Check for updates

Modelling the role of internal flexibility and inter-state couplings modulated by an explicit environment in the absorption spectrum of coelenteramide†

Mohammad Aarabi, [iD] ‡[ab] Samira Gholami, [iD] ‡§*[a] Samuele Giannini, [iD] ¶[b] Marco Garavelli, [iD] [a] Giacomo Prampolini [iD] [b] and Fabrizio Santoro [iD] *[b]

Recently developed mixed quantum classical (MQC) approaches, which integrate classical molecular dynamics (MD) driven by accurate quantum-mechanically derived force fields with vibronic models, were employed to investigate the absorption spectra in benzene solution of coelenteramide, the chromophore of different photoproteins. Our results confirm that light absorption in this medium is only due to the protonated neutral species **2H**. Neglecting the large-amplitude motions of the most flexible modes and adopting local harmonic expansions of the potential energy surfaces around the global minimum, the simulation of the spectrum is straightforward and it even leads to decent agreement with the experiment, indicating only moderate nonadiabatic effects. The complexity of the molecule becomes however evident when trying to accurately describe the effect of molecular flexibility, a key property expected to be very sensitive to the protein environment. MD sampling in fact shows that at room temperature the **2H** molecule frequently assumes conformations where its excited electronic states are almost degenerate and hence strongly mix, with drastic redistribution of the oscillator strengths. We document that such mixings are triggered by both stiff vibrations and soft flexible modes and that, in such a complex scenario, methods that do not properly describe nonadiabatic interactions at the quantum vibronic level introduce artifacts in the simulation of the spectrum. We also show that such problems can be cured, at least approximately, by resorting to the parameterization of linear vibronic coupling models specific for the different configurations sampled from the classical MD. A critical discussion of all achieved results is eventually addressed, hence setting up a robust grounding for further development of MQC protocols for electronic spectroscopy of flexible dyes in the condensed phase.

1 Introduction

Bioluminescent organisms have long captivated scientific interest due to their ability to emit light through biochemical

[a] *Dipartimento di Chimica Industriale "Toso Montanari", Universitá di Bologna – Alma Mater Studiorum, Via Piero Gobetti 85, 40129 Bologna, Italy. E-mail: samira.semsury@kit.edu*
[b] *Consiglio Nazionale delle Ricerche, Istituto di Chimica dei Composti Organo Metallici (ICCOM-CNR), I-56124 Pisa, Italy. E-mail: fabrizio.santoro@pi.iccom.cnr.it*
† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d5cp01696g
‡ These authors contributed equally to this work and should be considered the co-first authors.
§ Present address: Institute of Nanotechnology, Karlsruhe Institute of Technology (KIT), Kaiserstraße 12, 76131 Karlsruhe, Germany.
¶ Present address: Department of Chemistry and Industrial Chemistry, University of Pisa, Via Giuseppe Moruzzi 13, Pisa, 56124, Italy.

reactions, a process central to many biological functions and technological applications. Aequorin[1] and obelin,[2] two well-characterized calcium-activated photoproteins, have served as model systems for studying the molecular mechanisms of bioluminescence.[3] At the core of this process is coelenteramide (Coel), a chromophore that is non-covalently embedded within the protein hydrophobic cavity and plays a central role in light emission.[4–6] Depending on its environment, Coel can exist in neutral (protonated, **2H**) or anionic (deprotonated, **2O⁻**) forms,[7] both displayed in Fig. 1. These two species can interconvert particularly *via* excited-state proton transfer,[5,8] a process crucial to Coel's photophysics. Early experimental[7,9] and theoretical studies[5,6,10] identified potential emissive species, with growing evidence supporting the phenolate anion as the dominant emitter, especially in polar embeddings or following electronic excitation. These findings underscore the importance of the local environment (protein or solvent) in

15946 | *Phys. Chem. Chem. Phys.*, 2025, **27**, 15946–15959

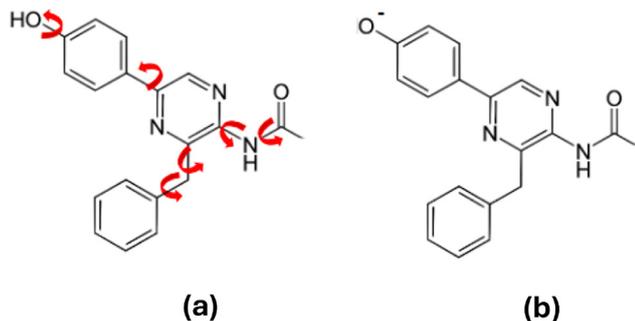This journal is © the Owner Societies 2025

Fig. 1 (a) Coelenteramide's protonated species, **2H**. Here, the most flexible internal coordinates are indicated in red; (b) coelenteramide's anionic species, **2O⁻**.

modulating the chromophore's structure and emission properties.[6,8,11] Indeed, the spectroscopic signature of chromophores embedded in complex environments such as protein matrices or solvated systems reflects the delicate interplay between the intrinsic properties of the absorbing species and its interactions with the embedding. At the molecular level, key factors include the chromophore's electronic structure, its internal conformational flexibility, the coupling of the electronic transition to fast vibrations, and the presence of closely-lying potentially coupled excited states. All these features are further modulated by the surrounding medium, which may impose steric constraints, enable specific non-covalent interactions such as π–π stacking or hydrogen bonds (HBs), or dynamically alter the local electrostatic environment.[12–15] Whether in simple solvents or in more structured protein scaffolds, this solute–environment coupling can in turn significantly influence spectral features by inducing shifts in excitation energies, modifying vibronic progressions, and broadening spectral bands. In this framework, computational approaches represent an essential tool to flank the experiment in interpreting the experimental signal, due to the unique possibility to unravel the separate contributions to the spectral shape and their complex interplay. Nonetheless, capturing all these effects with predictive accuracy necessitates setting up consistent and robust protocols that can simultaneously describe the interplay between electronic transitions, nuclear dynamics, and environmental fluctuations.

Due to its pronounced environmental sensitivity and conformational flexibility, Coel provides an ideal platform for addressing the aforementioned computational challenges associated with the simulation of electronic spectral shapes. In fact, within protein cavities, Coel is stabilized through a complex network of non-covalent interactions, including hydrogen bonding, steric confinement, and electrostatic contacts, which can significantly modulate its electronic structure and transition energies. In solvated systems, additional complexity may arise from dynamic solute–solvent interactions, solvent-induced electrostatic effects, and possible excited-state proton transfers, all of which influence its photophysical behavior. Understanding the extent to which these interrelated factors modulate Coel's absorption spectra is essential for evaluating and advancing computational protocols capable of capturing

vibrationally resolved transitions and environment-specific spectral modulations in realistic molecular systems. Recently, the adiabatic molecular dynamics generalized vertical Hessian (Ad-MD|gVH) protocol was introduced by some of us to simulate the vibronic spectroscopy of flexible molecular systems in complex environments.[14,16,17] The method relies on the adiabatic separation of nuclear degrees of freedom into slow soft modes, associated with large-amplitude and low-frequency motions (treated classically *via* molecular dynamics) and fast stiff modes, corresponding to high-frequency intramolecular vibrations (treated quantum-mechanically within a vibronic model). For each configuration sampled along a classical molecular dynamics (MD) trajectory, reduced-dimensional harmonic Hamiltonians are constructed in the stiff-mode subspace to compute configuration-specific vibronic spectra. These are eventually ensemble-averaged along the MD sample to generate the final spectral profile. Importantly, the use of accurate quantum-mechanically derived force fields (QMD-FFs), parameterized at the same level of theory employed for the vibronic calculations, ensures a consistent and physically grounded description of both intramolecular flexibility and solute–environment interactions. It should be noted that this protocol provides access to vibrationally resolved absorption spectra without the need for phenomenological broadening and has proven effective in capturing vibronic features and large-amplitude conformational motions across a range of solvated systems.[16,18] The Ad-MD|gVG protocol[17] (where VG stands for the vertical gradient) is a variation of the original Ad-MD|gVH scheme, assuming that the excited state has the same normal modes and frequencies of the ground state. This variant is an easy option which allows for avoiding significant artifacts which may arise in VH, due to possible imaginary frequencies caused by inter-state couplings. However, since the VG model neglects the latter couplings, when they are deemed to be important it is necessary to devise alternative strategies, able to properly account for these effects. The most straightforward choice consists in exploiting properly-parameterized linear vibronic coupling (LVC) models, which can be integrated in the Ad-MD|g procedure, hence giving rise to the Ad-MD|gLVC scheme.[15,17] In this case, vibronic spectra for each snapshot of the slow-coordinates need to be computed by numerically propagating nuclear wavepackets, describing the fast coordinates, on the coupled surfaces through quantum dynamics (QD) simulations. Depending on which set of the LVC parameters are made explicitly dependent on the fluctuation of the slow coordinates, a whole family of Ad-MD|gLVC can be set up, thus providing a powerful extension of the original implementation.[15]

In the present work, we systematically evaluate the capability of the above-mentioned computational strategies to accurately simulate the absorption spectral lineshape of Coel in benzene, a molecular system that poses significant challenges to current methodologies. Our goal is to disentangle the individual contributions of internal flexibility, vibronic effects, inter-state couplings, and specific solute–solvent interactions to the absorption spectral lineshape and vibronic structure. In particular, when dynamical effects are taken into account, we show

This journal is © the Owner Societies 2025

Phys. Chem. Chem. Phys., 2025, **27**, 15946–15959 | **15947**

that Ad-MD|gVG faces challenges in accurately predicting Coel's relative spectral intensities, and shows artifacts, notwithstanding its proven success across a wide range of solvated systems.[14,16,18] The worse performances here registered for Coel can be traced back to the fact that both fast and slow degrees of freedom are able to trigger strong coupling and mixing between the electronic states. On these grounds, in this contribution we test a new extension of the Ad-MD|gLVC approach, designed to overcome these limitations. The insights provided by this study underscore the need to critically assess the applicability of widely used methods and pave the way for developing more accurate and generalizable models for condensed-phase vibronic spectroscopy.

2 Methods and computational details

2.1 QM calculations

All electronic quantum mechanical (QM) calculations were carried out using density functional theory (DFT) for the ground electronic state and time-dependent DFT (TD-DFT) for the excited states, as implemented in the Gaussian16 software package.[19] In order to parametrize the QMD-FF and to identify the global minimum structure, a full geometry optimization as well as ground-state relaxed scans along selected flexible torsional coordinates (see ESI† for details) were performed using the CAM-B3LYP functional in combination with the 6-31G(d,p) basis set and Grimmes D3 dispersion correction. Furthermore, as required by the Joyce protocol,[18,20] the Hessian matrix was explicitly computed at the minimum energy geometry.

Absorption spectra were computed at the same level of theory and, to evaluate the influence of the exchange–correlation functional, they were also obtained using the excited state properties calculated using single-point B3LYP calculations on CAM-B3LYP optimized geometries. Additionally, the reliability of B3LYP and CAM-B3LYP functionals in describing Coel's excited-state properties was benchmarked applying multireference RASSCF/RASPT2 as well as coupled cluster STEOM-DLPNO-CCSD methods, as detailed in Sections S1 and S2 of the ESI,† respectively.

2.2 Adiabatic and nonadiabatic vibronic spectra in a static approximation

Considering the global-minimum structure and neglecting structural dynamics in solution (static approximation), the absorption spectrum of Coel was simulated with the adiabatic vertical gradient (FC|VG) approach implemented in the code $\mathcal{FC}classes$3,[21,22] at both B3LYP and CAM-B3LYP levels of theory. For the **2H** species, the possible impact of nonadiabatic couplings on the absorption lineshape was also evaluated *via* performing QD wavepacket propagations on the grounds of a LVC Hamiltonian model

$$
\mathbf{H} = \sum_{\eta} \left(K + V_{\eta\eta}^{\text{dia}}(\mathbf{q}) \right) |d_\eta\rangle\langle d_\eta|
$$
$$
+ \sum_{\eta,\zeta > \eta} V_{\eta\zeta}^{\text{dia}}(\mathbf{q}) \left(|d_\eta\rangle\langle d_\zeta| + |d_\zeta\rangle\langle d_\eta| \right) \quad (1)
$$

where $|d_\eta\rangle$ are the diabatic electronic states and $\mathbf{q}$ and $\mathbf{p}$ are the ground state dimensionless normal coordinates and associated momenta.

The kinetic (K) and potential (V) energy terms are

$$
K = \frac{1}{2}\mathbf{p}^{\text{T}}\mathbf{\Omega}\mathbf{p} \quad (2)
$$

$$
V_{\eta\eta}^{\text{dia}}(\mathbf{q}) = E_\eta^0 + \boldsymbol{\lambda}_{\eta\eta}^{\text{T}}\mathbf{q} + \frac{1}{2}\mathbf{q}^{\text{T}}\mathbf{\Omega}\mathbf{q} \quad (3)
$$

$$
V_{\eta\zeta}^{\text{dia}}(\mathbf{q}) = E_{\eta\zeta}^0 + \boldsymbol{\lambda}_{\eta\zeta}^{\text{T}}\mathbf{q} \quad (4)
$$

where $\mathbf{\Omega}$ is the diagonal matrix of the vibrational frequencies of the ground state g, E_η^0 represents the η-th excited-state energy at the g equilibrium geometry, $\boldsymbol{\lambda}_{\eta\eta}$ is the energy gradient of state η and $\boldsymbol{\lambda}_{\eta\zeta}$ is the gradient of the inter-state coupling $V_{\eta\zeta}^{\text{dia}}(\mathbf{q})$.

The diabatic states $|d_\eta\rangle$ were defined to be as similar as possible to reference states, which are the adiabatic states at the ground-state global minimum geometry. The LVC Hamiltonian was parameterised from B3LYP and CAM-B3LYP TD-DFT calculations in the gas phase and PCM/benzene through a maximum-overlap diabatization technique implemented in the open-source code OVERDIA.[23,24] For further details we refer to Section S3 of the ESI.† The nonadiabatic QD of the vibronic wavepacket was subsequently carried out in a vibronic space with reduced dimensionality. This was obtained by adopting a set of 45 selected effective vibrational modes, corresponding to the ones carrying the largest intra-molecular or inter-molecular linear couplings (see Fig. S1 and Section S4 of the ESI† for details). The vibronic wavepackets were propagated by the efficient ML-MCTDH method,[25–27] as implemented in the QUANTICS package,[28] which was also employed to compute the correlation functions required for the computation of nonadiabatic vibronic spectra as done in ref. 29.

2.3 Ad-MD vibronic spectra

2.3.1 QMD-FF parametrization and MD simulation. The MQC approaches used in this work – namely Ad-MD|gVG and its generalization, Ad-MD|gLVC – both rely on an accurate description of the classical dynamics of the solute and solvent degrees of freedom. As previously demonstrated,[18,30] QMD-FFs are beneficial in this context, as the accuracy of the force field parameters and related molecular dynamics time evolution can be tailored to the underlying quantum mechanical method. We note here in passing that the parametrization was performed by partitioning the total energy in: (i) an intramolecular part, subdivided into bonds, angles, stiff and flexible dihedrals (see Fig. 1) energy contribution and non-bonded atom pairs between atoms of the same molecules; and (ii) an intermolecular part. In this work, the intramolecular term of QMD-FF is derived against the reference CAM-B3LYP/6-31G(d,p) data, computed for the isolated **2H** species. The parametrization procedure employed the JOYCE 3.0 software[31,32] and the related protocol[18,20,33] is thoroughly described in Section S5 of the ESI.† Regarding the intermolecular component of Coel's QMD-FF, the Lennard-Jones (LJ) parameters were directly adopted from standard force-fields.[34]

15948 | *Phys. Chem. Chem. Phys.*, 2025, **27**, 15946–15959

This journal is © the Owner Societies 2025

To capture the electrostatic interactions with the embedding, the point charges for the **2H** species were derived through the RESP approach[35] from the electronic density of the dye's optimized geometry, calculated with PCM[36] adopting the appropriate dielectric constant for benzene. As far as the latter is concerned, all inter-molecular parameters for benzene were consistently taken from the OPLS database,[37] whereas the internal degrees of freedom were also parametrized using the JOYCE protocol. We note that the parameters and MD input files used in the work are freely available.[38]

The QMD-FF described above was later used in MD simulations, carried out with the GROMACS package[39] on a system consisting of one **2H** chromophore, solvated with around 1000 explicit benzene molecules. The production runs were carried out for a total of 10 ns, and uncorrelated frames were extracted each 100 ps, hence collecting a reliably thermalized sample which constitutes the statistical ensemble required by the Ad-MD protocols for the calculation of the vibronic spectra. Further details of the MD simulation settings are given in Section S5.4 of the ESI.†

2.3.2 Ad-MD|gVG spectra. The impact of the environment and, more in general, of soft modes on electronic spectra can be reliably included resorting to mixed quantum-classical (MQC) approaches. Concretely, we here employ the Ad-MD|gVG procedure,[14,17] which is rooted, as all Ad-MD|g schemes, in partitioning the total systems' (solute + solvent) degree of freedoms into soft and stiff modes. According to this protocol, the soft modes, which include the solute's flexible degrees of freedom and solvent coordinates, are sampled through classical MD, while the stiff modes are treated at the QM level using harmonic potentials parameterized with QM/MM energies, gradients (and Hessians for the Ad-MD|gVH variant) computed specifically for each configuration of the soft modes. The final Ad-MD|gVG spectrum is obtained by averaging the QM-level signals computed for the stiff coordinates over the ensemble of MD-generated configurations, explicitly incorporating solvent effects under thermodynamic conditions (room temperature and 1 atm). This approach represents a significant advancement over conventional methods such as classical ensemble average of vertical energies (CEA-VE), which simply takes the vertical energies and intensities at the different snapshots, assign them a small, yet phenomenological Gaussian bandwidth before averaging them. The present MQC scheme correctly accounts for homogeneous broadening and the spectral structure induced by high-frequency vibrational progressions. By incorporating these effects from first principles, the Ad-MD|gVG method eliminates the need for empirical parameters to reproduce spectral shapes, providing a more accurate and physically consistent description of electronic spectra.

Vertical energies and gradients are computed at each of the 100 uncorrelated snapshots purposely sampled from the MD trajectory, using an ONIOM QM/MM scheme, where the Coel is treated at the QM level. Solvent molecules within 7.5 Å of the Coel were described at the MM level, while the remaining solvent atoms within 15 Å were included as point charges (adopting the same partial charges used in the MD simulation).

The inclusion of this first MM sphere enhances the accuracy of gradient and Hessian calculations, as previously discussed in ref. 14. The resulting data were used for both the CEA-VE (vertical energies only) and Ad-MD|gVG approaches. Within the CEA-VE scheme, spectra were computed by considering the six lowest electronic transitions, across the extracted snapshots. Each transition was weighted by the squared transition dipole moment and subsequently broadened using Gaussian convolution with the HWHM = 0.08 eV to obtain the final spectral lineshape. VG vibronic computations were performed using the *FCclasses*3 code,[21] exploiting the time-dependent (TD) approach at 300 K and projecting out all the six flexible torsional modes. The Franck–Condon (FC) approximation was applied for the electric transition dipole, and a very small broadening of HWHM = 0.01 eV was selected.

2.3.3 Nonadiabatic Ad-MD|gLVC spectra. As discussed in the next section, the electronic states of **2H** species are found to be coupled through both the vibrations of high-frequency nuclear modes and the dynamical disorder brought in by the solute's soft coordinates and the solvent motions. As a consequence, in several snapshots extracted from the MD trajectory, the computed excited states are almost degenerate and their electronic character is strongly mixed, with a drastic exchange of oscillator strengths. These findings point to an extremely complex scenario. On one hand, quasi-degeneracies, mixing of the states and couplings due to the fast modes, indicate the necessity of a non-adiabatic vibronic treatment. On the other hand, those inter-state couplings induced by the solute's soft modes and the solvent dynamics call for a proper description of the slow, large-amplitude motions. At the state of the art, there are no well assessed methodologies for dealing with such challenging cases and a rigorous solution is outside the scope of the present work. With these premises in mind, in order to qualitatively assess the role of these couplings, here we resort to an approximate approach, by devising a new variant of the Ad-MD|gLVC approach, which combines the Ad-MD scheme with a generalized LVC model.[15] According to this protocol, a reliable statistical ensemble of large amplitude soft modes is sampled through MD simulations, while a vibronic LVC Hamiltonian takes care of the description of the role of fast degrees of freedom. The coupling between soft and stiff modes is introduced by making some of the parameters of the LVC Hamiltonian dependent on the MD snapshot. The final spectrum is eventually computed as a ensemble average of the nonadiabatic spectra. Concretely, since the LVC linear terms due to the fast degrees of freedom ($\lambda_{\eta\eta}$ and $\lambda_{\eta\zeta}$ in eqn (3) and (4)) are not expected to be significantly affected by the dynamic disorder, we considered them as equal to those of the LVC parameterized at the ground-state geometry for the static approach. In contrast, the effect of the slow degrees of freedom is taken into account by re-computing both vertical energies (E_{η}^0, see eqn (3)) and constant couplings ($E_{\eta\zeta}^0$, eqn (4)) at every sampled MD frame, hence obtaining a set of LVC models, specific for the instantaneous configuration of the system. To this end, we performed a new "single-point" diabatization at each snapshot like we did in the original version of the method designed for

This journal is © the Owner Societies 2025

Phys. Chem. Chem. Phys., 2025, **27**, 15946–15959 | **15949**

π-stacked dimers.[15] The difference here is that the definition of the diabatic states at each configuration is more complex. We decided to use as reference states those expressed by the same TD-DFT vectors that define the four lowest energy adiabatic states of **2H** species at the equilibrium geometry. This choice is based on the assumption that the MOs do not significantly change during the MD (which is constrained as explained below), so the fact that the TD-DFT vectors are the same implies that the electronic states are the same, as requested in a diabatic picture. The reliability of this approximation was tested analyzing in detail the MOs of several snapshots, and it was found acceptable for both considered functionals, yet better for B3LYP. The parameters of the LVC Hamiltonian were obtained for each snapshot by projecting the diabatic states on a set of 50 adiabatic states computed at the TD-DFT level, with the code OVERDIA.[23] It is noteworthy that, following this procedure, the couplings $E^0_{\eta\zeta}$, which were null at the ground-state geometry, now do not vanish, reflecting the mixing of the electronic states induced by the fluctuations of the slow coordinates. In the next section we will show that a further consequence of such mixing is a remarkable variation across the sampled ensemble of oscillator strengths of the adiabatic states, another effect calling for a proper non-adiabatic treatment. Therefore, at each snapshot we also reconstruct through the diabatization procedure the transition dipoles of the diabatic states. The fact that their modules always resemble those of the reference states at the ground-state geometry (where we have two bright and two dark states, see the Results section) is indirect proof of the validity of the overall diabatization procedure.

In order to disentangle as much as possible the couplings due to fast and slow degrees of freedom, the Ad-MD|gLVC protocol was applied along a trajectory where all bond stretching modes were constrained during the dynamics through the LINCS algorithm. Although described by the QMD-FF through similar harmonic potentials, imposing restraints on angle bendings or stiff torsions is less straightforward, as it would challenge energy conservation and MD reliability. Therefore, a further approximation of our scheme is related to the treatment of these soft modes. This is becuase their motion is both accounted at the classical level by MD, yet also partially described by the reduced set of 45 normal coordinates entering in the LVC model. In order to reduce the impact of this problem, we did not include any mode with a frequency <500 cm^{-1} among the 45 modes of the LVC Hamiltonian, so that at least stiff torsions and low-frequency bendings are only sampled by classical MD. For the rest of the modes, for instance high-frequency bendings, we assume that their role in the spectra is small; hence, double-counting their effect should not introduce significant artifacts. Finally, like for nonadiabatic static spectra, the time correlation functions needed to compute the individual nonadiabatic spectra were computed numerically through the QD propagation of nuclear wavepackets on the coupled PESs, again using the ML-MCTDH method as implemented in the QUANTICS code (see Section S4 in the ESI† for details).

3 Results

3.1 Excited state properties of 2H and 2O⁻ species

The absorption spectrum of Coel in benzene lies in the 3.35–4.30 eV range, peaking at 3.70 eV and 4.20 eV. Table 1 presents the main features of the five lowest-energy excited states of the neutral **2H** and the anionic **2O⁻** species, computed at the TD-DFT level, in their gas-phase global minimum structures (see Fig. 2). Further details, including natural transition orbitals (NTOs), are reported in Section S6 of the ESI.† At both B3LYP and CAM-B3LYP levels, the results for **2H** predict two bright states (S_1 and S_3) with considerable (and comparable) oscillator strengths, associated with $\pi\pi^*$ electronic transitions (see NTO's in Fig. S5 and S6, ESI†). With the former functional, these states lie at 3.76 and 4.28 eV, in excellent agreement with experiment, whereas CAM-B3LYP overestimates both transition energies, resulting in a systematic blue shift of ~ 0.5 eV for all states. Nevertheless, this functional accurately predicts the experimental spectral spacing of 0.5 eV associated with the S_1 and S_3 bright states. The results of our calculations also show that the S_2 and S_4 states are almost dark, being characterized by remarkably low oscillator strengths and a mixed $n\pi^*/\pi\pi^*$ character. While S_2 is slightly bright according to CAM-B3LYP, it remains dark according to B3LYP due to a greater contribution from $n\pi^*$ transitions, which are typically less intense and have an evident charge transfer (CT) character. The MO transitions contributing to the S_5 state were also found to be dependent on the employed functional: while CAM-B3LYP predicts a $\pi\pi^*$ transition, B3LYP predicts a mixed $n\pi^*/\pi\pi^*$ character, and consequently, a lower oscillator strength for this state.

For the **2O⁻** species the situation is completely different. CAM-B3LYP predicts that the first low-lying S_1 and S_2 states have $\pi\pi^*$ character and are bright, with a significant oscillator strength that is remarkably larger for the second state (see Fig. S7 and S8, ESI†). These states are found at 2.98 and 3.58 eV, respectively, thus falling significantly outside Coel's experimental absorption spectral range, with a large red-shit of 0.7 eV. Additionally, the S_3–S_5 states are dark with almost zero oscillator strengths, being slightly larger for S_5 than the other dark states. Notably, while S_3 has a significant $n\pi^*$ contribution, S_4 and S_5 have a marked $\pi\pi^*$ character, but remain dark. Employing the B3LYP functional yields seemingly different results: all first three states show a $\pi\pi^*$ character, yet with negligible (S_1 and S_3) or null (S_2) oscillator strengths. S_4 is a dark $n\pi^*$ state, while S_5 is the only state expected to show a significant absorption. Nonetheless, similarly to CAM-B3LYP, also at the B3LYP level, both the S_1/S_3 energies (2.32 and 2.91 eV) and the transition toward S_5 (3.33 eV) exhibit a remarkable red-shift (~ 1.4 eV and ~ 0.4 eV, respectively, similarly to the CAM-B3LYP level) compared to the experimental absorption spectrum. The significant difference between the excited state energies computed for **2O⁻** and the experimental absorption range indicate that the anionic form should not be present in benzene solution, as otherwise its signal would have been seen in the experiment. In contrast, the excitation energies of the first two bright states (S_1 and S_3) of

15950 | *Phys. Chem. Chem. Phys.*, 2025, **27**, 15946–15959

This journal is © the Owner Societies 2025

Table 1 CAM-B3LYP (and B3LYP, presented in parentheses) excitation energies (E, in eV), oscillator strengths (f), and their predominant character in terms of NTOs. All data were computed for the five lowest excited states of **2H** and **2O⁻**, at the FC point associated with the gas-phase global minimum structure

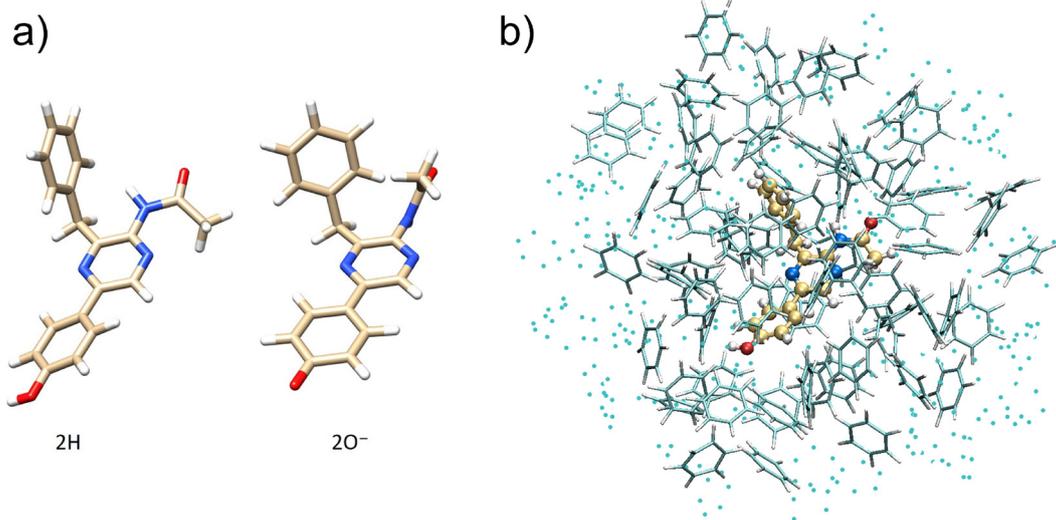| State	2H					2O⁻				
	E	f	Main transitions	Character	$W\%$	E	f	Main transitions	Character	$W\%$
S_1	4.259	0.451	H → L	π/π^*	84.76	2.983	0.195	H → L	π/π^*	84.24
			H−1 → L	π/π^*	3.09			H → L+3	π/π^*	11.89
			H → L+1	π/π^*	2.90					
	(3.756)	(0.297)	(H → L)	(π/π^*)	(91.94)	(2.323)	(0.045)	(H → L)	(π/π^*)	(86.86)
			(H → L+1)	(π/π^*)	(6.17)			(H → L+2)	(π/π^*)	(6.60)
								(H → L+3)	(π/π^*)	(5.09)
S_2	4.471	0.024	H−6 → L	n/π^*	24.92	3.578	0.796	H → L+1	π/π^*	67.05
			H−1 → L	$n, \pi/\pi^*$	24.08			H → L+3	π/π^*	21.00
			H−5 → L	π/π^*	9.42					
	(4.119)	(0.005)	(H−1 → L)	$(n, \pi/\pi^*)$	(72.00)	(2.688)	(0.001)	(H → L+1)	(π/π^*)	(91.40)
			(H−6 → L)	(n/π^*)	(14.91)			(H → L+2)	(π/π^*)	(8.08)
S_3	4.784	0.438	H → L+1	π/π^*	82.18	3.938	0.000	H−1 → L	n/π^*	44.46
			H → L	π/π^*	4.17			H−1 → L+1	n/π^*	28.43
			H → L+3	π/π^*	3.40			H−1 → L+7	n/π^*	7.09
	(4.284)	(0.550)	(H → L+1)	(π/π^*)	(88.98)	(2.906)	(0.052)	(H → L+2)	(π/π^*)	(51.01)
			(H → L)	(π/π^*)	(6.51)			(H → L+3)	(π/π^*)	(43.38)
								(H → L+1)	(π/π^*)	(5.10)
S_4	5.081	0.008	H−1 → L+1	$n, \pi/\pi^*$	25.78	4.089	0.001	H → L+2	π/π^*	95.22
			H−6 → L+1	n/π^*	14.26					
	(4.535)	(0.016)	(H−1 → L+1)	$(n, \pi/\pi^*)$	(65.67)	(2.996)	(0.000)	(H−1 → L)	(n/π^*)	(75.15)
			(H−6 → L+1)	(n/π^*)	(10.31)			(H−1 → L+2)	(n/π^*)	(19.34)
S_5	5.135	0.062	H → L+3	π/π^*	38.72	4.263	0.016	H → L+3	π/π^*	61.61
			H → L+4	π/π^*	15.35			H → L+1	π/π^*	29.34
	(4.717)	(0.021)	(H → L+3)	(π/π^*)	(29.95)	(3.331)	(0.783)	(H → L+3)	(π/π^*)	(50.60)
			(H−2 → L)	$(n, \pi/\pi^*)$	(24.78)			(H → L+2)	(π/π^*)	(33.78)

Fig. 2 (a) Global-minimum molecular structure of **2H** and **2O⁻** in the gas phase, obtained at the CAM-B3LYP/6-31G(d,p) level of theory; (b) QM/MM model used in this study for **2H**, involving both explicit solvent shells and point charges.

the **2H** species align well with experimental observations (at least at the B3LYP level).

Focusing on the MO transitions involved in the two bright states of **2H** species, it appears that S_1 and S_3 are mixed with a combination of H → L and H → L+1 transitions with $\pi\pi^*$ character. Notably, S_1 is predominantly characterized by the H → L transition, while S_3 is mainly governed by the H → L+1 transition. Moreover, while B3LYP predicts a relatively pure configuration for the bright states, CAM-B3LYP reveals a more complex mixture of electronic transitions, yet yielding nearly

This journal is © the Owner Societies 2025

Phys. Chem. Chem. Phys., 2025, **27**, 15946–15959 | **15951**

identical oscillator strengths for S_1 and S_3, in apparent better agreement with experiment, with respect to B3LYP. Aiming to a deeper insight, we calculated the excited state characteristics of the **2H** species in benzene, accounted for at the PCM level. The results shown in Table S4 of the ESI† indicate that while the nature and composition of the NTOs involved in the transitions remained unchanged (see Fig. S9 and S10 for B3LYP and CAM-B3LYP, respectively, ESI†), compared to those obtained in the gas-phase, the oscillator strengths of the lowest-energy bright state S_1 have increased with both functionals. Remarkably, the two bright states S_1 and S_3 feature now more similar oscillator strengths at the B3LYP level, apparently in better agreement with the experimental relative intensities.

This dependence of the peak intensities ratio on the applied functional, registered at the TD-DFT level, suggests that the residual difference can be due to electronic structure inaccuracies. Therefore, we calculated the excited state characteristics of the **2H** species at the FC point at the RASSCF/RASPT2 level in the gas-phase, by progressively extending the size of the active space (from 16e–16o to 22e–19o, where e stands for the electron and o stands for the orbital), adopting single-state (SS) and multi-state (MS) approaches. The results presented in Section S1 of the ESI† exhibit a significant variability depending on the SS and MS approach and on the active space. SS-RASPT2 results obtained with smaller active spaces provide an overall improved prediction of the peak spacing, but the MS-RASPT2(22,19) outcomes guarantee a more balanced description of the excitation energies and their relative intensities. In a further attempt, we also tried coupled cluster methods. The results of the STEOM-DLPNO-CCSD calculations, presented in Section S2 of the ESI,† provide a balanced description of the relative excitation intensities. However, they overestimate the peak spacing by ~ 0.2 eV. Notwithstanding our efforts and the associated computational burden, none of the methods employed provided results being in good agreement with experiment concerning, at the same time, both the energy separation and the intensity ratio. In conclusion, these findings support the idea that in order to reproduce such experimental observations the inclusion of vibronic effects may be crucial, and this is what we investigate in the following sections at the TD-DFT level.

3.2 Static vibronic spectra

To unravel the separate contributions of fast nuclear vibrations and slow thermal fluctuations, we first neglect the effect of the latter and compute static vibronic spectra, *i.e.* spectral shapes obtained describing the associated vibronic progressions through a simple harmonic expansion of the ES PES at the ground-state equilibrium geometry. More specifically, we used the global minimum structures obtained for **2H** and **2O**$^-$ in the gas phase.

3.2.1 Adiabatic spectra. The top panels of Fig. 3 compare the experimental spectrum recorded for Coel in benzene, with the adiabatic vibronic FC|VG spectra, computed for the isolated **2H** and **2O**$^-$ at room temperature, considering the six lowest singlet excited states, with either B3LYP or CAM-B3LYP. In

order to dissect the contribution of different states to the total spectral shape, for **2H** species the FC|VG spectra associated with each single state are also presented. The bottom panels report analogous results in benzene, described using the PCM. All the flexible torsional motions (rotations around single bonds), corresponding to low-frequency anharmonic modes, were projected out,[14,22] as they are usually associated with the large amplitude motions whose movements cannot be approximated with a harmonic expansion. The effect of such flexible modes becomes even more relevant at room temperature, where the conformational space accessible to the flexible coordinates further deviates from the one expected by a harmonic approximation, and it will be analysed using the Ad-MD|gVG approach in the following sections.

With both functionals, the overall absorption spectral features of **2H** appears to be much more consistent with the experiment than those obtained for the **2O**$^-$ species. On the one hand, the **2H** vibronic spectrum computed at the B3LYP level in the gas phase shows two main bands, lying at energies that perfectly match the experimental ones, while CAM-B3LYP excitation energies exhibit a consistent blue shift of 0.5 eV. On the other hand, the experimental ratio between the intensities of the two main peaks is not properly reproduced with the former functional, while CAM-B3LYP accurately captures the relative intensities and energy separation of the two main peaks. As anticipated by the excited state analysis reported in Table 1, the main spectral contributions originate from transitions to the S_1 and S_3 states. It is worth noticing that according to the vibronic computations at both B3LYP and CAM-B3LYP levels, the second band, positioned at higher energies (~ 4.25 eV in the experiment), arises from two separate peaks, which do not appear in the experiment. In the computed spectra, they arise from vibronic progressions of the S_3 state, as evident from the relevant individual FC|VG vibronic spectrum. All the computed spectra are slightly red-shifted in PCM/benzene, and the relative intensity of the two main bands is altered, with a slight increase of the first band intensity (more evident with B3LYP), which improves the agreement with experiment for B3LYP and partially deteriorates it using CAM-B3LYP. Finally, in any case considered, the vibronic spectra of **2O**$^-$ exhibit a significant red shift with respect to experiment, a result consistent with pure electronic vertical excitation trends, indicating that **2H** is the only species contributing to the absorption spectrum in solution. Therefore, in the following, we will focus on the absorption spectra of **2H**.

3.2.2 Nonadiabatic spectra. The static vibronic spectra approximated by the FC|VG model in the previous section are "adiabatic", in the sense that they ignore the possible couplings arising between close-lying excited-states. Remarkably, according to both functionals the four lowest excited states of **2H** do reside in a relatively narrow energy window of ~ 0.8 eV (see Table 1), a situation where inter-state nonadiabatic couplings are likely to be active. To investigate their effect, as detailed in the Methods, we adopted a LVC Hamiltonian. It is worth stressing that LVC collapses into the FC|VG model if inter-state couplings are switched off. The resulting LVC
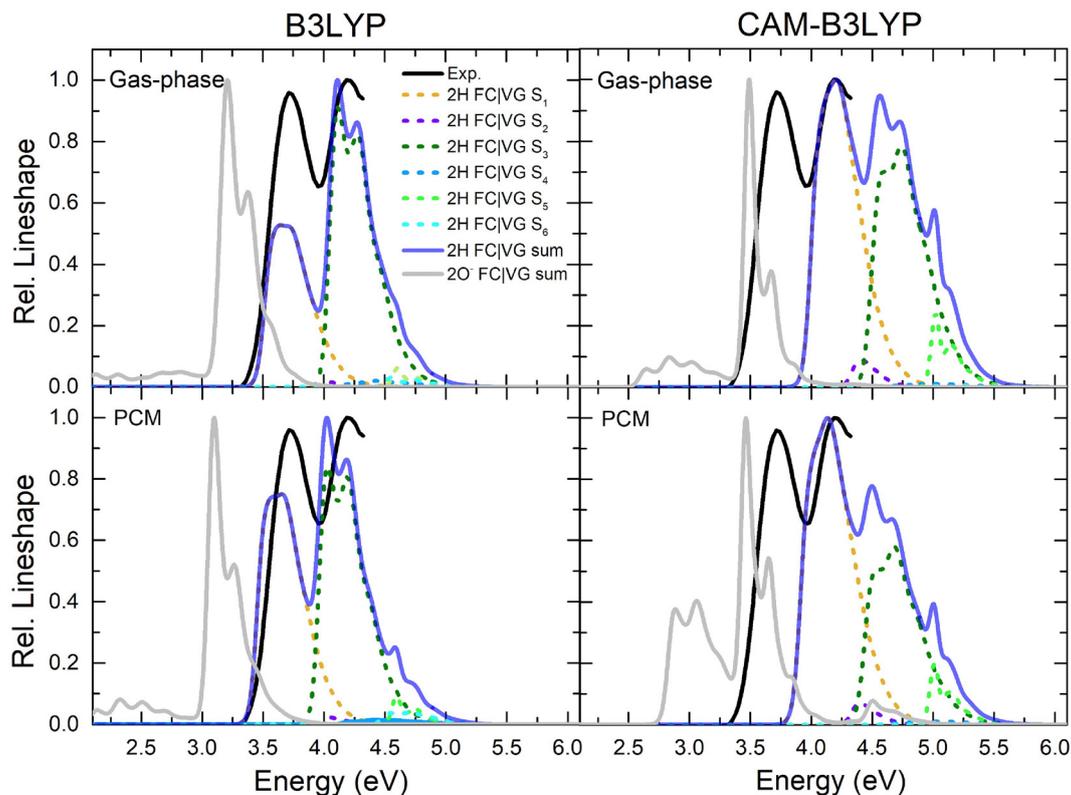
15952 | *Phys. Chem. Chem. Phys.*, 2025, **27**, 15946–15959

This journal is © the Owner Societies 2025

Fig. 3 Comparison of adiabatic FC|VG vibronic spectra computed at the global-minimum molecular structures of **2H** and **2O⁻** in the gas phase (top panels) or PCM/benzene (bottom panels), with either the B3LYP (left panels) or CAM–B3LYP (right panel) functionals. The contribution of different states to the total vibronic FC|VG spectra of the **2H** species is also displayed with dashed lines. To facilitate the comparison, all the spectra are normalized at the maximum peak height intensity of the main peak. No shift has been applied to the computed spectra, and all vibronic transitions are computed at $T = 300$ K and further convoluted with a Gaussian of HWHM = 0.04 eV.

spectra, computed at 0 Kelvin with or without accounting for the couplings in the LVC Hamiltonian, are compared in Fig. 4 with the experiment. It appears that, according to both functionals, accounting for interstate coupling erases the two-peak structure of the higher energy band, hence improving the agreement with the experimental signal. In general, LVC spectra appear to be moderately smoother when the couplings are switched on. It is also evident that accounting for inter-state couplings, in particular in PCM/benzene, modulates the intensity ratio between the first and second bands, leading to a slight improvement of the B3LYP predictions and a deterioration at the CAM-B3LYP level. As a final comment we notice that the comparison between the FC|VG spectra in Fig. 3, computed at 300 K considering all coordinates, and those displayed in Fig. 4, obtained at 0 K with the LVC model, switching off the inter-state couplings and considering only 45 coordinates, allows for assessing the only moderate smoothing induced on the spectral shape by the thermal excitation and the contribution of the less displaced modes (included only in FC|VG spectra). This finding supports the reliability of the approximation made reducing the number of normal modes to 45 to make LVC computations faster. The QD propagations required to compute the non-adiabatic spectra displayed in Fig. 4 also give access to the time evolution of the electronic populations after excitation to

any of the diabatic states of the model. Results are reported and discussed in detail in the ESI† (see Section S7 and Fig. S11, S12). In brief, according to both functionals, after an excitation to any of the four lowest excited states, the population is found mainly on S_1 in dozens of femtoseconds.

3.3 Modeling of spectral lineshapes in an explicit solvent model

3.3.1 Including dynamical disorder: CEA-VE and Ad-MD|gVG models. In order to capture the inhomogeneous broadening and spectral modulations arising from large-amplitude motions of the flexible modes of the solute and induced by specific solute/solvent interactions, we resorted to MD conformational sampling, hence abandoning the static approximation and instead computing the spectra in either the CAE-VE or Ad-MD|gVH schemes. In both approaches, the contribution of the first 4 states was taken into account.

The left panels of Fig. 5 compare Coel's experimental spectrum with those predicted by applying the CEA-VE protocol, where the vertical transitions computed along the MD trajectory were convoluted with a Gaussian with a HWHM of 0.08 eV. The separate contribution of the different states to the total spectral shape of **2H** species can be evaluated by looking at the CEA-VE spectra associated with each of the first four excited
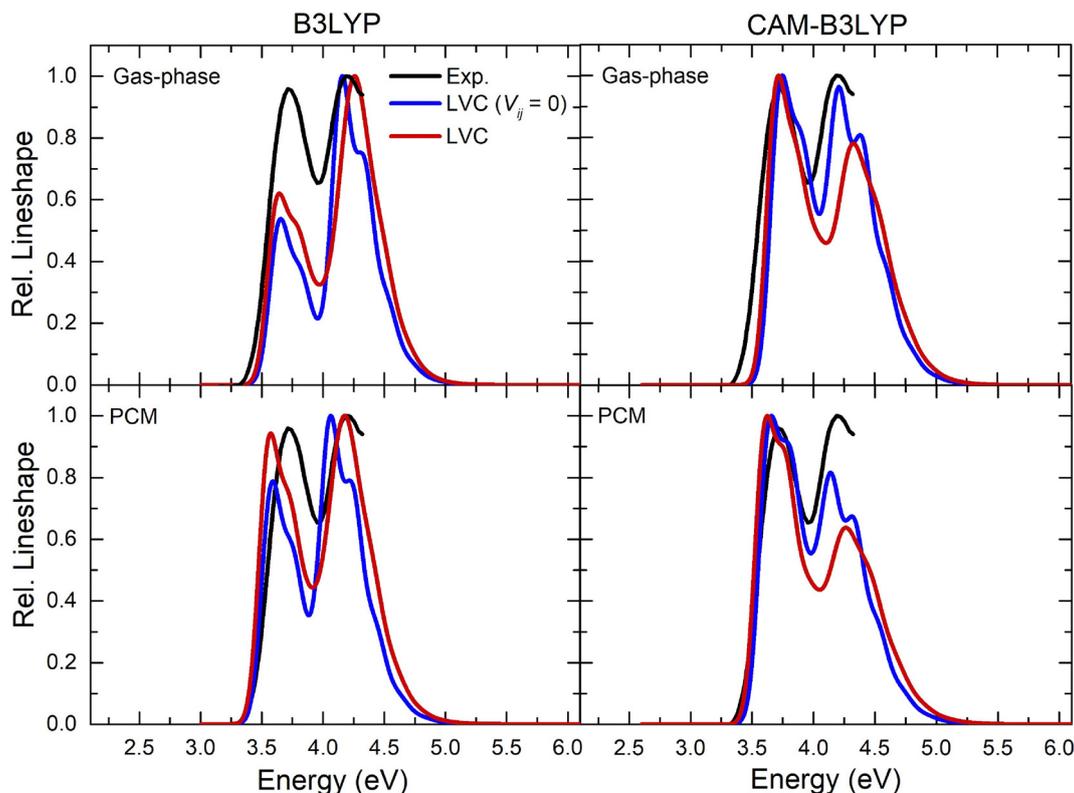
This journal is © the Owner Societies 2025

Phys. Chem. Chem. Phys., 2025, **27**, 15946–15959 | **15953**

Fig. 4 Comparison of the experimental absorption spectrum, recorded for Coel in benzene,[7] with the nonadiabatic vibronic LVC spectra, obtained at the B3LYP (left panels) or CAM-B3LYP (right panels) level, after turning on (red lines) or off (blue lines) the interstate couplings. All spectra were computed at $T = 0$ K, in a static approximation, $i.e.$ considering only the **2H** global-minimum structure, either in the gas phase (top panels) or in (PCM) benzene solution (bottom panels) including only the most important 45 normal modes. The vibronic transitions were convoluted with a Gaussian function with HWHM = 0.08 eV, normalizing the resulting spectral signal to match the maximum peak height intensity. To facilitate the comparison, gas and PCM spectra calculated with CAM-B3LYP have been red-shifted by 0.4 eV.

states, also displayed in Fig. 5 together with their sum. At the B3LYP level, the spectrum is predominantly shaped by large contributions arising from transitions to the S_1 and S_3 states, similarly to what obtained at the static approach in the gas-phase. These two transitions exhibit the largest intensities and are associated with the two main bands in the experimental spectrum. Interestingly, whereas both S_2 and S_4 states that are almost dark according to the static approach, they provide a substantial contribution to the spectrum at the CEA-VE level, indicating that internal motions trigger remarkable mixing of the electronic states (and therefore exchange of their oscillator strengths). Turning to CAM-B3LYP results, in agreement with previous observations, they systematically overestimate the position of CEA-VE spectra by ~ 0.45 eV. However, the most interesting finding in these computations is that the effect of the exchange of oscillator strength between the states is even larger than with B3LYP, leading to a three-peak shape of the spectrum as a result of the enhanced contribution of S_2, which features comparable intensity with the peaks associated with S_1 and S_3 states. M062X and ωB97-XD CEA-VE spectra computed on top of the same molecular structures, not reported for the sake of brevity, give similar results to CAM-B3LYP, predicting an even more intense third (central) peak due to S_2. Finally, the

CEA-VE spectra displayed in Fig. 5, obtained by convoluting vertical excitation energies computed taking into account first neighbor shells included at the QM/MM level, are compared in Fig. S13 of the ESI† with those obtained accounting for the solvent via the PCM model. The latter dieletric continuum model introduces mutual solute–solvent polarization effects. Similar to what observed for static vibronic spectra, the computed CEA-VE spectra in the PCM exhibit with both functionals a pronounced enhancement of the S_1 band. This leads to improved agreement with the experiment at the B3LYP level, while slightly worsening CAM-B3LYP performances. Nevertheless, the overall spectral shape remains largely similar with the predictions obtained using the explicit (QM/MM) approach. Interestingly, this comparison also suggests that the rise of the third peak is therefore more connected to internal motions rather than to solvent fluctuations. In summary, at the CAM-B3LYP level although the S_1 and S_3 transitions appear with comparable intensities, the additional contribution of the S_2 state results in a three-peak spectrum, leading to a poor agreement with the experimental spectrum.

The central and right panels of Fig. 5 present the Ad-MD|gVG spectra, respectively in benzene and in the gas phase. The comparison with CEA-VE spectra indicates that taking into
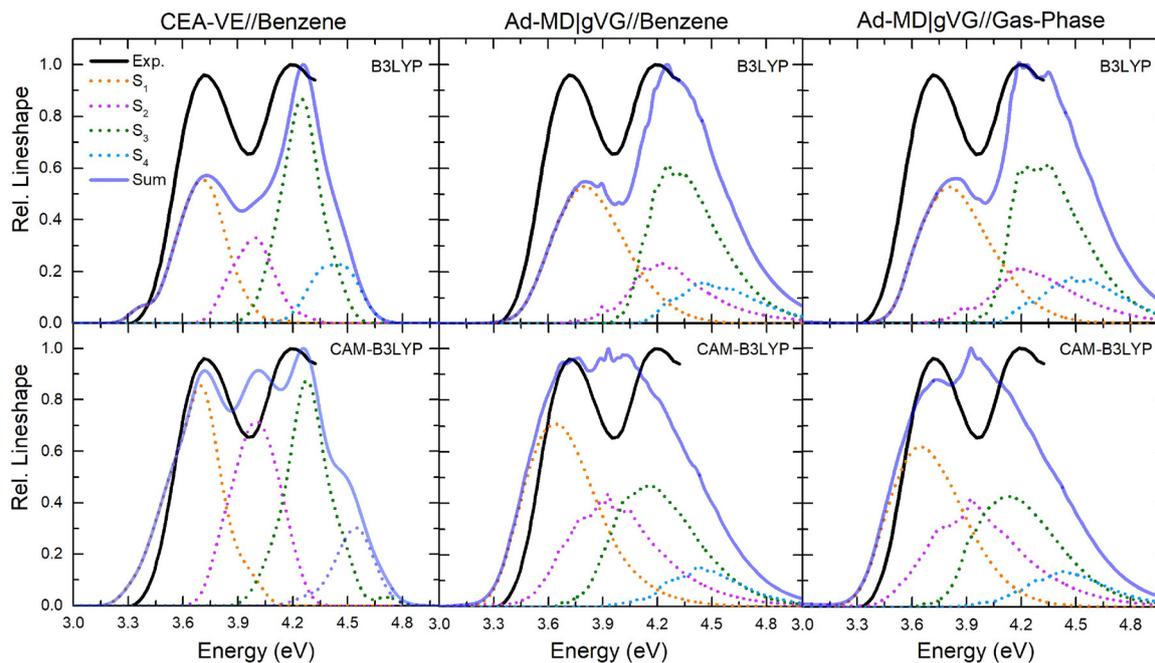
Fig. 5 Comparison between the Coel's experimental spectrum[7] with the MQC spectra computed for **2H** at either the B3LYP (top) or the CAM-B3LYP (bottom) level, using CEA-VE (left panels) and Ad-MD|gVG (middle and right panels) protocols. To isolate the effect of molecular flexibility from the embedding electrostatic effects, the Ad-MD|gVG spectra were computed both in solvent (middle panels), following the QM/MM scheme illustrated in Fig. 2, and for the isolated species in the gas phase (right panels). The contribution of each excited state (S_1 to S_4) is also shown with dashed lines. To facilitate comparison with the experimental spectrum, the theoretical CEA-VE and Ad-MD|gVG spectra calculated at the CAM-B3LYP level were red-shifted by 0.44 and 0.66 eV, respectively. All the computed spectra were normalized to match the maximum peak height intensity and all the transitions were convoluted applying a Gaussian function with a HWHM of 0.08 and 0.01 eV, respectively, for the CEA-VE and Ad-MD|gVG models.

account quantum vibronic effects leads to a noticeable broadening of the spectral lineshapes. It is important to remark that such result is achieved notwithstanding CEA-VE spectra obtained with a relatively large phenomenological broadening (0.08 eV), whereas a very small one is used within the Ad-MD|gVG scheme (the latter does not alter the overall width of the spectra, but is simply introduced to correct inaccuracies in the sampling due to the limited number of snapshots).[14] This increased bandwidth improves the agreement with experiment at the B3LYP level. However, while the peaks corresponding to the S_1 and S_3 states now appear with comparable intensities, the S_2 state is blue-shifted by 0.2 eV, overlapping with the peaks associated with S_3 and S_4, eventually giving in the total spectrum a larger intensity to the second peak with respect to the first one. The inclusion of vibronic effects at the CAM-B3LYP level instead results in a broad and almost structureless total spectrum. Therefore, although the Ad-MD|gVG approach captures the overall width of the experimental spectrum, it cannot provide a simulated spectral shape in good agreement with the experiment. A more detailed inspection of the contributions from individual excited states reveals a number of compensating effects. On the one hand, compared to the static vibronic spectra obtained with the same functional, all Ad-MD|gVG spectra show an increased intensity of the S_1 transition relative to S_3, but at least at the B3LYP level this increase corresponds to an improved agreement with experiment. On the other hand, transitions to S_2 and S_4, which are negligible at the FC|VG level,

gain intensity in Ad-MD|gVG, worsening agreement with experiment according to both functionals. To further decouple the contribution of the vibronic effects and those connected to the flexible degrees of freedom of the solute from those due to the embedding solvent, the Ad-MD|gVG spectra were also computed for the same Coel's molecular structures obtained by the MD in benzene, yet removing all solvent molecules from each considered snapshot. As could be expected considering the low polarity of benzene, the resulting spectra, displayed in the right-hand panels of Fig. 5, are remarkably similar to those obtained from the Ad-MD|gVG calculations in explicit QM/MM solvent, shown in the middle panels of the same figure. These findings suggest that polarization effects arising from solute–solvent interactions have only a marginal impact on the spectral lineshape, and the enhancement of the S_2 and S_4 contributions which deteriorates the agreement with experiment is essentially due to the intra-molecular motions.

A detailed analysis presented in Section S7 of the ESI† further investigates and disentangles the role of fast and slow modes in modulating the excited-state properties of Coel. In practice we analyzed the variations in excitation energy and oscillator strength of the four lowest singlet excited states (S_1–S_4) along MD-generated snapshots, using both B3LYP and CAM-B3LYP functionals. The results, presented in Fig. S14 and S15 of the ESI,† reveal that during the MD run the system explores regions of the configurational space where excited states may become quasi-degenerate and therefore strongly

mix. This effect is more pronounced in the CAM-B3LYP simulations, where S_2 displays non-negligible oscillator strength across a broader segment of the MD trajectory. Conversely, at the B3LYP level this state remains nearly dark over a larger portion of the sampled ensemble.

In order to try to unravel the separate contributions to the excited states coupling of the fast and slow degrees of freedom, we carried out several benchmarks. First, a search of crossing points within the LVC model allowed us to individuate molecular structures with two nearly degenerate bright excited states arising from a equally-weighted combination of H → L and H → L+1 transitions with ππ* character. This analysis confirms that even not considering molecular flexibility, fast vibrational modes can trigger effective mixing of the excited states. Second, as detailed in Section S10 of the ESI† (see Fig. S18–S23) we monitored the variation of the excitation energies and oscillator strengths of the four lowest excited states along relaxed scans of the flexible torsional modes (δ_1–δ_6) of the molecule. The results reveal that although state mixing can occur at specific distortions along these modes, the corresponding molecular structures, in which a single soft mode is strongly distorted whereas the other ones are not, are not populated during the MD simulation. Therefore, to investigate the impact of a more realistic simultaneous displacement of the soft modes and disentangle it from the effect of the high-frequency modes, we ran additional MD simulations in which all bond stretchings were selectively frozen. The distribution of vertical excitation energies and oscillator strengths for the four excited states along such a constrained MD trajectory shows that, freezing stretching modes partially attenuates state mixing, reducing the oscillator strengths of S_2 and S_4 states at both the B3LYP and CAM-B3LYP levels (see Fig. S16 and S17 in the ESI†). However, many configurations still exhibit significant mixing, indicating that overall even only slow nuclear motions can trigger effective inter-state mixing.

3.3.2 Including inter-state couplings: a generalized Ad-MD|gLVC approach. The results in the previous section show that CEA-VE and Ad-MD|gVG cannot provide a quantitative reproduction of the experimental spectral shapes. In contrast, they introduce artifacts not present in static vibronic spectra, at the CAM-B3LYP level, like the rise of a third peak or the appearance of a broad structureless band. Both features are due to the strong coupling and mixing of the states occurring along the MD trajectory, which makes adiabatic approaches not fully satisfactory. In this complex scenario, as an approximate solution, we apply here the generalized nonadiabatic Ad-MD|gLVC protocol described in Section 2.3.3.

The final Ad-MD|gLVC spectra were obtained from an average over the individual nonadiabatic vibronic spectra computed employing a HWHM of 0.04 eV along the usual statistical ensemble. It should be noticed that the larger value of the broadening with respect to the 0.01 eV adopted in Ad-MD|gVG was due to the lack of thermal effects and to the fact that, in order to save computational time, wave packets were propagated numerically only for 100 fs, a time-window too small to allow the computation of higher-resolution spectra. The resulting spectral signals are compared in Fig. 6 with the experimental, the adiabatic CEA-VE and Ad-MD|gVG and spectra (a recapitulatory picture comparing all the computed spectra shown here and in previous figures is reported in Fig. S25 in the ESI†). With respect to the latter schemes, only a moderate improvement appears at the B3LYP level. Differences are much more visible with CAM-B3LYP, where Ad-MD|gLVC clearly corrects the artifacts of Ad-MD|gVG and CEA-VE calculations. In fact, the Ad-MD|gLVC spectrum does not predict a third peak like CEA-VE and at the same time does not lead to a complete structureless band like Ad-MD|gVG. Conversely, the Ad-MD|gLVC spectrum is characterized by two well-resolved bands, in good agreement with experiment. Furthermore, although the second band is less intense than the first one,
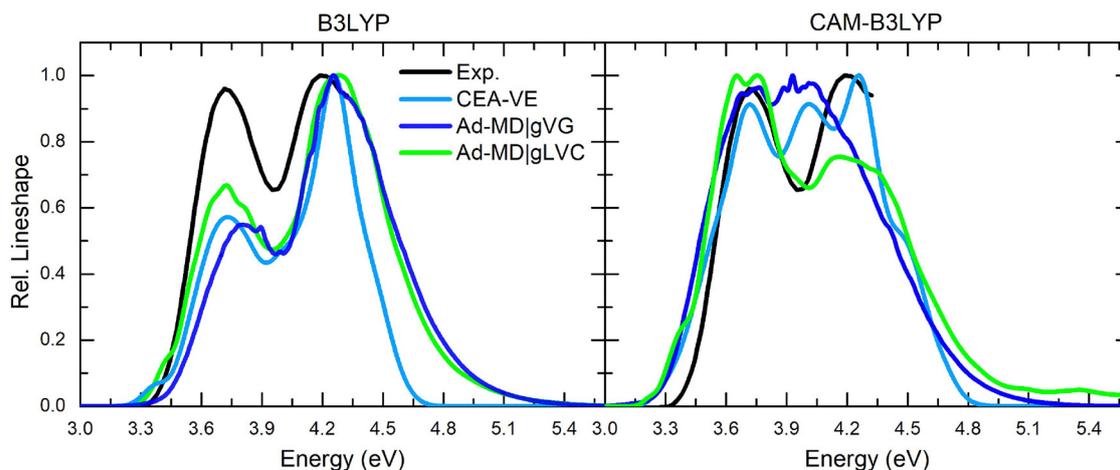

Fig. 6 Comparison of Coel's experimental absorption in benzene,[7] with the spectra simulated with the different MQC approaches considered in this work. All the computed spectra were normalized to match the maximum peak height intensity. The CEA-VE and Ad-MD|gVG and Ad-MD|gLVC spectra calculated at the CAM-B3LYP level have been red-shifted by 0.44 and 0.66 and 0.63 eV, respectively, to facilitate comparison with the experimental spectrum, and all the transitions were convoluted applying a Gaussian function with a HWHM of 0.08, 0.01 and 0.04 eV, respectively.

15956 | *Phys. Chem. Chem. Phys.*, 2025, **27**, 15946–15959

This journal is © the Owner Societies 2025

as also observed in the static LVC spectra in Fig. 4, the relative ratio is moderately improved.

The analysis of the Ad-MD|gLVC spectra obtained for the individual snapshots, displayed in Fig. S24 of the ESI,† provides further details. At the B3LYP level, for a substantial number of structures in the sample, the intensity computed for the first peak is actually comparable or even larger than that of the second peak. However, since the first peak is generally narrower than the second one, its intensity still appears underestimated in the averaged spectrum. In contrast, at the CAM-B3LYP level, most snapshots deliver a first peak with higher intensity compared to the second one, eventually yielding an underestimated intensity for the latter.

4 Conclusions

Coel is the chromophore of the photoprotein aequorin and its flexibility and intricate manifold of electronic states make it an interesting test case for modern computational methodologies for the simulation of electronic spectral shapes. In this contribution, we focused on Coel absorption spectroscopy in solution, namely benzene, a necessary step before investigating the interaction with the proteic scaffold. A standard investigation at a pure electronic level or with static vibronic approaches apparently provides a decent description of the spectrum. Our results allow us to confirm that the only molecular species contributing to the light absorption is the neutral form **2H**, whereas the anionic **2O⁻** form has spectral properties not compatible with experiment.[7] At this level of theory the two bands of the **2H** spectrum are attributed solely to the two bright $\pi\pi^*$ transitions, respectively with H → L and H → L+1 character. B3LYP and CAM-B3LYP reproduce the energy spacing reasonably well, though CAM-B3LYP systematically overestimates the position of both bands by approximately 0.4–0.5 eV. In contrast, B3LYP places them accurately but significantly underestimates the intensity of the low-energy band. Static non-adiabatic vibronic computations reveal that interstate couplings are active between the two bright states and with two nearby $n\pi^*$ states, so that all four states contribute to the spectra. However, non-adiabatic effects mainly result in a smoothing of vibronic progressions and have only a moderate impact on the relative intensities. A PCM description of solute/solvent mutual polarization improves the intensity ratio of the two bands for B3LYP but worsens it for CAM-B3LYP. Attempts to achieve a more accurate description of the band positions and intensities at the pure electronic level using RASPT2 and STEOM-DLPNO-CCSD proved very challenging and ultimately unsuccessful.

The true complexity of the Coel molecule is revealed when attempting to provide a description of the spectroscopic consequences of molecular flexibility—a key property that is particularly sensitive to interactions with the protein scaffold. Indeed, reliable MD sampling of the configurational space shows that distortions along both high-frequency vibrations and soft modes induce significant couplings between electronic states. The molecule frequently adopts geometries where the two $\pi\pi^*$ and two $n\pi^*$ states become nearly degenerate (*i.e.*, approach conical intersections), leading to strong mixing and redistribution of oscillator strength among them. In such a complex scenario, our results demonstrate that applying methods which neglect inter-state couplings—whether purely classical, like CEA-VE, or mixed quantum-classical, like Ad-MD|gVG—can result in significant artifacts in the spectral shapes, such as the appearance of a fictitious third band in the CAM-B3LYP spectrum. In our view, these issues stem fundamentally from the fact that state mixing is a non-adiabatic phenomenon, mediated by vibrational wavefunctions, and thus cannot be adequately captured by even a statistically sound ensemble of vertical excitation energies and oscillator strengths. Furthermore, adiabatic vibronic approaches like Ad-MD|gVG may face additional challenges, as they rely on computing excited-state energy gradients that, near conical intersections, may not yield realistic Taylor expansions of the local harmonic potential energy surfaces.

One of the main results of this work is indeed to highlight such an intrinsic limitation of current and widely used methodologies for simulating condensed-phase electronic spectroscopy. We also provide a partial solution, showing that, although approximate, the Ad-MD|gLVC approach corrects the artifacts introduced by the CEA-VE and Ad-MD|gVG adiabatic methods, restoring a more reasonable description of the spectral shape. We should note that the residual discrepancies with respect to the experiment may also depend on inaccuracies of the electronic structure methods, as suggested by the different predictions of the different levels of theories explored in this work.

Despite this partial success, we believe it is also important to critically discuss the potential limitations of the Ad-MD|gLVC approach, as these highlight the need for further methodological developments. We identify two main challenges that currently limit the general applicability of Ad-MD|gLVC: (i) the stiff/soft partitioning of vibrational modes can be problematic, and (ii) the parameterization of non-adiabatic LVC Hamiltonians at each snapshot is computationally demanding, with practical shortcuts not always available.

As far as the partition of modes, the example of Coel shows that running a full-coordinate MD and then projecting out the soft-modes when building vibronic Hamiltonians (the core of Ad-MD|g approaches) might not be the best solution, when it is necessary to disentangle the role of different soft and fast degrees of freedom in the coupling of electronic states. Alternative protocols, such as using Monte Carlo moves, may prove more versatile. As for the definition of LVC Hamiltonians, fully re-parameterizing them at each MD snapshot provides a robust but computationally too expensive solution. On the other hand, any attempt to transfer LVC parameters between snapshots implicitly assumes that the nature of the diabatic states remains consistent across different soft-mode configurations—a requirement that is difficult to guarantee in general cases.

In principle, a possible radical solution is to abandon MQC approaches, adopting fully quantum dynamical descriptions of

This journal is © the Owner Societies 2025

Phys. Chem. Chem. Phys., 2025, **27**, 15946–15959 | **15957**

the systems like those guaranteed by ensemble of interacting Gaussian wavepackets,[40–43] pionereed by Heller.[44] Variational equations of motions derived from quantum mechanics are available[42,43] and implemented in distributed codes.[28] Their application to the spectroscopy of middle-size chromophores (100 vibrational modes) in the condensed phase in the framework of QM/MM methods is however not mature yet.

Conflicts of interest

There are no conflicts to declare.

Data availability

Further data for this article including: (i) molecular dynamics (MD) input files (equilibrated configuration, topologies based on quantum-mechanics-derived force fields, and production-ready input files); (ii) optimized DFT structures used for the static analysis of colenteramide; and (iii) LVC input files are available at the Zenodo database under the accession code 15341644.[38]

Acknowledgements

References

1 O. Shimomura and F. H. Johnson, Chemical nature of bioluminescence systems in coelenterates, *Proc. Natl. Acad. Sci. U. S. A.*, 1975, **72**, 1546–1549.

2 E. Vysotski, V. Bondar and V. Letunov, Extraction and purification of obelin, the Ca²⁺-dependent photoprotein from the hydroid Obelia longissima, *Biokhimiya*, 1989, **54**, 965–973.

3 E. S. Vysotski and J. Lee, Ca²⁺-regulated photoproteins: structural insight into the bioluminescence mechanism, *Acc. Chem. Res.*, 2004, **37**, 405–415.

4 Z. J. Liu, G. A. Stepanyuk, E. S. Vysotski, J. Lee, S. V. Markova, N. P. Malikova and B.-C. Wang, Crystal structure of obelin after Ca²⁺-triggered bioluminescence suggests neutral coelenteramide as the primary excited state, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 2570–2575.

5 S. F. Chen, N. Ferré and Y. J. Liu, QM/MM study on the light emitters of aequorin chemiluminescence, bioluminescence, and fluorescence: a general understanding of the bioluminescence of several marine organisms, *Chem. – Eur. J.*, 2013, **19**, 8466–8472.

6 Z. S. Li, X. Zhao, L. Y. Zou and A. M. Ren, The dynamics simulation and quantum calculation investigation about luminescence mechanism of coelenteramide, *Photochem. Photobiol.*, 2013, **89**, 849–855.

7 Y. Imai, T. Shibata, S. Maki, H. Niwa, M. Ohashi and T. Hirano, Fluorescence properties of phenolate anions of coelenteramide analogues: the light-emitter structure in aequorin bioluminescence, *J. Photochem. Photobiol., A*, 2001, **146**, 95–107.

8 Y. Min, T. Y. Kim, Y. M. Kim, J. Kim, Y. I. Rhee and C. H. Choi, Computational Photochemistry for the Design of Bioluminescent Systems: Theoretical Insights into the Mechanism of Light Emission from Coelenteramide, *ChemPhysChem*, 2016, **17**, 2128–2135.

9 O. Shimomura and K. Teranishi, Light-emitters involved in the luminescence of coelenterazine, *Luminescence*, 2000, **15**, 51–58.

10 D. Roca-Sanjuán, M. G. Delcey, I. Navizet, N. Ferre, Y. J. Liu and R. Lindh, Chemiluminescence and fluorescence states of a small model for coelenteramide and cypridina oxyluciferin: a CASSCF/CASPT2 study, *J. Chem. Theory Comput.*, 2011, **7**, 4060–4069.

11 R. R. Alieva, F. N. Tomilin, A. A. Kuzubov, S. G. Ovchinnikov and N. S. Kudryasheva, Ultraviolet fluorescence of coelenteramide and coelenteramide-containing fluorescent proteins. Experimental and theoretical study, *J. Photochem. Photobiol., B*, 2016, **162**, 318–323.

12 F. Segatta, L. Cupellini, M. Garavelli and B. Mennucci, Quantum chemical modeling of the photoinduced activity of multichromophoric biosystems: focus review, *Chem. Rev.*, 2019, **119**, 9361–9380.

13 T. J. Zuehlsdorff, S. V. Shedge, S.-Y. Lu, H. Hong, V. P. Aguirre, L. Shi and C. M. Isborn, Vibronic and environmental effects in simulations of optical spectroscopy, *Annu. Rev. Phys. Chem.*, 2021, **72**, 165–188.

14 J. Cerezo, D. Aranda, F. J. Avila Ferrer, G. Prampolini and F. Santoro, Adiabatic-molecular dynamics generalized vertical Hessian approach: a mixed quantum classical method to compute electronic spectra of flexible molecules in the condensed phase, *J. Chem. Theory Comput.*, 2019, **16**, 1215–1231.

15 A. Segalina, D. Aranda, J. A. Green, V. Cristino, S. Caramori, G. Prampolini, M. Pastore and F. Santoro, How the Interplay among Conformational Disorder, Solvation, Local, and Charge-Transfer Excitations Affects the Absorption Spectrum and Photoinduced Dynamics of Perylene Diimide Dimers: A Molecular Dynamics/Quantum Vibronic Approach, *J. Chem. Theory Comput.*, 2022, **18**, 3718–3736.

16 J. Cerezo, J. Gierschner, F. Santoro and G. Prampolini, Explicit Modelling of Spectral Bandshapes by a Mixed

15958 | *Phys. Chem. Chem. Phys.*, 2025, **27**, 15946–15959

This journal is © the Owner Societies 2025

Quantum-Classical Approach: Solvent Order and Temperature Effects in the Optical Spectra of Distyrylbenzene, *ChemPhysChem*, 2024, **25**, e202400307.

17 H. Moumene, G. Prampolini, C. García-Iriepa, J. Cerezo, I. Navizet and F. Santoro, Deciphering the Luminescence Spectral Shape of an Oxyluciferin Analogue through a Mixed Quantum-Classical Approach, *J. Phys. Chem. B*, 2025, **129**, 2829–2844. PMID: 40063837.

18 S. Giannini, P. M. Martinez, A. Semmeq, J. P. Galvez, A. Piras, A. Landi, D. Padula, J. G. Vilhena, J. Cerezo and G. Prampolini, JOYCE3.0: A General Protocol for the Specific Parametrization of Accurate Intramolecular Quantum Mechanically Derived Force Fields, *J. Chem. Theory Comput.*, 2025, 1–56.

19 M. J. Frisch, *et al.*, *Gaussian ∼ 16 Revision D.02*, Gaussian Inc., Wallingford CT, 2016.

20 I. Cacelli and G. Prampolini, Parametrization and validation of intramolecular force fields derived from DFT calculations, *J. Chem. Theory Comput.*, 2007, **3**, 1803–1817.

21 F. Santoro and J. Cerezo, FCclasses3, a code for vibronic calculations. Available from **https://www.iccom.cnr.it/en/fcclasses**, last accessed on 20 June 2025.

22 J. Cerezo and F. Santoro, FCclasses3: Vibrationally-resolved spectra simulated at the edge of the harmonic approximation, *J. Comput. Chem.*, 2023, **44**, 626–643.

23 F. Santoro and J. A. Green, *Overdia 01, a Fortran 90 code for parametrization of model Hamiltonians based on a maximum-overlap diabatisation*. Available at **https://www.iccom.cnr.it/en/overdia-en**; 2022, last accessed on 20 June 2025.

24 J. A. Green, M. Yaghoubi Jouybari, H. Asha, F. Santoro and R. Improta, Fragment diabatization linear vibronic coupling Model for quantum dynamics of multichromophoric systems: population of the charge-transfer state in the photoexcited guanine-cytosine pair, *J. Chem. Theory Comput.*, 2021, **17**, 4660–4674.

25 H. Wang, X. Liu and J. Liu, Accurate calculation of equilibrium reduced density matrix for the system-bath model: A multilayer multiconfiguration time-dependent Hartree approach and its comparison to a multi-electronic-state path integral molecular dynamics approach, *Chin. J. Chem. Phys.*, 2018, **31**, 446–456.

26 H. Wang, Multilayer Multiconfiguration Time-Dependent Hartree Theory, *J. Phys. Chem. A*, 2015, **119**, 7951–7965. PMID: 26020459.

27 O. Vendrell and H. D. Meyer, Multilayer multiconfiguration time-dependent Hartree method: Implementation and applications to a Henon-Heiles Hamiltonian and to pyrazine, *J. Chem. Phys.*, 2011, **134**, 044135.

28 G. Worth, Quantics: A general purpose package for Quantum molecular dynamics simulations, *Comput. Phys. Commun.*, 2020, **248**, 107040.

29 D. Aranda and F. Santoro, Vibronic Spectra of π-Conjugated Systems with a Multitude of Coupled States: A Protocol Based on Linear Vibronic Coupling Models and Quantum Dynamics Tested on Hexahelicene, *J. Chem. Theory Comput.*, 2021, **17**, 1691–1700.

30 J. Cerezo, G. Prampolini and I. Cacelli, Developing accurate intramolecular force fields for conjugated systems through explicit coupling terms, *Theor. Chem. Acc.*, 2018, **137**, 1–15.

31 G. Prampolini, J. Cerezo, S. Giannini, N. De Mitri and I. Cacelli, Joyce3.0, intra-molecular force field parameterization software, available free of charge at **https://www.iccom.cnr.it/en/joyce-2/**, last consulted on 20 June 2025, 2024.

32 P. M. Martinez, A. Piras, S. Giannini, A. Semmeq, J. P. Galvez, D. Padula, J. Cerezo, J. G. Vilhena and G. Prampolini, Joyce website: **https://joyce-documentation.gitlab.io/**, last consulted December, 2024.

33 J. G. Vilhena, L. Greff da Silveira, P. R. Livotto, I. Cacelli and G. Prampolini, Automated Parameterization of Quantum Mechanically Derived Force Fields for Soft Materials and Complex Fluids: Development and Validation, *J. Chem. Theory Comput.*, 2021, **17**, 4449–4464.

34 A. W. Sousa da Silva and W. F. Vranken, ACPYPE – AnteChamber PYthon Parser interfacE, *BMC Res. Notes*, 2012, **5**, 367.

35 C. I. Bayly, P. Cieplak, W. Cornell and P. A. Kollman, A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model, *J. Phys. Chem.*, 1993, **97**, 10269–10280.

36 J. Tomasi, B. Mennucci and R. Cammi, Quantum Mechanical Continuum Solvation Models, *Chem. Rev.*, 2005, **105**, 4708.

37 L. S. Dodda, I. Cabeza de Vaca, J. Tirado-Rives and W. L. Jorgensen, LigParGen web server: an automatic OPLS-AA parameter generator for organic ligands, *Nucleic Acids Res.*, 2017, **45**, W331–W336.

38 DOI: **10.5281/zenodo.15341644**.

39 M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindahl, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, *SoftwareX*, 2015, **1–2**, 19–25.

40 B. F. E. Curchod and T. J. Martínez, Ab Initio Nonadiabatic Quantum Molecular Dynamics, *Chem. Rev.*, 2018, **118**, 3305–3336.

41 M. Ben-Nun, J. Quenneville and T. J. Martínez, Ab Initio Multiple Spawning: Photochemistry from First Principles Quantum Molecular Dynamics, *J. Phys. Chem. A*, 2000, **104**, 5161–5175.

42 G. A. Worth, M. A. Robb and I. Burghardt, A novel algorithm for non-adiabatic direct dynamics using variational Gaussian wavepackets, *Faraday Discuss.*, 2004, **127**, 307–323.

43 G. W. Richings, I. Polyak, K. E. Spinlove, G. A. Worth, I. Burghardt and B. Lasorne, Quantum dynamics simulations using Gaussian wavepackets: the vMCG method, *Int. Rev. Phys. Chem.*, 2015, **34**, 269–308.

44 E. J. Heller, Time-dependent approach to semiclassical dynamics, *J. Chem. Phys.*, 1975, **62**, 1544–1555.

This journal is © the Owner Societies 2025

Phys. Chem. Chem. Phys., 2025, **27**, 15946–15959 | **15959**