Towards Accurate Simulations of Two-Dimensional Electronic Spectroscopy

Javier Segarra-Martí,[a] *Shaul Mukamel*[b]*, Marco Garavelli**[c]*, Artur Nenov**[c] *and Ivan Rivalta**[a]

[a]Université de Lyon, École Normale Supérieure de Lyon, CNRS, Université Claude Bernard Lyon 1, France.
[b]Department of Chemistry, University of California, Irvine, California 92697-2025, United States.
[c]Dipartimento di Chimica Industriale, Università degli Studi di Bologna, Viale del Risorgimento 4, I-40136 Bologna, Italy.

Abstract

In this chapter, we introduce the basic concepts of 2D electronic spectroscopy (2DES) and a general theoretical framework adopted to calculate, from first principles, the non-linear response of multichromophoric 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 simulations of their 2DUV spectra. The roadmap for accurate 2DUV spectroscopy simulations is illustrated starting with benchmark of the excited states manifold of the chromophoric units in vacuum, which could be used for building exciton Hamiltonians for large scale applications or as 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 dynamical 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, we observe a remarkable agreement between the predicted 2DUV spectra of the pyrene molecule and experiment. These results stimulated theoretical studies of the excited states dynamics in a solvated dinucleoside system, showing that spectroscopic fingerprints of long-living excited states 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 simulations of 2DES spectra in revealing complex physico-chemical properties of fundamental biological systems and should trigger further theoretical developments as well as new experiments.

1. Introduction

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 on several fields, with first handover to the infrared (IR) regime (2DIR) now being 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 electronic couplings by 2D electronic spectroscopy,[4] has become feasible thanks to [5,6] advances in ultrafast optical techniques that allow the targeting of electronic transitions in the visible range. 2DES in the visible (2DVis) has become increasingly popular in the last decade, showcasing its great potential by deciphering energy transfer processes in photosynthesis.[7-10] The desirable extension to the UV domain (2DUV), where many fundamental biomolecules display strong absorption bands, has been so far slowed down by technical difficulties associated to the use of UV laser pulses.[11] Recent progress,[12-18] attaining interferometric phase stability and sufficient laser bandwidth, granted access to the first examples of experimental 2DUV spectra.[19-22] Simulation studies have demonstrated that 2DUV of aromatic residues in the near UV and the backbone in the far UV can effectively probe protein secondary structure.[23,24] Among the most recent development in multidimensional spectroscopy related to 2DES, it is important to highlight the advances of the electronic-vibrational spectroscopy [25], the 2D Stark spectroscopy for characterizing dark charge transfer states[26], along with the impressive potential of techniques developing in the X-ray regime.[27]

The signals recorded in 2DES correlation plots refer to the third-order nonlinear response of the sample, 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 call for the implementation of computational tools based on a combination of *ab initio* electronic structure and the nonlinear response formalism. Accurate simulations of 2DUV electronic spectroscopy might, in fact, lead to both interpretation and prediction of nonlinear spectra, enhancing the potential application of this technique in various fields. In this chapter, we draw a route towards accurate simulations of 2DUV spectra of fundamental biological systems. After describing the general 2DES technique and the target UV-active bio-chromophores in this Section 1, a basic background is given in Section 2 aimed at providing the theoretical foundations of our simulation protocols and showing the approximations adopted. In Section 3 the main results of our developments and applications are illustrated for both nucleic acids and proteic model systems, showing the potential of 2DUV in characterizing the ground state conformational space and excited states dynamics in these target systems. Finally an outlook summary and a perspective view are provided in Section 4.

1.1 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 (Figure 1). 2DES major advantage is the higher spectral resolution with respect to 1D-PP time-resolved techniques, with the nonlinear signal containing information on system dynamics and electronic couplings being spread over two frequency axes (pump and probe, Ω_1 and Ω_3, respectively), see Fig. 1b. The spectral resolution in two dimensions allows accurate characterization of inhomogeneous and homogeneous broadening processes, with the ability to separate these two contributions (not possible with 1D techniques) giving information about solvent reorganization timescales, as well as allowing 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[30,5] 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 ($k_{LO} = -k_1 + k_2 + k_3$) and non-rephasing K_{II} ($k_{LO} = +k_1 - k_2 + k_3$) 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 non-linear response signal

which is heterodyned by the probe pulse itself (self-heterodyning). The PCPP experiment has the drawback of strong background signal and it intrinsically provides the combined $K_I + K_{II}$ "quasi-absorptive" response, with lost information on specific rephasing and non-rephasing signals, while removal of slowly decaying dispersive contributions due to phase cancelation of the concurring K_I and K_{II} signals might help resolving weak signals.

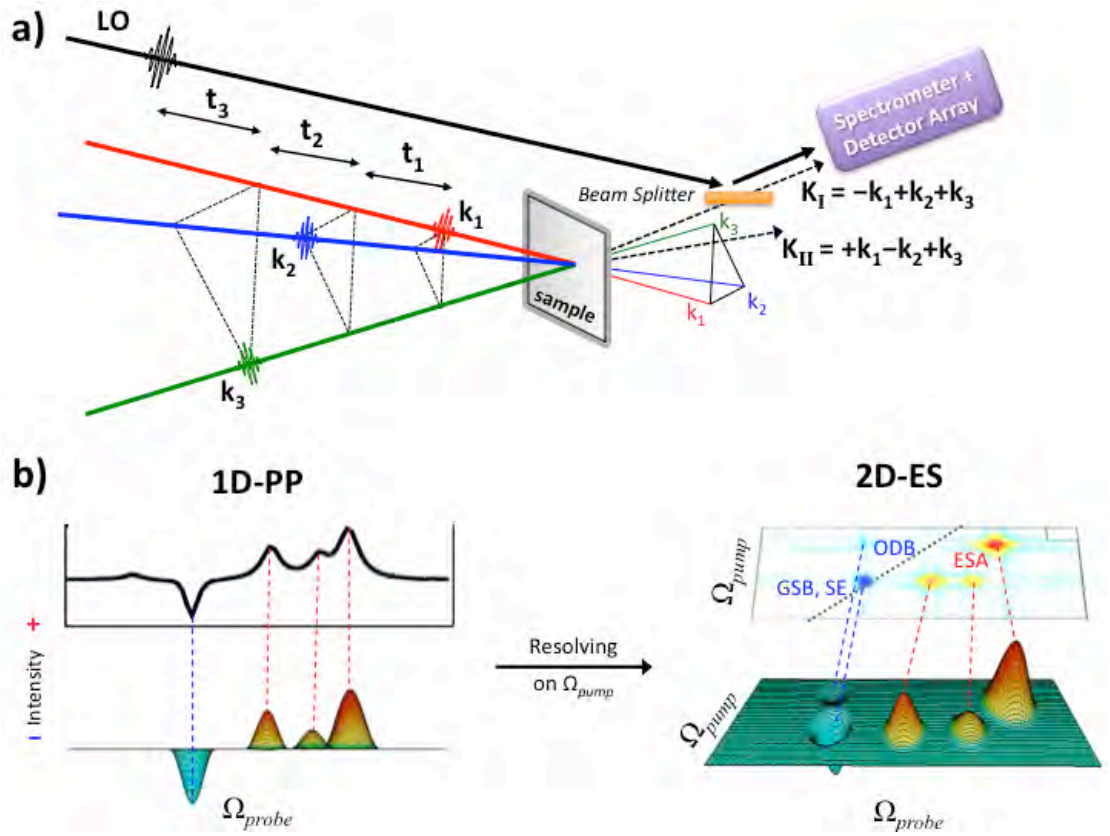

Figure 1. Schematic representation of (a) heterodyne detected three-pulse photon echo (3PPE) 2DES experiment (LO: local oscillator) showing both rephasing K_I ($k_{LO} = -k_1 + k_2 + k_3$) and non-rephasing K_{II} ($k_{LO} = +k_1 - k_2 + k_3$) phase-matching directions, and (b) 1D pump-probe (1D-PP) and analogous 2D electronic spectra (2D-ES) at a fixed value of the waiting time t_2, showing how the different types shapes and signs of the non-linear 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 function of "excitation frequency" Ω_1 (Ω_{pump}) and "detection frequency" Ω_3 (Ω_{probe}) by Fourier transforming with respect to the t_1 and t_3 time intervals. Thus, the nonlinear spectral signatures of the ground state are obtained by setting the time t_2 to zero, while the state-specific response along excited states dynamics can be monitored for $t_2 > 0$.

As illustrated in Figure 1b, the induced non-linear polarization measured in a 2DES map, at a certain population time, is richer than that of a 1D-PP spectrum as by adding the spectral resolution along Ω_{pump} is it possible to disentangle the contributions arising from different pump frequencies and the line broadenings along both pump and probe frequencies. In general, a 2DES spectrum at t_2=0 contains diagonal peaks that correspond to ground-to-excited states absorptions (ground state bleaching, GSB) and stimulate emissions (SE), symmetric off-diagonal signals associated with coupling between ground-to-excited states electronic excitations (off-diagonal bleaching, ODB) and asymmetric off-diagonal peaks (with opposite sign with respect to diagonal peaks) due to excited state absorptions (ESAs) to the high-lying states. For $t_2 > 0$, the excited states 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 more details on 2DES experimental techniques and 2DES maps features, we recommend to refer to the review of Maiuri and Brazard present in this series.

1.2 2DES in the Ultraviolet: a new light in photobiology

The emergence of novel UV pulse technologies[14] has allowed the use of pump and/or probe laser pulses in the UV window. These efforts might have a particular impact in monitoring biologically relevant processes.[33] Indeed, cyclic aromatic groups are often the UV-active chromophores that play a major role in photoinduced events in biological systems.[34] UV light, in fact, is absorbed by an incredibly large portion of biomolecules, ranging from proteins to nucleic acids. As showed in Figure 2, both proteic and nucleic acid backbones absorb 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 acids sidechains and adenine (Ade), guanine (Gua), cytosine (Cyt), thymine (Thy) and uracil (Ura) nucleobases, feature their most intense absorptions absorb in the same spectral window of their backbones. However, all these aromatic units (except histidine) have distinctive absorption bands associated to $\pi \rightarrow \pi^*$ 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 near-UV range around 300 and 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 cite some. This evidence suggests that the low-energy UV window represents a rather promising option to monitor many photoinduced biological phenomena involving aromatic units in protein or DNA/RNA and eventually coenzymes.

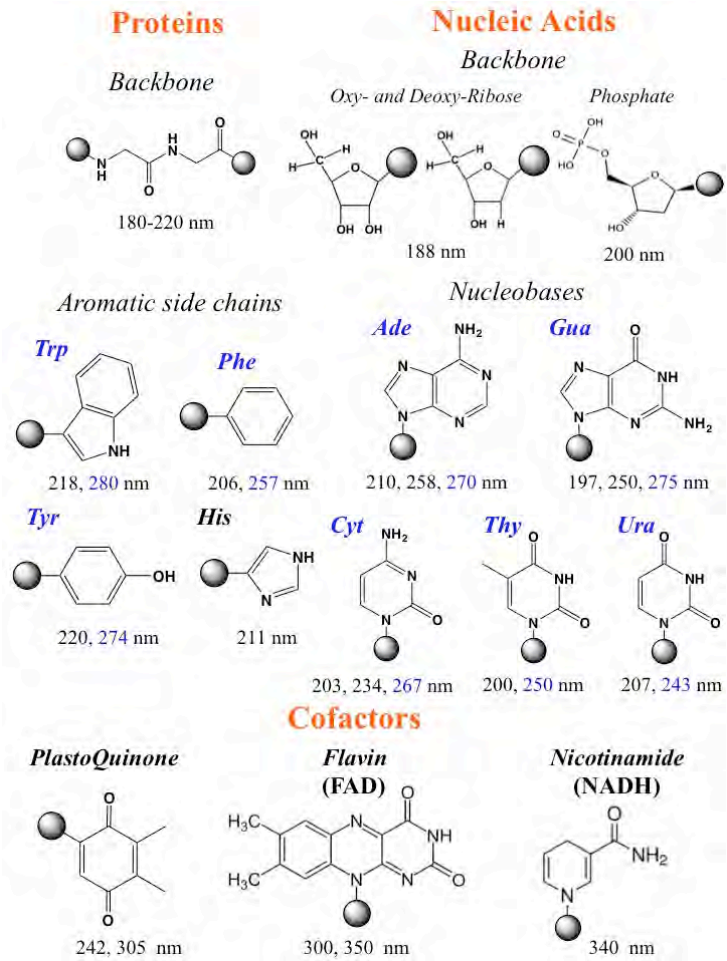

Figure 2. UV-absorbing bio-chromophores in proteins and nucleic acids together with associated absorption wavelengths. States involved in absorptions out of the spectral region of backbones are highlighted in blue for the UV-chromophores reported in this review. Their respective backbone units, absorbing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

Thus, since 2DES experiments might provide unique information on electronic coupling and related electron/energy transfer processes (see section 1.1), applications of this nonlinear spectroscopy

technique in the UV range could shed light on a vast number of biological processes, from various enzymatic redox catalysis to photosynthesis, and it could be employed to obtain either structural information (in the ground state) or to track photo-induced excited state events. Notably, given its high spectral and temporal resolutions, 2DUV spectroscopy can be used to entangle complex decay processes as in DNA/RNA polynucleotides, where multiple chromophores exhibit various competitive deactivation pathways with time scales from sub-100 femtoseconds to nanoseconds. In this context, developing computational protocols for accurate simulations of 2DUV spectroscopy is highly desirable, as coupling with current advancements of experimental techniques it would finally lead to a comprehensive understanding.
In the following sections, after providing the basic theoretical background of 2DES and the approximation frameworks adopted in our studies, we review our recent efforts on computational developments for simulating 2DUV spectra of two main biological macromolecules, i.e. proteins and nucleic acids, focusing on the chromophoric units that provide absorptions distinct from their backbones.

2. Theoretical background

2.1 Density matrix formalism and 2DES response functions

An unified theory for nonlinear optical spectroscopy [2] has been developed using the 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 we have mentioned above, the source of the signal field recorded in 2DES experiments is the induced non-linear polarization, which is defined as the expectation value of the dipole operator $\hat{\mu}$

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

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 in general molecules)

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

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

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

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,

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

where $\hat{H}$ is the system Hamiltonian. In 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

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

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

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

where the superscript denotes the nth 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 nth order polarization $P^{(n)}(t) = \left\langle \hat{\mu}\hat{\rho}^{(n)}(t) \right\rangle = Tr\left[\hat{\mu}\hat{\rho}^{(n)}(t)\right]$ becomes in third-order

$$P^{(3)}(t) = Tr\left[\hat{\mu}\hat{\rho}^{(3)}(t)\right] = \\ = \int_0^{\infty} dt_3 \int_0^{\infty} dt_2 \int_0^{\infty} dt_1 R^{(3)}(t_3, t_2, t_1) \times \\ \times E(t - t_3) E(t - t_3 - t_2) E(t - t_3 - t_2 - t_1) \tag{7}$$

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

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

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

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

that are associated with different physical processes occurring in the system during the interaction with the external fields (see GSB, SE and ESA in section 1.1) that can be represented by Feynman diagrams. Commonly, 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 non-linear 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 obscuration of any coherent dynamics, which occurs on the same time scale as the duration of the pulses, is avoided and the highest possible temporal resolution is achieved. The impulsive limit has been adopted throughout this chapter.

Simulation protocols for non-linear optical spectroscopy deal with solving eq. (7), or more in general eq. 8 when working with realistic pulse shapes. Evidently, 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 provided knowledge of the multi-dimensional potentials associated with each electronic state involved in the evolution of the density matrix, allowing to propagate populations and coherences. Non-adiabatic couplings and transition dipole moments along the potential energy surfaces (PES) would be required, in this case, to enable population transfer and the interaction with the field, respectively. It would be then 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 wavepacket decoherences and recoherences. Furthermore, arbitrary pulses could be straightforwardly incorporated in the simulations giving the opportunity to study the effects of pulse's central frequency, bandwidth and polarization.[35,36]
The drawback of quantum dynamics simulations is associated with the high cost of pre-computing the

PES, a task that quickly becomes unfeasible with 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.[37] Methods like multi-configuration time-dependent Hartree have facilitated simulations on multi-dimensional surfaces.[38]

Another approach toward solving eqs. (7) and (8) involves the stochastic modeling of bath fluctuations in a semi-classical fashion (stochastic Liouville equations, SLE)[39,40]. At the heart of SLE there is the inclusion of collective bath modes explicitly in the system's Hamiltonian. The Hamiltonian is formally time-independent, but depends on some classical time-dependent stochastic variables *σ(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 time-step following it is evaluated as

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

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

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

Since the 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 the ensemble averaging. Trajectory based-approaches face the challenges of the computational cost associated with the propagation of swarm of trajectories (particularly those in the excited states) 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 the past years by Vanicek,[41] Jansen,[42] Fingerhut,[43] and Subotnik.[44]

For systems with classical bath following Gaussian statistics and linear system-bath coupling, eq. (8) can be solved using the second-order cumulant expansion, i.e. the cumulant expansion of Gaussian fluctuations (CGF) method.[45] This method allows calculating the shapes of electronic transition bands coupled to a bath (for fluctuations with arbitrary time scales) by using the formalism of lineshape 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 allowing partitioning the nonlinear response (eq. 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 [35]), the population ($e'=e$ during delay time t_2) contributions read

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

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

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

where ε_a, with $a\in\{g,e,f\}$, are the energies of the eigenstates of the system's Hamiltonian H_0, their energies differences ($\varepsilon_a-\varepsilon_b$) being the transition energies (TEs), μ_{ab} are the associated transition dipole moments (TDMs), $\varphi_{e'fe}(\tau_4,\tau_3,\tau_2,\tau_1)$ 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 thus depend on the level of sophistication adopted for describing the system vibrational dynamics. [46,2] Anyway, the phase function is based on the lineshape functions, two examples of phase functions build from lineshape functions will be given in sections 3.3 and 3.4 (eqs. 16 and 19, respectively).

2.2 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, representing in general a great challenge even for the most advanced computational techniques, despite a plethora of quantum-chemical (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 large number of excited states in the target manifolds, might already border the limits of what is currently computationally feasible. Then treating at *ab initio* level multiple interacting chromophores, i.e. supermolecular aggregates, becomes quite rapidly prohibitive with increasing number or size of the chromophoric units. At the same time, as showed in eq. 12-13, 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 to cope with these limitations is to adopt Frenkel exciton models that make simulations of 2DES spectra of realistic model systems computationally feasible.[35,47] Exciton modeling works well for vibrational excited states absorptions in 2DIR spectroscopy (namely overtone absorptions), since anharmonicities for overtones of a chromophore (local overtones) or between coupled chromophores can be reliably computed perturbatively (without explicit knowledge of the high-lying energy levels 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. Then, the electronic couplings in multichromophoric systems can 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 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 associated 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 of the spectral window of interest. However, broadband transient absorption [48] and, more recently, also 2D [49,50,16] 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, allow studying selectively 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,[51,36,52-61] as well as double quantum coherences spectroscopy.[52,62,63]

As it will be illustrated along section 3, obtaining the electronic structure of biologically relevant chromophores from first principles, with the accuracy required to compare 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 section 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 to envision problem-driven experimental setups. To this aim, we take a step back in the theoretical descriptions given in section 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 reduces to a $e^{-i\varepsilon_f t}$ phase factor (here named *static picture*).

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

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

and

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

All dephasing processes are, thus, condensed into the phenomenological dephasing constants γ_{ab}, which induce homogeneous broadening of all $a\rightarrow b$ electronic transitions. Eqs. 14 and 15 allow simulating the nonlinear responses for systems with hundreds of excited states assuming all the corresponding TEs (ε_b-ε_a) and TDMs (μ_{ab}), a protocol known as the sum-over-states (SOS). Figure 3 shows more in details, for the specific case of an ESA signal recorded in the rephasing K_I 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).

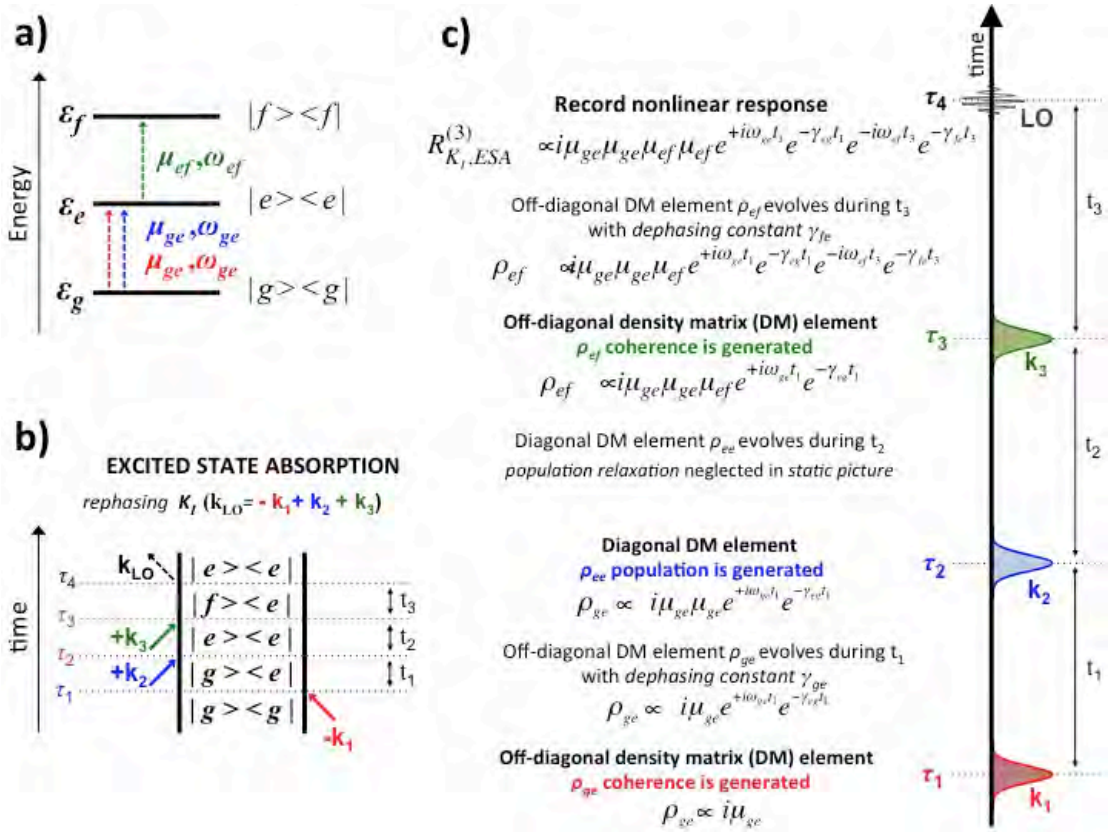

Figure 3. (a) A three-level system, showing transition dipole moments (μ) and frequencies (ω) associated to the ground (g) and excited (e,f) states and their energies ($\varepsilon_{g,e,f}$). (b) Feynman diagram of an ESA signal recorded in the rephasing K_I phase-matching direction. (c) Schematic representation 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).

Recently,[51] we have combined this quasi-static (as all dynamical effects have been neglected) protocol with a hybrid scheme combining QM electronic structure with a molecular mechanics (MM) treatment of environmental effects (i.e. the SOS//QM/MM 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 3, mainly evidencing how a number of bright excited states in the Vis-UV spectral range often represent the major contributions to 2DES maps and could embody specific spectroscopic fingerprints for tracking ground or excited state dynamics. Accurate characterization of the excited states manifold by advanced electronic structure computations thus represents a pillar for reliable simulations of 2D electronic spectra.

3. Results and discussion

The minimal theoretical background for understanding how to compute non-linear electronic spectra has been illustrated in Section 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 side, practical use of quantum dynamics allows modeling spectral dynamics accurately, yet only relying 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 end, the electronic structure complexity can be straightforwardly incorporated in a static approach, recovering ESA contributions. However, as dynamics is neglected, this approach is not able to provide realistic spectral line shapes. Trajectory based approaches offer a compromise by permitting to adjust the level of precision applied to molecular dynamics and PES sampling. In the following, we begin our discussion by utilizing the static approach to outline a protocol for the accurate computation of transition energies and dipole moments (sec. 3.1), as well as for benchmarking low-cost methods. Section 3.2 and 3.3 demonstrate how the latter can be employed to study ground state conformational dynamics of oligopeptides and DNA nucleobase dimers within the framework of the SOS//QM/MM protocol. Subsequently, we go beyond the static approximation and re-introduce dynamical features in eqs. 14 and 15 within the CGF framework for the study of excited states. In particular, coherent intramolecular vibrational dynamics is discussed in sec. 3.3, while sec. 3.4 covers population transfer.

3.1 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 intra-molecular) charge transfer characters and thus involving covalent as well as ionic (valence-bond like) excited states. Wavefunctions approaches introduced in the 80's, based on the combination of CASSCF multiconfigurational wavefunctions[64] and PT2 perturbative energy corrections[65] (CASSCF//CASPT2), currently represent the most widely used methodology to handle such variety of excited states at equal footing,[66] generally providing good quantitative predictions of TEs and TDMs, with expected errors 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 during the years, including the implementation of the restricted active space (RASSCF) methodology,[67] efficient approximations for two-electron integrals 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 just PT2) has been likely preferred to the more expensive (and in most of the cases unfeasible) multistate PT2 approach. However, the accuracy of CASSCF//PT2 predictions strongly depends on the choice of active spaces (AS) and basis sets, two parameters that encompass a critical increase of the computational costs. 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 multiconfigurational wavefunctions, with a consequent further increase of the computational costs. To cope with the computational feasibility of the multiconfigurational treatments, we have generally adopted the RASSCF methodology, as its flexibility allows either improving the electronic wavefunction of the largest CASSCF computation feasible, usually involving the full valence π-orbital space, or reducing the overall computational cost by limiting the CAS treatment to a relatively small active space (RAS2).

In the RASSCF approach, indeed, extra virtual orbitals can be included in the active space where a restricted number of electrons is allowed (RAS3 space), increasing the number of electronic excitations and expanding the multiconfigurational character of the electronic wavefunctions with respected 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 scope we generally attempted to make a direct comparison between vertical excitations in the Franck-Condon (FC) region, i.e. at the ground state equilibrium geometries, with experimentally recorded cross sections in vacuum, when available. The choice of the gas phase is due to the necessity of having a comparison in absence of environmental effects that would complicate the computational modelling strategy as well as reducing the experimental spectral resolution (as line broadenings are larger in solution), aiming 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 associated to the Phe and Tyr amino acid aromatic side chains (see Figure 2). As showed in Figure 4, CASSCF//PT2 computations of the $S_0 \rightarrow S_N$ vertical excitations (in the FC region) indicate how the full valence active spaces of benzene and phenol, i.e. CAS(6,6) and CAS(8,7), respectively, are not accurate enough if compared with experimental cross sections values in vacuum. [58] Instead, increasing the active space by allowing single and double excitations to additional extravalence RAS3 orbitals, allows an appropriate description of the dynamical σ-π polarization and significantly improves the accuracy, reaching convergence with 8 extra orbitals and a quantitative agreement with experiments. It is worth mentioning that these computations have been performed applying a computational recipe to eliminate Rydberg contamination in the valences states after adding a set of diffuse and uncontracted basis functions in the centre of charge of the molecules (ANO-L-aug), [71,72] which has been further used to simulations of dimers of amino acid sidechains and other chromophores.

While the $S_0 \rightarrow S_N$ excitations do not cover all the possible excitations involved in a 2DUV experiment, as the $S_N \rightarrow S_M$ excitations would strongly contribute in the 2D spectra as ESA signals, the energy range considered in this benchmark study involve 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 centred 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 level. This outcome suggests that theoretical predictions of vertical excitations in the UV can be reliably obtained by extended RAS schemes. Notably, TDMs showed convergence as well as the TEs, thus providing robust theoretical and reference RAS schemes to employed in further studies.

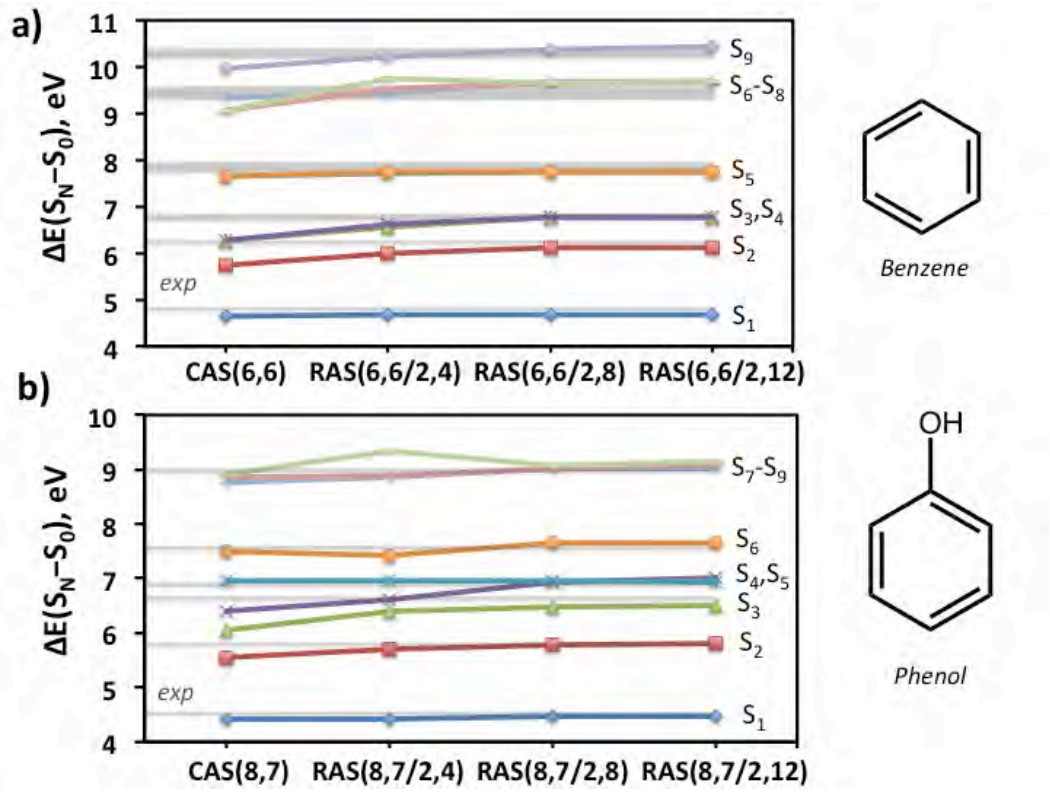

Figure 4. CASSCF//PT2 and RASSCF//PT2 vertical $S_0 \rightarrow S_N$ excitation energies (in eV) of benzene and phenol monomers in vacuum obtained using the large ANO-L(432,21)-aug basis set at different levels of theory. Minimal active spaces CAS(X,Y) results are compared to restricted active spaces

(RAS2/RAS3) calculations with increasing number of virtual orbitals (4, 8 and 12) and experimental values (grey bars, indicating ranges of experimental absorption maxima recorded). Reproduced from data reported in Ref [58].

When extending the benchmark study of benzene and phenol monomers [Ref] to the remaining aromatic proteic 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 gives a direct visualization of the influence of the level of theory adopted on the resulting 2D map, which is shaped by both TEs and TDMs computed values.

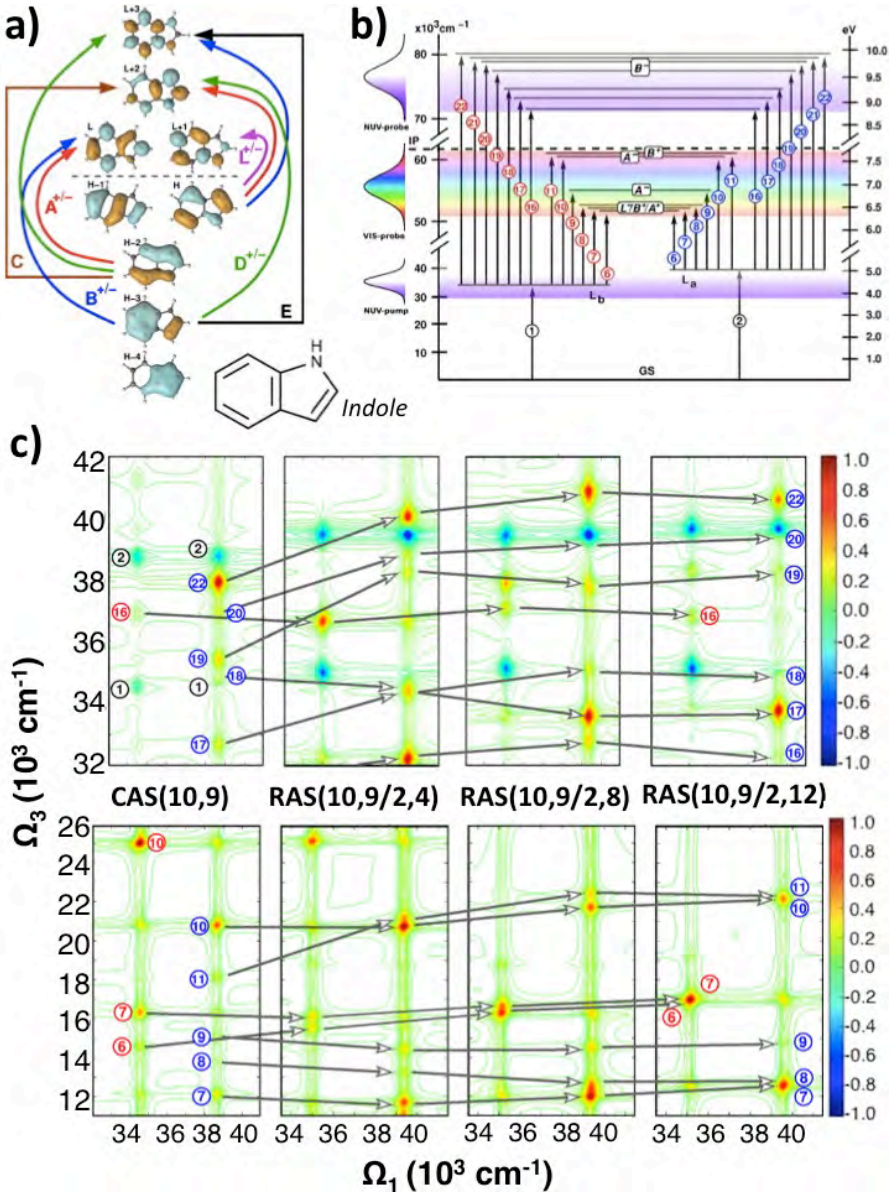

Figure 5. Benchmark computations on indole: (a) molecular orbitals in the CAS(10,9) active space displaying a labelling scheme for (b) the various type of transitions associated to ESAs in the Vis and UV probing (Ω_3 frequency) windows, as arising from the primary ππ* (L_b and L_a) states (in the pump frequency range Ω_1=34000-40000 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 associated to bleaching while positive (red) peaks arise from $S_{1,2} \rightarrow S_M$ ESAs. Reproduced from data reported in Ref [57].

Figure 5 summarizes the benchmark study of indole[52] 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 8 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 Figure 5c). Both one-color (UV-pump and UV-probe, 2DUV-UV) and two-colour (UV-pump and Vis-probe, 2DUV-Vis) 2D spectra are reported in Figure 5c, showing a broad set of ESA and bleaching signals that vary in positions and intensities as function of the level of theory. Overall in the 2DUV spectra of indole, two signal traces associated to the first two (ππ*) excited states, labelled 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 Figure 5) feature off-diagonal bleaching signals (ODB, see Fig. 1) associated 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 single-exciton energies sum as due to 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 position 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 of the 2D maps according to the theoretical level employed. This work is particularly tedious as it requires labelling all relevant transitions (Fig. 5b) according to the molecular orbitals (MO) involved in the electronic excitations.
Figure 6 shows the benchmarks computations for the Ura, Thy and Cyt pyrimidine nucleobases and the effect of the active space enlargement 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 sidechains (Fig. 4-5). While some ESA signals show more AS-dependent fluctuations than others, the overall the 2DUV-UV spectra converge with the addition of 8 extravalence orbitals. Cyt is particularly interesting as the first excited state is significantly bright but is not the brightest low-lying ππ* 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 associated to the two low-lying ππ* states (as in indole and purine bases) and the performances of the various levels of theory in producing reliable 2DUV-UV maps.

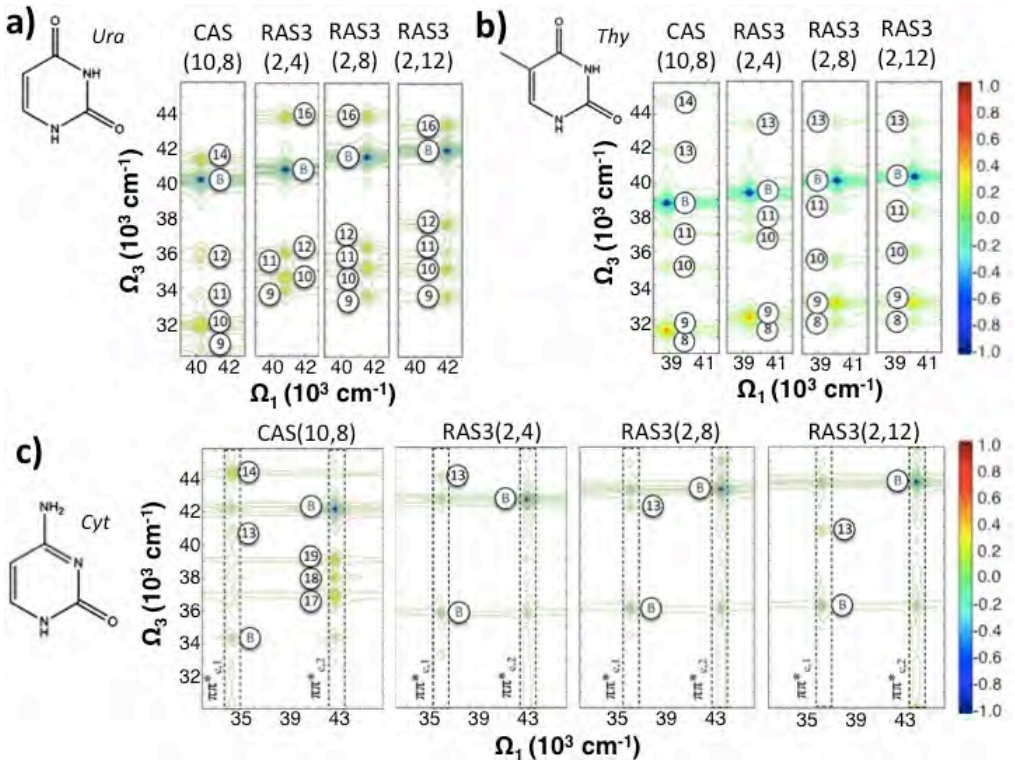

Figure 6. Benchmark computations on pyrimidine nucleobases (a) uracil, (b) thymine and (c) cytosine, showing the 2DES maps in the UV-pump/UV-probe window (2DUV-UV) as computed at the CAS(10,8), RAS(10,8|2,4), RAS(10,8|2,8) and RAS(10,8|2,12) levels of theory and using the ANO-L(432,21)-aug basis set.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

Adenine and Guanine are the canonical purine nucleobases, featuring larger aromatic heterocyclic structure than pyrimidine bases, with the pyrimidine ring fused to an imidazole ring. Adenine has just one π-orbital in the valence AS more than indole (MOs depicted in Fig. 5a), i.e. CAS(12,10), and similar MOs that give rise to analogous electronic transitions (Figure 5 and 7). As indole, adenine features two close-lying ππ* excited states but in the purine base the GS→L_a transition has higher TDM than GS→L_b. As showed in Figure 7b, the benchmark of 2DUV-UV maps shows remarkable changes when the AS is increased with respect to the full valence space, with positions of the Ω_1 traces being more sensible to the level of theory than in indole, while convergence is still reached when 8 extra-valence orbitals are added in a RAS(12,10|2,8) scheme.

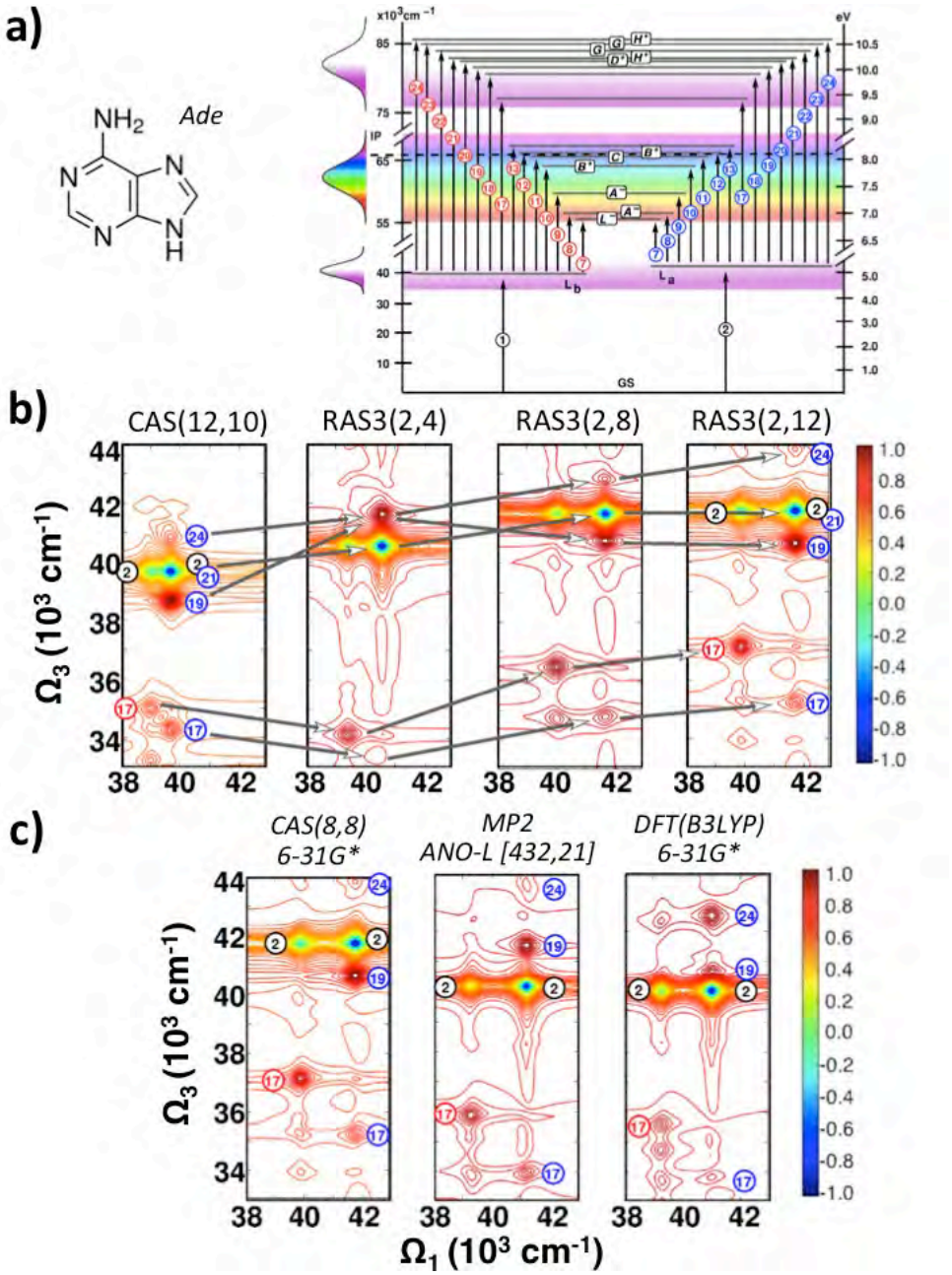

Figure 7. Benchmark computations of the adenine excited states manifold (a), showing the 2DES maps in the UV-pump/UV-probe window (2DUV-UV) as computed (b) at the CAS(12,10), RAS(12,10|2,4), RAS(12,10|2,8) and RAS(12,10|2,12) levels of theory and using the ANO-L(432,21)-aug basis set or (c) at the RAS(12,10|2,12) level with 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].

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] with respect to other aromatic systems. Therefore, for this nucleobase the benchmark study was extended to the investigation on the role of the geometry optimization methodology used to define the ground state 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 Møller–Plesset perturbation theory (MP2) levels. As showed in Figure 7c, indeed, the choice of the geometry optimization method might 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.

So far, we have reviewed the benchmark studies of the aromatic proteic chromophores and the canonical nucleobases of DNA and RNA, except for Gua that is a work in progress, aiming at obtaining converged results among the highest (still computationally feasible) levels of multiconfigurational 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 simulations of larger (multi)chromophoric systems and/or several of their structural conformations. To this aim, we have used computational recipes to account for the σ-π polarization effects when dealing with reduced active spaces[80,72] (as necessary for system size increase). In the benchmark study reported above, we have observed that reduced AS schemes (i.e. lacking virtual orbitals with higher orbital momenta) can dramatically underestimate transition energies as it overestimates 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 side we observed that deleting a number of virtual extravalence π*-orbitals with higher angular momentum in the perturbation treatment the amount 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)[81,82] invoke a non-uniform decrease of the correlation contribution, much more pronounced for ionic that 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. [83]). 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 the similar shift parameter, such as homodimeric systems. We emphasize 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 gas-phase and for their corresponding amino acid sidechains (Phe and Tyr, respectively) in a model tetrapeptide in water solution.[58]

As showed in Figure 8, the reference TEs computed for benzene and phenol (see Figure 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 multichromophoric 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 reduction of the computational costs is thus to develop a recipe that at least allows using minimal (still full-valence) active spaces (*m*AS), i.e. CAS(6,6) and CAS(8,7) for benzene and phenol, and relatively small basis sets without compromising TEs and TDMs estimates. As showed in Figures 4 and 8, the full-valence *m*AS in combination with large basis set, i.e. ANO-L(432,21)-aug, provide large underestimation of several excited states energies (i.e. those with ionic characters) and this discrepancy is essentially due to σ-π polarization effects associated with the reduction of the AS 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 improving significantly the performances of the *m*AS (Figure 8). This AS refinement procedure allows obtaining 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 gain in computational cost that allows moving forward to the study of small multimeric systems.

Extending the refined-*m*AS approach from the benzene and phenol monomers to their dimer has been proved 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 (Figure 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. Simulations of the 2DUV-UV spectrum of the non-interacting dimer with the refined-*m*AS CAS(14,13)$^{\delta}$ provide a 2D map that is consistent with the sum of the 2D map 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

could be in turn used as reference spectrum for further reduction of 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. A small (or even empty) RAS2 space could be, in fact, 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 schemes have been tested to minimize the AS and the computational cost for the simulations of the non-interacting benzene and phenol dimer. Thus here we apply the RASSCF protocol to reduce the AS by limiting the number of excitations, instead to increase it as done for the reference computations.

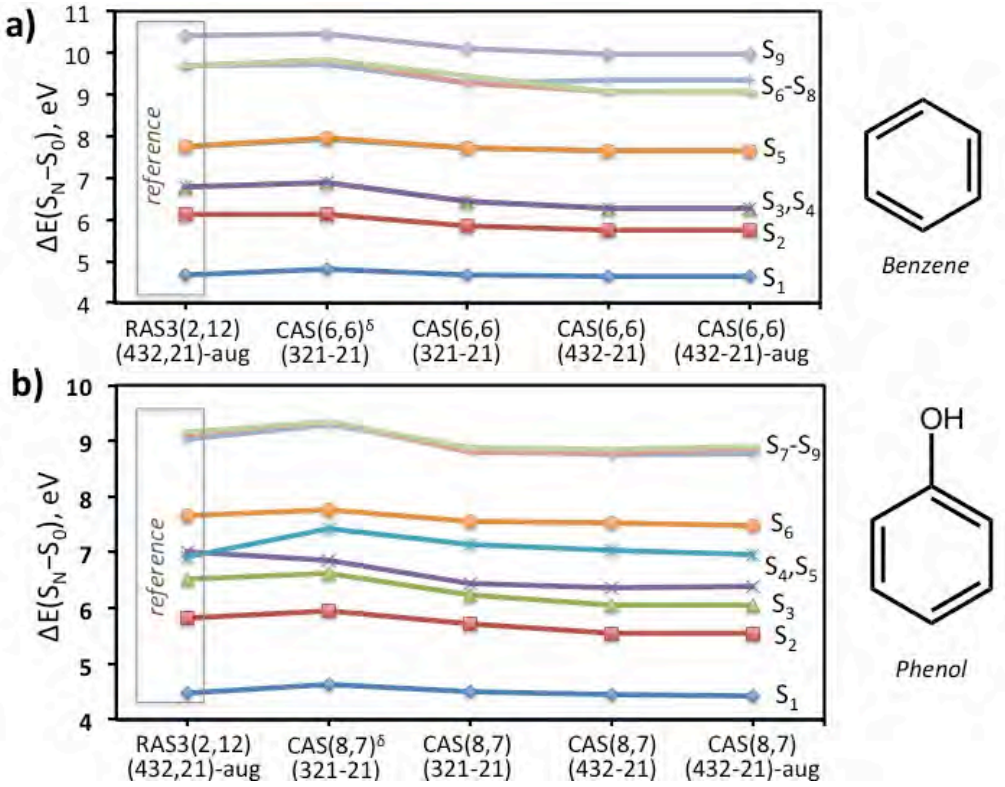

Figure 8. CASSCF//PT2 and RASSCF//PT2 vertical $S_0 \rightarrow S_N$ excitation energies (in eV) of benzene and phenol monomers in vacuum obtained at various levels of theory. Reference data are those computed at the RAS3(2,12)//PT2/ANO-L(432,21)-aug levels and are compared to cheaper computational approaches, 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 smallest ANO-L(321-21) basis set. Reproduced from data reported in Ref [58]

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 being associated to 219'048 and 52'641 CSF, respectively. As showed in Figure 9, the former RAS scheme provides reliable one-color spectrum when compared to the CAS(14,13)$^{\delta}$ one, while the latter could not properly describe some of the states located in the UV probing region (e.g. phenol signals associated to the $S_0 \rightarrow S_{7\text{-}9}$ transitions, namely 7P, 8P and 9P in Figure 9a) as well as mixed doubly excited states that give rise to off-diagonal signals (1B/1P or 1P/1B in Figure 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 energies sum of the localized single excitations. The electronic coupling between two interacting monomers, denoted as "quartic" coupling (Δ),[2,84] shifts the energies of mixed states with respect to the sum of the corresponding single excitations yielding to 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 much reduction of the AS deteriorates the description of the excited state manifold and the corresponding 2D maps.

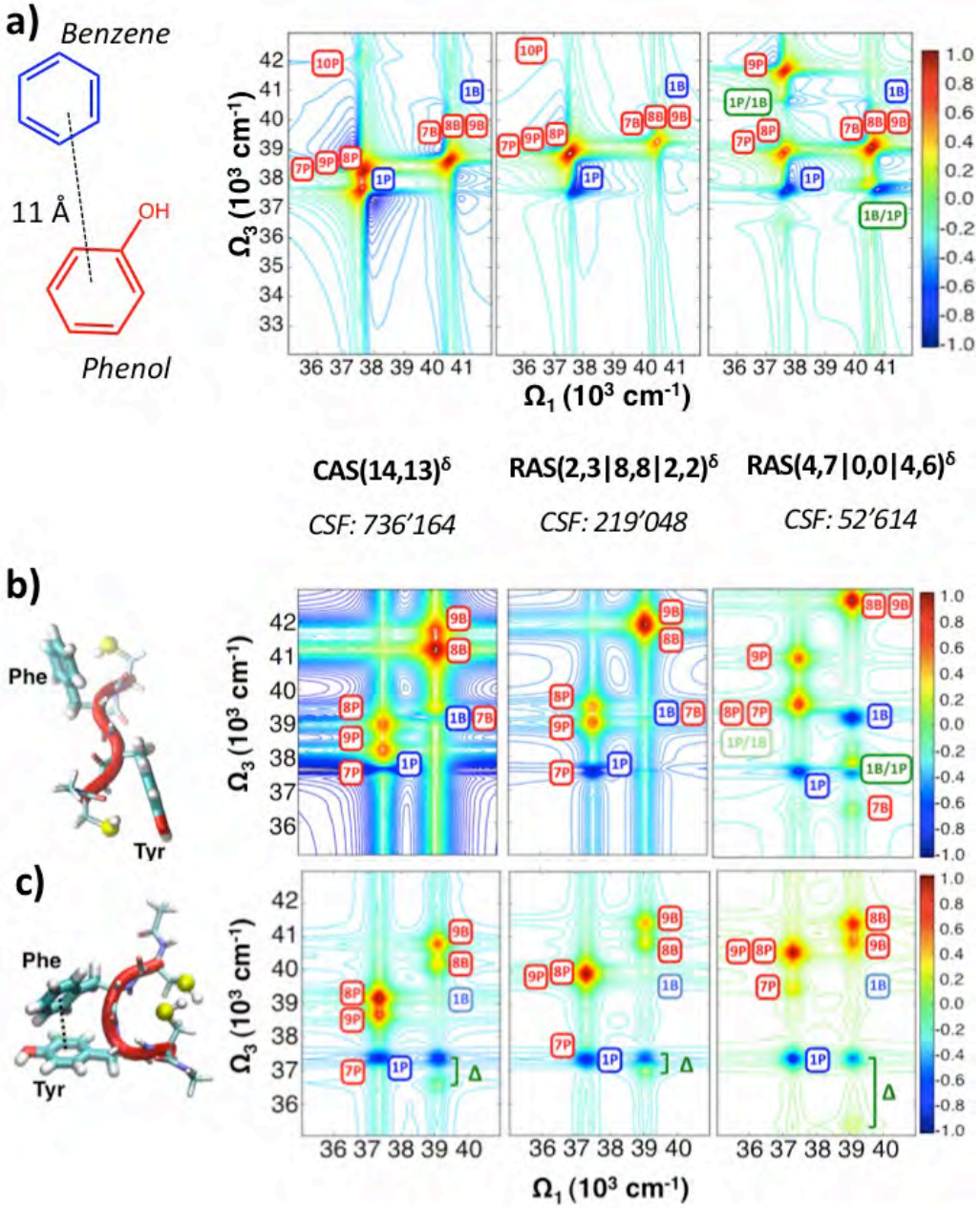

Figure 9. Comparison of simulated one-color 2DUV-UV spectra of benzene-phenol dimers with different active spaces, including CAS(14,13)$^{\delta}$ and RAS(2,3|8,8|2,2)$^{\delta}$ and RAS(4,7|0,0|4,6)$^{\delta}$ schemes and using the ANO-L(321,21) basis set. The numbers of CSFs are also indicated for each level of theory. Spectra are showed for (a) non-interacting dimer in gas-phase, (b) non-interacting and (c) stacked dimers in the water-solvated CFYC tetrapeptide model. The signals labels (*n*B, *n*P) refer to benzene (B) or phenol (P) S*n* 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].

In order to validate the benchmark study on the gas-phase non-interacting and to setup an efficient protocol for 2DUV spectra simulations of realistic systems, the cysteine-phenylalanine-tyrosine-cysteine (CFYC) tetrapeptide (Figure 9b-c) solvated in water solution, was considered as model proteic system.[53] As it will be showed in the next section, the folding/unfolding dynamics of this tetrapeptide involves configurations with interacting and non-interacting benzene and phenol chromophores, comprised as aromatic side chains of the phenylalanine (Phe, F) and tyrosine (Tyr, Y) amino acid residues respectively. In Figure 9, two representative structures of these configurations and their corresponding 2DUV-UV spectra, simulated with our SOS//QM/MM approach,[51] are reported. Notably, the comparison of various levels of theory showed for the non-interacting dimer in gas-phase (Figure 9a) is completely preserved for the non-interacting aromatic sidechains in the realistic solvated

tetrapeptide model, with the computed RAS(2,3|8,8|2,2)$^{\delta}$ spectrum showing same accuracy as the CAS(14,13)$^{\delta}$ one, while a cheaper RAS scheme yields artificial quartic coupling. When considering a configuration with interacting aromatic sidechain in a stacked conformation (Figure 9c), an analogous trend is found. In this case, a sizeable electronic quartic coupling is expected and it is equally found 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.
A successful benchmark study aimed at identifying efficient multiconfigurational/multireference approaches in terms computational cost/accuracy ratio has been showed in details here for benzene and phenol chromophores involved in proteic systems. Analogous studies have been performed for indole [83] and adenine[57] chromophores, while the remaining nucleobases are currently under investigation. Overall, we believe that such theoretical benchmark studies are going to provide a full set of parameters and computational recipes that will allow building excitonic model Hamiltonians for applications in large systems (such full proteins or large DNA/RNA sequences) and simulating 2DUV spectra of dimeric and small oligomeric systems in realistic conditions with unprecedented accuracy.
In the next sections, we will illustrate how accurate simulations of 2DUV spectroscopy could be used as powerful tool to investigate physico-chemical properties of biological systems.

3.2 2DES for tracking GS conformational dynamics

In section 3.1, we have illustrated for a model proteic system containing two aromatic sidechains how 2DUV spectroscopy holds the potential to resolve electronic couplings associated to the interaction between UV chromophores. 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, allowing discriminating 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 could 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 (Figure 10). QM/MM geometry optimization of the closed CFYC, solvated in water solution, yields a structure in which the two aromatic sidechains are rather near each other, in a so-called T-stacked conformation.[51] On the other hand, the open CFYC conformation is associated to unstacked (i.e. non-interacting) chromophoric units (Figure 10a).
As showed in Figure 10b, the excited states manifolds of both closed T-stacked and open unstacked QM/MM optimized structures have been fully characterized at the CAS(14,13)$^{\delta}$ level [51,36] 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 viceversa) with concomitant increase of the TDMs associated to the corresponding $S_{1\text{-}2}\rightarrow$CT transitions. The former would make off-diagonal (negative) signals to appear in 2DUV-UV the maps (see Figure 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 signals arising from the $S_{1\text{-}2}\rightarrow S_M$ 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 sidechains while they are accompanied by a large number of ESA signals in the T-stacked conformation, due to bright $S_{1\text{-}2}\rightarrow$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 showing here their relevance.

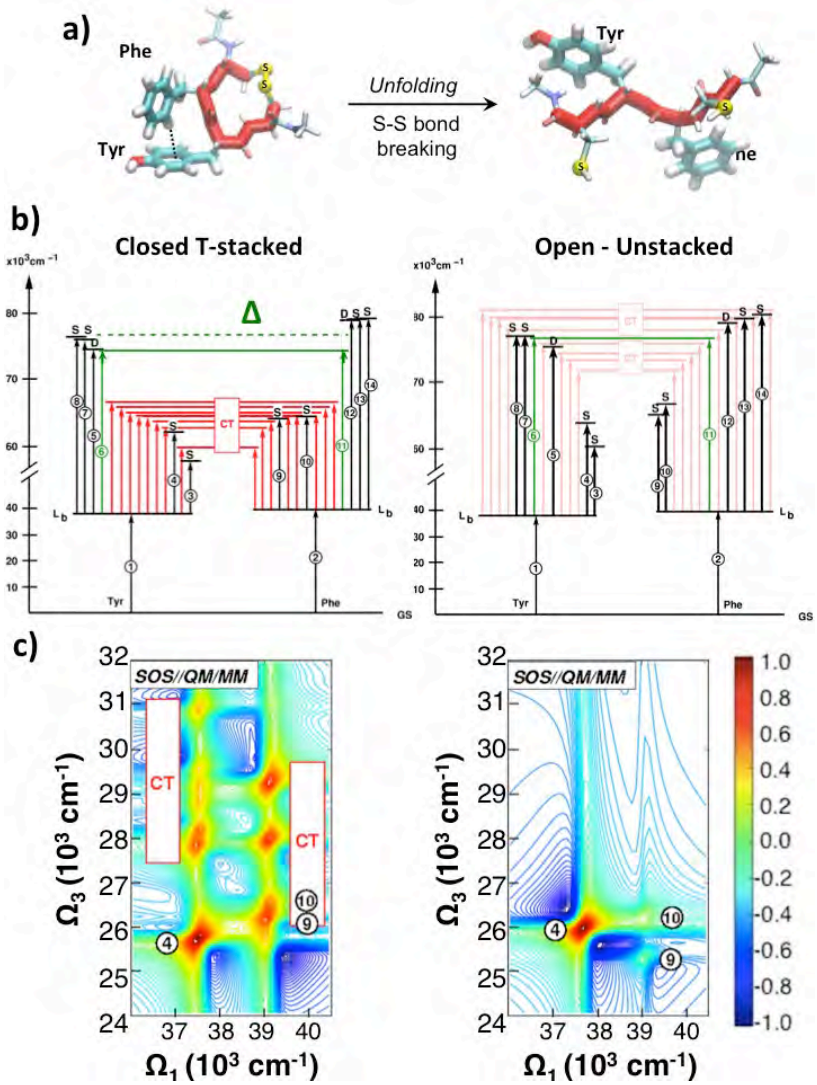

Figure 10. (a) Optimized QM/MM ground state geometries of the solvated CFYC tetrapeptide in the closed T-stacked and (upon disulfide bond breaking) open unstacked conformations. (b) Excited state manifolds of the two CFYC conformations computed at the CAS(14,13)$^{\delta}$//PT2/ANO-L(432-2) level, with 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-colours 2DUV-Vis spectra of both CFYC conformations, with ESAs labels as defined in the scheme of panel b. Reproduced from data reported in Ref [36].

In another example we benchmarked low-cost methods for indole.[83] The first two ($\pi\pi^*$) excited states (L_b and L_a) of indole (see Figure 5) have essentially opposite character: L_b is covalent, 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 that possess a single (covalent) state absorbing in the near-UV. Understandably, more 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 a reasonable agreement with the reference data obtained with RAS(10,9|2,12) level of theory (Figure 5c and 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 proteic model for studying protein folding/unfolding that contains Trp and Tyr sidechains. As for the Phe- and Tyr-containing CFYC tetrapeptide (Figure 10), two-colour 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 18000-24000 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 having the two

chromophores at about 5 Å distance. This outcome indicates that an experimentally detected enhancement of the ESA signal in Tyr- and Trp-containing peptides, in the (otherwise) signal-free visible spectral window between 18000-24000 cm^{-1}, should be regarded as a clear signature of chromophore-chromophore proximity of 5 Å or below.

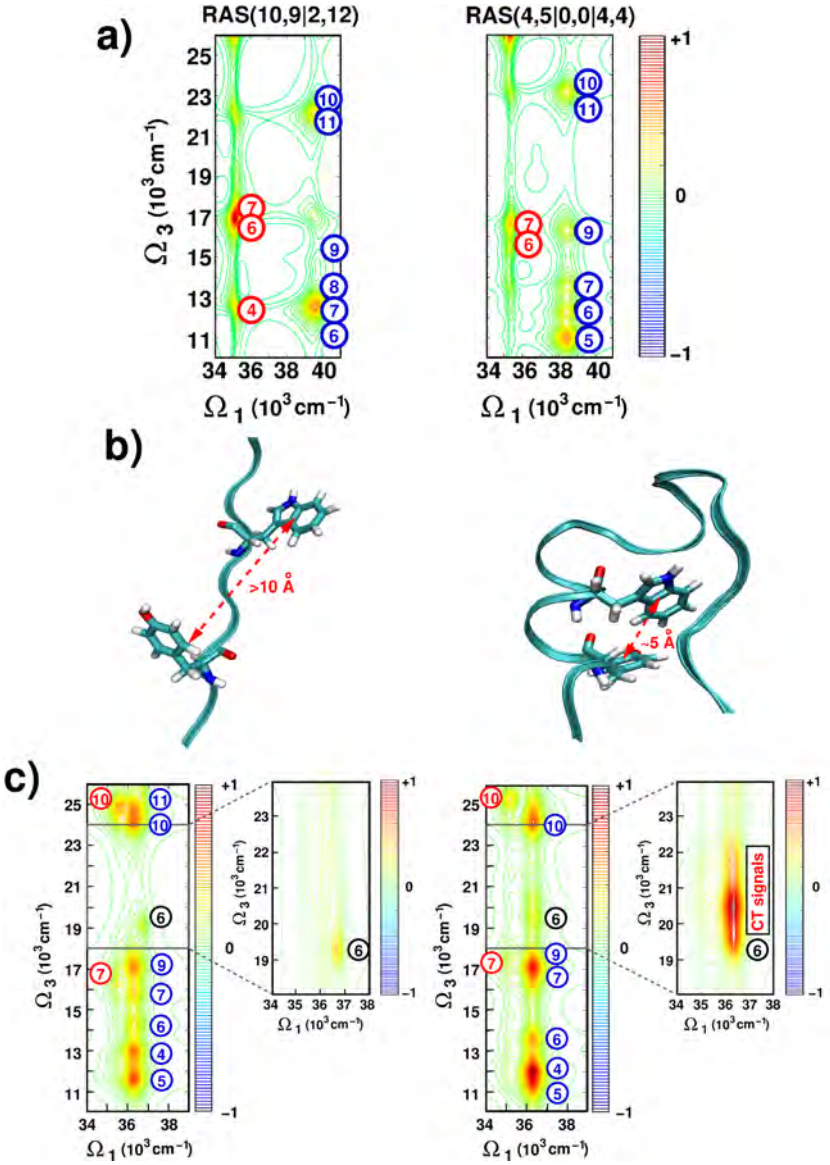

Figure 11. a) Comparison of simulated 2DUV-Vis spectra of indole in gas-phase, including the reference RAS(10,9|2,12) (left panel) and the low-cost RAS(4,9|0,0|4,7) level (right panel); Structures (b) and 2DUV-Vis spectra (c) of unfolded (left panel) and folded (right panel) Trp-cage from representative snapshots. ESAs associated with the indole sidechain of Trp (L_b in red, L_a in blue) or with the phenol sidechain of Tyr (black) are labelled. Probing spectral region around Ω_3=18000-24000 cm^{-1} is zoomed in to highlight the weak signatures of charge transfer states emerging in the stacked conformation. Labelling in (a) and (c) according to Figure 5b.

From the above examples, thus, we could conclude that 2DUV spectroscopy can be used to distinguish between folded and unfolded structures of proteic models with interacting and non-interacting UV-chromophores, respectively. The question arise whether, given the high temporal (femtoseconds) resolution of 2DUV experiments, this technique could be also used to monitor folding/unfolding dynamics of proteic model and, more in general, if 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.[53] 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),[85] 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 α = 40-60°). Notably, given the restricted flexibility of the peptide backbone, the two adjacent in sequence aromatic sidechains are never observed in a parallel (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 indicated a good correlation between the inter-chromophore distance and the quartic electronic coupling, with smaller d accompanied by larger shifts Δ (see Figure 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,[53] where the energetic positions 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 proteic aromatic chromophores described in the previous section, which provide converged results that are consistent with available experimental data, we could conclude that our predictions of two-dimensional electronic spectra strongly support the use of the 2DUV technique as an analytical tool for monitoring GS conformational dynamics of proteic systems containing aromatic UV-chromophores.

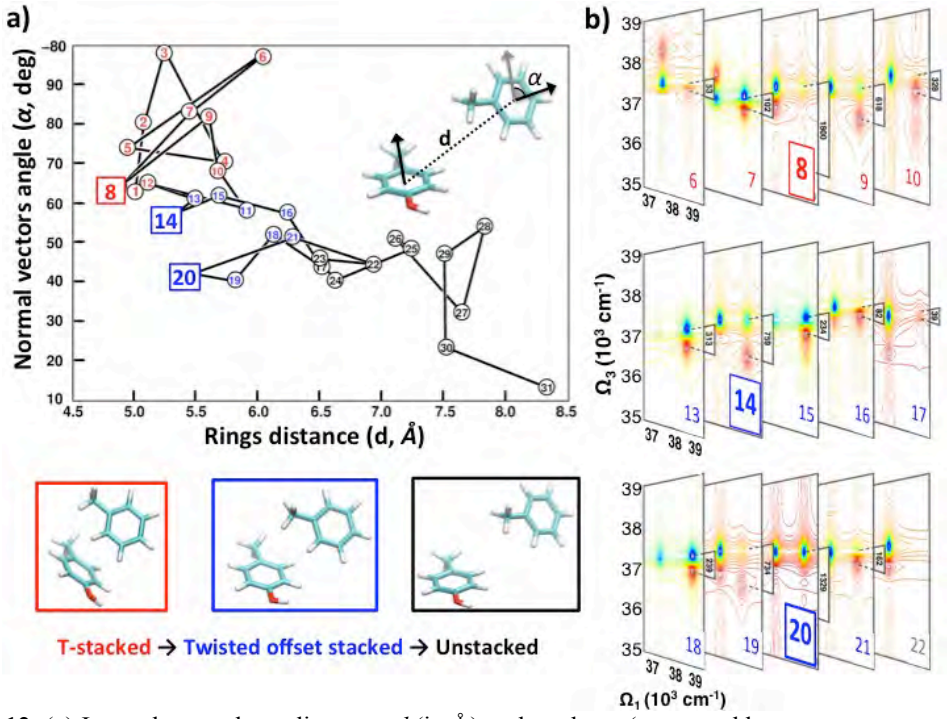

Figure 12. (a) Inter-chromophore distances d (in Å) and angles α (measured between vectors normal to the aromatic planes, in degree) for selected snapshots along the biased rMD simulation of the CYFC tetrapeptide unfolding dynamics. Colour code according to the three-step unfolding/unstacking process (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 being highlighted. Reproduced from data reported in Ref [53].

Adopting 2DUV to track GS conformational changes in nucleic acids is expected to be as suited as in proteins. Dinucleoside monophosphates, i.e. dimers of DNA/RNA nucleobases, represent excellent model systems to explore this hypothesis from our theoretical standpoint. In fact, they are realistic biomolecules that are treatable by our QM approaches[60,57] and assume various stacking conformations in water solution,[86] contemporarily representing model systems for studying the complex excited state dynamics of nucleic acids.[87] Very recently,[61] we have performed 2DUV spectra simulations of the adenine-adenine monophosphate dinucleoside (ApA) in water solution, as an example case study. 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 ground state, featuring various intermolecular arrangements that are known to deeply affect the photophysical properties of the adenine moieties.[88] Computations have been performed using the SOS//QM/MM approach and

adopting the RAS(4,12|0,0|4,6) active space, i.e. the cheapest (still reliable) multiconfigurational 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 states 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 as well as 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 an energy splitting in the main signal trace along Ω_1, i.e. that one associated with the fundamental $S_0 \rightarrow S_{3\text{-}4}$ 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 dominating ESA signals arising from excitations to adenine doubly excited state (named HL^2), lying at Ω_3 around 28000-30000 cm^{-1}.

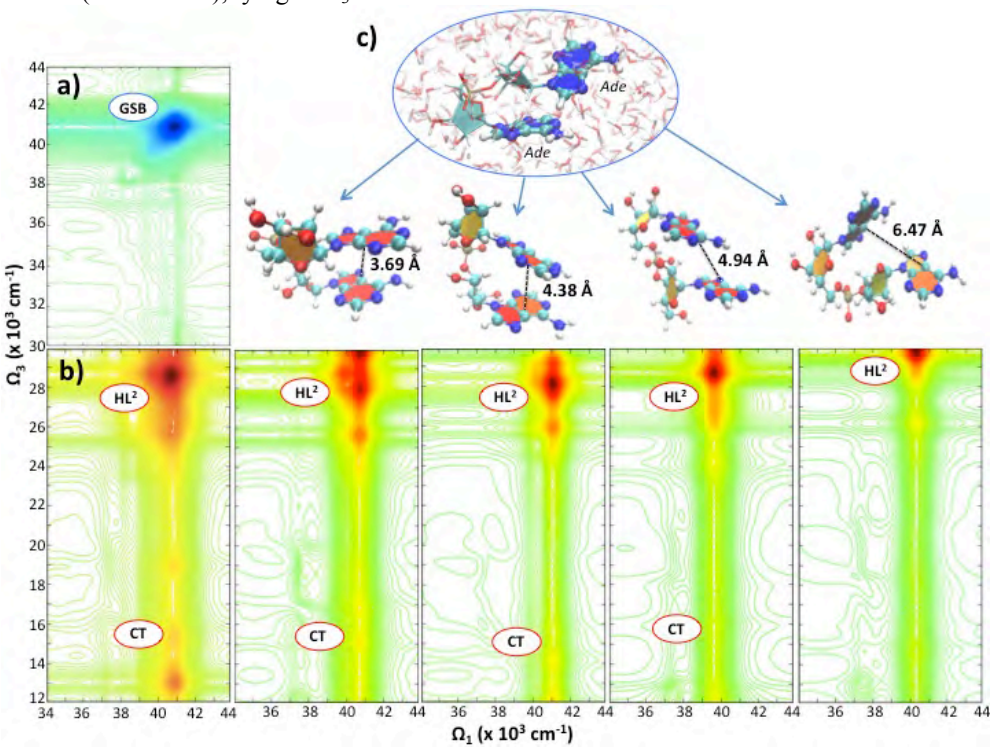

Figure 13. Simulated 2DUV-UV (a) and 2DUV-Vis (b) spectra of the water-solvated ApA dinucleoside monophosphate considering the conformational space sampled with unbiased MD in the ground state. Panel c) shows the molecular structure of the ApA system and a few selected representative conformations displaying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

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 the most representative ones, few of which are depicted in Figure 13c. Generally, short inter-chromophores distance in the 3.5-4.5 Å range (measured as C5-C5 distance) are associated to rotated π-stacked structures that yield in the 2D maps sizable broadening along Ω_1 (i.e. ^1L_a trace splitting) and concurrently feature CT signals in the low-energy visible probing window (at Ω_3 around 16000 cm^{-1}). At inter-chromophores distances around 5 Å, the ApA structures can be found in T-shaped conformations, with usually the amino group of one adenine pointing towards the 5-ring of the opposite adenine moiety. These T-stacked conformations can be generally 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 broadenings.
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, showing up as a powerful tool alternative or complementary to standard pump-probe experiments for elucidating the role of structural arrangements on nucleic acids photophysics and photochemistry. However, the spectral line shapes reported in the examples above are not realistic as they account just for partial contributions to the spectral lines broadening. Various effects (coupling to nuclear degrees of freedom, coupling to environment, finite excited state lifetimes) are the source of dephasing-induced broadenings. In real 2D electronic spectra broadenings could obscure the spectral fingerprints predicted by *ab initio* simulations. Therefore, inclusion of dynamical effects is indispensible to move towards accurate 2DES simulations and it is in the focus of the next two sections.

3.3 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 few hundred fs (for wagging, twisting and other slow vibrational modes). Nonlinear spectroscopies with ultra-short sub-10fs pulses (an order or magnitude shorter than most of the vibrational periods) are capable of detecting this vibrational dynamics, which manifests in coherent oscillations of the spectral signatures. These coherences provide insight into the photoactive vibrational modes of the system and help understand 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 $\tau_{ij}=\tau_i-\tau_j$. Eq. 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 intra-molecular vibrational modes. The line shape function for this model is formulated on the basis of the multi-dimensional uncoupled displaced harmonic oscillator (DHO)[2];

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

where the one-dimensional potential for the i^{th} state along each normal mode k is fully characterized by two parameters, frequency ω_k and relative displacement $\tilde{d}_{ik}$. In the simulations a composite line shape function is used, constructed by combining the DHO line shape function with the lineshape 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 time scale of few tens of fs 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 might induce additional signal broadening which in the present framework are treated phenomenologically.

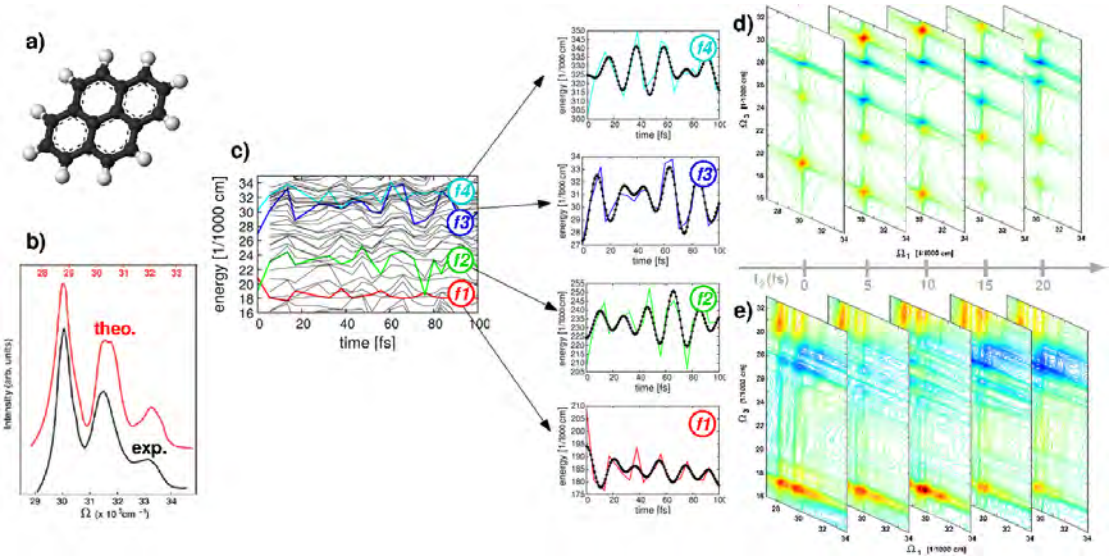

Figure 14. (a) structure of pyrene; (b) comparison of recorded (black) and simulated (red) linear absorption spectrum; (c) state density up to 34000 cm^{-1}, highlighting the energy profiles of the bright states f_1-f_4, fingerprints of the L_a state, with superimposed fits according to eq. 18 (black 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 Vis-to-UV region within the "quasi"-static picture (d) and with bath fluctuations included (e). Colour code: GSB (blue), ESA (red). Reproduced from data reported in Ref [89].

Pyrene (see Figure 14a) is a polycyclic aromatic hydrocarbon that has attracted attention for its prominent photophysical properties such as its remarkably long fluorescence lifetime and high fluorescent quantum yields[90]. It is used to probe solvent polarity[91] and constitutes the basis of commercial dyes, fluorescence probes for electro-chemiluminescence and many larger photoactive molecules.[92-94] Pyrene features well-separated absorption bands (see Fig. 14b) with clear Franck-Condon progressions in the near UV (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 near [16] and deep UV[18] therefore representing a perfect target to benchmark our theoretical protocols against 2D experimental data.[89]

Here, we review 2DES simulations which resolve the coherent spectral dynamics initiated immediately after photoexciting the first bright excited state (L_a) of pyrene, exhibiting a vibrational progression with fundamental (most intense) transition and two overtones, between 28500 cm^{-1} (~350 nm) and 33500 cm^{-1} (~300 nm) (see Fig. 14b). In particular, we aim at reproducing the spectra for waiting times 0-20 fs in a broadband probe window ranging from 16000 cm^{-1} (~600 nm) to 32000 cm^{-1} (~300 nm). These simulations require knowledge of: i) the electronic structure of pyrene up to 65000 cm^{-1} (ca. 8 eV), thus covering not only the L_a state (labelled *e*) but also the manifold of higher lying states (labelled *f*) that fall within the envelope of the probe pulse (i.e. 16000-32000 cm^{-1} above the L_a state); ii) the Franck-Condon active modes of the L_a state, source of the vibrational dynamics appearing as coherences in the 2D spectra; 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 active spaces comprise all valence π-orbitals of the molecule, allowing up to quadruple excitations only in the RAS1/RAS3 spaces. Coherent vibrational dynamics was described with the DHO model. The frequency (ω_k) and displacement ($\tilde{d}_{ik}$) parameters used to construct the lineshape 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 Schroedinger equation, while applying classical Newtonian dynamics to the nuclei. The dynamics was initiated in the L_a state of pyrene at the Franck-Condon point without initial kinetic energy. The choice of a reference Hamiltonian acting on the

classical bath during the coherence propagation, as it is the case for the delay times t_1 and t_3, is not unique [95] 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 if 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)

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

with $h \in \{g, f\}$ which allowed extracting 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 ground state are zero per definition). In the adopted DHO framework the spectral dynamics of the ESA is 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 states PES with respect to the PES of the photoactive state and can induce positive or negative frequency correlations.[96] 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 wavefunction 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 wavefunction 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 compare 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 dynamical 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 30000 cm^{-1} (shown in blue); b) the oscillatory dynamics of the SE between 30000 cm^{-1} and 25000 cm^{-1} with a period of ca. 20 fs (shown in blue); c) the aforementioned three ESA contributions (shown in red) around 19000 cm^{-1} (f_1), 24000 cm-1 (f_2) and 32000 cm^{-1} (f_4); d) the time-dependent fluctuation of the intensities, a consequence of the coordinate-dependence of the TDM. However, the static approach fails to describe correctly the spectral lineshapes. 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 (28500 cm^{-1}) and two overtones. As a consequence, the GSB and SE adopt a characteristic checkerboard pattern.[97,98] 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 a reduced intensity compared to the static spectrum. Overall, the ESA associated with f_1 (Ω_3~17500 cm-1) constitutes the most characteristic signature of the L_a state in the visible. It has been observed in fact in transient pump-probe spectra.[44,99] The weaker ESA around 25000 cm^{-1} associated with state f_2 has been observed in 1D-PP experiments.[100] To the best of our knowledge no transient spectra pumping the L_a state and probing beyond 32000 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.

3.4. Spectral characterization of long-lived ES intermediates

Excited states have finite lifetimes. Decay to the ground state 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 potential energy surfaces of electronic states intersect. Population transfer has a two-fold effect on the appearance of the 2D electronic spectra: i) homogeneous line broadening and ii) decay and simultaneous build up 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.[101,102] 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). Instead, when bath fluctuations decay slower compared to the coherence dephasing but faster compared to 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 reading

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

where, in the limit $t_2 \rightarrow \infty$ (i.e. time scales 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}. Evidently, 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 of the initial coherence state *g-e* during t_1 and in the final coherence state *e′-f* during t_3 are treated formally exactly but the dynamics are not coupled. More sophisticated treatments are possible when bath fluctuations and transport occur on the same time scale in order to retain memory effects during t_2 and to describe ultrafast population transfer events.[46]

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 time scale 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 lineshape 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 living 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 drawing the focus entirely on the electronic structure, thereby providing invaluable insight in the spectral fingerprints of excited states minima.

From experimental observations, it is known that in pyrene, after L_a excitation, an ultrafast population transfer toward 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 ns time scale before it eventually decays to the ground state through fluorescence with a high quantum yield.[103,104] Recently, excited state molecular dynamics simulations[89] revealed the mechanism of non-adiabatic $L_a \rightarrow L_b$ internal conversion. As showed in Figure 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 couplings between the two states, the wavepacket bifurcates, 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.[100]

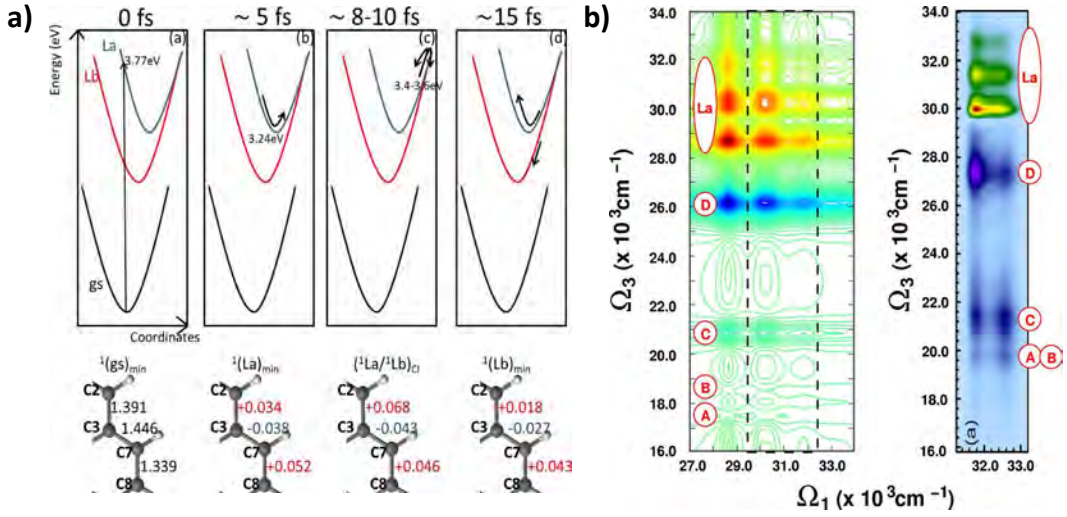

Figure 15. (a) Schematic representation of the photophysics of pyrene based on a quantum-classical dynamics. The dominant bond deformations are given, with the GS equilibrium geometry used as a reference. (b) Theoretical (left panel) and experimental (right panel) quasi-absorptive 2D electronic spectra 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 Vis-to-UV region. Colour code: GSB (red), ESA (blue). Note that the colour code is inverted with respect to other figures as to have an easy comparison with the experimentally reported spectrum. Reproduced from data reported in Ref [89].

Considering the mechanism described above 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 28000 cm^{-1} and 32000 cm^{-1}, as well as two distinct ESA peaks at 21000 cm^{-1} (labelled B) and 26000 cm^{-1} (labelled C, D). Further less intense peaks (A, B) appear at around 18000 cm^{-1}. To underline that different higher lying excited states are probed with respect 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 identical peak structure with a progressively decreasing intensity of the overtones. The physical interpretation behind is that, on a ps time scale, excess energy acquired by exciting the overtones would be dissipated and probing would become insensitive to the pump wavelength. Riedle and co-workers reported recently 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, centred at around 32000 cm^{-1}, i.e. at the second and third vibrational bands of the L_a absorption, and a super-continuum probe pulse covering the Vis-to-UV spectral window (16000-38000 cm^{-1}).[89] This result represents one of the few experimental examples employing Vis-UV broad pulses for 2D spectroscopy. The spectrum is shown in Fig. 15b and compared to our simulations (note the inverted colour code with respect to the usual colour coding adopted in other Figures to cope with appropriate visual comparison with the experimental map). A remarkable agreement between theory and experiment regarding the GSB and the positions and relative intensities of the ESA signals is achieved. 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 are noticeably elongated along Ω_3, which could be rationalized by the short lifetimes of the higher excited states. The pyrene example, where simulated 2DUV spectrum adopting the excited state equilibrium geometry has been successfully compared with the experimental spectrum recorded at a waiting time of 1 ps, demonstrates that even when the equilibration has not been carried out completely (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 states minima. This outcome paves the way for extending simulations of 2DUV spectra through our SOS//QM/MM to other biologically relevant

systems, pursuing the idea that 2DUV fingerprints of excited states minima can be computed at a reasonable computational cost and could be directly compared with experimental data recorded at appropriate t_2 delay times.

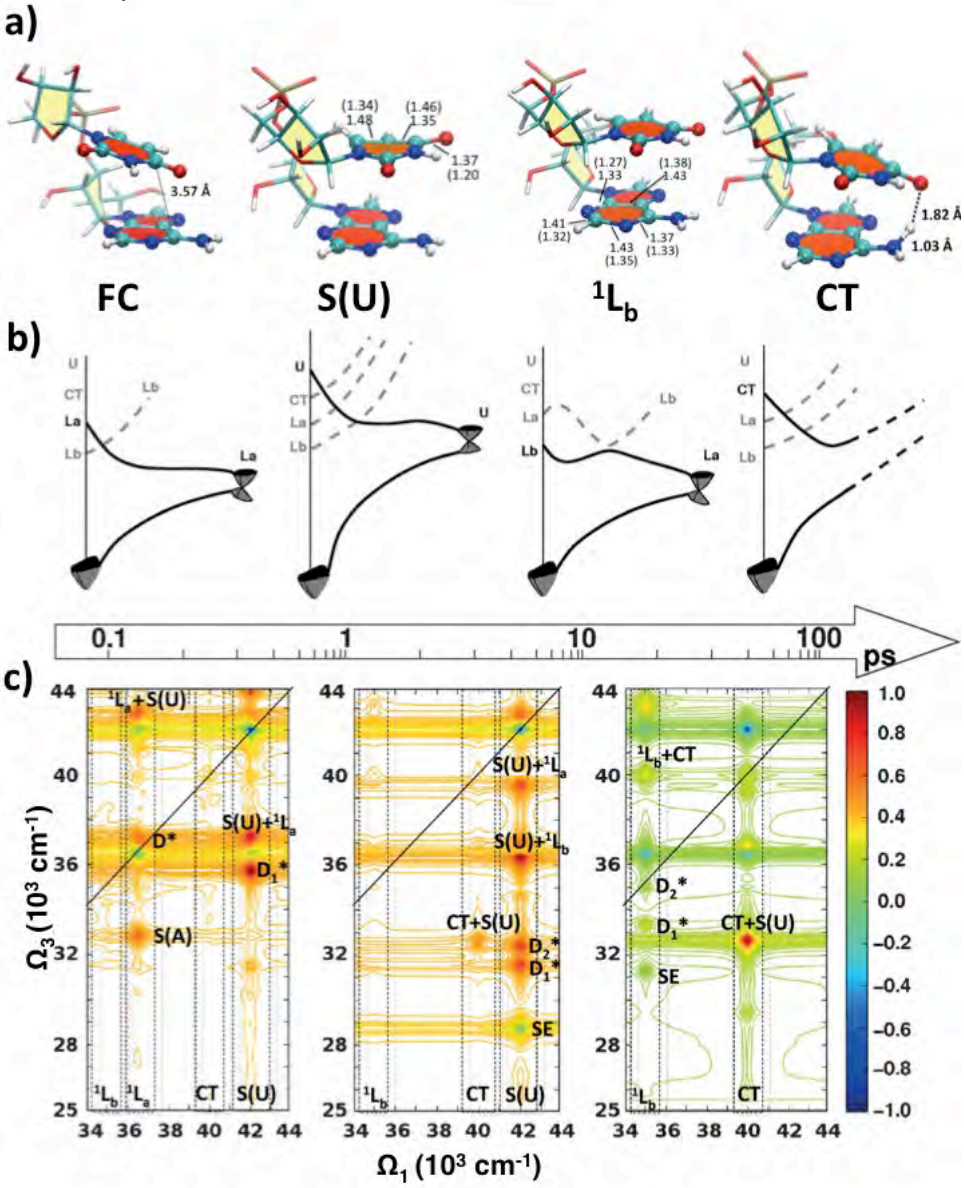

Figure 16. (a) Ground-state (FC) and excited states optimized structures of the solvated ApU dinucleoside monophosphate, including S(U), ^1L_b and CT states. Representative inter-bases or bond distances (in Å) are reported, with GS values in parentheses for excited states minima. (b) Schematic representation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

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 made to obtain the first experimental 2DUV spectra of DNA nucleobases monomers [22,21] indicate the concrete possibility of using this technique to understand the photophysics of DNA/RNA more realistic models, e.g. multimeric species in short single or double strands. As we have mentioned in section 3.2, dinucleoside monophosphates, i.e. dimers of DNA/RNA nucleobases, represent excellent model systems to start exploring nucleic acids photophysics from both theoretical and experimental standpoints. In fact, the presence of a neighbouring base largely affects the excited state dynamics of the monomeric systems due to excimer/exciplex formation and a range of charge/proton/hydrogen transfer/recombination events.[105-108] These have been shown to be

extremely difficult to separate and assess with standard 1D pump-probe techniques, and their particular contributions to the overall de-excitation process are key for understanding photo-induced lesions and mutations.[109,110] Figure 16 shows the simulated 2DES spectra of water-solvated adenine-uracil dinucleoside (ApU), a system exhibiting different excited states relaxation mechanisms that would be intricately obscured in 1D-PP experiments. Early experiments had shown how adenine and uracil separately would feature ultrafast signals,[111-113] as well as long-lived ones arising from their dark $^1n\pi^*$ states being partially populated,[114] but that an additional signal would arise in dimeric species, probably due to inter-nucleobases interactions.[88] This signal could not be clearly disentangled but it was associated to 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 to start using 2DUV spectroscopy for characterizing with high temporal and spectral resolution the spectroscopic fingerprints of various excited state minima along the complex photoinduced pathways.

A first 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 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 states geometry optimizations from the FC region, associated to a ground state optimized geometry with vicinal interacting nucleobases (Figure 16a). Transition states and conical intersections along the excited states decay channels were also characterized, providing an estimate of the time scale 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, leading to the simulated 2DES maps depicted in Fig 16c. As it can be seen, in the FC region (at zero or ultrashort t_2 waiting time, sub-100fs) 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 being rapidly depleted due to ultrafast deactivation and CT signals being weak and 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,[115] in the sub-10ps timescale it was predicted that most spectral contributions 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 34000-36000 cm^{-1}) and a particularly intense fingerprints of the CT minimum, which trace (at Ω_1 around 40000 cm^{-1}) would be well separated from the ^1L_b trace, allowing to quantify the population of ApU conformations with accessible CT states. Thus, these simulations could help assign distinctive signals arising from different excited states decay channels that would possibly be recorded in upcoming 2DUV experiments, providing a state-specific characterization of the ApU excited states dynamics. With such example, here we showed how 2DUV might be extremely powerful for elucidating nucleic acids photophysics in model systems.

Outlook and perspectives

A route towards simulations of third-order nonlinear electronic spectroscopy with an accuracy compatible with current experiments is outlined in this review. 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 protein and nucleic acids, the building blocks of life (section 1). 2DUV experiments feature high spectral and temporal resolution, allowing to follow the evolution of individual chromophores though their spectral fingerprints, providing direct signatures of electronic couplings, charge and energy transfers, and characteristic lineshapes and dynamics, which give information on PES topology and solvent reorganization timescales. The density matrix formalism (section 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 drew the 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 allows 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 lineshapes and the dynamical evolution of the nonlinear signals. This route is covered by illustrating some practical examples (section 3), starting with the benchmark studies of excited state manifolds of the target bio-molecular systems (section 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 multiconfigurational wavefunctions combined with multireference 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, geometrical parameters and demonstrate its sensitivity to these parameters. CASSCF/CASPT2 allows adaptive protocols for obtaining converged energies and TDMs, reaching an accuracy of 0.2-0.25 eV that is considered state-of-art in the field of photochemistry, yet corresponding to ~2000 cm^{-1}, a significant deviation from a spectroscopy point of view. Therefore, reaching this upper-bound limit of accuracy is mandatory for reliable theoretical spectroscopy. We have shown how such accuracy has been obtained for the TEs and TDMs of monomeric units of aromatic protein sidechains and nucleobases in the gas-phase, demonstrating how the (commonly used) full valence active spaces yield large discrepancy with respect 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 a 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, allowed 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-chromophores interactions on the nonlinear response (section 3.2).

By embedding multiconfigurational/multireference computations within hybrid QM/MM schemes allows extending the SOS simulations (SOS//QM/MM) of 2D electronic spectra to realistic model systems of proteins and nucleic acids. This allowed demonstrating how 2DUV spectroscopy holds the potential of resolving inter-chromophores interactions occurring during GS dynamics of nucleic acids and proteic systems. In particular, simulations of 2DUV spectra of the CFYC tetrapeptide and Trp-cage proteic model indicated how both electronic (quartic) coupling shifts and charge-transfer states yield clear spectroscopic fingerprints (mainly associated to ESAs) in the one-color 2DUV-UV and two-colours 2DUV-Vis spectra, respectively. This significant outcome stimulated investigations on the potential of 2DUV spectroscopy for tracking the protein folding/unfolding processes of proteic models with high temporal resolution and for the quantitative analysis of nucleobases 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 the 2DUV technique as a powerful tool alternative (or complementary) to standard transient absorption experiments. These results highlight the importance of accurate electronic structure computations in 2DES simulations, as to account for appropriate descriptions of the ESAs signals. However, the massive computational costs of *ab initio* computations restrict the applications to small (still realistic) dimeric systems. Current developments in the field of multiconfigurational, density-based and linear-response techniques[116-124] would certainly provide beneficial tools for treating large (multi)chromophoric systems, providing they accomplish completeness in the description of excited state manifolds (e.g. including doubly excited states).[125] Efficient QM methods employed in hybrid QM/MM schemes would also allow accounting for signal broadenings due to solvent rearrangements, as showed in the ApA case.

Various types of dephasing-induced broadenings (due to coupling to nuclear degrees of freedom and environment, finite excited state lifetimes) can shape the 2D maps. When neglected, simulated signal lineshapes and positions would differ significantly from their experimental counterparts. We moved a step forward towards a more realistic description of spectral lineshapes by inclusion of vibrational dynamics in the simulations (section 3.3). The high temporal resolution featured in 2DUV experiments with ultra-short sub-10fs pulses allows for detection of vibrational features, which manifests 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, the first direct comparison of theoretical predictions with experimental data, available from broadband 2DES spectra of the pyrene molecule, recorded along its excited state relaxation process,[16] became possible (section 3.4).

Comparison of theoretical and experimental 2DUV spectra, showed 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

endorsed considering extension of 2DUV spectra simulations for resolving the spectroscopic fingerprints of long-living excited state minima along the complex photoinduced decay pathways of DNA/RNA model systems. In particular, our predictions of state-specific fingerprints of excited states dynamics in solvated ApU dinucleoside monophosphate exemplifies the potential impact of accurate simulations of 2DES spectra in revealing complex physico-chemical properties of fundamental biological systems.

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

Acknowledgments

I.R. acknowledges support from the French Agence National de la Recherche (FEMTO-2DNA, ANR-15-CE29-0010). J.S.M. thanks Dr. Lara Martínez-Fernández for useful discussions. M.G. acknowledges support from the European Research Council Advanced Grant STRATUS (ERC-2011-AdG No. 291198). S.M. gratefully acknowledges the support of the National Science Foundation (grant CHE-1361516), and of 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 absence of an external field ($\hat{H}=\hat{H}_0$), the free evolution of an unperturbed density matrix, $\hat{\rho}^{(0)}(t)$, reads

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

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

In the perturbative scheme described in section 2.1(eq. 5), the third-order density matrix reads

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

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-van Neumann equation (eq. 4), leading to the Lindblad equation

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

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

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

with *K* the rate matrix with elements $K_{ee,e'e'}$ depicting the population transfer rate 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 lineshape

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 Ref. [46,2]). The main building block is the lineshape function $g_{ij}(t)$, which is the integral transformation of the autocorrelation function of bath fluctuations

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

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 lineshape function of the semi-classical Brownian oscillator (OBO).[126]

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

with λ_{ij} and Λ the system-bath coupling strength and fluctuation time scale, 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)
11. Fuller FD, Ogilvie JP (2015) Experimental implementations of two-dimensional fourier transform electronic spectroscopy. Annual Review of Physical Chemistry, Vol 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. Tseng C-h, Matsika S, Weinacht TC (2009) Two-dimensional ultrafast fourier transform spectroscopy in the deep ultraviolet. Optics 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. Optics 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. New Journal of Physics 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 Phenomena 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. The Journal of Physical Chemistry Letters 3 (18):2575-2581
22. Prokhorenko VI, Picchiotti A, Pola M, Dijkstra AG, Miller RJD (2016) New insights into the photophysics of DNA nucleobases. The Journal of Physical Chemistry Letters 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. Proceedings of the National Academy of Sciences 111 (28):10061-10066
26. Loukianov A, Niedringhaus A, Berg B, Pan J, Senlik SS, Ogilvie JP (2017) Two-dimensional electronic stark spectroscopy. Journal of Physical Chemistry Letters 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
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. Optics Express 15 (25):16681-16689
33. Brańczyk AM, Turner DB, Scholes GD (2014) Crossing disciplines - a view on two-dimensional optical spectroscopy. Annalen der Physik 526 (1-2):31-49
34. Son M, Schlau-Cohen GS Ultrabroadband 2d electronic spectroscopy as a tool for direct visualization of pathways of energy flow. In, 2017. SPIE, p doi: 10.1117/1112.2273417
35. 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
36. 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
37. 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
38. 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
39. Sanda F, Mukamel S (2008) Stochastic liouville equations for coherent multidimensional spectroscopy of excitons. J Phys Chem B 112 (45):14212-14220
40. Tanimura Y (2006) Stochastic liouville, langevin, fokker–planck, and master equation approaches to quantum dissipative systems. J Phys Soc Jpn 75 (8):082001
41. 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
42. Tempelaar R, van der Vegte CP, Knoester J, Jansen TLC (2013) Surface hopping modeling of two-dimensional spectra. The Journal of Chemical Physics 138 (16):164106
43. 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
44. 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
45. Mukamel S (1983) Non-impact unified theory of 4-wave mixing and 2-photon processes. Phys Rev A 28 (6):3480-3492
46. Abramavicius D, Valkunas L, Mukamel S (2007) Transport and correlated fluctuations in the nonlinear optical response of excitons. Epl 80 (1)
47. Jiang J, Mukamel S (2010) Two-dimensional ultraviolet (2duv) spectroscopic tools for identifying fibrillation propensity of protein residue sequences. Angewandte Chemie-International Edition 49 (50):9666-9669
48. 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
49. 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

50. 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
51. 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
52. Nenov A, Rivalta I, Mukamel S, Garavelli M (2014) Bidimensional electronic spectroscopy on indole in gas phase and in water from first principles. Computational and Theoretical Chemistry 1040:295-303
53. 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
54. 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 Discussions 177:345-362
55. 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
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. The Journal of Physical Chemistry 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. The Journal of Chemical Physics 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. Chemistry – A European Journal 22 (22):7497-7507
61. Segarra-Marti J, Jaiswal VK, Pepino AJ, Giussani A, Nenov A, Mukamel S, Garavelli M, Rivalta I (2017) Two-dimensional electronic spectroscopy as a tool for tracking molecular conformations in DNA/rna aggregates. Faraday Discussions:in press, DOI:10.1039/C1037FD00201G
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. In: Advances in chemical physics. John Wiley & Sons, Inc., pp 399-445. doi:10.1002/9780470142943.ch7
65. Andersson K, Malmqvist PA, Roos BO, Sadlej AJ, Wolinski K (1990) 2nd-order perturbation-theory with a casscf 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 Interdisciplinary Reviews-Computational Molecular Science 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)
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. The Journal of Physical Chemistry 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. The Journal of Chemical Physics 17 (5):484-495
80. 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. The Journal of chemical physics 98 (4):3151-3162
81. 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. Journal of Molecular Structure: THEOCHEM 388 (0):257-276
82. Roos BO, Andersson K (1995) Multiconfigurational perturbation-theory with level shift - the cr-2 potential revisited. Chem Phys Lett 245 (2-3):215-223
83. 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
84. Hamm P, Zanni M (2011) Concepts and methods of 2d infrared spectroscopy. Cambridge University Press, Cambridge, U.K.
85. 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)
86. 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
87. Crespo-Hernandez CE, Cohen B, Hare PM, Kohler B (2004) Ultrafast excited-state dynamics in nucleic acids. Chem Rev 104 (4):1977-2019
88. 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. Proceedings of the National Academy of Sciences 105:10285-10290
89. 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
90. Ruzicka P, Kral T (2013) Pyrene: Chemical properties, biochemistry applications and toxic effects. Chemistry research and applications. Nova Science Publishers, New York
91. Reichardt C (1994) Solvatochromic dyes as solvent polarity indicators. Chem Rev 94 (8):2319-2358
92. 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. Biosensors & Bioelectronics 87:89-95
93. Figueira-Duarte TM, Mullen K (2011) Pyrene-based materials for organic electronics. Chem Rev 111 (11):7260-7314
94. 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. Scientific Reports 6

95. 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
96. 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
97. 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
98. Butkus V, Valkunas L, Abramavicius D (2012) Molecular vibrations-induced quantum beats in two-dimensional electronic spectroscopy. J Chem Phys 137 (4)
99. 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
100. 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
101. 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
102. Meier T, Chernyak V, Mukamel S (1997) Femtosecond photon echoes in molecular aggregates. J Chem Phys 107 (21):8759-8780
103. 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
104. 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. The Journal of Physical Chemistry 99 (19):7439-7445
105. Chen J, Zhang Y, Kohler B (2015) Excited states in DNA strands investigated by ultrafast laser spectroscopy. In: Barbatti M, Borin AC, Ullrich S (eds) Photoinduced phenomena in nucleic acids ii, vol 356. Topics in current chemistry. Springer International Publishing, pp 39-87. doi:10.1007/128_2014_570
106. 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
107. Schreier WJ, Gilch P, Zinth W (2015) Early events of DNA photodamage. Annu Rev Phys Chem 66 (1):497-519
108. 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
109. Cadet J, Grand A, Douki T (2015) Solar uv radiation-induced DNA bipyrimidine photoproducts: Formation and mechanistic insights. In: Barbatti M, Borin AC, Ullrich S (eds) Photoinduced phenomena in nucleic acids ii, vol 356. Topics in current chemistry. Springer International Publishing, pp 249-275. doi:10.1007/128_2014_553
110. 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
111. 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
112. 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
113. 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. The Journal of Physical Chemistry B 106 (43):11367-11374
114. 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. Proceedings of the National Academy of Sciences 104:435-440
115. Pepino AJ, Segarra-Martí J, Nenov A, Improta R, Garavelli M (2017) Resolving ultrafast photoinduced deactivations in water-solvated pyrimidine nucleosides. The Journal of Physical Chemistry Letters 8 (8):1777-1783
116. 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

117. 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
118. 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
119. Casida ME, Huix-Rotllant M (2012) Progress in time-dependent density-functional theory. Annu Rev Phys Chem 63:287-323
120. Dreuw A, Wormit M (2015) The algebraic diagrammatic construction scheme for the polarization propagator for the calculation of excited states. Wiley Interdisciplinary Reviews: Computational Molecular Science 5 (1):82-95
121. Sneskov K, Christiansen O (2012) Excited state coupled cluster methods. Wiley Interdisciplinary Reviews: Computational Molecular Science 2 (4):566-584
122. Nemeth A, Milota F, Mančal 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. The Journal of Chemical Physics 133 (9):094505
123. Chan GK-L, Sharma S (2011) The density matrix renormalization group in quantum chemistry. Annu Rev Phys Chem 62 (1):465-481
124. Thomas RE, Sun Q, Alavi A, Booth GH (2015) Stochastic multiconfigurational self-consistent field theory. J Chem Theory Comput 11 (11):5316-5325
125. 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:doi:10.1021/acs.jctc.1027b01208
126. 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