). The vibrational progression in the L_a spectrum (Fig. 14b) translates into three traces along Ω_1 associated with the fundamental transition (28,500 cm^{-1}) and two overtones. As a consequence, the GSB and SE adopt a characteristic checkerboard pattern [96, 97]. Another remarkable difference is that the oscillations of the ESA peaks become less distinct. Signals showing more pronounced energy gap fluctuations (f_4) appear broadened along Ω_3 and, as a consequence of this broadening, exhibit reduced intensity compared to the static spectrum. Overall, the ESA associated with f_1 ($\Omega_3 \sim 17{,}500$ cm^{-1}) constitutes the most characteristic signature of the L_a state in the Vis range. In fact, it has been observed in transient pump–probe spectra [45, 98]. The weaker ESA around 25,000 cm^{-1} associated with state f_2 has been observed in 1D-PP experiments [99]. To the best of our knowledge, no transient spectra pumping the L_a state and probing beyond 32,000 cm^{-1} for short (< 100 fs) waiting times has been reported in the literature; therefore, the absorption associated with f_4 is yet to be detected experimentally and it is predicted by our simulations.

4.4 Spectral Characterization of Long-Lived ES Intermediates

Excited states have finite lifetimes. Decay to the GS occurs either via light radiation or non-radiatively. In the second scenario, population transfer between electronic states is facilitated through non-adiabatic (e.g. singlet–singlet transfer) or spin–orbit (e.g. singlet–triplet transfer) couplings, which are coordinate-dependent and become large in areas where the PES of electronic states intersect. Population transfer has a twofold effect on the appearance of the 2D electronic spectra: (i)

homogeneous line broadening, and (ii) decay and simultaneous buildup of spectral signatures, the former process being associated with the population decrease in the initial state, the latter with the population growth in the final state. The CGF framework offers a prescript for incorporating population transfer, i.e. the so-called doorway–window (DW) factorization [100, 101]. When dephasing is much faster compared to transport, the secular Redfield approximation can be applied, providing separate expressions for coherences (Eq. 13) and populations (Eq. 12). Correlated bath dynamics in the coherence is described with the phase function $\varphi_{e'fe}(\tau_4, \tau_3, \tau_2, \tau_1)$ introduced in the previous section (Eq. 16). In contrast, when bath fluctuations decay more slowly than the coherence dephasing but faster than population transport, the population term (in Eq. 13) can be treated by applying the Markovian approximation to the time intervals between the pulses, with the phase function expressed as

$$\begin{aligned}\varphi_{e'fe}(t_3, t_1) = &-g_{ee}(t_1) - g_{e'e'}(t_3) - g^*_{ff}(t_3) + g_{fe'}(t_3) + g^*_{e'f}(t_3) \\ &+ 2i(\lambda_{e'e'} - \lambda_{fe'})t_3\end{aligned} \tag{19}$$

where, in the limit $t_2 \rightarrow \infty$ (i.e. timescales for which memory is lost and the spectral dynamics is dominated by population transfer), the *e'*–*f* coherence evolution during t_3 introduces a broadening along Ω_3, which is a function of the reorganization energy λ_{ij}. Clearly, Eq. 19 depends on both states (*e* and *e'*) involved in the population transfer, while *f* denotes the manifold of excited states coupled to state *e'*. Fluctuations in the initial coherence state *g*–*e* during t_1 and in the final coherence state *e'*–*f* during t_3 are formally treated exactly, but the dynamics are not coupled. More sophisticated treatments are possible when bath fluctuations and transport occur on the same timescale in order to retain memory effects during t_2 and to describe ultrafast population transfer events [47].

A limiting, and quite interesting, case for Eq. 19 is associated with the trapping of population in an excited state during the decay process on a timescale facilitating the dissipation of excess vibrational energy in the environment. In such case, the spectrum is dominated by the electronic structure of the equilibrium geometry of the excited state in which the population is trapped. Within this framework, the dynamics during t_3 can be calculated with the *e'* Hamiltonian as reference, thus effectively eliminating all terms depending on *e'*, with the phase function taking the simplified form

$$\varphi_{fe}(t_3, t_1) = -g_{ee}(t_1) - g^*_{ff}(t_3) \tag{20}$$

which depends only on the line shape function $g_{ee}(t_1)$ of the initial state and $g^*_{ff}(t_3)$ of the states coupled to the probe pulse. The 2D electronic spectra of long-lived intermediates in photochemical and photophysical processes can thus be obtained, in this case, at a reasonably low computational cost, as they do not require quantum dynamics simulations. The reduced computational effort would allow the focus to be directed entirely on the electronic structure, thereby providing invaluable insight into the spectral fingerprints of excited-state minima.

From experimental observations, it is known that in pyrene, after L_a excitation, an ultrafast population transfer towards a lower-lying state (namely L_b) takes place with a time constant of 85 fs [16]. This state is spectroscopically dark; therefore, the population remains trapped on a nanosecond timescale before it eventually decays to the GS through fluorescence with a high quantum yield [102, 103]. Recent excited-state molecular dynamics simulations [93] have revealed the mechanism of non-adiabatic $L_a \rightarrow L_b$ internal conversion. As shown in Fig. 15, after irradiation, carbon–carbon stretching modes with an oscillation period of 20 fs are activated. The momentum accumulated in these modes drives the system beyond the L_a minimum to a turning point on the PES, which lies in the vicinity of the L_a/L_b crossing region. Low-frequency out-of-plane vibrations induce a finite non-adiabatic coupling between the two states, the wave packet bifurcates, and a part continues its dynamics on the L_b surface while the rest remains on the L_a surface, oscillating back and returning to the crossing region after ca. 20 fs. Dissipation of excess vibrational energy (cooling) in the L_b state was found to occur with a time constant of 4 ps [99].

Considering the above-described mechanism and employing the RASSCF(4,8|0,0|4,8)//PT2 level of theory mentioned above for computing TEs and TDMs at the optimized geometry on the L_b PES, i.e. named $^1(L_b)_{min}$, we generated a 2D spectrum representative of waiting times in the ps regime, i.e. longer than the L_a lifetime but shorter than the L_b lifetime, which justify the use of the CGF protocol with the phase function presented in Eq. 20. Figure 15 shows the simulated 2DUV spectrum of pyrene, exhibiting a structured checkerboard pattern attributed to the GSB (L_b is dark, so there is no SE signal) between 28,000 cm^{-1} and 32,000 cm^{-1}, as well as two distinct ESA peaks at 21,000 cm^{-1} (labeled B) and 26,000 cm^{-1} (labeled

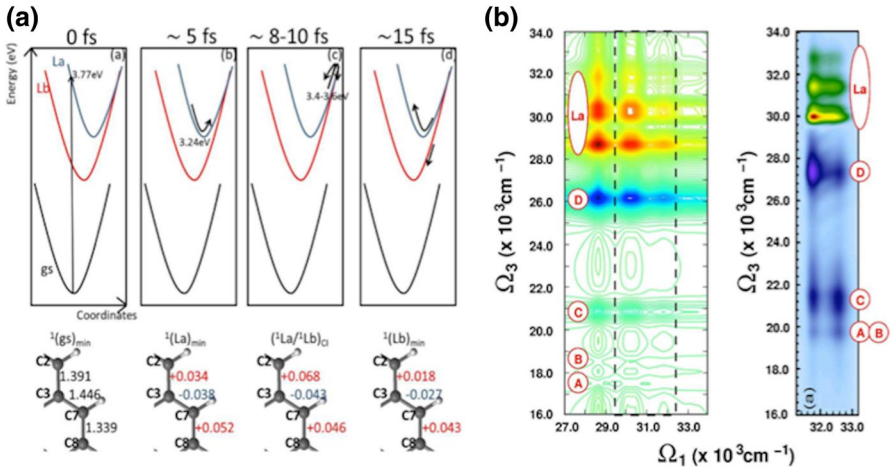

Fig. 15 **a** Schematic representation of the photophysics of pyrene based on quantum–classical dynamics. The dominant bond deformations are given, with the GS equilibrium geometry used as a reference. **b** Theoretical (left panel) and experimental (right panel) quasi-absorptive 2D electronic spectra of pyrene for a delay time t_2 in the picosecond range ($t_2 = 1$ ps in the experiment) obtained through pumping at the frequency of the L_a transition and supercontinuum probing in the UV–Vis region. Color code: GSB (red), ESA (blue). Note that the color code is inverted with respect to other figures so as to enable easy comparison with the experimentally reported spectrum. Reproduced from data reported in Ref. [93]

C, D). Further, less intense peaks (A, B) appear at around 18,000 cm^{-1}. To underscore that different higher-lying excited states are probed relative to those previously reported for the 2DUV spectra at the FC, i.e. $t_2 = 0$ (see Fig. 14d, e), we use letter-code (A–D) to label the ESA peaks. As a consequence of the Markovian approximation, the three traces along Ω_1 adopt an identical peak structure with a progressively decreasing intensity of the overtones. The physical interpretation behind this is that, on a picosecond timescale, excess energy acquired by exciting the overtones would be dissipated, and probing would become insensitive to the pump wavelength. Riedle and co-workers recently reported an experimental 2DUV spectrum of pyrene in methanol at time delay $t_2 = 1$ ps, recorded in a collinear pump–probe set-up, utilizing a narrowband pump–pulse pair, centered at around 32,000 cm^{-1}, i.e. at the second and third vibrational bands of the L_a absorption, and a supercontinuum probe pulse covering the UV–Vis spectral window (16,000–38,000 cm^{-1}) [93]. This result represents one of the few experimental examples employing UV–Vis broad pulses for 2D spectroscopy. The spectrum is shown in Fig. 15b and compared to our simulations (note the inverted color code with respect to the usual color coding adopted in other figures, in order to facilitate appropriate visual comparison with the experimental map). Remarkable agreement is achieved between theory and experiment with regard to the GSB and the positions and relative intensities of the ESA signals. Few differences could be found: (i) through the use of a narrowband pump–pulse pair, the fundamental (0–0) transition was suppressed along Ω_1 in the experiment; (ii) the ESA peaks were noticeably elongated along Ω_3, which could be rationalized by the short lifetimes of the higher-excited states. The pyrene example, where a simulated 2DUV spectrum adopting the excited-state equilibrium geometry was successfully compared with the experimental spectrum recorded at a waiting time of 1 ps, demonstrates that even when equilibration has not been completely carried out (time constant is 4 ps), the spectrum is dominated by the electronic structure of the excited-state equilibrium geometry, and theoretical simulations can predict/interpret 2D spectroscopic fingerprints of excited-state minima. This outcome paves the way for extending simulation of 2DUV spectra through our SOS//QM/MM to other biologically relevant systems, pursuing the idea that 2DUV fingerprints of excited-state minima can be computed at a reasonable computational cost and could be directly compared with experimental data recorded at appropriate t_2 delay times.

In this context, we have focused our efforts on 2DUV simulations for tracking the complex photophysics of nucleic acids, based on the extended benchmark studies illustrated in the previous sections. Recent efforts to obtain the first experimental 2DUV spectra of DNA nucleobase monomers [21, 22] indicate the concrete possibility of using this technique to understand the photophysics of more realistic DNA/RNA models, e.g. multimeric species in short single or double strands. As mentioned in Sect. 3.2, dinucleoside monophosphates, i.e. dimers of DNA/RNA nucleobases, represent excellent model systems to start exploring nucleic acid photophysics from both a theoretical and experimental standpoint. In fact, the presence of a neighboring base significantly affects the excited-state dynamics of monomeric systems due to excimer/exciplex formation and a range of charge/proton/hydrogen transfer/recombination events [104–107]. These have been shown to be extremely difficult to separate and assess with standard

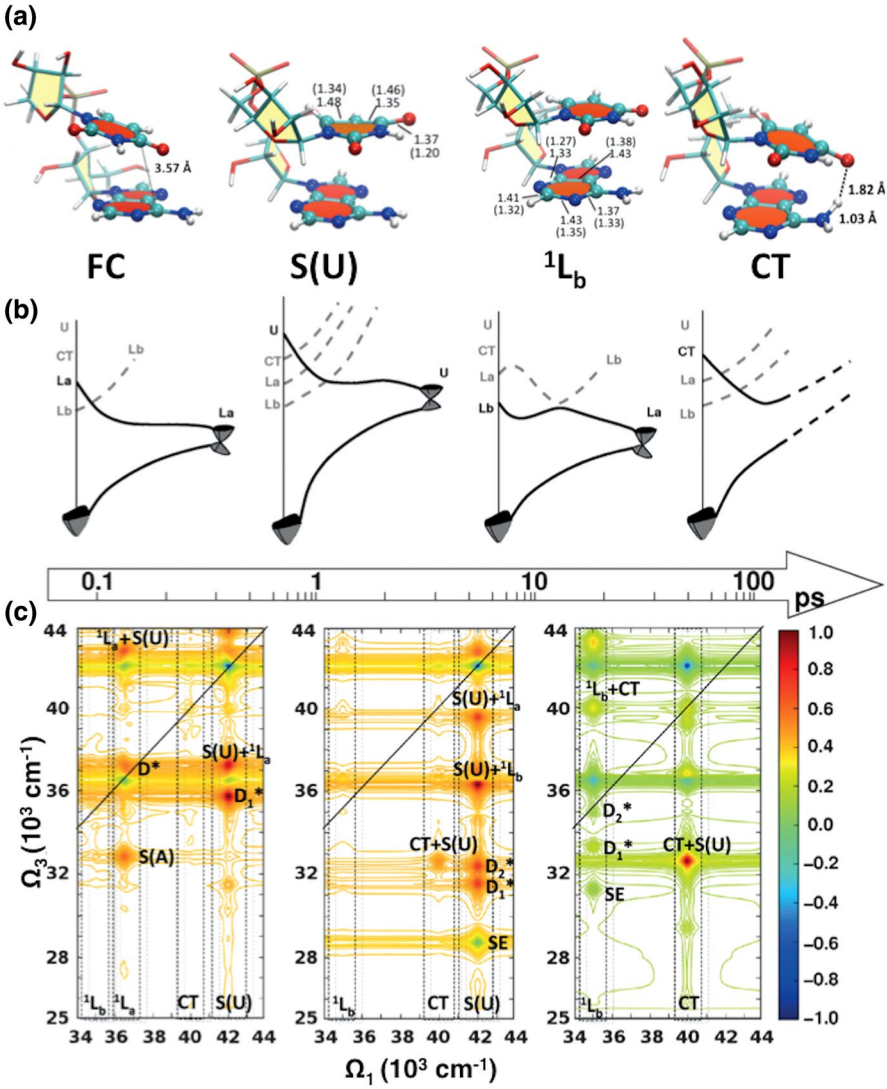

Fig. 16 **a** Ground- (FC) and excited-state optimized structures of the solvated ApU dinucleoside monophosphate, including S(U), ^1L_b and CT states. Representative inter-bases or bond distances (in Å) are reported, with GS values in parentheses for excited-state minima. **b** Schematic representation of the PES involved in excited-state decay channels. **c** Predicted time evolution of the 2DUV–UV spectra for stacked ApU conformations. The time arrow marks the approximate scale at which the spectra will evolve according to the potential energy calculations. Reproduced from data reported in Ref. [60]

1D pump–probe techniques, and their particular contributions to the overall de-excitation process are key to understanding photo-induced lesions and mutations [108, 109]. Figure 16 shows the simulated 2DES spectra of water-solvated

adenine–uracil dinucleoside (ApU), a system exhibiting different excited-state relaxation mechanisms that would be intricately obscured in 1D-PP experiments. Early experiments had shown how adenine and uracil separately would feature ultrafast signals [110–112], as well as long-lived signals arising from partial population of their dark $^1n\pi^*$ states [113], but that an additional signal would arise in dimeric species, probably due to inter-nucleobase interactions [87]. This signal could not be clearly disentangled, but it was associated with the ability of the dimeric nucleosides/tides to attain stacking conformations that would lead to the formation of excimer/exciplex or charge-transfer events. In order to ascertain this possibility, we have proposed the use of 2DUV spectroscopy for characterizing, with high temporal and spectral resolution, the spectroscopic fingerprints of various excited-state minima along the complex photoinduced pathways.

An initial theoretical study was carried out for the solvated ApU dinucleoside [60], where each of the decay channels exploited by the monomeric units were investigated by means of CASSCF excited-state optimizations and conical intersection characterizations. Figure 16 shows the equilibrium geometries of the localized excited states, named here as ^1L_b for adenine and S(U) for uracil, as well as delocalized CT states as originated by excited-state geometry optimizations from the FC region, associated with a GS optimized geometry with vicinal interacting nucleobases (Fig. 16a). Transition states and conical intersections along the excited-state decay channels were also characterized, providing an estimate of the timescale of the relaxation times, as schematically reported in Fig. 16b. To disentangle among these relevant stationary points involved in the ApU photophysics, broadband 2DES spectra were computed on top of their structures, resulting in the simulated 2DES maps depicted in Fig. 16c. As can be seen, in the FC region (at zero or ultrashort t_2 waiting time, sub-100 fs), the main signals are expected to arise from the ^1L_a and S(U) traces, with peaks from the spectroscopic ^1L_a state of adenine rapidly depleted due to ultrafast deactivation and weak CT signals that are present only for those conformations where such non-localized states are bright. Due to the presence of a sizeable energy barrier for the deactivation of S(U) of uracil in ApU, in contrast to those recently computed for uridine in water [114], it was predicted that most spectral contributions in the sub-10-ps timescale would come from the intense S(U) trace. At longer waiting times, in the 10–100-ps range, the 2D maps would feature signals on the ^1L_b trace (at Ω_1 around 34,000–36,000 cm^{-1}) and a particularly intense fingerprint of the CT minimum, whose trace (at Ω_1 around 40,000 cm^{-1}) would be well separated from the ^1L_b trace, allowing quantification of the population of ApU conformations with accessible CT states. Thus, these simulations could help assign distinctive signals arising from different excited-state decay channels that might be recorded in upcoming 2DUV experiments, providing a state-specific characterization of the ApU excited-state dynamics. Here, with such an example, we demonstrated how 2DUV could be an extremely powerful tool for elucidating nucleic acid photophysics in model systems.

5 Outlook and Perspectives

In this review, we have outlined a route towards simulation of third-order nonlinear electronic spectroscopy with accuracy compatible with current experiments. The focus on 2D electronic spectroscopy is motivated by the huge impact that its application in the UV region could have on the study of molecular systems with absolute biological relevance, such as proteins and nucleic acids, the building blocks of life (Sect. 1). 2DUV experiments feature high spectral and temporal resolution, allowing us to follow the evolution of individual chromophores though their spectral fingerprints, providing direct signatures of electronic couplings, charge and energy transfers, and characteristic line shapes and dynamics, which provide information on PES topology and solvent reorganization timescales. The density matrix formalism (Sect. 2) provides an appropriate platform to describe the state-of-matter evolution in the Liouville space when interacting with weak electric fields, accessing the nonlinear response recorded in 2DES experiments.

We outlined a route towards simulation and prediction of 2DES spectra from first principles first by employing "drastic" (static) approximations to the working equations within the global eigenstates basis (SOS approach), which enables us to ascertain the two fundamental ingredients to be accurately computed, i.e. the electronic transition energies and dipole moments (TEs and TDMs), and then recovering the missing contributions that shape 2D electronic spectra, comprising the factors that determine the line shapes and the dynamic evolution of the nonlinear signals. This route is covered by illustrating some practical examples (Sect. 3), starting with the benchmark studies of excited-state manifolds of the target biomolecular systems (Sect. 3.1) by means of accurate electronic structure computations, providing reliable TEs and TDMs estimates. For accurate calculations of these fundamental ingredients, we rely on the multi-configurational wave functions combined with multi-reference perturbation theory energy corrections (the CASSCF//PT2 level of theory), which in our opinion offers the optimal compromise between accuracy/completeness and computational cost. We addressed its performance with respect to active space size, number and nature of simultaneously computed states, basis set and geometric parameters, and demonstrate its sensitivity to these parameters. CASSCF/CASPT2 enables adaptive protocols for obtaining converged energies and TDMs, reaching accuracy of 0.2–0.25 eV, which is considered state of the art in the field of photochemistry, yet corresponding to ~ 2000 cm^{-1}, a significant deviation from a spectroscopic point of view. Therefore, reaching this upper-bound limit of accuracy is essential for reliable theoretical spectroscopy. We have shown how such accuracy has been obtained for the TEs and TDMs of monomeric units of aromatic protein side chains and nucleobases in the gas phase, demonstrating how the (commonly used) full-valence active spaces yield large discrepancies relative to available experimental cross-sections, and large, as computationally quite expensive, approaches are required for quantitative comparisons. These high-level computations yielded important parameters that could be implemented in Hamiltonian models (within a quasi-particle representation) in order to extend the simulation to large scale, but

become impracticable when working within an SOS approach, even for the smallest conceivable chromophore aggregates, i.e. the dimeric species. Computational recipes based on semi-empirical parameterization against our reliable reference data enabled accurate predictions of the excited-state manifold of dimeric species at a reasonable computational cost, paving the way to the first applications of 2DUV simulation protocols for estimating the effects of inter-chromophore interactions on the nonlinear response (Sect. 3.2).

By embedding multi-configurational/multi-reference computations within hybrid QM/MM schemes, it was possible to extend the SOS simulation (SOS//QM/MM) of 2D electronic spectra to realistic model systems of proteins and nucleic acids. This allowed us to demonstrate how 2DUV spectroscopy holds the potential to resolve inter-chromophore interactions occurring during GS dynamics of nucleic acid and protein systems. In particular, simulations of 2DUV spectra of the CFYC tetrapeptide and Trp-cage protein model revealed how both electronic (quartic) coupling shifts and charge-transfer states yield clear spectroscopic fingerprints (mainly associated with ESAs) in the one-color 2DUV–UV and two-color 2DUV–Vis spectra, respectively. This significant outcome prompted investigations on the potential of 2DUV spectroscopy for tracking the protein folding/unfolding processes of protein models with high temporal resolution and for the quantitative analysis of nucleobase stacking in nucleic acid models. In this regard, extensive studies on the unfolding dynamics of the CFYC tetrapeptide and on the conformational space characterization of the water-solvated ApA dinucleoside monophosphate indicated that the 2DUV technique represented a powerful alternative (or complement) to standard transient absorption experiments. These results highlight the importance of accurate electronic structure computations in 2DES simulation, to account for appropriate descriptions of the ESAs signals. However, the massive computational costs of ab initio computations restrict its application to small (still realistic) dimeric systems. Current developments in the field of multi-configurational, density-based and linear-response techniques [95, 115–122] would certainly provide beneficial tools for treating large (multi)chromophoric systems, provided they accomplish completeness in the description of excited-state manifolds (e.g. including doubly excited states) [123]. Efficient QM methods employed in hybrid QM/MM schemes would also make it possible to account for signal broadening due to solvent rearrangement, as shown in the ApA case.

Various types of dephasing-induced broadening (due to coupling to nuclear degrees of freedom and environment, finite excited-state lifetimes) can shape the 2D maps. When neglected, simulated signal line shapes and positions would differ significantly from their experimental counterparts. We moved a step towards a more realistic description of spectral line shapes through the inclusion of vibrational dynamics in the simulation (Sect. 3.3). The high temporal resolution featured in 2DUV experiments with ultrashort sub-10-fs pulses allows for the detection of vibrational features, which manifest in coherent oscillations of the spectral signatures. These features can be computed by combining quantum–classical excited-state dynamics simulations that account for discrete high-frequency intramolecular vibrational modes, with electronic structure calculations of the manifold of higher-lying excited states along the dynamics. By accounting for population transfer processes

in the theoretical treatment, it became possible, for the first time, to directly compare theoretical predictions with experimental data, available from broadband 2DES spectra of the pyrene molecule, recorded along its excited-state relaxation process [16] (Sect. 3.4).

A comparison of theoretical and experimental 2DUV spectra revealed that the 2D map of pyrene recorded at waiting times on a timescale that allows dissipation of excess vibrational energy in the environment (i.e. when population is trapped in an excited state) is dominated by the spectroscopic signatures of the excited-state equilibrium geometry, as predicted by our computations. This outcome supports the extension of 2DUV spectra simulation for resolving the spectroscopic fingerprints of long-lived excited-state minima along the complex photoinduced decay pathways of DNA/RNA model systems. In particular, our predictions of state-specific fingerprints of excited-state dynamics in solvated ApU dinucleoside monophosphate exemplifies the potential impact of accurate simulations of 2DES spectra in revealing complex physicochemical properties of fundamental biological systems.

The accuracy of the theoretical treatments proposed here can be improved considerably on several fronts, from more efficient and reliable ab initio characterization of excited-state manifolds, to inclusion of non-Gaussian bath fluctuations, to trajectory-based approaches for handling bath fluctuations and population transfers concurring in the ultrafast timescale, just to mention a few developments that can be envisioned in this field.

Acknowledegements Ivan Rivalta acknowledges support from the French Agence National de la Recherche (FEMTO-2DNA, ANR-15-CE29-0010). Javier Segarra-Marti thanks Dr. Lara Martínez-Fernández for useful discussions. Marco Garavelli acknowledges support from the European Research Council STRATUS Advanced Grant (ERC-2011-AdG No. 291198). Shaul Mukamel gratefully acknowledges the support of the National Science Foundation (Grant CHE-1361516) and the Chemical Sciences, Geosciences, and Biosciences Division, Office of Basic Energy Sciences, Office of Science, U.S. Department of Energy.

Appendix: Retarded Green's function and third-order density matrix

In the absence of an external field ($\hat{H} = \hat{H}_0$), the free evolution of an unperturbed density matrix, $\hat{\rho}^{(0)}(t)$, is stated as

$$\hat{\rho}^{(0)}(t) = G(t)\hat{\rho}(0) = \Theta(t)e^{-\left(\frac{i}{\hbar}\right)\hat{H}_0 t}\hat{\rho}(0)^{\left(\frac{i}{\hbar}\right)\hat{H}_0 t} \tag{A.1}$$

where $\hat{\rho}(0) = |g\rangle\langle g|$ is the density matrix of the system in the GS equilibrium (g), and $\Theta(t) = \int_{-\infty}^{t} d\tau\delta(\tau)$ is the Heaviside step-function ensuring causality.

In the perturbation scheme described in Sect. 2.1 (Eq. 5), the third-order density matrix is stated as

$$\begin{aligned}\hat{\rho}^{(3)}(t) = G(t)\hat{\rho}(0) + \left(\frac{i}{\hbar}\right)^3 \int_0^t d\tau_3 \int_0^{\tau_3} d\tau_2 \cdots \int_0^{\tau_2} d\tau_1 \\ G(t-\tau_3)[\mathrm{H}'(\tau_3)G(\tau_3-\tau_2)[\mathrm{H}'(\tau_2)G(\tau_2-\tau_1)[\mathrm{H}'(\tau_1)G(\tau_1)\rho(0)]]]\end{aligned} \tag{A.2}$$

Lindblad equation and population transfer

In the CGF approach, population transfer is assumed to arise from fast (and thus memoryless) bath fluctuations, i.e. characterized by rapidly decaying correlation functions. Population transfers can be included phenomenologically by adding fluctuation and dissipation terms to the Liouville–von Neumann equation (Eq. 4), leading to the Lindblad equation

$$\dot{\hat{\rho}} = \frac{i}{\hbar}[\hat{H}, \hat{\rho}] + \sum_{\alpha} \left(\hat{V}_{\alpha} \hat{\rho} \hat{V}_{\alpha}^{\dagger} - \frac{1}{1} \hat{V}_{\alpha} \overset{\dagger\wedge}{V}_{\alpha} \hat{\rho} - \frac{1}{2} \hat{\rho} \hat{V}_{\alpha} \overset{\dagger\wedge}{V}_{\alpha} \right) \tag{A.3}$$

where $\hat{V}_{\alpha}$ are operators describing system–bath couplings in the most general form. Applying the secular approximation to the Green's function, i.e. discarding the fast oscillating coherence terms that decay before population transfer is activated, results in the Pauli master equation describing the population relaxation

$$\frac{\mathrm{d}\rho_{\mathrm{ee}}(\mathrm{t})}{\mathrm{d}t} = -\sum_{e'} K_{ee,e'e'} \rho_{e'e'}(t) \tag{A.4}$$

where K is the rate matrix, with elements $K_{ee,e'e'}$ depicting the rate of population transfer from state e into state e'. The solution of the differential equation is formally given by the population Green's function

$$\rho_{e'e'}(t) = -\sum_{e} G_{e'e',ee}(t) \rho_{ee}(0) \tag{A.5}$$

, and the elements of the matrix $G_{e'e',ee}(t)$ act as time-dependent weighting factors in Liouville pathways, where populations evolve in the excited state, like those of ESAs and SEs (in Eq. 9).

Phase functions and line shape

The phase functions can assume different forms depending on the level of sophistication applied to describe the vibrational dynamics of the system (an overview is given in Refs. [2, 47]). The main building block is the line shape function $g_{ij}(t)$, which is the integral transformation of the autocorrelation function of bath fluctuations

$$g_{ij}(t) = \frac{1}{2\pi} \int \frac{C_{ij}(\omega)}{\omega^2} \left[\coth\left(\frac{\hbar\omega}{2k_B T} \right) (1 - \cos\ \omega t) + i \sin\ \omega t - i\omega t \right] d\omega \tag{A.6}$$

It can be obtained from MD simulations or in closed-form expressions derived from different models. For example, the homogeneous (anti-diagonal) broadening of the spectral signals arising due to coupling to a continuum of fast-decaying low-frequency modes can be expressed by the line shape function of the semi-classical Brownian oscillator (OBO) [124]:

$$g_{ij}^{OBO}(t) = \frac{\lambda_{ij}}{\Lambda}\left(\frac{2k_BT}{\hbar\Lambda} - i\right)(e^{-\Lambda t} + \Lambda t - 1) \tag{A.7}$$

where λ_{ij} and Λ are the system–bath coupling strength and fluctuation timescale, respectively.

References

1. Aue WP, Bartholdi E, Ernst RR (1976) 2-dimensional spectroscopy - application to nuclear magnetic-resonance. J Chem Phys 64(5):2229–2246
2. Mukamel S (1995) Principles of nonlinear optical spectroscopy. O.U.P, New York
3. Zanni MT, Hochstrasser RM (2001) Two-dimensional infrared spectroscopy: a promising new method for the time resolution of structures. Curr Opin Struct Biol 11(5):516–522
4. Jonas DM (2003) Two-dimensional femtosecond spectroscopy. Annu Rev Phys Chem 54:425–463
5. Cowan ML, Ogilvie JP, Miller RJD (2004) Two-dimensional spectroscopy using diffractive optics based phased-locked photon echoes. Chem Phys Lett 386(1–3):184–189
6. Brixner T, Mancal T, Stiopkin IV, Fleming GR (2004) Phase-stabilized two-dimensional electronic spectroscopy. J Chem Phys 121(9):4221–4236
7. Brixner T, Stenger J, Vaswani HM, Cho M, Blankenship RE, Fleming GR (2005) Two-dimensional spectroscopy of electronic couplings in photosynthesis. Nature 434(7033):625–628
8. Collini E, Wong CY, Wilk KE, Curmi PMG, Brumer P, Scholes GD (2010) Coherently wired light-harvesting in photosynthetic marine algae at ambient temperature. Nature 463(7281):644–U669
9. Mukamel S, Abramavicius D, Yang L, Zhuang W, Schweigert IV, Voronine DV (2009) Coherent multidimensional optical probes for electron correlations and exciton dynamics: from nmr to x-rays. Acc Chem Res 42(4):553–562
10. Mukamel S, Bakker HJ (2015) Preface: special topic on multidimensional spectroscopy. J Chem Phys 142(21):212101
11. Fuller FD, Ogilvie JP (2015) Experimental implementations of two-dimensional Fourier transform electronic spectroscopy. Annu Rev Phys Chem 66(66):667–690
12. Selig U, Schleussner C-F, Foerster M, Langhojer F, Nuernberger P, Brixner T (2010) Coherent two-dimensional ultraviolet spectroscopy in fully noncollinear geometry. Opt Lett 35(24):4178–4180
13. C-h Tseng, Matsika S, Weinacht TC (2009) Two-dimensional ultrafast Fourier transform spectroscopy in the deep ultraviolet. Opt Express 17(21):18788–18793
14. Varillas RB, Candeo A, Viola D, Garavelli M, De Silvestri S, Cerullo G, Manzoni C (2014) Microjoule-level, tunable sub-10 fs UV pulses by broadband sum-frequency generation. Opt Lett 39(13):3849–3852
15. Borrego-Varillas R, Oriana A, Ganzer L, Trifonov A, Buchvarov I, Manzoni C, Cerullo G (2016) Two-dimensional electronic spectroscopy in the ultraviolet by a birefringent delay line. Opt Express 24(25):28491–28499
16. Krebs N, Pugliesi I, Hauer J, Riedle E (2013) Two-dimensional Fourier transform spectroscopy in the ultraviolet with sub-20 fs pump pulses and 250–720 nm supercontinuum probe. N J Phys 15(8):085016
17. Baum P, Lochbrunner S, Riedle E (2004) Tunable sub-10-fs ultraviolet pulses generated by achromatic frequency doubling. Opt Lett 29(14):1686–1688

18. Prokhorenko VI, Picchiotti A, Maneshi S, Miller RJD (2015) Broadband electronic two-dimensional spectroscopy in the deep UV. Ultrafast Phenom Xix 162:432–435
19. Tseng C-H, Sandor P, Kotur M, Weinacht TC, 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(11):2654–2661
20. West BA, Womick JM, Moran AM (2011) Probing ultrafast dynamics in adenine with mid-UV four-wave mixing spectroscopies. J Phys Chem A 115(31):8630–8637
21. West BA, Moran AM (2012) Two-dimensional electronic spectroscopy in the ultraviolet wavelength range. J Phys Chem Lett 3(18):2575–2581
22. Prokhorenko VI, Picchiotti A, Pola M, Dijkstra AG, Miller RJD (2016) New insights into the photophysics of DNA nucleobases. J Phys Chem Lett 7(22):4445–4450
23. Jiang J, Mukamel S (2011) Two-dimensional near-ultraviolet spectroscopy of aromatic residues in amyloid fibrils: a first principles study. Phys Chem Chem Phys 13(6):2394–2400
24. Jiang J, Golchert KJ, Kingsley CN, Brubaker WD, Martin RW, Mukamel S (2013) Exploring the aggregation propensity of gamma s-crystallin protein variants using two-dimensional spectroscopic tools. J Phys Chem B 117(46):14294–14301
25. Oliver TAA, Lewis NHC, Fleming GR (2014) Correlating the motion of electrons and nuclei with two-dimensional electronic–vibrational spectroscopy. Proc Natl Acad Sci 111(28):10061–10066
26. Loukianov A, Niedringhaus A, Berg B, Pan J, Senlik SS, Ogilvie JP (2017) Two-dimensional electronic Stark spectroscopy. J Phys Chem Lett 8(3):679–683
27. Kowalewski M, Fingerhut BP, Dorfman KE, Bennett K, Mukamel S (2017) Simulating coherent multidimensional spectroscopy of nonadiabatic molecular processes: from the infrared to the x-ray regime. Chem Rev 117(19):12165–12226
28. Scholes GD, Fleming GR, Chen LX, Aspuru-Guzik A, Buchleitner A, Coker DF, Engel GS, van Grondelle R, Ishizaki A, Jonas DM, Lundeen JS, McCusker JK, Mukamel S, Ogilvie JP, Olaya-Castro A, Ratner MA, Spano FC, Whaley KB, Zhu XY (2017) Using coherence to enhance function in chemical and biophysical systems. Nature 543(7647):647–656
29. Mukamel S (2000) Multidimensional femtosecond correlation spectroscopies of electronic and vibrational excitations. Annu Rev Phys Chem 51:691–729
30. Brixner T, Stiopkin IV, Fleming GR (2004) Tunable two-dimensional femtosecond spectroscopy. Opt Lett 29(8):884–886
31. Tian P, Keusters D, Suzaki Y, Warren WS (2003) Femtosecond phase-coherent two-dimensional spectroscopy. Science 300(5625):1553–1555
32. Grumstrup EM, Shim S-H, Montgomery MA, Damrauer NH, Zanni MT (2007) Facile collection of two-dimensional electronic spectra using femtosecond pulse-shaping technology. Opt Express 15(25):16681–16689
33. Maiuri M, Brazard J (2018) Electronic couplings in (bio-) chemical processes. Top Curr Chem 376(2):10
34. Brańczyk AM, Turner DB, Scholes GD (2014) Crossing disciplines - a view on two-dimensional optical spectroscopy. Ann der Phys 526(1–2):31–49
35. Son M, Schlau-Cohen GS (2017) Ultrabroadband 2D electronic spectroscopy as a tool for direct visualization of pathways of energy flow. Proc SPIE 10(1117/1112):2273417
36. Abramavicius D, Palmieri B, Voronine DV, Sanda F, Mukamel S (2009) Coherent multidimensional optical spectroscopy of excitons in molecular aggregates; quasiparticle versus supermolecule perspectives. Chem Rev 109(6):2350–2408
37. 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(4):767–771
38. Johnson PJM, Farag MH, Halpin A, Morizumi T, Prokhorenko VI, Knoester J, Jansen TLC, Ernst OP, Miller RJD (2017) The primary photochemistry of vision occurs at the molecular speed limit. J Phys Chem B 121(16):4040–4047
39. Bruggemann B, Persson P, Meyer HD, Maya V (2008) Frequency dispersed transient absorption spectra of dissolved perylene: A case study using the density matrix version of the MCTDH method. Chem Phys 347(1–3):152–165
40. Sanda F, Mukamel S (2008) Stochastic Liouville equations for coherent multidimensional spectroscopy of excitons. J Phys Chem B 112(45):14212–14220
41. Tanimura Y (2006) Stochastic Liouville, Langevin, Fokker–Planck, and master equation approaches to quantum dissipative systems. J Phys Soc Jpn 75(8):082001

42. Zimmermann T, Vanicek J (2014) Efficient on-the-fly ab initio semiclassical method for computing time-resolved nonadiabatic electronic spectra with surface hopping or Ehrenfest dynamics. J Chem Phys 141(13):134102
43. Tempelaar R, van der Vegte CP, Knoester J, Jansen TLC (2013) Surface hopping modeling of two-dimensional spectra. J Chem Phys 138(16):164106
44. Richter M, Fingerhut BP (2016) Simulation of multi-dimensional signals in the optical domain: quantum-classical feedback in nonlinear exciton propagation. J Chem Theory Comput 12(7):3284–3294
45. Petit AS, Subotnik JE (2014) Calculating time-resolved differential absorbance spectra for ultrafast pump-probe experiments with surface hopping trajectories. J Chem Phys 141(15):154108
46. Mukamel S (1983) Non-impact unified theory of 4-wave mixing and 2-photon processes. Phys Rev A 28(6):3480–3492
47. Abramavicius D, Valkunas L, Mukamel S (2007) Transport and correlated fluctuations in the nonlinear optical response of excitons. Epl 80(1):17005
48. Jiang J, Mukamel S (2010) Two-dimensional ultraviolet (2DUV) spectroscopic tools for identifying fibrillation propensity of protein residue sequences. Angew Chem Int Ed 49(50):9666–9669
49. Stuhldreier MC, Temps F (2013) Ultrafast photo-initiated molecular quantum dynamics in the DNA dinucleotide d(apg) revealed by broadband transient absorption spectroscopy. Faraday Discuss 163:173–188; discussion 243–175
50. Ostroumov EE, Mulvaney RM, Cogdell RJ, Scholes GD (2013) Broadband 2D electronic spectroscopy reveals a carotenoid dark state in purple bacteria. Science 340(6128):52–56
51. Dean JC, Rafiq S, Oblinsky DG, Cassette E, Jumper CC, Scholes GD (2015) Broadband transient absorption and two-dimensional electronic spectroscopy of methylene blue. J Phys Chem A 119(34):9098–9108
52. 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
53. 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 1040:295–303
54. Nenov A, Beccara SA, 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. ChemPhysChem 15(15):3282–3290
55. Nenov A, Segarra-Marti J, Giussani A, Conti A, Rivalta I, Dumont E, Jaiswal VK, Altavilla SF, Mukamel S, Garavelli M (2015) Probing deactivation pathways of DNA nucleobases by two-dimensional electronic spectroscopy: First principles simulations. Faraday Discuss 177:345–362
56. 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
57. Nenov A, Giussani A, Segarra-Martí J, Jaiswal VK, Rivalta I, Cerullo G, Mukamel S, Garavelli M (2015) Modeling the high-energy electronic state manifold of adenine: calibration for nonlinear electronic spectroscopy. J Chem Phys 142(21):212443
58. 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
59. Giussani A, Segarra-Martí J, Nenov A, Rivalta I, Tolomelli A, Mukamel S, Garavelli M (2016) Spectroscopic fingerprints of DNA/RNA pyrimidine nucleobases in third-order nonlinear electronic spectra. Theor Chem Acc 135(5):1–18
60. Li Q, Giussani A, Segarra-Martí J, Nenov A, Rivalta I, Voityuk AA, Mukamel S, Roca-Sanjuán D, Garavelli M, Blancafort L (2016) Multiple decay mechanisms and 2D-UV spectroscopic fingerprints of singlet excited solvated adenine-uracil monophosphate. Chem—A Eur J 22(22):7497–7507
61. Segarra-Marti J, Jaiswal VK, Pepino AJ, Giussani A, Nenov A, Mukamel S, Garavelli M, Rivalta I (2018) Two-dimensional electronic spectroscopy as a tool for tracking molecular conformations in DNA/RNA aggregates. Faraday Discuss 207:233–250
62. Kim J, Mukamel S, Scholes GD (2009) Two-dimensional electronic double-quantum coherence spectroscopy. Acc Chem Res 42(9):1375–1384
63. Li Z, Abrarnavicius D, Mukamel S (2008) Probing electron correlations in molecules by two-dimensional coherent optical spectroscopy. J Am Chem Soc 130(11):3509–3515

64. Roos BO (1987) The complete active space self-consistent field method and its applications in electronic structure calculations. Adv Chemical Phys. https://doi.org/10.1002/9780470142943.ch7
65. Andersson K, Malmqvist PA, Roos BO, Sadlej AJ, Wolinski K (1990) 2nd-order perturbation-theory with a cassef reference function. J Phys Chem 94(14):5483–5488
66. 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
67. 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
68. Aquilante F, Lindh R, Pedersen TB (2007) Unbiased auxiliary basis sets for accurate two-electron integral approximations. J Chem Phys 127(11):114107
69. 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 (2016) Molcas 8: new capabilities for multiconfigurational quantum chemical calculations across the periodic table. J Comput Chem 37(5):506–541
70. Avila Ferrer FJ, Cerezo J, Stendardo E, Improta R, Santoro F (2013) Insights for an accurate comparison of computational data to experimental absorption and emission spectra: beyond the vertical transition approximation. J Chem Theory Comput 9(4):2072–2082
71. Serrano-Andrés L, Merchán M, Nebot-Gil I, Lindh R, Roos BO (1993) Towards an accurate molecular-orbital theory for excited-states—ethene, butadiene, and hexatriene. J Chem Phys 98(4):3151–3162
72. Lorentzon J, Malmqvist P-Å, Fülscher M, Roos BO (1995) A caspt2 study of the valence and lowest rydberg electronic states of benzene and phenol. Theor Chim Acta 91(1):91–108
73. Serrano-Andrés L, Roos BO (1996) Theoretical study of the absorption and emission spectra of indole in the gas phase and in a solvent. J Am Chem Soc 118(1):185–195
74. Barbatti M, Aquino AJA, Lischka H (2010) The UV absorption of nucleobases: semi-classical ab initio spectra simulations. Phys Chem Chem Phys 12(19):4959–4967
75. 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
76. 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
77. Yamada T, Fukutome H (1968) Vacuum ultraviolet absorption spectra of sublimed films of nucleic acid bases. Biopolymers 6(1):43–54
78. 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
79. Platt JR (1949) Classification of spectra of cata-condensed hydrocarbons. J Chem Phys 17(5):484–495
80. Roos BO, Andersson K, Fülscher MP, Serrano-Andrés L, Pierloot K, Merchán M, Molina V (1996) Applications of level shift corrected perturbation theory in electronic spectroscopy. J Mol Struct 388:257–276
81. Roos BO, Andersson K (1995) Multiconfigurational perturbation-theory with level shift - the cr-2 potential revisited. Chem Phys Lett 245(2–3):215–223
82. Giussani A, Marcheselli J, Mukamel S, Garavelli M, Nenov A (2017) On the simulation of two-dimensional electronic spectroscopy of indole-containing peptides. Photochem Photobiol 93(6):1368–1380
83. Hamm P, Zanni M (2011) Concepts and methods of 2D infrared spectroscopy. Cambridge University Press, Cambridge
84. Camilloni C, Broglia RA, Tiana G (2011) Hierarchy of folding and unfolding events of protein G, CI2, and ACBP from explicit-solvent simulations. J Chem Phys 134(4):045105
85. Ezra FS, Lee CH, Kondo NS, Danyluk SS, Sarma RH (1977) Conformational properties of purine-pyrimidine and pyrimidine-purine dinucleoside monophosphates. Biochemistry 16(9):1977–1987
86. Crespo-Hernandez CE, Cohen B, Hare PM, Kohler B (2004) Ultrafast excited-state dynamics in nucleic acids. Chem Rev 104(4):1977–2019

87. Takaya T, Su C, Harpe KdL, Crespo-Hernández CE, Kohler B (2008) UV excitation of single DNA and RNA strands produces high yields of exciplex states between two stacked bases. Proc Natl Acad Sci 105:10285–10290
88. Ruzicka P, Kral T (2013) Pyrene: chemical properties, biochemistry applications and toxic effects. Chemistry research and applications. Nova Science Publishers, New York
89. Reichardt C (1994) Solvatochromic dyes as solvent polarity indicators. Chem Rev 94(8):2319–2358
90. Kwon J, Park SK, Lee Y, Lee JS, Kim J (2017) Tailoring chemically converted graphenes using a water-soluble pyrene derivative with a zwitterionic arm for sensitive electrochemiluminescence-based analyses. Biosens Bioelectron 87:89–95
91. Figueira-Duarte TM, Mullen K (2011) Pyrene-based materials for organic electronics. Chem Rev 111(11):7260–7314
92. Niko Y, Didier P, Mely Y, Konishi G, Klymchenko AS (2016) Bright and photostable push-pull pyrene dye visualizes lipid order variation between plasma and intracellular membranes. Sci Rep 6:18870
93. 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
94. Petit AS, Subotnik JE (2014) How to calculate linear absorption spectra with lifetime broadening using fewest switches surface hopping trajectories: a simple generalization of ground-state Kubo theory. J Chem Phys 141(1):014107
95. Nemeth A, Milota F, Mancal T, Pullerits T, Sperling J, Hauer J, Kauffmann HF, Christensson N (2010) Double-quantum two-dimensional electronic spectroscopy of a three-level system: experiments and simulations. J Chem Phys 133(9):094505
96. Butkus V, Zigmantas D, Valkunas L, Abramavicius D (2012) Vibrational vs. Electronic coherences in 2D spectrum of molecular systems. Chem Phys Lett 545:40–43
97. Butkus V, Valkunas L, Abramavicius D (2012) Molecular vibrations-induced quantum beats in two-dimensional electronic spectroscopy. J Chem Phys 137(4):044513
98. Raytchev M, Pandurski E, Buchvarov I, Modrakowski C, Fiebig T (2003) Bichromophoric interactions and time-dependent excited state mixing in pyrene derivatives. A femtosecond broad-band pump-probe study. J Phys Chem A 107(23):4592–4600
99. Krebs N (2013) New insights for femtosecond spectroscopy: from transient absorption to 2-dimensional spectroscopy in the UV spectral domain. PhD Dissertation, Faculty of Physics, Ludwig-Maximilians-University Munich
100. Zhang WM, Meier T, Chernyak V, Mukamel S (1998) Exciton-migration and three-pulse femtosecond optical spectroscopies of photosynthetic antenna complexes. J Chem Phys 108(18):7763–7774
101. Meier T, Chernyak V, Mukamel S (1997) Femtosecond photon echoes in molecular aggregates. J Chem Phys 107(21):8759–8780
102. Neuwahl FVR, Foggi P (1999) Direct observation of s2–s1 internal conversion in pyrene by femtosecond transient absorption. Laser Chem 19(1–4):375–379
103. Foggi P, Pettini L, Santa I, Righini R, Califano S (1995) Transient absorption and vibrational relaxation dynamics of the lowest excited singlet state of pyrene in solution. J Phys Chem 99(19):7439–7445
104. 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. Topics in current chemistry. Springer International Publishing, Berlin, vol 356, pp 39–87. https://doi.org/10.1007/128_2014_570
105. Bucher DB, Kufner CL, Schlueter A, Carell T, Zinth W (2016) UV-induced charge transfer states in DNA promote sequence selective self-repair. J Am Chem Soc 138(1):186–190
106. Schreier WJ, Gilch P, Zinth W (2015) Early events of DNA photodamage. Annu Rev Phys Chem 66(1):497–519
107. Vayá I, Gustavsson T, Douki T, Berlin Y, Markovitsi D (2012) Electronic excitation energy transfer between nucleobases of natural DNA. J Am Chem Soc 134(28):11366–11368
108. 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. Topics in current chemistry. Springer International Publishing, Berlin, vol 356, pp 249–275. doi:10.1007/128_2014_553
109. 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

110. Pecourt JML, Peon J, Kohler B (2001) DNA excited-state dynamics: ultrafast internal conversion and vibrational cooling in a series of nucleosides. J Am Chem Soc 123(42):10370–10378
111. Peon J, Zewail AH (2001) DNA/RNA nucleotides and nucleosides: Direct measurement of excited-state lifetimes by femtosecond fluorescence up-conversion. Chem Phys Lett 348(3–4):255–262
112. Onidas D, Markovitsi D, Marguet S, Sharonov A, Gustavsson T (2002) Fluorescence properties of DNA nucleosides and nucleotides: a refined steady-state and femtosecond investigation. J Phys Chem B 106(43):11367–11374
113. Hare PM, Crespo-Hernández CE, Kohler B (2007) Internal conversion to the electronic ground state occurs via two distinct pathways for pyrimidine bases in aqueous solution. Proc Natl Acad Sci 104:435–440
114. Pepino AJ, Segarra-Martí J, Nenov A, Improta R, Garavelli M (2017) Resolving ultrafast photoinduced deactivations in water-solvated pyrimidine nucleosides. J Phys Chem Lett 8(8):1777–1783
115. Segarra-Marti 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
116. Vogiatzis KD, Li Manni G, Stoneburner SJ, Ma D, Gagliardi L (2015) Systematic expansion of active spaces beyond the casscf limit: a gasscf/splitgas benchmark study. J Chem Theory Comput 11(7):3010–3021
117. Li Manni G, Carlson RK, Luo S, Ma D, Olsen J, Truhlar DG, Gagliardi L (2014) Multiconfiguration pair-density functional theory. J Chem Theory Comput 10(9):3669–3680
118. Casida ME, Huix-Rotllant M (2012) Progress in time-dependent density-functional theory. Annu Rev Phys Chem 63:287–323
119. Dreuw A, Wormit M (2015) The algebraic diagrammatic construction scheme for the polarization propagator for the calculation of excited states. Wiley Interdiscip Rev 5(1):82–95
120. Sneskov K, Christiansen O (2012) Excited state coupled cluster methods. Wiley Interdiscip Rev 2(4):566–584
121. Chan GK-L, Sharma S (2011) The density matrix renormalization group in quantum chemistry. Annu Rev Phys Chem 62(1):465–481
122. Thomas RE, Sun Q, Alavi A, Booth GH (2015) Stochastic multiconfigurational self-consistent field theory. J Chem Theory Comput 11(11):5316–5325
123. Segarra-Martí J, Zvereva E, Marazzi M, Brazard J, Dumont E, Assfeld X, Haacke S, Garavelli M, Monari A, Léonard J, Rivalta I (2018) Resolving the singlet excited state manifold of benzophenone by first-principles simulations and ultrafast spectroscopy. J Chem Theory Comput 14(5):2570–2585
124. Li BL, Johnson AE, Mukamel S, Myers AB (1994) The Brownian oscillator model for solvation effects in spontaneous light-emission and their relationship to electron-transfer. J Am Chem Soc 116(24):11039–11047