Journal Name

ARTICLE TYPE

Cite this: DOI: 00.0000/xxxxxxxxxx

Unraveling the contributions to the spectral shape of flexible dyes in solution: insights on the absorption spectrum of an oxyluciferin analogue.†

Javier Cerezo,*[a,b] Cristina García-Iriepa,*[c,d] Fabrizio Santoro,[b] Isabelle Navizet,[e] and Giacomo Prampolini*[b]

Received Date
Accepted Date

DOI: 00.0000/xxxxxxxxxx

We present a computational investigation of the absorption spectrum in water of 5,5-spirocyclopropyl-oxyluciferin (5,5-CprOxyLH), an analogue of the emitter compound responsible for the bioluminescence in fireflies. Several factors concur to 5,5-CprOxyLH's spectral shape: i) the contribution of the four close-energy excited states, which show significant non-adiabatic couplings, ii) the flexible molecular structure and iii) the specific interactions established with the surrounding protic solvent, which strongly couple the solvent dynamics with the dye's spectral response. To tackle the challenge to capture and dissect the role of all these effects we preliminarly investigate the role of non-adiabatic couplings with quantum dynamics simulations and a Linear Vibronic Coupling model in gas phase. Afterwards we account for both molecular flexibility and solvent interactions by resorting to a mixed quantum classical protocol, named Adiabatic Molecular Dynamics generalized Vertical Gradient (Ad-MD|gVG), which is built on a method recently proposed by some of us. It is rooted in the partition between stiff degrees of freedom of the dye, accounted for at vibronic level within the harmonic approximation, and flexible degrees of freedom of the solute (and of the solvent, if present), described classically through a sampling based on Molecular Dynamics (MD). Ad-MD|gVG avoids spurious effects arising in the excited state Hessians due to non-adiabatic couplings, and can be therefore applied to account for the contributions of the first four excited states to 5,5-CprOxyLH absorption spectrum. The final simulated spectrum is in very good agreement with the experiment, especially when the MD is driven by a refined quantum-mechanically derived force-field. More importantly, the origin of each separate contribution to the spectral shape is properly accounted for, paving the way to future applications of the method to more complex systems or alternative spectroscopies, as emission or circular dichroism.

[a] *Departamento de Química and Institute for Advanced Research in Chemical Sciences (IAdChem), Universidad Autónoma de Madrid, 28049 Madrid, Spain. E-mail: javier.cerezo@uam.es*

[b] *CNR–Consiglio Nazionale delle Ricerche, Istituto di Chimica dei Composti Organo Metallici (ICCOM-CNR), SS di Pisa, Area della Ricerca, via G. Moruzzi 1, I-56124 Pisa, Italy. E-mail: giacomo.prampolini@pi.iccom.cnr.it*

[c] *Universidad de Alcalá, Departamento de Química Analítica, Química Física e Ingeniería Química, Grupo de Reactividad y Estructura Molecular (RESMOL), 28806 Alcalá de Henares (Madrid), Spain. E-mail: cristina.garciai@uah.es*

[d] *Universidad de Alcalá, Instituto de Investigación Química "Andrés M. del Río" (IQAR), 28806 Alcalá de Henares (Madrid), Spain*

[e] *Univ Gustave Eiffel, Univ Paris Est Creteil, CNRS, UMR 8208, MSME, F-77454 Marne-la-Vallée, France*

† Electronic Supplementary Information (ESI) available: Details on LVC theory, QMD-FF parameterization, 5,5-CprOxyLH QMD-FF parameters, MD simulations and characterization of the first four excited states. See DOI: 00.0000/00000000.

1 Introduction

The analysis of the absorption and emission spectra of a given compound is a essential first step to characterise and understand its photophysical and photochemical properties. In this regard, the synergy between experiments and calculations is essential for building a complete image of the system. The simulation of the spectra of a target chromophore embedded in a specific environment remains indeed a quite challenging task, because of the large number of variables and factors governing it. In particular, the shape and width of electronic spectra arise from the interplay between molecular motion and solvent fluctuations, which are coupled with the electronic transitions of the specific dye. They thus encode detailed information about the system that can be only unveiled with computational simulations, using models that account for the main physical effects. The characteristics of the system under study mainly determines which computational approach would be the most suitable. For instance, semi-rigid molecules in homogeneous environments, where no specific and local interactions are expected to settle between the solvent and the solute, can be treated at a nearly full quantum mechanical (QM) level. In fact, the contribution of molecular vibration can be included at quantum, vibronic level, exploiting the possibility to rely on the harmonic approximation, and on the recent developments of effective implementations with either a Time Independent (TI) sum-over-states approach[1–3] or a Time Dependent (TD) one, based on the the Fourier transform of a correlation function[4–8]. Moreover, solvation effects can be conveniently accounted for through continuum models as PCM[9], which can be used also to include solvent broadening effects,[10] following an approach based on Marcus's theory[11]. Nonetheless, when either i) large-amplitude displacements of the flexible degrees of freedom (DoF, such as dihedral rotations around σ-bonds) are expected for the solute or ii) specific interactions with the solvent (e.g. hydrogen bonds (HBs), π-stacking, etc.) or with the complex embedding (polymers matrices, metal-organic frameworks, biomolecules, etc.) are taken into account, a first principle simulation of the absorption spectrum becomes more challenging and the development of new methods and protocols is necessary.

The quest for reliable yet effective models to simulate the shape of electronic spectra in solution has resulted in many contributions from different research groups in the recent years. Such mixed quantum-classical (MQC) strategies are generally based on MD samplings, followed by effective protocols to introduce vibronic effects, including temperature, and the environment[12–16]. Alternative formulations, based on a cumulant expansion of the energy gap fluctuations or spectral densities, have been also adopted to simulate both linear[17,18] and non-linear[19,20] spectroscopies. The successful application of these strategies mainly rely on two factors. On the one hand, the quality of the MD sampling is determined by the force-field (FF), and specifically tailored FFs parameterised against QM data have been exploited along with classical treatment of nuclear degrees of freedom[21,22]. On the other hand, solute solvent interaction has benefited from the development of efficient solvent models[23] and QM/MM strategies able to capture electrostatic and non-electrostatic effects[24] accurately.

In this context, some of us have recently introduced a MQC methodology, named Ad-MD|gVH (Adiabatic Molecular Dynamics generalised Vertical Hessian)[25], which aims at simulating absorption and emission lineshapes for flexible molecules in complex environments. In short, flexible modes and the fluctuations of the environment are treated classically, through reliable quantum mechanically derived force-fields (QMD-FF) and molecular dynamics (MD) sampling, whereas the stiff modes of the dye are treated at the QM vibronic level, using harmonic expansions of the potential energy surface (PES), specific for each instantaneous arrangement of the DoFs, which also account for Duschinsky mixings. The method has shown promising results, being able to provide spectral shapes for flexible molecules in different solvents in very good agreement with the experiment[25,26]. Yet, in its original implementation, Ad-MD|gVH relies on the Born-Oppenheimer (BO) approximation, so it faces some issues when non-adiabatic couplings between electronic states are significant.

A proper treatment of the impact of non-adiabatic interactions on the spectra can be achieved by resorting to quantum dynamics (QD) wavepacket propagation over the coupled excited states[27]. Despite the impressive advances in this field during the past years, three main issues hamper the straightforward application of these approaches for our scopes, namely the usual restriction to harmonic surfaces, the lack of thermal effects, and the difficulties to account for the coupling with the environment. A recent strategy to account for the environment and, to some extent, anharmonicities, proposed by Zuehlsdorff and coworkers[28], involves the use of LVC Hamiltonians parameterised from energy gap fluctuations, through spectral densities extracted from classical MD simulations. Moreover, the adoption of the thermalised time-evolving density operator with orthogonal polynomials algorithm (T-TEDOPA)[29], allows QD propagation including temperature. While promising, it is unclear how such strategy would effectively account for static disorder and dynamical effects of the solvent, which may be relevant when specific interactions are established[30,31]. Another alternative approach able to effectively include such effects, also accounting for temperature and interaction with the environment, is based on the same quantum-classical framework invoked in Ad-MD|gVH. Namely, it is achieved by substituting the harmonic PES of the model (gVH) with a Linear vibronic coupling (LVC) Hamiltonian for coupled states, obtaining the so-called Ad-MD|gLVC[32].

Application of these methods is computationally cumbersome, so that simplifications are often looked for and, when possible, nonadiabatic couplings are neglected. However, the application of methods that account for Duschinsky mixings like Ad-MD|gVH to systems where non-adiabatic couplings are significant may be impractical and lead to artefacts. In fact, inter-state couplings often perturb significantly specific regions of the PES, severely modifying selected vibrational frequencies, hence undermining the definition of a meaningful harmonic surface for the stiff vibrations.

In this work, we aim at simulating the absorption spectrum in water of 5,5-spirocyclopropyl-oxyluciferin (5,5-CprOxyLH), an analogue of the emitter compound responsible of the biolumines-

cence in fireflies, commonly known as oxyluciferin[33,34]. The experimental spectrum of the natural form of oxyluciferin is broad, but far from a simple bell shape, being strongly asymmetric and featuring a clear shoulder on the high-energy side of the maximum. These characteristics may be due to the interplay between vibronic progressions along high-frequency modes and the intrinsic flexibility of the chromophore, but also to the strong interactions with the surrounding water solvent[35–37].

The scenario is however further complicated by the fact that shape and width of the spectrum could be originated by several optically bright excited states, which lye close in energy and can potentially be coupled. As displayed in Figure 1, oxyluciferin can coexist in six different chemical forms[35,36,38–41], which may be populated through keto-enol tautomerisation equilibrium or protonation/deprotonation reactions[40–43]. It is therefore evident that it is impossible to ensure experimentally that the recorded spectra belong to exclusively one specific form. For this reason, a synthetic derivative designed to block the keto-enol tautomerisation has been proposed in the recent past, and its photophysical properties characterised (i.e. absorption and emission spectra), mimicking the ones of the natural phenol-keto form[35]. Hence, for experimental-computational comparison purposes, in this study we simulate the complex absorption spectrum of such oxyluciferin synthetic derivative, namely 5,5-CprOxyLH (depicted in Figure 1). Besides its biological and technological relevance[44–47], this system was chosen to evaluate the performance of MQC approaches in a case where several low lying close electronic excited states are responsible for the observed spectral signature. The aim is to provide a simulation where the main contributions to the spectral shape are analysed from first principles hence, allowing us to unravel the role of each effect concurring to the broadening.

2 Methods

In this section, we briefly review the different strategies adopted in this work to simulate the absorption spectrum accounting for the effect of the chromophore vibrations and of the surrounding medium, the main players in homogeneous and inhomogeneous broadening, respectively.

2.1 Full Quantum Mechanical (QM) approaches

In principle, the most accurate way to describe the role of the solute nuclear motion is retrieving the spectral shape from the sum of transitions between vibronic states $|e;\mathbf{v}_i\rangle$. Within the BO approximation, the electronic and vibrational components decouple, $|e;\mathbf{v}_i\rangle = |e\rangle|\mathbf{v}_i\rangle$, and the calculation of the spectrum can be carried out taking into account the vibrational states at the initial and final (adiabatic) electronic states. When the adiabatic PESs are described at the harmonic approximation, analytical expressions for both TI and TD formulations of the vibronic spectrum have been derived, and are implemented in our code $\mathscr{FCclasses3}$[48,49]. TD approaches are especially efficient in getting fully converged spectra at arbitrary temperatures, hence we adopt such formulation in all calculations conducted in this work. The harmonic model implies a quadratic expansion of the initial and final PESs, which is performed at the minimum for the initial state (the so-called Franck-Condon point), and either at the minimum (Adiabatic models) or the Franck-Condon geometry (Vertical models) for the final state. In this work, we resort to vertical models. Moreover, for reasons that will become clear in the following, we will make the further assumption that the Hessians of the excited states coincide with the one of the ground state, adopting the so-called Vertical Gradient (VG) model. Briefly, this choice is a convenient way to avoid the artefact due to the presence of spurious frequencies due to non-adiabatic couplings among excited electronic states, and it allows us to skip the time-consuming frequency calculations in the final states. For specific cases, we explicitly addressed the role of non-adiabatic couplings among the final electronic states. For such calculations, analytical expressions cannot be derived, and an effective way to compute the spectrum is adopting a TD formulation (at 0 K) by numerical propagation of the ground state wavepacket over the coupled (diabatic) excited states[27] employing Quantum Dynamics (QD). Such QD simulations are here carried out with the Multi-Layer version of the Multiconfigurational Time-Dependent Hartree (ML-MCTDH) method as implemented in the QUANTICS[50] code. The coupled electronic states are described in the diabatic representation, with quadratic expansions of the diabatic potentials, adopting a LVC model, which extends the VG model by adding linear potential couplings among the states[51].

The main drawback of the full QM approaches discussed above stands in their computational cost, which restricts their application to a limited number of DoFs. Consequently, they have been mostly applied only considering isolated chromophores, making them especially suited for gas phase calculations. To account for the role of the environment embedding the dye, additional approximations must be invoked. In the case of spectra in simple solution, the average effect of the surrounding solvent can be recovered by applying a broadening function. For non-protic solvents, according to Marcus[11], the Gaussian inhomogeneous broadening is connected to the solvent reorganisation energy, and it has been shown that the latter one can be estimated with State-specific PCM calculations[10]. Nonetheless, when the solvent establishes strong and specific interactions, as for instance HBs, or when the solvent is embedded in more complex environments (e.g. a protein or a polymer matrix), the linewidth of the broadening function is often determined phenomenologically, i.e. fitted to match the experimental shape. In these latter cases, a detailed rationale of the mechanisms leading to the inhomogenueos contribution to the spectral broadening can still be derived by a first principles approach, yet resorting to explicit representations of the environment and its fluctuations[52]. In the following, we therefore describe two different approaches, designed to explicitly account for the embedding environment and include its effect on the spectral shapes.

2.2 Classical approach

A straightforward way to explicitly account for the effect of the solute and environment fluctuations is obtaining the spectrum as the distribution of vertical energies computed over a reliable ensemble of frames, which include the coordinates of both the sol-

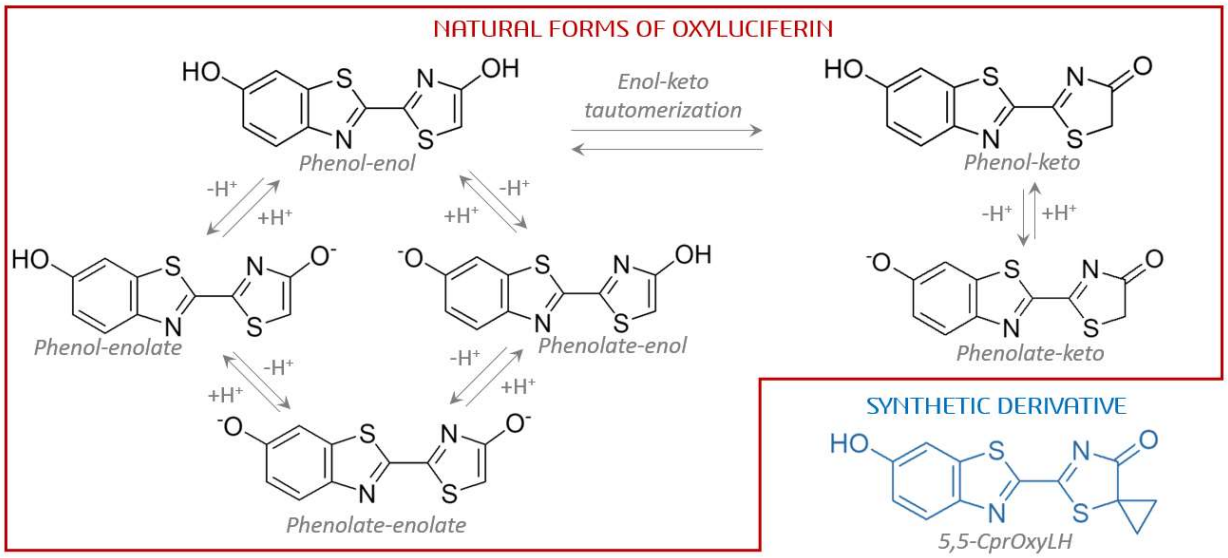

Fig. 1 Chemical structure of the six possible forms of oxyluciferin, depicting the equilibria connecting them, and of the synthetic derivative under study.

vent and the surrounding solvent molecules. Such solute+solvent configurations can be for instance extracted from a reliable classical molecular dynamics (MD) trajectory. The vertical energy can be conveniently computed by adopting a QM/MM method: such a protocol, which relies on the classical FC principle, is here referred to as Classical Ensemble Average of Vertical Energies (CEA-VE). Notwithstanding the ease of its implementation and the computational convenience, CEA-VE protocols have three main limitations, which may undermine the accuracy of the resulting spectra:

i) the reliability of the frame collections sampled from the MD trajectory heavily depends on the adopted force-field description, and an accurate refinement is often required [21,22,53].

ii) in MD runs, the nuclear dynamics is propagated at a classical level, implying a wrong sampling of the high-frequency modes, which in turn is reflected in an underestimation of the spectral width [53].

iii) vibronic progressions, which only arise in a quantum mechanical treatment, are lost.

Within CEA-VE approaches, the second of the above drawbacks is often circumvented by including a large enough phenomenological broadening to the distribution, yet we still need to fit a parameter to reproduce the experimental spectrum, which means that some relevant effects, namely those due to the quantum nature of the high-frequency molecular vibrations, are not accounted for in our calculation, and are re-introduced in a phenomenological way. Conversely, it might be worth noticing that a benefit of CEA-VE approach consists in being able to account also for flexible systems, whereas, as discussed above, the aforementioned QM vibronic approaches are confined to describe semi-rigid systems.

2.3 Mixed Quantum Classical approach

With the aim of including the contribution of the embedding and of the flexible DoFs of the solute, while not renouncing to account for important vibronic effects, electronic spectra are further simulated in this work, adopting a Mixed Quantum Classical (MQC) strategy recently introduced by some of us and named Ad-MD|gVH [25]. The protocol is based on partitioning the coordinates of the total system (solute+medium) in soft and stiff modes: the first block includes the flexible DoFs of the solute and solvent modes and are accounted for through a classical MD sampling, whereas the role of stiff modes is described at vibronic level with harmonic potentials specific for each configuration of the soft modes [25]. The final Ad-MD|gVH spectrum is obtained from the averaged signals computed at the QM level for the stiff coordinates over the configurations collected along the MD trajectory, which include explicitly the solvent, equilibrated at given thermodynamic conditions (here, room temperature and 1 atm). It is important to remark that such a strategy involves a significant

step forward compared to CEA-VE, as the present MQC approach correctly considers the homogeneous broadening and the spectral structure induced by the progressions over high-frequency modes. In this way, the Ad-MD|gVH method can accurately account for all the primary sources of broadening from first principles, thus avoiding using empirical parameters to reproduce the observed spectral shapes.

3 Computational details

3.1 Quantum Dynamics simulations

The possible impact of non-adiabatic couplings on the spectra of 5,5-CprOxyLH was evaluated with QD propagations on the grounds of a LVC model (whose mathematical form is given in the Electronic supplementary Information, ESI) comprising the four lowest electronic states. The LVC Hamiltonian was parameterised from TD-DFT calculations through a maximum-overlap diabatization technique implemented in our code Overdia, that is freely distributed[54]. In practice, first, the molecular structure is optimised at the ground state, and its normal modes and frequencies are computed. The gradients over the diabatic surfaces and the linear inter-state couplings are computed by numerical differentiation, taking backward and forward displacement over each normal mode, with a step in dimensionless shift of $\Delta q_i = \pm 0.02$. At each displaced geometry, the diabatization requires the computation of the overlap of the adiabatic states with respect to those at the ground-state geometry taken as reference. Such overlap is obtained as the scalar product of TD-DFT transition densities[55]. QD propagation was carried out accounting for all internal degrees of freedom, adopting the efficient ML-MCTDH method as implemented in QUANTICS[50], which was also employed to compute the correlation functions required for the non-adiabatic spectra.

3.2 Classical conformational sampling

Three different steps are employed to retrieve a reliable sampling: i) the parameterisation of a specific QMD-FF[56,57], ii) running classical MD trajectories of the solvated system (composed by the chromophore and $>$ 1000 water molecules), using the QMD-FF for the solute in its initial electronic state (ground state, GS, for absorption), and iii) extracting from the MD equilibrated trajectories a set of uncorrelated snapshots.

As far as the QMD-FF is concerned, it was explicitly parameterised for 5,5-CprOxyLH as described in detail in the ESI. Concretely, the intramolecular term (S9), ruling the dye's flexibility, was parameterised with the JOYCE code[58], following the JOYCE standard procedure[56,57,59], based on the DFT optimised geometry, Hessian matrix and relaxed torsional energy profiles along the most flexible coordinates. Furthermore, the interaction of the dye with the (water) solvent was also specifically refined against the same QM description. Concretely, the point charges defining the intermolecular FF term in Eq. (S14) are fitted with the RESP procedure over the 5,5-CprOxyLH electronic density computed for the isolated molecule with the B3LYP functional and accounting for the solvent at C-PCM level[60]. The choice of this functional is in line with previous computational studies on the optical properties of oxyluciferin analogues[36,37,61–64]. Further details can be found in ESI, together with a complete list of all QMD-FF parameters. All QM calculations either employed in the QMD-FF parametrization or in the absorption spectra simulation (*vide infra*) were performed with the GAUSSIAN16 suite of programs[65] resorting to density functional theory (DFT) or its time-dependent implementation (TD-DFT). Unless otherwise stated, in all calculations, the B3LYP functional was used with the 6-311G(2d,p) basis set. This basis set has shown to perform well for this system[37], and we verified that addition of diffuse functions does not appreciably change the results.

MD simulations were carried out with the GROMACS code[66] employing the parameterised QMD-FF, either on the isolated dye (gas phase) or on a system consisting of one dye (5,5-CprOxyLH) and $\sim$ 1200 explicit water molecules. Simulations were performed in the NVT ensemble at 298 K for the isolated molecule in the gas phase and at constant pressure and temperature (NPT, 1 atm and 298 K) in solution. In the first case, 100 dye conformations were stored by sampling the MD trajectories every 50 ps, while the same number of configurations were extracted from the solvated systems, but sampling every 100 ps in the longer run. Further computational details on MD simulations can be found in the ESI.

3.3 Spectra calculation over conformational ensembles

At each extracted snapshot, the vertical energy and the gradient/Hessian at the ground state and the gradient of all the selected excited states are computed, adopting the ONIOM QM/MM approach[67] in which the 5,5-CprOxyLH molecule is included in the QM region, treated at DFT/TD-DFT with B3LYP/6-311g(2d,p), while the solvent molecules around 15 Å from the dye are treated at MM level, using the same FF adopted in the MD simulation. The resulting data are used for both the CEA-VE (vertical energies only) and Ad-MD|gVG strategies. With the CEA-VE scheme, the spectra are computed from all stick transitions, including the four lowest excited states ($i = 1-4$), along the extracted snapshots, weighted by the squared transition dipole moment, $|\mu_{gi}|^2$. Each transition is broadened by convolution with a Gaussian lineshape.

According to the Ad-MD|gVG and Ad-MD|gVH schemes, reduced-dimensionality gradient/Hessian were built (yet for each considered excited state), where soft modes (solvent and flexible internal coordinates) are projected out of the nuclear coordinate space[25]. Namely, apart from the solvent, the torsion of the hydroxyl group and that around the C-C bond connecting the two heterocycles are removed (δ_{R} and δ_{OH} respectively as depicted in Figure 2). Such flexible coordinates are defined as a linear combination, with the same weights, of all the dihedrals involved (2 for hydroxyl and 4 for C-C). Using the energy and the reduced-dimensionality gradient and Hessian, a harmonic model PES is expanded over the remaining stiff coordinates adopting a generalised vertical model, in which both the initial and final states are, in general, at a non-stationary point. The here-employed Ad-MD|gVG protocol represents a variation of the original Ad-MD|gVH scheme, where VG model replaces the VH model, i.e. the Hessian of the excited states are considered equal to those of

the ground state.

The application of the VG model to the investigated system leads to no imaginary frequencies arising from the vibrational analysis once the flexible DoFs are removed, suggesting that a proper choice was made in the partition of the solute's DoFs between the classical and the quantum sets. Such a harmonic model is used to calculate the QM spectrum at the Frank-Condon level (i.e., assuming the transition dipole moment as constant) at the target temperature, resorting to the efficient TD formulation[68] implemented in $\mathscr{FC}classes$. The lineshapes obtained with reduced spaces for each snapshot are eventually convoluted with a narrow Gaussian function (HWHM=0.01 eV). The final spectrum is obtained by averaging over those computed at each snapshot, thus eventually recovering the effect on the spectrum of the soft modes excluded from the reduced QM Hessians.

It should be remarked that the employment of the VG model does not account for the non-adiabatic couplings. It simply has the benefit of avoiding significant artefacts in VH due to the imaginary frequencies caused by the non-adiabatic couplings. The effect of non-adiabatic couplings could be accounted for following an extension of the Ad-MD|gVH method recently proposed by some of us and named Ad-MD|gLVC[32]. Still, in the context of the present contribution, we skip such a costly calculation and limit ourselves to investigating non-adiabatic effects in the gas phase.

4 Results and Discussion

4.1 QMD-FF parametrization and validation

As detailed in Section 2.2 in the ESI, the intramolecular term of the QMD-FF was parameterised through the JOYCE procedure[56,57] based on the QM data specifically computed for 5,5-CprOxyLH. The parametrisation was carried out on a set of 125 redundant internal coordinates (see Tables S2 to S5 in ESI), fitting 89 independent variables with a final standard deviation of 0.08 kJ/mol. As summarised in Figure 2, the resulting QMD-FF was validated by comparing selected features with their QM counterpart. The nice overlap between the QM optimised structure and the one obtained with a minimisation of the QMD-FF appearing in Figure 2.a, is quantitatively confirmed by the RMSD reported in detail for each coordinate in Table 1. Similarly, the good agreement between QM and QMD-FF vibrational frequencies and normal modes displayed in Figure 2.b supports the quality of the proposed FF to reliably mimic small oscillation around 5,5-CprOxyLH equilibrium structure. Finally, by looking at the QM relaxed torsional profiles shown in Figure 2.c, it is evident that the dihedrals δ_R and δ_{OH} displayed in the insets (see also Figure S1) are expected to allow for larger distortions of the molecular shape. This supports the choice to describe such flexible coordinates in the QMD-FF beyond the harmonic models (see Section 2 in the ESI for further details), allowing for the remarkable agreement with the QM curves appearing in Figure 2.c.

A final validation of the quality of the QMD-FF can be achieved by testing its accuracy over a larger portion of the PES, which was explored through the MD simulations carried out on the isolated molecule with the refined FF. Figure 3 displays the population distribution of selected 5,5-CprOxyLH internal coordinates computed along the 1 ns MD trajectory. As far as δ_R and δ_{OH} are concerned, a different behaviour is registered between the two flexible dihedrals. On the one hand, the imposed temperature (298 K) allows δ_{OH} to overcome the barrier at 90° (see Figure 2.c) and to populate the two degenerate conformers at 0° and 180° , yet with significant oscillations (±50° , see also Figure S2 in the ESI). On the other hand, only the global minimum at 180° (i.e. with the two sulphur atoms in *trans* conformation) is populated for δ_R, while the local one at 0° is never explored, as indicated by the single, less broadened (±30°) distribution peak. It is worth mentioning that such oscillations on δ_R distort the co-planarity of the two rings and are therefore expected to play a relevant effect on the absorption spectrum. Further distortions of the equilibrium structure, which are expected to significantly alter the electronic density and, consequently the spectral properties, are connected with the stiff dihedrals χ_{ring} and χ_O, which respectively rule the planarity of the two rings and the out-of-plane vibration of the attached carboxyl group, both displayed in the insets of Figure 3. In this case, all χ dihedrals do show much more peaked distributions with respect to δ_R and δ_{OH}, hence supporting our choice to treat them as stiff (harmonic) coordinates.

4.2 Gas Phase Absorption Spectra

4.2.1 Harmonic models

We first approach the simulation of 5,5-CprOxyLH absorption spectral shape in gas phase, where the vibrational shapes mainly arise from the contribution of the system nuclear movements coupled with the electronic transitions to the excited states that lie within the energy region between ~ 3 and ~ 4 eV. Concretely, we consider up to the fourth lowest singlet excited state (see Table F and Figure S3 in the ESI), with vertical excitation energies at the optimised ground state geometry 3.33 eV (S1), 3.46 eV (S2), 3.62 (S3) and 3.71 eV (S4). At this geometry, S1, S3 and S4 correspond to $\pi-\pi^*$ transitions (A' in C_s symmetry), which we label respectively $\pi\pi_1^*$, $\pi\pi_2^*$, $\pi\pi_3^*$, with S3 ($\pi\pi_2^*$) displaying moderate Charge-Transfer (CT) character, while S2 is a $n-\pi^*$ transition, $n\pi_1^*$, which belongs to A" irreducible representation and has an evident CT character. The oscillator strength follows the expected trend according to the nature of the states, i.e. $\pi\pi_1^* > \pi\pi_3^* > \pi\pi_2^* >> n\pi_1^*$, with $n\pi_1^*$ being practically dark. Once the involved excited states have been determined, the contribution of the nuclear motion of the solute to the spectral shape can be introduced.

The left panel of Figure 4 shows 5,5-CprOxyLH's absorption spectra computed at FC|VG level considering the four lowest states. As expected $n\pi_1^*$ (S2 at ground state minimum) gives a negligible contribution. All other spectra exhibit some vibronic progressions which survive also in the total spectrum obtained by the sum of them. To gain a preliminary insight on the possible couplings between the low energy states contributing to the spectrum, we carried out a number of tests which confirmed the presence of multiple crossings between the diabatic electronic states. Geometry optimisations in C_s symmetry, starting from S3 or S2 states at the FC point, ended in different stationary points, both on the adiabatic S1 PES, but respectively with $\pi\pi_2^*$ and $n\pi_1^*$ char-

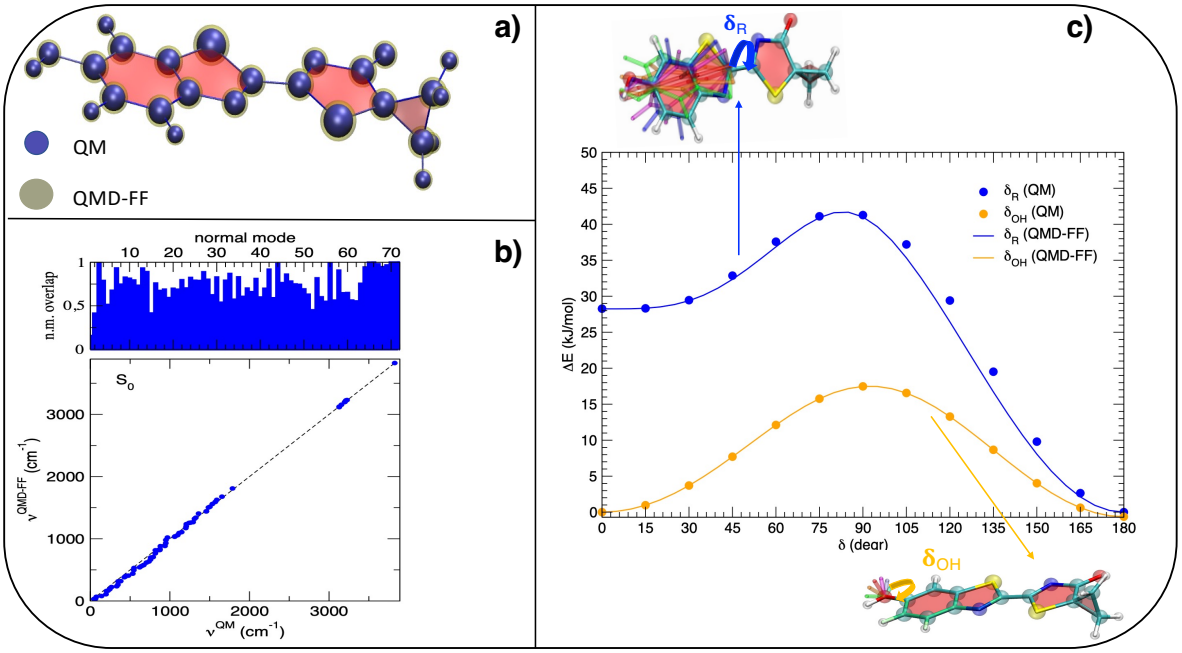

Fig. 2 Summary of QMD-FF parametrization. a) overlap of the QM (dark blue spheres) and MM (transparent greenish spheres) optimised structures, obtained at DFT level and with the QMD-FF, respectively. b) Bottom panel: correlation between QM and MM vibrational frequencies. Top: QM vs MM normal mode overlap. c) DFT (solid circles) and QMD-FF (lines) relaxed torsional profiles of the δ_R (blue) and δ_{OH} (orange) flexible dihedral evidenced in the insets.

	Bond lengths (Å)	Bending angles (deg)	Dihedrals (deg)	Structure (Å)
RMSD	0.000	0.008	0.013	0.001

Table 1 RMSD between QM and MM optimised geometries in terms of bond lengths, bending angles, dihedrals and from the superimposed structures shown in Figure 2.a.

acters (the latter being the most stable). Similarly, an optimisation started from the S1 ($\pi\pi^*_1$) minimum falls into the same $\pi\pi^*_2$ stationary point. All these stationary points, feature imaginary frequencies connected to A" modes, which can, by symmetry, carry linear couplings between $\pi\pi^*$ and $n\pi^*$ states. These imaginary frequencies challenge the employment of Vertical Hessian (VH) and Adiabatic Hessian (AH) models to investigate the effects of frequency changes and Duschinsky mixings. Distortion along these modes allow to locate the global S1 minimum that is dark, features mainly a $n\pi^*_1$ character and is 16 cm^{-1} more stable than the C_s stationary point with the same character.

These indications brought us to explore the possible effect of inter-state couplings adopting a LVC model comprising the four diabatic states $\pi\pi^*_1$, $n\pi^*_1$ $\pi\pi^*_2$, $\pi\pi^*_3$. It should be remembered that LVC collapses into the FC|VG model if inter-state couplings are switched off. According to the LVC Hamiltonian, the two ground-state normal modes carrying the largest couplings between $n\pi^*_1$ and both $\pi\pi^*_1$ and $\pi\pi^*_2$ are A" modes 4 (76 cm^{-1}) and 10 (245 cm^{-1}), corresponding respectively to a twisting and out-of-plane bending of the cyclopropane group with respect to the molecular plane. Couplings among $\pi\pi^*$ states are on the contrary triggered by higher-frequency A' modes (in-plane ring distortions). Interestingly, mode 10 is exactly the mode featuring an imaginary frequency in the $n\pi^*_1$ minimum on S1 (i45 cm^{-1}). These findings support the thesis that such an imaginary frequency, and the consequent distortion of the molecular structure that removes the C_s symmetry in the global S1 minimum, are due to the coupling of the $n\pi^*_1$ state with the close-lying $\pi\pi^*$ states. The comparison of LVC and FC|VG (sum) spectra, also shown in the left panel of Figure 4, shows that inter-state coupling introduce a remarkable smoothing of the vibronic progressions, without altering however the general width of the spectrum. Noticeable changes in the spectral shape are nevertheless observed, with a smoother rise in the low-energy wing.

The QD propagations also give access to the time evolution of the electronic population after an impulsive excitation to any of the diabatic states of the model. The results in Figure 5 document the existence of ultrafast internal conversions starting from any of the 4 diabatic states in the gas phase. About 100 fs after the excitation, in all cases most of the population is on $n\pi^*_1$ state, a finding in line with the preliminary geometry optimisations, which assessed the nπ^*_1 character of the most stable minimum on the S1 surface. In particular $\pi\pi^*_2$ and $\pi\pi^*_3$ are shown to undergo an ultrafast decay, so much that $\pi\pi^*_3$ is completely depopulated in $\sim$ 40 fs. It is worthy to notice that $n\pi^*_1$ is dark and therefore its initial excitation is not realistic. Nonetheless the time

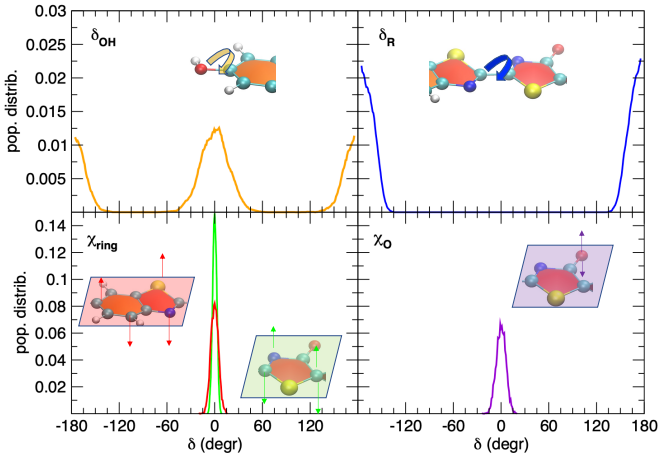

Fig. 3 Conformational analysis of the 5,5-CprOxyLH dye in gas phase at 298 K in terms of flexible (δ_{OH} and δ_R, top panels) and stiff (χ_{ring} and χ_O, bottom panels) dihedral angles. The rotation around C-O and C-C bonds are evidenced in the top panels with coloured arrows, whereas the out-of-plane distortions of the heavy atoms with respect to the first (χ_{ring1}, left) and second (χ_{ring2},right) 5,5-CprOxyLH aromatic group and of the oxygen atom of the keto group with respect to the second group (χ_O) are evidenced with red, green and violet arrows, respectively.

evolution of the initial population on $n\pi_1^*$ is useful to confirm that (although some moderate internal conversion takes place), once reached $n\pi_1^*$, the electronic population mainly remains in such state. Couplings of the $n\pi_1^*$ with the bright states mainly involve out-of plane bending and torsional motions of the cyclopropane substituent with respect to the adjacent ring. The first mode is the one featuring the largest imaginary frequencies in the adiabatic minima located with Gaussian, indicating that the existence of the non-planar global minimum of S1 is actually connected to the $n\pi^*/\pi\pi^*$ coupling. On the contrary, couplings between bright states are triggered by higher-frequency in-plane distorsions of the conjugated rings.

A further confirmation of the existence of crossings and couplings between the states at molecular structures relevant for the electronic spectra come from the classical MD exploration. In fact, when QM calculations are carried out along the MD trajectory in the gas phase, more than 80% of the sampled snapshots displayed a large number of imaginary frequencies on S1, which can be directly connected with the remarkable inter-state couplings. This is even clearer when looking at Figure S4 in the ESI, where the energy and oscillator strength computed for the first two excited states are compared: for many snapshots S1 and S2 are quite close in energy and actually exchange the electronic character. On the other hand, it should be emphasised that the non adiabatic approaches discussed above only allow to access spectral shapes at 0 K, thus completely lacking thermal effects. The latter might become crucial especially in presence of flexible coordinates, as the ones revealed in Joyce QMD-FF parametrisation for 5,5-CprOxyLH (see Figure 2.c). In fact, the large amplitude distortions from the chromophore equilibrium structure induced by the rotation of δ_{OH} and, most importantly, δ_R are expected to induce a large effect on the spectral shape. Therefore, introducing thermal effects and flexible coordinates while neglecting the couplings among excited states may represent a balanced strategy.

4.2.2 Computations with Ad-MD|gVG approach

Recognising that torsions are normally properly described at a classical level, the calculation of the lineshape could be in principle approached adopting the MQC method that we have recently developed, the Ad-MD|gVH[25], which involves reconstructing the adiabatic potential energy surfaces at different snapshots extracted along an MD trajectory at room temperature. In practice, the large number of imaginary frequencies found for all excited states along the MD trajectories and discussed in the previous paragraphs, suggest to resort to the generalised version of the simpler VG model (*g*VG), where we assume that each excited state has the same normal modes and frequencies of the ground-state. As in the original Ad-MD|gVH approach, a thermally averaged spectrum is again computed from the MD sampling, but with the simplified VG model, yet including the contributions of the 4 lowest excited states. Spectra are computed treating 2 torsions as classical DoFs. One, is the rotation around the CC bond between the two rings (δ_R depicted in Figure 2), defined as the combination of the 4 dihedrals of the quadruplet. The other, is the rotation of the OH group (δ_{OH} depicted in Figure 2), defined from the two dihedrals involved. No phenomenological broadening is necessary in this approach and spectra for all snapshots were computed with a narrow Gaussian linewidth with a HWHM=0.01 eV which does not alter the spectral width but is useful to avoid numerical issues in the application of the TD methods to compute the spectral shapes.

Results displayed in the right panel of Figure 4 for each state can be compared with the corresponding FC|VG spectra shown in the left panel. Proper accounting for the effect of molecular flexibility makes all spectra broader and erases most of their vibronic structure. Moreover there is a remarkable redistribution of the intensity among the different states which show how the fluctuations of the molecular structures mix the electronic states. It is particularly noteworthy that S2 which is dark in the FC position, gives a significant contribution to the spectrum at Ad-MD|gVG level. Moreover, the bands corresponding to S1 and S2 states span the same frequency range, which further indicates that these states mix. A similar effect is observed for S3 and S4. Comparing the total Ad-MD|gVG spectrum with the LVC one, shown in right and left panels, respectively, of Figure 4, it is interesting to notice that they are not so different, apart from a larger spread of the LVC spectrum in the red wing, which is already observed comparing the LVC spectrum with the sum of the FC|VG spectra. The overall similarity of Ad-MD|gVG and LVC spectra indicates that fluctuations of the flexible torsions and non-adiabatic couplings have a similar impact on the spectrum, washing out vibronic resolution, and shifting the maximum toward the blue. Since the latter effect is clearly visible also in Figure 4, left panel, we can deduce that it is mainly due to a redistribution of the intensity from S1 to higher states.

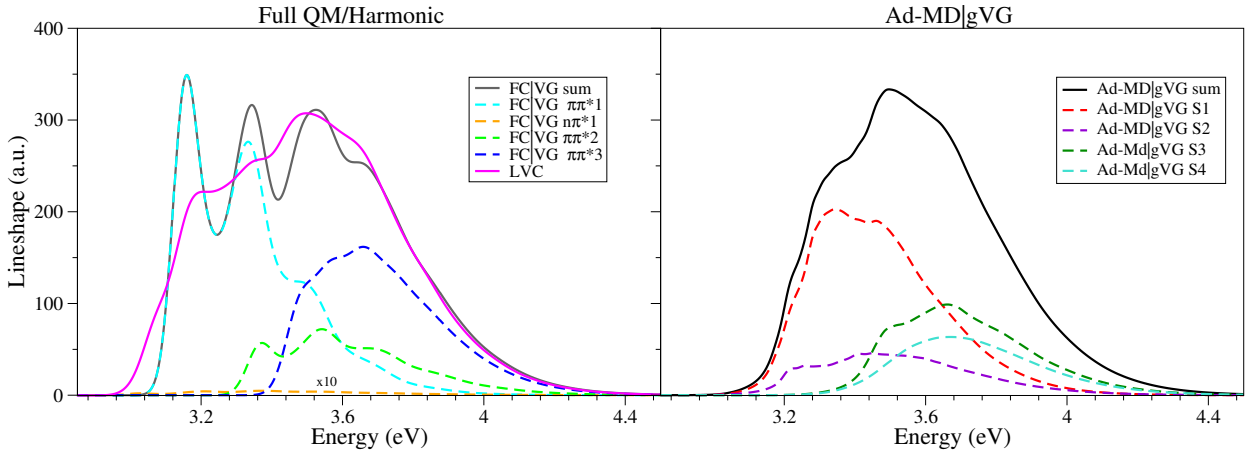

Fig. 4 Vibronic spectra of 5,5-CprOxyLH in the gas phase computed with fully QM methods on harmonic models (left) or including flexibility with the Ad-MD|gVG method (right). For the harmonic models, the spectrum at T=0 K computed at nonadiabatic level with an LVC model including the four lowest state is compared with the spectra obtained switching off the couplings, i.e. at FC|VG level, and their sum. All spectra convoluted with a Gaussian with HWHM=0.04 eV. The Ad-MD|gVG spectra are computed at room temperature applying a Gaussian convolution with HWHM=0.01 eV.

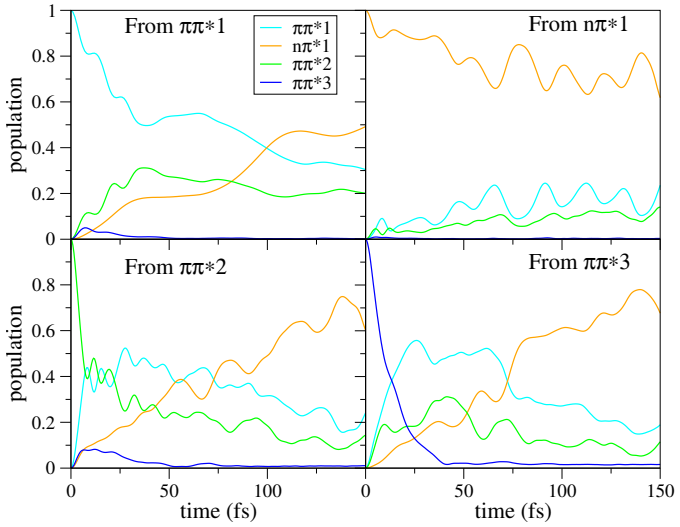

Fig. 5 Evolution of the electronic populations along QD propagation starting from the four different coupled diabatic excited states.

4.3 5,5-CprOxyLH Absorption Spectrum in Water

The significant effect on the spectral shapes exerted by the nuclear motion of the dye discussed in the previous section does not include the contribution of the environment, which usually induces a significant inhomogeneous broadening [52,69]. Effects of environment can be accounted for using implicit solvation models, which allow the application of the rigorous QD treatment adopted in the gas phase. This implies the parametrization of LVC models against QM data computed with the PCM model [70]. The simulations for our system using LVC parameterised including the effect of water solvent through PCM (see Figures S9 and S10 in the ESI), show that the population dynamics are notably affected by the destabilisation of the $n\pi_1^*$ state, which favours the population of $\pi\pi_1^*$ state. As a result, while ultrafast population transfer is still being detected, it mainly goes to $\pi\pi^*$ states, and the effect of non-adiabatic couplings on the spectrum is also significantly reduced. Moreover, these simulations are far from reproducing the spectral shape in solution, confirming the relevance of the factors neglected in LVC like molecular flexibility and specific solute-solvent interactions. as mentioned in the Introduction, in the case of aqueous solutions, if a specific interaction (HB) can be established with 5,5-CprOxyLH, the solvent should be accounted for at molecular level, thus abandoning the PCM description. Moreover, strong solute-solvent interactions could modify the behaviour of the more flexible coordinates, like δ_R or δ_{OH}, which could in turn alter further the spectral shape. Concretely, we here apply the Ad-MD|gVG protocol, already employed in the gas phase, which is able to simultaneously take into account both the presence of the solvent and solute's flexible coordinates. To verify the existence of stable HBs and further investigate the solvation structure around the chromophore, MD simulations on the solvated system are carried out explicitly accounting for the solvent, by considering $\sim$ 1200 water molecules surrounding the 5,5-CprOxyLH solute.

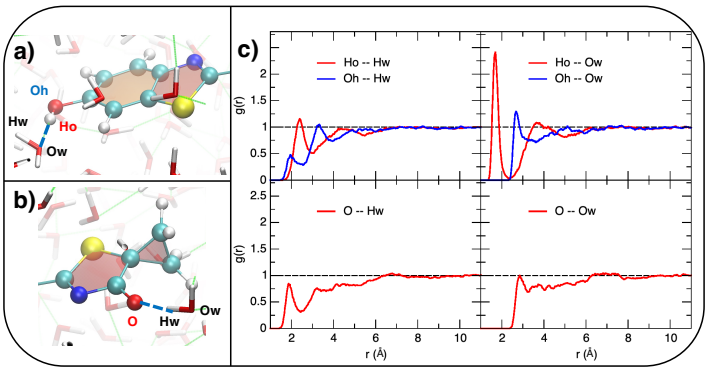

Fig. 6 Pair correlation functions between selected 5,5-CprOxyLH atom pairs (Oh, Ho and O) and water atoms (Ow, Hw). a) Definition of Oh and Ho atoms of the phenolic hydroxyl; b) definition of the keto oxygen atom (O); c) pair correlation functions, g(r) computed for hydroxyl (top) and keto (bottom) atoms along the MD trajectories.

The analysis of the solvent structure around the solute displayed in Figure 6 confirms the establishment of specific water-5,5-CprOxyLH HBs, as already reported in ref.[36]. As shown in panels a) and b) of Figure 6, here we identify HBs from pair correlation functions g(r) between the solvent and H-bond sites in the solute, namely the two atoms of the hydroxyl group (Oh and Ho) and the oxygen atom of the keto group (O). As far as the g(r) functions of the Ho and Oh atoms are concerned, in the top panels of Figure 6.c an intense and well structured peak appears between the Ho atom and the water oxygen (Ow) at distances less than 2 Å, clearly indicating that rather strong HBs are established between the Ho-Ow pair. By looking at the top left panel, it is also clear that the Oh-Hw peak is less intense than the Ho-Ow one, suggesting that the dye's phenolic substituent behaves as a HB donor rather than as an acceptor. Conversely, the O atom attached to the keto moiety accepts water hydrogen atoms, but the peak is less intense and somewhat shifted to larger distances (Figure 6.c, lower panels). Notwithstanding the other hetero-atoms present on both rings do not reveal strong HB peaks (see Figure S5 in the ESI), the HB network around 5,5-CprOxyLH is rather stable, with an average number of ~ 4.5 HB being settled during the MD run. More in detail, as shown in Figure S6, the HBs are more frequent on the 5,5-CprOxyLH's moiety bearing the phenolic group, and in general settled at smaller distances.

The well structured HB network established by 5,5-CprOxyLH with the neighbouring water molecules revealed by MD suggests that such strong intermolecular interactions could significantly alter the conformational behaviour of the dye, in particular in the distribution of the most flexible coordinates. To investigate this effect, in Figure 7 we compare the population distribution achieved for δ_{OH} and δ_R along the MD trajectories computed in the gas phase or in solution. We emphasise that the adoption of a QMD-FF for the intramolecular part ensures a PES consistent with the (TD-)DFT method employed to compute the Hessian and provides an accurate description of the flexible degrees of freedom. For both dihedrals the distributions in water show a more intense and less broadened maximum peak, indicating that the solvent slows down the dye's conformational dynamics. More specifically, δ_{OH} interconverts more slowly between its two degenerate minima at 0° and 180° (see Figure 2.c), whereas δ_R shows reduced oscillations around the planar *trans* conformation. Interestingly, similar results compared to the gas phase are also obtained with another specifically tailored FF, here referred to as Abr-FF, based on AMBER parameters and previously refined by some of us[36] to successfully describe the same molecular system. However, remarkable differences arise between the two FFs, in particular for δ_R, whose distribution is much narrower for the Abr-FF, suggesting significantly smaller distortions from the planar equilibrium structure with respect to those expected with the Joyce. Such a different behaviour stems from the different strategies adopted to obtain the FF parameters to describe the torsion. Namely, while for Abr-FF the standard GAFF parameters were adopted, those for Joyce are specifically fitted against the torsional profile evaluated at QM level.

To build a configurational ensemble able to account for both the slow conformational dynamics of the dye and the embedding solvent structure, 100 snapshots were extracted from two MD production trajectories, carried out with either the Abr-FF or the present Joyce QMD-FF. Indeed, even though the previous parametrization already provided satisfactory results, we here highlight the need of QMD-FF in order to deliver a unified description of the PES across different, QM and MM, calculations. In each snapshot, the solute and the solvent were included as described in Figure S7 in the ESI and in the QM/MM scheme outlined in the following. All solute atoms were included in the QM layer, i.e. (TD-)B3LYP/6-311g(2d,p), whereas solvent molecules within 4 Å from a solute's atom were treated at MM level, transferring the parameters for non-bonded interaction with the solute from the MD simulations. The remaining solvent atoms within a sphere of 15 Å centred on the solute were treated as point charges. The snapshots extracted along the Abr-FF and QMD-FF trajectories, are eventually used to compute the spectra with both the CEA-VE and Ad-MD|gVG methods.

By looking at the left panel of Figure 8, it can be noticed as already at the CEA-VE level, some differences are observed between the two FFs, both in the shape and intensities of the contributions of the four states. When vibronic effects are included through the Ad-MD|gVG method, shapes become more similar. This is due to both the larger broadening and the procedure to estimate, for each snapshot, the minimum along the stiff (quantum) coordinates. As observed in Figure 8, the main difference of Ad-MD|gVG spectra is the larger intensity of (mainly) S2 and S3-S4 observed for Abr-FF snapshots. In practice the MD trajectory guided by Abr-FF visits regions of the configurational space where S1 is more mixed to the other states. To explore this hypothesis, we computed the distributions along the normal modes[71] that predominantly couple $n\pi_1^*$ (S2 at the FC position) with the close-lying $\pi\pi^*$ states (namely modes 4 and 10 discussed above), showing appreciable differences between Joyce and Abr-FF, with the latter displaying larger deviations from the equilibrium structure (see Figure S8 in the ESI).

In Figure 9, we compare the simulated spectra with the experimental one[36]. The agreement with experiment is generally very good (notice that no artificial displacement has been added). It is noteworthy that although the width of the classical CEA-VE and mixed quantum classical (Ad-MD|gVG) spectra seem similar, CEA-VE were convoluted with a much larger phenomenological Gaussian (0.05 eV vs 0.01 eV). Ad-MD|gVG spectra better reproduce the structureless broad band (residual peaks are only due to an insufficient sampling). Moreover, as the vibronic contributions are also included, also the behaviour of the high-energy wing is captured much better by the Ad-MD|gVG method. The best agreement is obtained in combination with QMD-FF parameterised with JOYCE in this work. The main difference with the results achieved with the former Abr-FF is the higher intensity in the low-energy part of the band around 3.1 eV. As discussed above, this is due to the fact that the Joyce QMD-FF trajectory visits configurations where S1 is less coupled to the higher energy states and therefore it redistributes to a smaller extent its absorption intensity, as discussed above.

We recall again that the Ad-MD|gVG simulations are able to account, from first principles, the main physical processes that

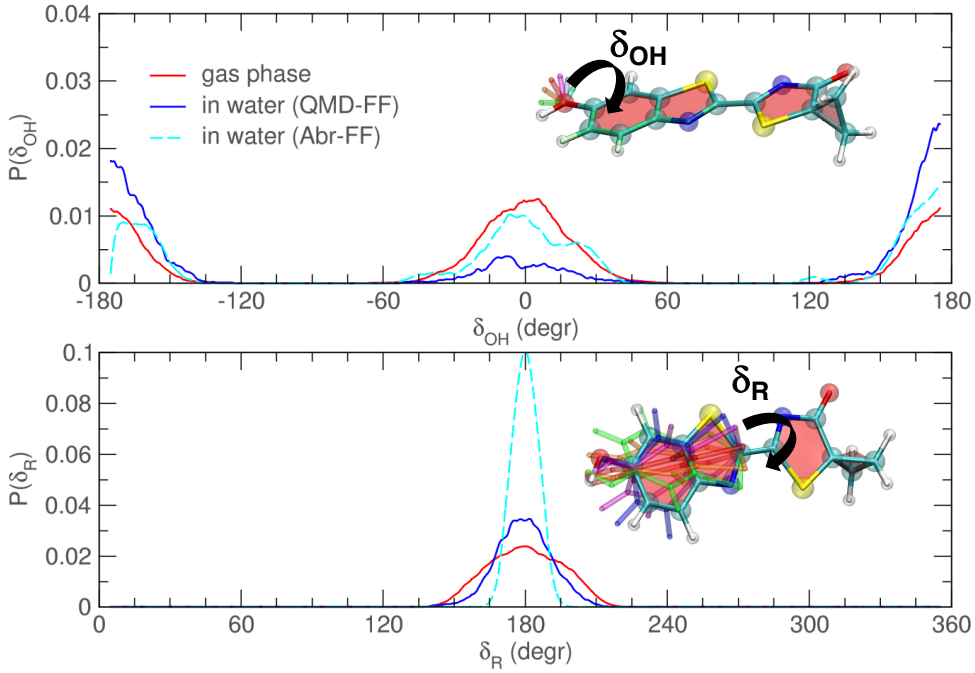

Fig. 7 Population distributions P(δ) of 5,5-CprOxyLH flexible dihedrals δ_{OH} (top) and δ_R (bottom) computed along MD trajectories in the gas phase (red lines) or water solution (blue lines), carried out with the Joyce QMD-FF. Results obtained with the Abr-FF [36] (cyan lines) are also reported for comparison.

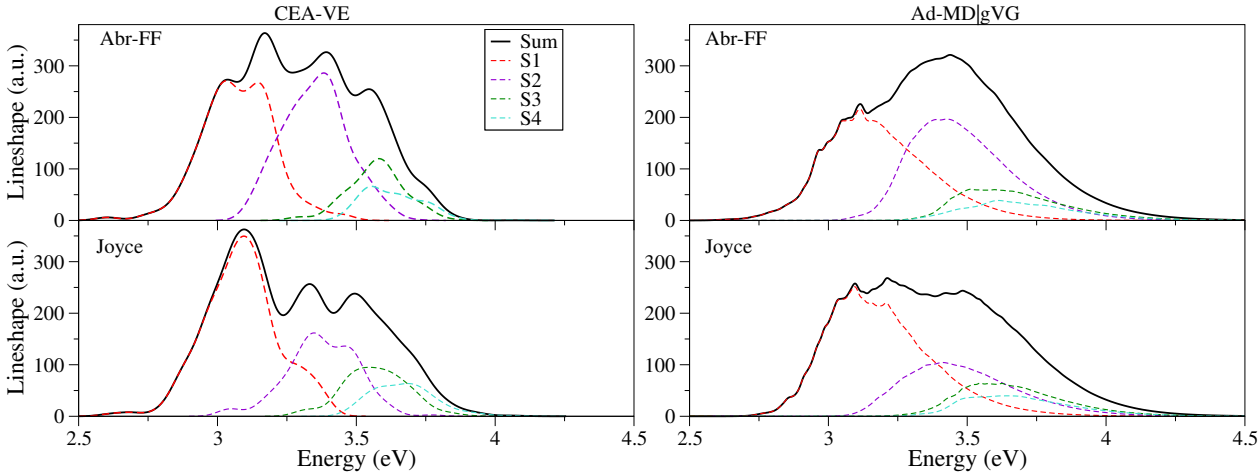

Fig. 8 Total lineshapes (solid black lines) computed from trajectories generated with either Abr-FF or Joyce using both CEA-VE (left) and Ad-MD|gVG (right) methods. The contribution of each single state (from S1 to S4) is also shown (dashed spectra). A Gaussian broadening with HWHM=0.05 eV is applied for CEA-VE spectra and with HWHM=0.01 eV for Ad-MD|gVG ones.

result in the observed lineshape. The only phenomenological parameter adopted was a very narrow Gaussian convolution, which is only required to ensure numerical stability when calculating the spectra with the TD approach, but does not alter the spectral shape. Moreover, such a small broadening can also correct, at least partially, the limitation of the sampling and the lack of non-adiabatic couplings in the Ad-MD|gVG simulations. Beside a slight overestimation of the intensity of the shoulder at $\sim$ 3.5-3.6 eV, the other significant discrepancy with respect to the experiment is observed at low-energies where the computed spectra fall too rapidly. This might be due to the lack of non-adiabatic couplings, although their effect on the spectrum is expected to be reduced in solution according to the simulated spectrum with the LVC model parameterised accounting for the solvent with PCM (also included in Fig. 9). In fact, in Figure 4 we show that, in the gas-phase, the introduction of their effect with LVC model broadens this part of the spectrum, while such an effect nearly vanishes in solution (see also Fig S10 in the ESI). Figure 9 also allows to compare the CEA-VE and Ad-MD|gVG spectra computed in water solution with those obtained in gas-phase. Just limiting to the data obtained with JOYCE FF, it is evident that with both strategies the solvent remarkably broadens the spectrum, mostly

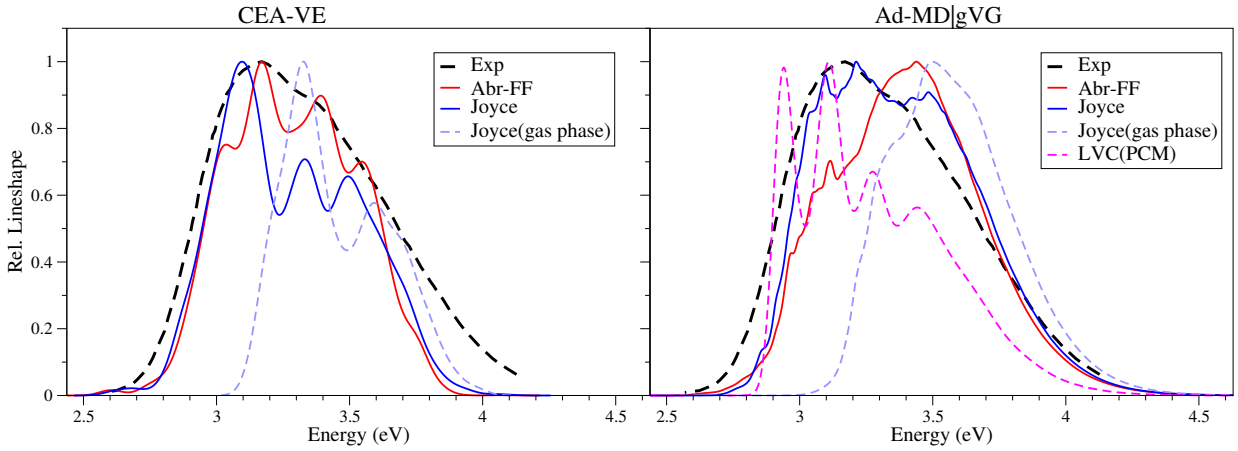

Fig. 9 Lineshapes computed with CEA-VE (left) and Ad-MD|gVG (right) methods from trajectories generated with either Abr-FF or Joyce. A Gaussian broadening with HWHM=0.05 eV is applied for CEA-VE spectra and with HWHM=0.01 eV for Ad-MD|gVG ones. The experimental spectrum is included for reference, as well as the spectra computed from trajectories in the gas phase (with Joyce) and the spectrum computed from the QD propagation using the LVC model parameterised with vertical energies computed in water with PCM

.

extending it toward the red, with a consequent red-shift of the absorption maximum by ~ 0.4 eV.

5 Conclusions

In this paper, we have reported the simulation of the light absorption of a luminescent dye, 5,5-CprOxyLH, accounting for the main factors responsible for its structured spectral shape within a first-principle ab initio framework. Namely, the here-applied Ad-MD|gVG procedure, which relies of a mixed quantum-classical method, is able to account for the effects of the coupling between molecular motions at room temperature and the electronic transitions, including both the effects of molecular flexibility and vibronic contributions. Concretely, stiff modes of the solute are described at vibronic level resorting to the harmonic approximation to describe the PESs, while the effect of flexible coordinates and that of the solvent are included at classical level by means of unconstrained MD sampling.

A total of 4 excited electronic states are considered, including 3 $\pi-\pi^*$ and one $n-\pi^*$ states which lie in an energy range of ~ 0.5 eV. We have further assessed the impact of non-adiabatic couplings among such nearby electronic states by computing the spectra from wavepacket dynamics, propagated over the four coupled states at 0 K both in the gas phase and in solution described with LVC models. On the one hand, the results in gas phase show a moderate effect of the inter-state couplings, which mainly result in a substantial blurring of the spectral features[51], while the effect is drastically reduced in solution. On the other hand, we document the occurrence of ultrafast (<100 fs) internal conversions among these states, with the state that mainly gets the populated being $n\pi_1^*$ in gas phase and $\pi\pi_1^*$ in solution. The blurring effect of the nonadiabatic interactions in the gas phase is similar to the one of molecular flexible modes, as predicted by the Ad-MD|gVG method also applied in the gas phase. On these grounds, notwithstanding some promising yet computationally expensive strategies to perform non-adiabatic QD calculations in the condensed phase have been recently proposed by some of us[32], in the computation of the spectrum in water solution we here neglected non-adiabatic couplings, hence adopting the simpler and faster Ad-MD|gVG approach.

On the balance, the agreement with the experiment is very good. Still, we should take in mind that the inclusion of such couplings may lead to an additional blurring of the spectral shapes, which would slightly improve the agreement with the experiment. Our Ad-MD|gVG approach relies, in part, on the accuracy of the MD sampling. To that end, the adoption of QMD-FFs reveals an essential step in the whole procedure. Comparison with a more standard parametrization, based on AMBER FF, shows some differences on the sampling of flexible coordinates, which eventually result in appreciable differences in the spectral shapes, with the one computed using JOYCE being closer to the experiment. In any case, it is worth noticing that the Ad-MD|gVG approach partly corrects the inaccuracies of the FFs regarding the stiff modes of the solute, which are extrapolated to the QM minimum with the gVG approach at each snapshot.

This work confirms that, at the state-of-the-art, it is possible to simulate spectral band shapes of flexible chromophores in condensed phases without the need of any phenomenological parameter. This implies that all sources of broadening can be unveiled by computational simulations opening the door to a more detailed analysis of experimental signals. In perspective, these approaches should reveal very interesting for application to more complex and structured environments, such as polymer matrices and, in the specific case, the proteins involved in bioluminiscent systems.

Conflicts of interest

There are no conflicts to declare.

Acknowledgments

I.N and C.G.I. acknowledge the French Agence Nationale de la Recherche, grant ANR-BIOLUM ANR-16-CE29-0013. J.C. thanks

Ministerio de Universidades, Plan de Recuperación, Transformación y Resiliencia and UAM for funding the research stay in Pisa with a requalification program (CA2/RSUE/2021-00890) and the MICINN Project PID2019-110091GB-I00 for financial support. Computational resources provided by ICCOM and the Centro de Cálculo Científico at Universidad Autónoma de Madrid (CCC-UAM) are also acknowledged.

Notes and references

1 M. Dierksen and S. Grimme, *J. Chem. Phys.*, 2004, **120**, 3544–3554.

2 H.-C. Jankowiak, J. L. Stuber and R. Berger, *J. Chem. Phys.*, 2007, **127**, 234101.

3 F. Santoro, A. Lami, R. Improta and V. Barone, *J. Chem. Phys.*, 2007, **126**, 184102.

4 J. Tang, M. T. Lee and S. H. Lin, *J. Chem. Phys.*, 2003, **119**, 7188–7196.

5 R. Ianconescu and E. Pollak, *J. Phys. Chem. A*, 2004, **108**, 7778–7784.

6 Q. Peng, Y. Yi, Z. Shuai and J. Shao, *J. Chem. Phys.*, 2007, **126**, 114302.

7 R. Borrelli and A. Peluso, *Phys. Chem. Chem. Phys.*, 2011, **13**, 4420–4426.

8 A. Baiardi, J. Bloino and V. Barone, *J. Chem. Theory Comput.*, 2013, **9**, 4097–4115.

9 J. Tomasi, B. Mennucci and R. Cammi, *Chem. Rev.*, 2005, **105**, 2999–3094.

10 F. J. Avila Ferrer, R. Improta, F. Santoro and V. Barone, *Phys. Chem. Chem. Phys.*, 2011, **13**, 17007–17012.

11 R. Marcus, *J. Phys. Chem.*, 1989, **93**, 3078–3086.

12 M. D'Alessandro, M. Aschi, C. Mazzuca, A. Palleschi and A. Amadei, *J. Chem. Phys.*, 2013, **139**, 114102.

13 M. D'Abramo, M. Aschi and A. Amadei, *J. Chem. Phys.*, 2014, **140**, 164104.

14 T. J. Zuehlsdorff and C. M. Isborn, *J. Chem. Phys.*, 2018, **148**, 024110.

15 T. J. Zuehlsdorff and C. M. Isborn, *Int. J. Quantum Chem.*, 2019, **119**, e25719.

16 A. Borrego-Sánchez, M. Zemmouche, J. Carmona-García, A. Francés-Monerris, P. Mulet, I. Navizet and D. Roca-Sanjuán, *J. Chem. Theory Comput.*, 2021, **17**, 3571–3582.

17 D. Loco and L. Cupellini, *Int. J. Quant. Chem.*, 2019, **119**, e25726.

18 T. J. Zuehlsdorff, A. Montoya-Castillo, J. A. Napoli, T. E. Markland and C. M. Isborn, *J. Chem. Phys.*, 2019, **151**, 074111.

19 T. J. Zuehlsdorff, H. Hong, L. Shi and C. M. Isborn, *J. Chem. Phys.*, 2020, **153**, 044127.

20 S. V. Shedge, T. J. Zuehlsdorff, A. Khanna, S. Conley and C. M. Isborn, *J. Chem. Phys.*, 2021, **154**, 084116.

21 N. De Mitri, S. Monti, G. Prampolini and V. Barone, *J. Chem. Theory Comput.*, 2013, **9**, 4507–4516.

22 O. Andreussi, I. G. Prandi, M. Campetella, G. Prampolini and B. Mennucci, *J. Chem. Theory Comput.*, 2017, **13**, 4636–4648.

23 B. Mennucci, *Int. J. Quantum Chem.*, 2015, **115**, 1202–1208.

24 S. Gómez, T. Giovannini and C. Cappelli, *ACS Phys. Chem. Au*, 2022, **in press**, DOI:10.1021/acsphyschemau.2c00050.

25 J. Cerezo, D. Aranda, F. J. Avila Ferrer, G. Prampolini and F. Santoro, *J. Chem. Theory Comput.*, 2020, **16**, 1215–1231.

26 A. Segalina, J. Cerezo, G. Prampolini, F. Santoro and M. Pastore, *J. Chem. Theor. Comput.*, 2020, **16**, 7061–7077.

27 D. Aranda and F. Santoro, *J. Chem. Theory Comput.*, 2021, **17**, 1691–1700.

28 A. J. Dunnett, D. Gowland, C. M. Isborn, A. W. Chin and T. J. Zuehlsdorff, *J. Chem. Phys.*, 2021, **155**, 144112.

29 D. Tamascelli, A. Smirne, J. Lim, S. F. Huelga and M. B. Plenio, *Phys. Rev. Lett.*, 2019, **123**, 090402.

30 J. Cerezo, Y. Liu, N. Lin, X. Zhao, R. Improta and F. Santoro, *J. Chem. Theory Comput.*, 2018, **14**, 820–832.

31 F. Santoro, J. A. Green, L. Martinez-Fernandez, J. Cerezo and R. Improta, *Phys. Chem. Chem. Phys.*, 2021, **23**, 8181–8199.

32 A. Segalina, D. Aranda, J. A. Green, V. Cristino, S. Caramori, G. Prampolini, M. Pastore and F. Santoro, *J. Chem. Theory Comput.*, 2022, **18**, 3718–3736.

33 K. Gomi and N. Kajiyama, *J. Biol. Chem.*, 2001, **276**, 36508–36513.

34 P. Naumov, Y. Ozawa, K. Ohkubo and S. Fukuzumi, *J. Am. Chem. Soc.*, 2009, **131**, 11590–11605.

35 A. Ghose, M. Rebarz, O. V. Maltsev, L. Hintermann, C. Ruckebusch, E. Fron, J. Hofkens, Y. Mély, P. Naumov, M. Sliwa and P. Didier, *J. Phys. Chem. B*, 2015, **119**, 2638–2649.

36 C. García-Iriepa, P. Gosset, R. Berraud-Pache, M. Zemmouche, G. Taupier, K. D. Dorkenoo, P. Didier, J. Léonard, N. Ferré and I. Navizet, *J. Chem. Theory Comput.*, 2018, **14**, 2117–2126.

37 C. García-Iriepa, M. Zemmouche, M. Ponce-Vargas and I. Navizet, *Phys. Chem. Chem. Phys.*, 2019, **21**, 4613–4623.

38 S.-F. Chen, Y.-J. Liu, I. Navizet, N. Ferré, W.-H. Fang and R. Lindh, *J. Chem. Theory Comput.*, 2011, **7**, 798–803.

39 P. Naumov and M. Kochunnoonny, *J. Am. Chem. Soc.*, 2010, **132**, 11566–11579.

40 M. Rebarz, B.-M. Kukovec, O. V. Maltsev, C. Ruckebusch, L. Hintermann, P. Naumov and M. Sliwa, *Chem. Sci.*, 2013, **4**, 3803–3809.

41 R. Berraud-Pache, C. Garcia-Iriepa and I. Navizet, *Front. Chem.*, 2018, **6**, 116.

42 T. Hirano, Y. Hasumi, K. Ohtsuka, S. Maki, H. Niwa, M. Yamaji and D. Hashizume, *J. Am. Chem. Soc.*, 2009, **131**, 2385–2396.

43 N. Manuel de Almeida Barbosa, M. Zemmouche, P. Gosset, C. García-Iriepa, V. Ledentu, I. Navizet, P. Didier and N. Ferré, *ChemPhotoChem*, 2019, **3**, 1219–1230.

44 A. J. Syed and J. C. Anderson, *Chem. Soc. Rev.*, 2021, **50**, 5668–5705.

45 C.-G. Min, Y. Leng, X.-K. Yang, S.-J. Huang and A.-M. Ren, *Dyes Pigm.*, 2016, **126**, 202–208.

46 A. Ghose, O. V. Maltsev, N. Humbert, L. Hintermann, Y. Arntz, P. Naumov, Y. Mély and P. Didier, *J. Phys. Chem. B*, 2017, **121**, 1566–1575.

47 K. P. Asare, M. Zniber, M. Zouheir, L. Wang, X. Wang and T.-P. Huynh, *MRS Commun.*, 2022, **12**, 90–94.

48 J. Cerezo and F. Santoro, *J. Comput. Chem.*, 2022, **in press**, DOI: 10.1002/JCC.27027.

49 F. Santoro and J. Cerezo, *FCclasses3, a code for vibronic calculations, visit:* `http://www.iccom.cnr.it/en/fcclasses`, *last consulted September*, 2022.

50 G. Worth, *Comput. Phys. Commun.*, 2020, **248**, 107040.

51 Y. Liu, L. Martínez-Fernández, J. Cerezo, G. Prampolini, R. Improta and F. Santoro, *Chem. Phys.*, 2018, **515**, 452 – 463.

52 J. Cerezo, F. J. A. Ferrer, G. Prampolini and F. Santoro, *J. Chem. Theory Comput.*, 2015, **11**, 5810–5825.

53 J. Cerezo, F. Santoro and G. Prampolini, *Theor. Chem. Acc.*, 2016, **135**, 143.

54 F. Santoro and J. A. Green, *Overdia. A code for diabatisation based on maximum overlap, Version 1.0, Visit http://www.iccom.cnr.it/overdia, last consulted September*, 2022.

55 M. Yaghoubi Jouybari, Y. Liu, R. Improta and F. Santoro, *J. Chem. Theory Comput.*, 2020, **16**, 5792–5808.

56 I. Cacelli and G. Prampolini, *J. Chem. Theory Comput.*, 2007, **3**, 1803–1817.

57 J. Cerezo, G. Prampolini and I. Cacelli, *Theor. Chem. Acc.*, 2018, **137**, 80.

58 I. Cacelli, J. Cerezo, N. De Mitri and G. Prampolini, JOYCE*2.10, a Fortran 77 code for intra-molecular force field parameterization. , available free of charge at http://www.iccom.cnr.it/en/joyce-2/, last consulted September 2022*.

59 V. Barone, I. Cacelli, N. De Mitri, D. Licari, S. Monti and G. Prampolini, *Phys. Chem. Chem. Phys.*, 2013, **15**, 3736–3751.

60 M. Cossi, N. Rega, G. Scalmani and V. Barone, *J. Comput. Chem.*, 2003, **24**, 669–681.

61 C.-G. Min, A.-M. Ren, J.-F. Guo, Z.-W. Li, L.-Y. Zou, J. D. Goddard and J.-K. Feng, *ChemPhysChem*, 2010, **11**, 251–259.

62 C.-G. Min, Y. Leng, Y.-Q. Zhu, X.-K. Yang, S.-J. Huang and A.-M. Ren, *J. Photochem. Photobiol. A*, 2017, **336**, 115–122.

63 M. Hiyama, M. Shiga, N. Koga, O. Sugino, H. Akiyama and Y. Noguchi, *Phys. Chem. Chem. Phys.*, 2017, **19**, 10028–10035.

64 M. Zemmouche, C. García-Iriepa and I. Navizet, *Phys. Chem. Chem. Phys.*, 2020, **22**, 82–91.

65 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox, *Gaussian~16 Revision D.02*, 2016, Gaussian Inc. Wallingford CT.

66 M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindahl, *SoftwareX*, 2015, **1-2**, 19 – 25.

67 S. Dapprich, I. Komáromi, K. S. Byun, K. Morokuma and M. J. Frisch, *J. Mol. Struct. THEOCHEM*, 1999, **461-462**, 1–21.

68 F. J. Avila Ferrer, J. Cerezo, J. Soto, R. Improta and F. Santoro, *Comput. Theoret. Chem.*, 2014, **1040–1041**, 328–337.

69 J. Cerezo, F. J. Avila Ferrer and F. Santoro, *Phys. Chem. Chem. Phys.*, 2015, **17**, 11401–11411.

70 Y. Liu, D. Aranda and F. Santoro, *Phys. Chem. Chem. Phys.*, 2021, **23**, 16551–16563.

71 A. Petrone, J. Cerezo, F. J. A. Ferrer, G. Donati, R. Improta, N. Rega and F. Santoro, *J. Phys. Chem. A*, 2015, **119**, 5426–5438.