Predicting Light Absorption Properties of Anthocyanidins in Solution: a Multi-Level Computational Approach

Ivo Cacelli,[a,b] Alessandro Ferretti,[*, b] Giacomo Prampolini[*, b]

date: 29/03/2019

a) Dipartimento di Chimica e Chimica Industriale, Università di Pisa, Via G. Moruzzi 13, I-56124, Pisa, Italy
b) Istituto di Chimica dei Composti OrganoMetallici (ICCOM-CNR), Area della Ricerca, via G. Moruzzi 1, I-56124 Pisa, Italy

Abstract

A multi-level computational protocol is devised to calculate the absorption spectra in ethanol solution of a series of anthocyanidins relevant for dye-sensitized solar cells. The protocol exploits the high accuracy of second order multi-reference perturbation theory to correct the results of the more feasible TD-DFT calculations, which were performed on hundreds of configurations sampled from Molecular Dynamics (MD) trajectories. The latter were purposely carried out with accurate and reliable force fields, specifically parameterized against Quantum Mechanical data, for each of the investigated dyes. Besides yielding maximum absorption wavelengths very close to the experimental values, the present approach was also capable to predict reliable band-shapes, even accounting for the subtle differences observed along the homologue series. Finally, the atomistic description achieved by MD simulations allowed for a deep insight into the different micro-solvation patterns around each anthocyanidin and their effects on the resulting dye's properties. This work can be considered as a step toward the implementation of a computational protocol able to simulate the whole system formed by the organic dye and its heterogeneous embedding, that constitutes dye-sensitized solar cells.

[*] corresponding author e-mail: ferretti@pi.iccom.cnr.it, giacomo.prampolini@pi.iccom.cnr.it

Introduction

Devices able to generate electricity from sunlight using charge transfer (CT) molecules and, in particular, Dye Sensitized Solar Cells (DSSC) were first investigated by O'Regan and Grätzel.[1] Their work has given the impulse to an intense activity aiming to engineer efficient solar cells based on molecular systems rather than on silicon. Transition metal complexes, mainly based on Ru(II), were first utilized for capturing sunlight [2-6], but disadvantages and costs have moved the interest to organic dyes [4,7-12] and in particular to those of natural origin, [13-15] which in spite of the low efficiency available today, offer low cost and reduced environmental impact. Such characteristics make these compounds much more appealing than transition metal complexes and big effort is still being devoted to improving the efficiency of sun light-electric energy conversion in these systems.

Among natural dyes, anthocyanidins are well known polyphenols of natural origin that are responsible for the intense color of red fruits, vegetables and flowers and are therefore present in many foods and beverages. For this reason, anthocyanidins and their colors are of wide interest in many aspects of biology, nutrition and enology.[16-20] For instance, anthocyanidins at pH values of wine and vegetables are not expected to have intense colors, thus the understanding of how the species are stabilized and how the final color can be tuned may contribute to improve wine making and commerce. Finally, in the past decade, their natural origin and the sensitivity of their optical properties to the embedding environment made anthocyanidins promising dyes for application in optoelectronic devices in general [17,21] and, specifically, in DSSC.[22-25] This latter aspect, although appealing for the wide availability and low environmental impact of anthocyanidins, after the first attempts, has not found the interest it deserved, mainly for the low efficiency obtained.

The color of anthocyanidins depends on many factors: the number of hydroxyl groups and their mutual orientation, the torsion around the main dihedral angle ϕ, which drives the reciprocal orientation of the two aromatic moieties constituting the whole anthocyanidin (A/C and B rings, see Figure 1), the embedding environment and the solvent's pH. In fact, owing to their electron withdrawing character the OH groups are capable to modify the aromatic ring electron density, the ϕ torsion may alter electron delocalization between the two moieties, while solvent and pH affect the withdrawing character of the OHs themselves. In natural substances we can also have bulk effects, such as stacking of two or more molecules of the same or different species, which can originate co-pigmentation.[26-28]

In this perspective, the achievement of an accurate understanding of factors and quantities governing the light absorption of these dyes is of basic importance and computational approaches may serve to guide the research in the field. In fact, several studies have been performed on the various members of the class of anthocyanidins, while few experimental investigations are also available. After the work by Harborne in 1958, [29] only Dai and Rabani in 2002 [30] reported on measured spectra for the three most representative species of the class: pelargonidin, cyanidin and delphinidin. Harborne discusses the observed absorption of all anthocyanidins in acid solution of methanol and ethanol, while Dai and Rabani, being motivated by the potential use of the dies in photovoltaic devices, also investigate aqueous solution at pH=1 and 2.8, both for simply solvated molecules (*i.e.* in H_2O or ethanol) as well as in presence of TiO_2 layers. From a

computational point of view, most of the work has been carried out for anthocyanidins and similar flavonoids species, exploiting density functional theory (DFT) and its time-dependent extension (TDDFT). [31-34,19,24,20,35-38] Post-Hartree-Fock methods have been considered only in two papers, [39,40] where the results of symmetry adapted cluster configuration interaction (SAC-CI) and multi reference configuration interaction (MRCI) calculations, complemented with second order perturbation theory (CI-MRPT2), were reported, respectively. Most of the TDDFT work is focused on the calculation of vertical excitation energies *in vacuo* [31-33] or in solution, modeling solute-solvent interactions by dielectric continuous models. [34,19,24,20,36,37]

On the one hand, notwithstanding the many computational efforts, the assessment of the most reliable DFT functional to compute anthocyanidins optical properties is still under debate.[20,33,36] As a matter of fact, despite the good performances of hybrid functionals (e.g. B3P86, PBE0 or B3LYP) recently reported for the flavonoid family,[20,34-36] a remarkable underestimation of the experimental absorption wavelength (λ_{max}) was systematically found in the case of charged anthocyanidins.[20,34] Yet, it is worth noticing that in most of these works the TDDFT vertical transition wavelengths are directly compared with the energies corresponding to the experimental λ_{max}, *i.e.* without taking into account the vibronic spectral shift and broadening.[41] On the other hand, reliable calculations of the absorption spectra requires a detailed [20,35,36,38] account for both molecular flexibility and specific solvation effects. In this framework, the dynamical evolution of the solvent surrounding some anthocyanins dyes has been recently accounted for through the Car-Parrinello (CP) method, [35,38] and their absorption spectra in water successfully evaluated at TDDFT level. Up to our knowledge, no attempt to extend the approach to anthocyanidins has been reported yet. However, both the CP computational cost (that rules out applications involving large hybrid systems as DSSC) and the possible TDDFT inaccuracies expected in the calculations of anthocyanidins spectra,[20,34] prompted us to adopt an alternative protocol described in the following.

With the present paper we aim to investigate the absorption spectrum of three anthocyanidins (pelargonidin, cyanidin and delphinidin; Figure 1) in ethanol solution, setting up a protocol which both allows for the reproduction of the whole spectral line-shape and can be safely extended to the simulation of more complex systems (*e.g.* DSSC). As already mentioned, the three molecules differ for the number of OH groups in the B ring (see Figure 1), thus solute-solvent interactions are expected to play a different role for the three systems. To the aim of investigating this aspect, since experimental results for all three species are to date available only in ethanol, we limit our study ~~here only~~ to this solvent. The need to account for the structural modifications on the dye's geometry induced by the aforementioned specific and local interactions and the intention to extend the approach to more complex and inhomogeneous systems, rule out the possibility to adopt standard vibronic approaches,[42,43,41,44-46] which could have also allowed for the simulation of spectral band shapes. Instead, a sequential molecular dynamics/quantum mechanics (MD/QM) approach[47,48] will be here employed, which has been successfully applied to large and complex systems[49,50] and is based on computing the absorption spectrum averaging on a set of configurations extracted from classical MD. Besides including specific solvent conformational and thermal

effects,[47,48,51] with this approach we also take into account at a classical level[52] dye's nuclear degrees of freedom, that also affect the spectrum, as shown for cyanidin in Ref. [[40]].

Concretely, after a detailed QM investigation, the computational route starts with the construction of an accurate force field (FF) for each molecule/solvent system. Indeed, the subtle structural differences among the three investigated homologues, as well as their not negligible influence on the optical behavior, prompted us to take into account their chemical specificity through the adoption of specifically tailored FFs. The FFs were parameterized according to the JOYCE protocol, [53,54] and thereafter employed in MD simulations. A set of geometrical arrangements of the dye and of the nearest solvent molecules is then sampled and transition energies and oscillator strengths computed for each set. For the calculation of these quantities we have chosen a TDDFT approach for its low computational cost. However, since we are aware that DFT, in the case of anthocyanidins, fails in reproducing quantitatively the position of the observed bands ~~is affected, in the case of anthocyanidins, by remarkable inaccuracies~~,[20,34] we have decided to calibrate the choice of the functional by means of accurate CI-MRPT2 calculations, performed on selected molecular geometries with the "in house" software BALOO. [55,56] By the comparison with CI-MRPT2 results, we have found that CAM-B3LYP, which is already well known to be suitable for the treatment of CT excitations,[57,58] behaves much better than B3LYP along the torsional profile of φ. However, in order to obtain an accurate overlap of TDDFT and CI-MRPT2 data, it is necessary to displace TDDFT by a constant value, to correct the systematic overestimation of the transition energies found for all tested functionals. CAM-B3LYP excitation energies and oscillator strengths, computed for each species for a set of 200 coordinates extracted from MD, show that our approach is able to reproduce the experimental data very well. Indeed, absorption maxima and band shape profiles are found very close to those observed experimentally, also accounting for the small difference measured among the three species. Furthermore, MD simulations allow an accurate investigation, at the classical level, of the different micro-solvation patterns around each anthocyanidins and their effects on the dye's properties.

2. Computational Details

2.1 General Scheme

All absorption spectral line shapes were obtained according to a sequential MD/QM multi-level protocol that can be summarized as follows:

i. A detailed QM investigation on the energy minima, conformational energy landscapes and vibrational frequencies is preliminarily carried out for each target molecule, collecting all optimized geometries, energies and the Hessian matrix in the minimum energy geometry.
ii. An accurate intramolecular FF is parameterized with the JOYCE [53,54] protocol exploiting the previously assembled QM data for each molecule.
iii. MD simulations are performed on systems composed by the target chromophore solvated by a large ensemble of solvent molecules, describing the solute through the JOYCE FF.
iv. A large collection of snapshots, containing the target molecule as well as all the solvent molecules within a fixed distance, is extracted from the MD trajectories and used to sample the conformational landscape explored by the system.
v. QM/MM calculations are performed on each snapshot to compute the vertical transition energies to a number of excited states. The solute is always accounted for at QM level, whereas the surrounding solvent molecules are included at different levels of accuracy.
vi. The final spectrum is obtained by averaging the resulting stick spectra, after convoluting each transition line with a Gaussian function of a given width.

Additional details regarding the implementation of the above general scheme can be found in Refs. [[55,48]].

2.2 QM calculations

All energy optimizations were performed at DFT level, employing the CAM-B3LYP functional with the 6-311+G(2d,2p) basis set. Minimum energy conformations were obtained by simultaneous optimization of all internal coordinates, whereas relaxed scans were carried out by optimizing all degrees of freedom, but the scanned coordinate. The Hessian matrix was computed for the ground state (GS) minimum geometry. To minimize the effect of the overall positive charge (+1) of the target compounds, all calculations above were performed in solution (ethanol), modeling solute solvent interaction by the Polarizable Continuum Model (PCM).[59]

TDDFT vertical transition energies were computed for the first three excited states employing several functionals (B3LYP, PBE0, M06-2X, CAM-B3LYP and ωB96XD) and basis sets, namely 6-31G(d), 6-311G(d), 6-311G(d,p), 6-311+G(2d,2p) and 6-311++G(2d,2p). The calculations were carried out for both the dye *in vacuo* and solvated in ethanol. In the latter the surrounding ethanol molecules within a radius of 15 Å from the solute were considered as point charges, taking the same charges as those previously employed in the MD simulations (*vide infra*). The whole system was then inserted in a single PCM cavity with a radius of about 15 Å. All DFT and TDDFT calculations were performed with the Gaussian09 package. [60]

CI-MRPT2 calculations were performed using the BALOO code. [56] The molecular orbital (MO) basis set is

first obtained by an Hartree-Fock (HF) calculation with the 6-311G(d) basis set using GAMESS [61] followed by a transformation from atomic to molecular basis of the one- and two-electron integrals. The CI-MRPT2 calculation is then carried out by four subsequent steps in which the dimension of the initial configurational space is gradually enlarged by a given factor, adding at each step those configurations with the largest second order perturbative energy correction to the CI states. In the present case, we have made 3 different calculations with dimension roughly 240-1200-4800-20500, 240-1700-9500-47600 and 240-2200-15200-90200, respectively (dimensions slightly depend from the species and from geometry [56]). Notice that the final variational space includes excitations up to 4 and the subsequent perturbative correction includes then excitation level up to 6. These four steps are followed by a last MRPT2 step which exploits the diagrammatic perturbation theory, [62] which is much more efficient than the traditional many-body perturbation theory.

2.3 FF parameterization and MD simulations

The total energy in MD simulations is expressed as a sum of an intra-molecular term, E^{intra}, and an inter-molecular contribution E^{inter}. The former is computed for each molecule (either the anthocyanidins dye or the ethanol solvent) as the usual[47,48,54,53] sum of stretching, bending, dihedral torsions and intra-molecular non-bonded contributions, whereas E^{inter}, which accounts for both solute-solvent and solvent-solvent interactions, is expressed by a standard atom-atom Lennard-Jones (LJ) term plus a Coulomb charge-charge contribution.

As far as the dyes are concerned, the E^{intra} term was specifically parameterized for each molecule with the JOYCE program,[54,53] using the energy and Hessian matrix in the GS minimum and relaxed energy scans computed at QM level. The energies of several conformers obtained by rotation of the flexible dihedrals displayed in Figure 1, have been computed separately for pelargonidin, cyanidin and delphinidin. For the E^{inter} contribution, the LJ parameters have been taken from the Optimized Potentials for Liquid Simulations (OPLS-AA)[63] FF, whereas the point charges were computed, in PCM solution (ethanol), through the CM5 protocol[64] at CAM-B3LYP/6-311+G(2d,2p) level on the optimized geometry. All parameters are reported in the Supporting Information, along with additional details of the parameterization procedure. Turning to the surrounding environment, the intermolecular FF parameters of the Cl^- counterion, included in all systems to neutralize the positive charge of the dyes, were taken from Ref. [[65]]. A full atomic description was adopted for ethanol, describing intra- and inter-molecular terms through the OPLS-AA FF, whose parameters were taken from the Virtual Chemistry database.[66]

All MD simulations, as well as molecular mechanics (MM) optimizations, were performed with the GROMACS4.5 engine.[67] All simulations were carried out in the NPT ensemble, on systems composed of one solute molecule (either pelargonidin, cyanidin or delphinidin), ~800 ethanol molecules and one Cl^- counterion, keeping temperature T (300 K) and pressure P (1 atm) constant through the v-rescale[68] and Parrinello-Rahman[69] schemes, using coupling constants of ps 0.1 and 1 ps, respectively. Due to the fast stretching vibrations, a time step of 0.25 fs was employed in all runs. A cut-off of 11 Å was employed for

both charge-charge and van der Waals terms, whereas long-range electrostatics was accounted through the particle mesh Ewald procedure. The starting conformations were created through the *genbox* tool[67] of the GROMACS package, by solvating the charged dye molecule and its counterion in a pre-equilibrated cubic box containing ~800 solvent molecules. The resulting systems were first minimized, then velocities were assigned to each atom according to a Maxwell-Boltzmann distribution at 300 K and each system was equilibrated for 2 ns. Thereafter production runs were performed for further 5 ns, collecting configurations every 1.25 ps and 25 ps, for a total of 4000 and 200 snapshots, to be used for trajectory analysis and spectra computation, respectively.

The equilibrated systems were characterized by computing the average bulk density (ρ), dye's solvation energy ΔE_{solv} (*i.e.* the total intermolecular energy between the considered anthocyanidin and all solvent molecules), and number of hydrogen bonds $<n_{HB}>$. The solvent structure embedding the dye, with particular attention on the HB network, was investigated by computing the number of HBs over time (n_{HB}), the distribution of selected geometrical patterns describing the HB network and the atom-atom pair correlation functions $g_{ij}(r)$ between atom *i* of the solute and atom *j* of the solvent.

A more accurate validation could be obtained by a direct comparison of QM and MM energies and structural data for various dye-ethanol clusters, but this would have required extremely long computations for the QM part and, also in light of the use of standard parameters for the solvent, we have decided to avoid such computationally expensive tests.

2.4 Absorption Spectra

Once the vertical transition energies and oscillator strengths for the first three excited states were obtained for each of the given snapshot as previously described, the resulting stick spectrum was convoluted with a Gaussian function with a Half Width at Half Maximum (HWHM) of 0.05. A simple average on the set of the convoluted spectra for all snapshots allows us to obtain the final absorption band shape profile, whose profile is connected to both solvent effects and solute vibrational dynamics.

3. Results and Discussion

3.1 Preliminary analysis of anthocyanidins geometry and flexibility

The three anthocyanidin dyes investigated in the present work are displayed in Figure 1. All homologues are composed by two building blocks, namely two condensed aromatic rings (A and C) and a poly-phenol cycle (B), connected by a single bond, with partial π character. Pelargonidin, cyanidin and delphinidin differ in the number of OH groups attached onto the B ring (one, two or three, respectively), whereas the structure of the A and C rings remains unaltered.

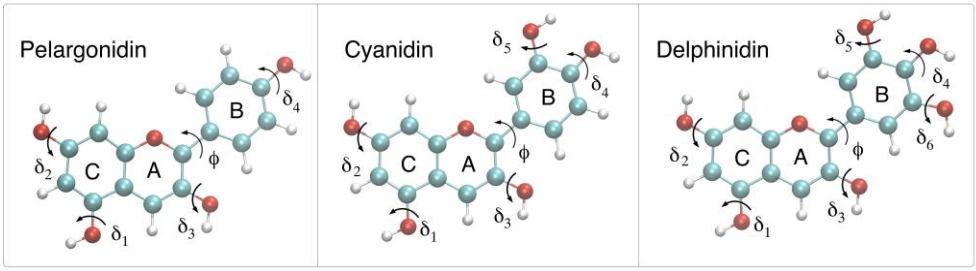

Figure 1 Equilibrium structures of the investigated anthocyanidin dyes. Flexible dihedrals are evidenced with black arrows. In the displayed conformers the dihedrals are: $\phi = 0°$; $\delta_1 = \delta_2 = 180°$; $\delta_3 = 0°$; $\delta_4 = \delta_5 = \delta_6 = 180°$.

The geometry of each dye was optimized at QM level, as described in the previous sections, resulting in a perfectly planar arrangement, in agreement with previous reports. [31-34,19,24,20,35] Vibrational calculations, yielded in all cases several low frequencies, suggesting that large amplitude slow motions can be expected for some internal coordinates. In fact, normal mode analysis revealed that all normal coordinates connected with low frequencies are essentially localized on a single (or a few) internal coordinate, *i.e.* the dihedral ϕ between the two aromatic moieties and/or the dihedrals $\delta_1 - \delta_6$ (see Figure 1), which drive the orientation of each hydroxyl group. All molecules are thus expected to exhibit a certain degree of flexibility, essentially located in the dihedrals evidenced in Figure 1. The ϕ dihedral drives the overall molecular shape and, most important for the optical properties, influences the electronic density distribution over the two aromatic moieties that constitute the chromophore. On the other hand, $\delta_1 - \delta_6$ dihedrals rule the orientation of the OH groups, which may become fundamental in the interaction with the neighboring (protic) ethanol molecules. Moreover, the absorption properties of the dyes could be affected by the different number of -OH substituents, since their electron withdrawing character may alter the electronic density on the aromatic moieties and in particular in the B ring. In conclusion, the not negligible role on absorption properties of both the expected flexibility and the possible interactions with the surrounding ethanol molecules, evidenced by this preliminary analysis, seems to suggest that a static approach, which only considers the equilibrium geometry of each dye, might be inadequate in the present case. For this reason, the multi-level approach described in the previous section was adopted.

3.2 CI characterization of the transition and DFT functional selection

In view of the large number of calculations that the chosen sequential MD/QM protocol requires, the vertical transition energies for each snapshot have to be computed with some cost-effective QM technique, among which TDDFT is certainly the most popular and efficient. Nonetheless, as recently reported in literature for anthocyanidins and, more in general, for similar flavonoid compounds,[20,36,34] care must be taken in choosing the functional, as the results may dramatically depend on it. Moreover, as previously mentioned, the choice of the ”best” functional should not be solely based on the comparison between experimental λ_{max} and the computed wavelength corresponding to the vertical transition from the GS minimum, but the comparison of the experimental and computed broadened spectral shapes should be also taken into account.[41] On the other side, since, due to the multiple OH groups, the coupling between solvent and

solute dynamics is not expected to be negligible, standard vibronic approaches, [42,43,41,44-46] which take into account only solute vibrational contributions were also not considered for the present study. Here we thus choose to follow a third route, which has also the advantage to preserve the predicting capabilities of QM and MD approaches, since no experimental data is exploited in the DFT functional choice. To this aim, high-level CI-MRPT2 calculations of the lowest four states are performed on selected geometries of each dye, and the DFT functionals benchmarked in their ability to reproduce the CI-MRPT2 trends. Assuming that both calculations are affected in the same way by solvation, the comparison between reference CI-MRPT2 and TDDFT results can be safely performed *in vacuo*. Besides the equilibrium geometry, the comparison was performed for several values of the torsional angle ϕ, which is expected to give a large contribution to the line shape. Moreover, the comparison between the two methods can be extended to the charge density of both the ground and excited states, in order to further check the quality of the several functionals for these specific systems.

Considering the pelargonidin at the optimized energy minimum as a reference case, we have first checked the convergence of CI-MRPT2 computations performing three different calculations with the CI spaces listed in section 2.2. The results are reported in Table 1.

Dimension of the CI spaces	**$S_0 \rightarrow S_1$**		**$S_0 \rightarrow S_2$**		**$S_0 \rightarrow S_3$**		**CPU time (days)**
	ΔE_{01}	**f**	**ΔE_{02}**	**f**	**ΔE_{03}**	**f**	
236-1170-4749-25272	2.35	0.858	2.84	0.061	3.73	0.007	8
240-1200-4800-20500	2.28	0.833	2.75	0.062	3.64	0.007	19
240-1700-9500-47600	2.33	0.868	2.83	0.065	3.71	0.008	37
240-2200-15200-90200	2.43	0.905	2.98	0.065	3.74	0.012	124

Table 1 CI-MRPT2 transition energies (ΔE_{0n}, eV) and oscillator strengths (f) computed for the first three singlet excited states of pelargonidin, increasing the CI space (first columns) using the 6-311G(d) basis set. In the last column the CPU time/processor on a 2 GHz Intel® Xeon® are reported. The calculations refer to the optimized geometry *in vacuo*.

Of the first three excitations from the ground state S_0 to the singlet excited states S_1, S_2 and S_3, only the first two are bright, with the second which is one order of magnitude less intense than the first one (Table 1). These are both $\pi \rightarrow \pi^*$ transitions which can be essentially assigned as HOMO$\rightarrow$LUMO and HOMO-1$\rightarrow$LUMO (Figure 2), respectively. S_3 corresponds again to a $\pi \rightarrow \pi^*$ transition, mainly HOMO-2$\rightarrow$LUMO. From the results reported in Table 1, it appears that for the strongest $S_0 \rightarrow S_1$ transition the difference between the calculations performed in the smallest and the largest CI space is 0.15 eV, being 0.1 eV the difference between the intermediate and the largest one. In order to have a good compromise between computational effort and accuracy, we then decided to adopt the intermediate CI space for the calculations at various values of the torsion ϕ (see next section), whereas the excitation energy obtained with the largest space was taken as reference value for DFT tuning (*vide infra*).

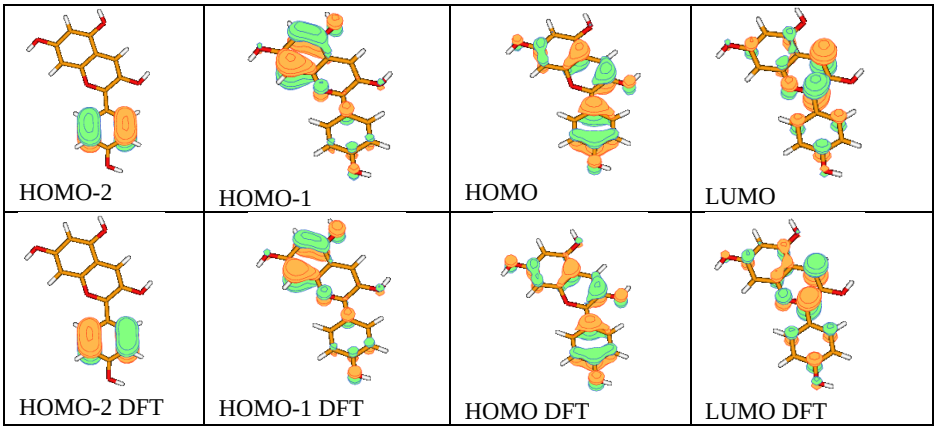

Figure 2 CI-MRPT2 (top) and DFT (bottom) MOs involved in the visible absorption of pelargonidin at the optimized minimum geometry *in vacuo*.

As far as the charge distribution in the A/C and B moieties of the molecule is concerned, the CI-MRPT2 Mulliken charge distributions computed with the intermediate CI space are reported in Table 2. In Table 2, are also shown the results of DFT computations which will be discussed later in the article.

Method	**Basis set**	**Localization**	**S_0**		**S_1**		**S_2**	
			A/C	B	A/C	B	A/C	B
CI-MRPT2	6-311G(d)	Mulliken	0.71	0.29	0.56	0.44	0.75	0.25
CAM-B3LYP	6-311G(d)	Mulliken	0.70	0.30	0.68	0.32	0.73	0.27
CAM-B3LYP	6-311+G(2d,2p)	Mulliken	0.46	0.54	0.45	0.55	0.50	0.50
CAM-B3LYP	6-311+G(2d,2p)	CM5[(a)]	0.77	0.23	-	-	-	-

[(a)] CM5 charges were computed only for the GS in ethanol PCM solution

Table 2 Mulliken charge distribution over the two fragments (A/C and B) of pelargonidin as computed by CI-MRPT2 and CAM-B3LYP level for the first three excited states *in vacuo*. In the last row, CM5 point charges, computed from the CAM-B3LYP density in ethanol PCM solution, are reported for the GS (S_0).

According to CI results, the GS electron density is more localized in the B fragment and a significant CT can be observed when going to the first excited state, where the charge results more delocalized over the two moieties. On the contrary, the excitation to S_2 does not show any significant CT character. This is in line with the assignments based on the single excitations and with the charge distribution shown by the frontier orbitals of Figure 2.

The CI-MRPT2 vertical transition energies (ΔE_{0n}, n=1-3) and their oscillator strengths (f) were then used to benchmark different DFT functionals and basis sets. Regarding the latter, it was found that for all tested functionals the 6-311+G(2d,2p) basis set was a good compromise between computational efficiency and accuracy, also because the addition of a second diffuse function (*i.e.* the 6-311++G(2d,2p) basis set) did not result in any appreciable change in the computed properties (see Table A in the Supporting Information for details). As far as the functional choice is concerned, literature results seem to indicate hybrid functionals as the most appropriate for flavonoids,[20,36,34] although, within this class of compounds, anthocyanidins were reported to be somewhat outside the good average trend.[34] Considering the CT character of the

investigated transition, an alternative choice could stand in long range corrected functionals, which are recommended in these cases.[57] ΔE_{0n} for the first three excited states and their f's, computed with several functionals belonging to the aforementioned classes and coupled with the 6-311+G(2d,2p) basis set are reported for pelargonidin in Table 3.

	$S_0 \rightarrow S_1$		$S_0 \rightarrow S_2$		$S_0 \rightarrow S_3$	
Method	ΔE_{01}	**f**	ΔE_{02}	**f**	ΔE_{03}	**f**
B3LYP	2.68	0.448	3.07	0.276	3.45	0.015
PBE0	2.75	0.499	3.18	0.253	3.61	0.013
M062X	2.91	0.670	3.55	0.136	4.10	0.012
CAM-B3LYP	2.89	0.671	3.56	0.128	4.07	0.013
ωB97XD	2.91	0.669	3.60	0.128	4.14	0.013
CI-MRPT2	2.43	0.905	2.98	0.065	3.74	0.012

Table 3 TDDFT transition energies (ΔE_{0n}, eV) and oscillator strengths (f) *in vacuo* computed for the first three excited states of pelargonidin, computed with different functionals with the 6-311+G(2d,2p) basis set at the equilibrium geometry.

The TD-DFT assignments of the three transitions do not differ from those of CI-MRPT2 and, for all functionals, the Kohn-Sham orbitals are graphically similar to the CI MO's displayed in Figure 2. Conversely, all functionals yield ΔE_{01} and ΔE_{02} blue shifted with respect to the CI-MRPT2 values. Nonetheless, the DFT excitation energies show a marked dependence on the functional, in particular for the highest state. In fact, the shift with respect to CI results seems associated with the nature of the functional: standard hybrids (B3LYP and PBE0) overestimate ΔE_{01} reference values by ~0.3 eV, whereas dispersion corrected (M062X) and range separated (CAM-B3LYP and ωB97XD) functionals are blue shifted by ~0.5 eV. The $S_0 \rightarrow S_1$ oscillator strengths are all smaller than the CI-MRPT2 ones, whereas those of the $S_0 \rightarrow S_2$ transition are higher and their sum is rather similar for TDDFT and CI-MRPT2. Thus it seems that in the TDFFT results the absorption intensity is distributed in the first two lowest transitions, whereas in the CI-MRPT2 case the intensity in mainly concentrated in the first transition. Finally, the S_3 state is not optically active according to all methods, but a peculiar trend inversion is worth mentioning: B3LYP and PBE0 underestimate or match the CI-MRPT2 ΔE_{03} energies, whereas, consistently with the lowest transitions, long range corrected functionals still overestimate the reference value.

The differences in the oscillator strength between CI-MRPT2 and TDDFT may find an explanation in the composition of the states in term of HF-MOs and Kohn-Sham MOs (KS-MOs). In fact, in both cases, S_1 and S_2 are made by combination of HOMO→LUMO and HOMO-1→LUMO excitations (with some HOMO-3→LUMO for S_2), but the contribution of the HOMO→LUMO excitation to S_1 is slightly larger for CI-MRPT2 than for TDDFT, while it is slightly smaller in the excitation at S_2 (the opposite holds for HOMO-1→LUMO). This seems to indicate that the HOMO→LUMO single excitation is the main source of intensity.

The DFT benchmarking was then extended to the other two homologues of the series. Given the similar results obtained within the two classes of functional (*i.e.* with or without the long range correction), in the following only one functional of each class, namely B3LYP and CAM-B3LYP, will be tested. In Table 4 the absorption spectrum computed at different levels of theory are reported for pelargonidin, cyanidin and delphinidin.

	$S_0 \rightarrow S_1$		$S_0 \rightarrow S_2$		$S_0 \rightarrow S_3$	
	ΔE_{01}	f	ΔE_{02}	f	ΔE_{03}	f
Pelargonidin						
CI-MRPT2	2.43	0.905	2.98	0.065	3.74	0.012
B3LYP	2.68	0.448	3.07	0.277	3.45	0.015
CAM-B3LYP	2.89	0.671	3.56	0.128	4.07	0.013
Cyanidin						
CI-MRPT2	2.34	0.907	2.92	0.022	3.24	0.018
B3LYP	2.62	0.557	2.91	0.011	3.11	0.161
CAM-B3LYP	2.83	0.727	3.47	0.045	3.65	0.040
Delphinidin						
CI-MRPT2	2.27	0.882	2.67	0.007	2.84	0.020
B3LYP	2.44	0.007	2.58	0.585	2.90	0.048
CAM-B3LYP	2.80	0.706	3.07	0.059	4.27	0.025

Table 4 CI-MRPT2 and TDDFT transition energies (ΔE_{0n}, eV) and relative oscillator strengths (f) computed for the first three excited states of the investigated anthocyanidins, *in vacuo* at the equilibrium geometry.

It is apparent that the small red shift along the series (from 2.43 to 2.34 to 2.27 eV, on going from pelargonidin to delphinidin) registered at CI level and in agreement with the experimental trend[29] is reproduced by both DFT functionals. However, for the B3LYP functional, f significantly varies along the homologue series, being 0.448 for pelargonidin, 0.557 for cyanidin and 0.007 for delphinidin. This latter value is in strong disagreement with both the CI-MRPT2 and CAM-B3LYP results, where the strength of the transition is nearly constant along the series. Such suspect behavior, which is rooted in the B3LYP inversion of the S_1 and S_2 states, prompted us to orient ourselves towards the CAM-B3LYP functional, which seems to yield a more consistent overall picture. However, since for the minimum energy geometry B3LYP vertical transitions are closer in energy to the reference CI ones, an additional comparison is carried out.

3.3 TD-DFT performances on conformers different from the minimum

Considering that many different conformers will be sampled from MD to compute the spectral broadening, the dependence on molecular geometry of the performances of both B3LYP and CAM-B3LYP representative of the two investigated classes should also be tested. To this aim, reference CI-MRPT2 calculations were performed for different pelargonidin and cyanidin conformers, obtained for each homologue through the DFT relaxed scan of the torsional angle ϕ. To avoid computational burden, these calculations were limited to the intermediate size CI space reported in Table 1. The same geometries were thereafter employed for the calculations with the two chosen functionals. The results are compared in Figure 3 for the first transition and in Figures A and B of the Supporting Information for the next two excited states.

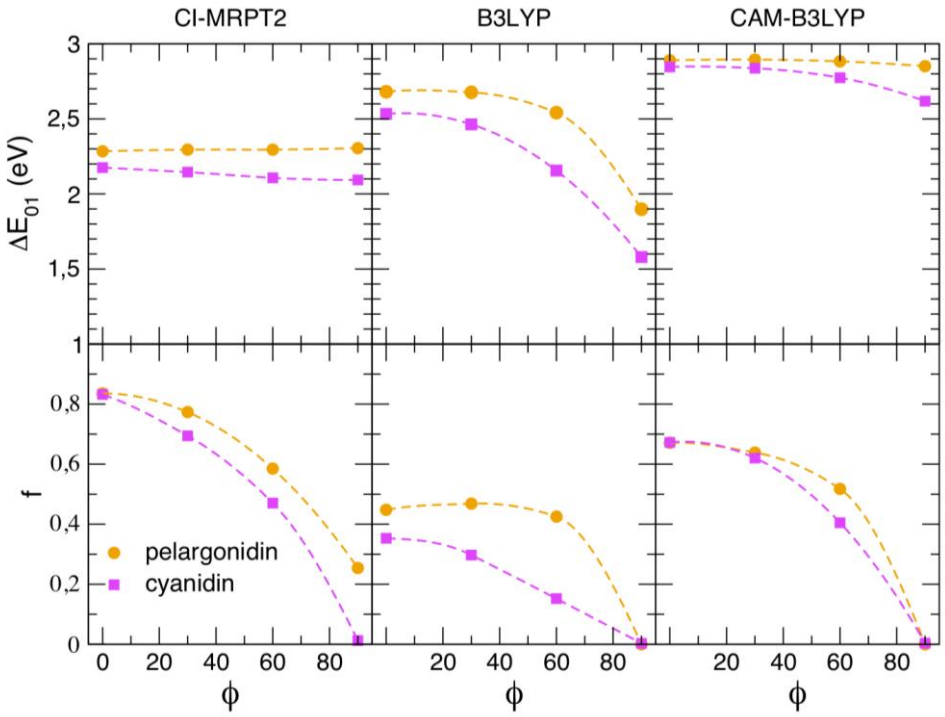

Figure 3 CI-MRPT2 and TDDFT *in vacuo* vertical transition energies (ΔE_{01}) and oscillator strengths (f) computed for the first excited state of pelargonidin (orange) and cyanidin (magenta) in the geometries obtained along the relaxed energy scan of dihedral ϕ (see Figure 1).

By looking at the top panels of Figure 3, it results that, despite B3LYP ΔE_{01} for the minimum energy conformation (ϕ = 0°) is closer to the reference data with respect to CAM-B3LYP, its dependence from conformational changes is qualitatively different, whereas CAM results seem to retain their shift with respect to the reference CI curve. The same observation holds for the oscillator strength, whose monotonic decrease for increasing ϕ angles is found for both CI-MRPT2 and CAM-B3LYP.

As a final test for the CAM-B3LYP functional, the distribution of Mulliken charges over the two fragments was computed for the minimum energy conformer of pelargonidin, and compared to the CI values in Table 2. The remarkable disagreement observed between the values obtained by localization of the CI and TDDFT electron density, computed for this latter with the 6-311+G(2d,2p) basis set, confirms that Mulliken population analysis is not the most appropriate algorithm to derive point charges for CT molecules[58] when diffuse functions are included in the basis set. In fact, by repeating the calculations with the smaller 6-311G(d) basis set (*i.e.* the same employed in the CI calculations), it is evident that the lack of charge separation between the two moieties suggested by the larger basis set is an artifact, introduced by the diffuse functions, and it disappears when either the diffuse functions are removed form the basis set or an alternative localization scheme, such as CM5 reported in Table 2 is adopted. It is important to note that CM5 charges, actually employed in MD simulations, closely respect the CI charge distributions.

In conclusion, the CAM-B3LYP functional will be adopted in the following, since it seems able to better catch the difference among homologues found with reference CI calculations, as well as the dependence of the absorption features for different conformations of the molecules. Furthermore, in the hypothesis that the significant shift Δ^{corr} (~0.5 eV), found in ΔE_{01} for each considered dye, might be nearly constant and almost

independent on the dye's geometry, the CAM-B3LYP transition energies will be corrected for such shift Δ^{corr} to yield more accurate predictions in the position of the absorption bands. In Table 5, the values computed at both CI-MRPT2 and TDDFT level for ΔE_{01} are reported together with the correction factor Δ^{corr}.

Anthocyanidin	ΔE_{01} CI-MRPT2 (eV)	ΔE_{01} CAM-B3LYP (eV)	Δ^{corr} (eV)
Pelargonidin	2.43	2.89	0.46
Cyanidin	2.34	2.83	0.49
Delphinidin	2.27	2.80	0.53

Table 5 CI-MRPT2 and CAM-B3LYP vertical energies (ΔE_{01}) computed for the first transition ($S_0 \rightarrow S_1$) in the investigated anthocyanidins.

3.4 FF parameterization and validation

Pelargonidin

The first FF was parameterized for the pelargonidin molecule. Following JOYCE standard protocol, [48,54,53] all internal coordinates, except flexible dihedrals, were represented by harmonic potentials (see Supporting Information for further details), whereas the dihedrals ϕ and δ_1 to δ_4 were parameterized using a sum of cosine functions with different multiplicity (n=0-6). The same FF torsional parameters were imposed for δ_2 and δ_4 (see Supporting Information). Finally, intra-molecular LJ interactions, between the H_3 atom (*i.e.* the one rotated by δ_3) and the B ring Hydrogens were also included in the FF, to take into account the coupling between the ϕ and δ_3 torsions. In fact, as shown in Figure 4, a δ_3 rotation alters the interaction of the O_3H_3 group with the B ring, which can readjust its position by a rotation of the ϕ dihedral, to further minimize the intramolecular energy. The JOYCE fitting was performed achieving a final standard deviation of 1.6 10^{-2} kJ/mol. To validate the FF, three different tests were conceived. First, both full optimization and relaxed scans were performed at MM level, optimizing the geometry according to the E^{intra} JOYCE potential. The resulting structures were then visually compared to their QM optimized counterparts (see Figures D and E in the Supporting Information), to check significant differences. Thereafter, on each MM optimized structure, the FF conformational energy was computed and compared to the QM energy obtained on the QM optimized counterpart. In particular, to ascertain the capability of the FF to mimic the coupling between the δ_3 and ϕ dihedrals, the ϕ values obtained during the QM and MM δ_3 scan were also compared, and displayed in Figure 4. Finally, the accuracy of the MM structures was further ascertained by computing, at TD-DFT level on the MM geometries, the vertical transition energies and their oscillator strengths, comparing them with the same quantities computed at the same level of theory on the QM optimized structures. By looking at all panels of Figure 5, it is evident that the agreement is again very good for all investigated conformers, supporting the conclusion that parameterized FF can be safely used to sample, through MD simulations, pelargonidin conformational space.

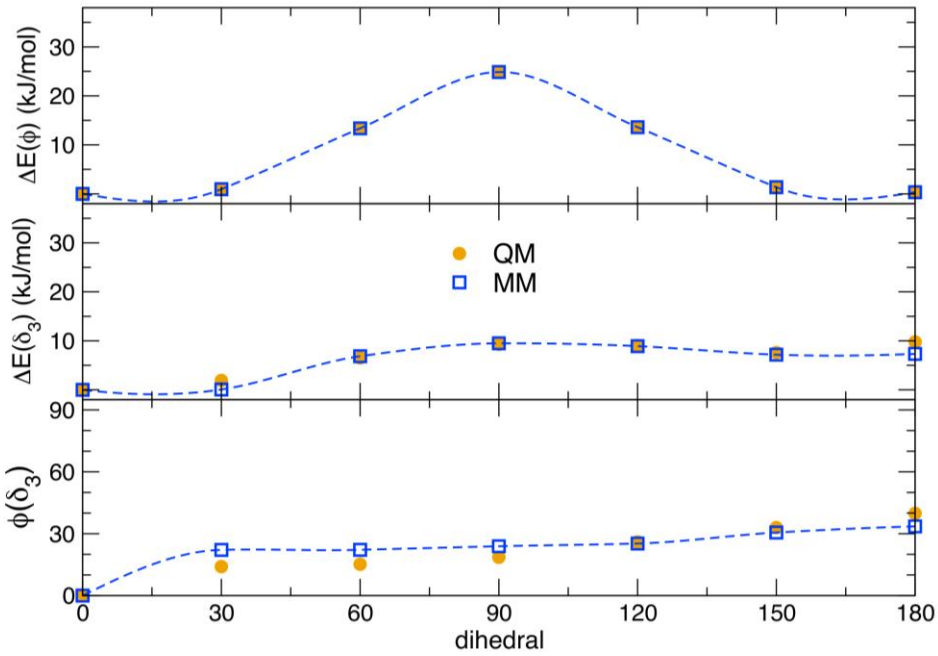

Figure 4 Comparison of QM (magenta circles) and MM (blue squares) internal torsional energies of pelargonidin, computed along the relaxed scans of ϕ (top panel) and δ_3 (middle panel) dihedrals, performed with DFT and Joyce FF, respectively. Bottom panel: values taken by the ϕ dihedral during the δ_3 relaxed scan, performed at QM (DFT, magenta) and MM (Joyce FF, blue), respectively.

Cyanidin

The FF for cyanidin was also parameterized through the JOYCE protocol (see Supporting Information for details) and validated by a careful check, as performed with pelargonidin, of the QM and MM torsional profiles and their relative optimized structures.

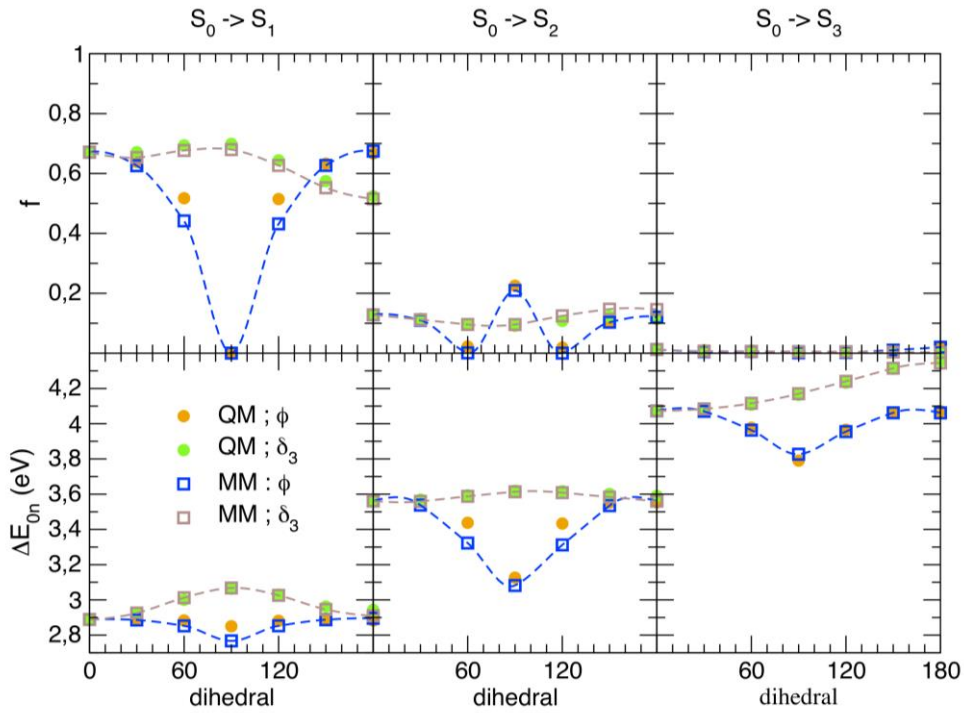

Figure 5 Vertical transition energies (bottom panels, ΔE_{0n}, n = 1,2,3) and oscillator strengths (top panel, f) for the first three electronic transitions, computed for different pelargonidin structures. Each structure was obtained through a

relaxed scans of the φ (orange and blue symbols) and the δ_3 (green and brown symbols) dihedrals, performing optimization either at QM (CAM-B3LYP/6-311+G(2d,2p), solid circles) or MM (JOYCE FF, empty squares) level.

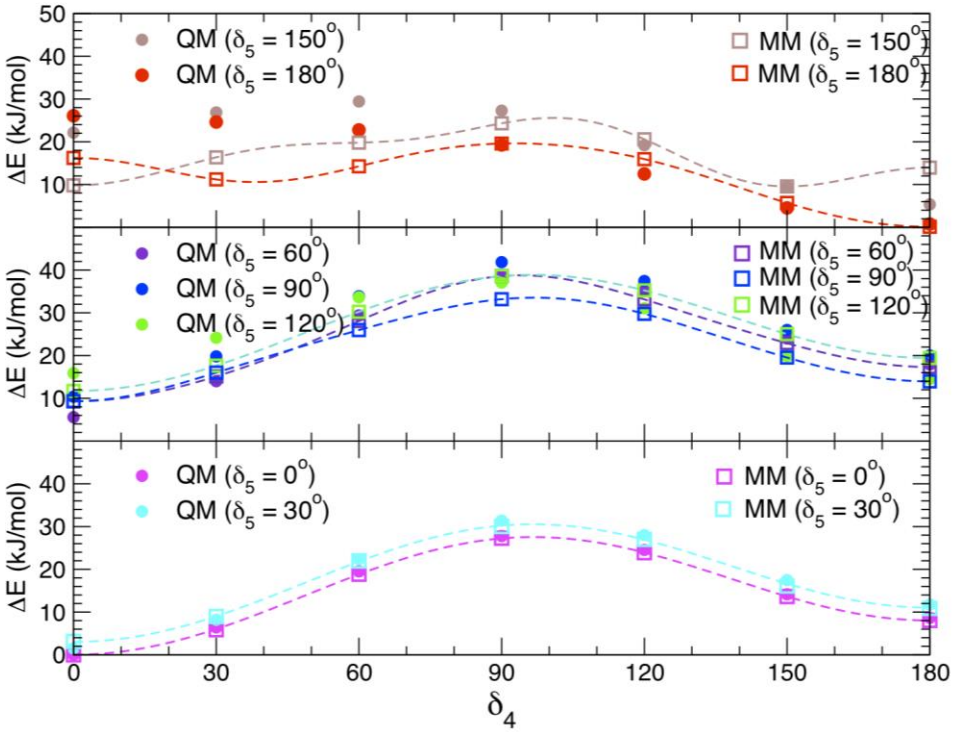

Figure 6 Comparison of QM (circles) and MM (squares) internal torsional energies of cyanidin, computed along relaxed scans of the δ_4 dihedral performed with the CAM-B3LYP functional or with the Joyce FF, respectively. Bottom panel: δ_4 torsional profiles with δ_5 constrained to 0° (magenta) or 30° (cyan). Middle panel: δ_4 torsional profiles with δ_5 constrained to 60° (violet), 90° (blue) or 120° (green). Top panel: δ_4 torsional profiles with δ_5 constrained to 150° (red) or 180° (brown).

Two additional tests were also performed, to verify: i) the capability of the FF to take into account the effect of the coupling between δ_4 and δ_5; ii) the accuracy of the MM optimized geometries in reproducing the TDDFT vertical energies yielded by the structures minimized at QM level with the same restraints. The first of these requirements is satisfied, as appears by looking at Figure 6, where the δ_4 MM scans at given δ_5 are compared to their QM counterparts. More in detail, an excellent agreement is registered for δ_5 angles less than 90°, whereas the FF accuracy diminishes for the less stable conformers. However, their weight in standard conditions is expected to be limited according to Boltzmann populations. Furthermore, the visual examinations of such repulsive conformers as well as the root mean squared deviations with respect their QM counterparts, reveal the agreement over the geometry seems better than that achieved on energies. To confirm these finding, TD vertical energies were computed for selected conformers and compared with those obtained with the analogous DFT optimized geometries. From the results summarized in the Table 6, it is evident that MM and QM structure are practically identical (as they give the same transition energies and f), with the exception of the most repulsive ϕ=90° conformer. Also in this case however, the effect of such kind of conformers on the final spectrum is expected to be negligible, both for its scarce probability to be populated and for its null oscillator strength.

ϕ	δ_4	δ_5	QM			MM		
			ΔE (kJ/mol)	**ΔE_{01}** (eV)	**f**	**ΔE** (kJ/mol)	**ΔE_{01}** (eV)	**f**
0	0	180	26.0	2.85	0.681	16.0	2.85	0.691
0	180	0	8.5	2.81	0.720	8.0	2.84	0.684
90	0	0	24.2	2.61	0.004	23.6	2.47	0.000
150	0	0	2.5	2.80	0.551	1.9	2.81	0.640

Table 6 Conformational energies (ΔE), vertical transition energies to the first excited state (ΔE_{01}) and oscillator strengths (f) for selected cyanidin conformations, obtained after a constrained optimization carried out by fixing φ, δ_4 and δ_5 dihedrals and performing minimization at either QM (CAM-B3LYP/6-311+G(2d,2p)) or MM (JOYCE FF) level.

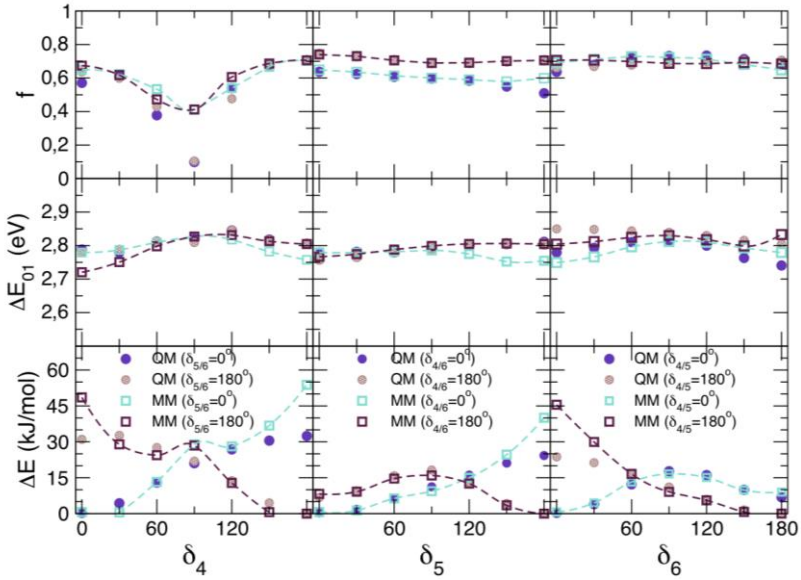

Figure 7 Conformational energies (ΔE), vertical transition energies to the first excited state (ΔE_{01}) and oscillator strengths (f) for selected delphinidin conformations, obtained after a constrained optimization carried out by fixing δ_4, δ_5 and δ_6 dihedrals and performing minimization at either QM (CAM-B3LYP/6-311+G(2d,2p)) or MM (Joyce FF) level.

Delphinidin

The FF parameterized for delphinidin (see Supporting Information for details) was validated with all the tests performed for pelargonidin and cyanidin.

By looking at Figure 7, it appears as most probable conformers are again well described by the FF, whereas more repulsive geometries (ΔE > 30 kJ/mol) are more difficult to be described. Nonetheless, as mentioned for cyanidin, the agreement with respect to reference QM data achieved for delphinidin geometries is definitely better than that found for energetics. In fact, as displayed in the top and middle panel of Figure 7, all TDDFT energies and oscillator strengths (with the exception of the δ_4 = 90° conformer) are reproduced with good accuracy, testifying the reliability of the subtending MM structures.

3.5 MD simulations

MD simulations were performed, exploiting the FFs parameterized for the three anthocyanidins, over systems composed by one dye and one Cl^- counterion, both solvated in ~800 ethanol molecules at 300 K and 1 atm. A preliminary analysis of some relevant thermodynamic properties, averaged over the 5 ns production runs performed for each dye, is reported in Table 7. All systems appear equilibrated to very

similar densities, not far from the values reported for pure ethanol at 298 K, namely 796 and 785 kg/m^3, obtained[66] by OPLS FF simulations and experiment, respectively. The solvation energy ΔE_{solv}, shows a clear trend along the series, increasing as expected with the number of OH groups. By looking at the different contributions to the solvation, *i.e.* to the total LJ and charge-charge interaction energies between the dye and the solvent, it is also evident that it is the Coulomb term that essentially determines such increase, while the LJ part is practically constant. This suggests that the augmented solvation energy might be connected to a larger number of HB, formed between the dye and the surrounding ethanol molecules. This hypothesis is confirmed by looking at the last row of Table 7, where the average number of HB n_{HB} is computed over the stored trajectories, based on geometrical factors:[67] n_{HB} increase from 4 to 6 along the series, pairing the number of OH groups in the dye.

	Pelargonidin	**Cyanidin**	**Delphinidin**
ρ (kg/m^3)	782	785	782
$\boldsymbol{\Delta E_{solv}}$ **(kJ/mol)**	299	336	360
$\boldsymbol{\Delta E_{solv}^{LJ}}$ **(kJ/mol)**	99	97	102
$\boldsymbol{\Delta E_{solv}^{Coul}}$ **(kJ/mol)**	200	239	258
$\boldsymbol{\langle n_{HB} \rangle}$	3.9	5.1	5.9

Table 7 Average density (ρ), number of HBs ($\langle n_{HB} \rangle$) and solvation energy ($\boldsymbol{\Delta E_{solv}}$), and its LJ and charge-charge contributions, $\boldsymbol{\Delta E_{solv}^{LJ}}$ and $\boldsymbol{\Delta E_{solv}^{Coul}}$) achieved in MD simulations at 300 K and 1 atm performed on systems composed by one anthocyanidin dye solvated with ~800 ethanol molecules

Considering the different local solvation environments, it can be expected that the dyes conformation "reacts" in different ways, to enhance the interaction with the surrounding ethanol molecules. This should reflect in conformational changes on the "softer" part of the molecules, *i.e.* in rotation around the identified flexible dihedrals. In Figure 8, the distributions P, computed over the stored MD trajectories, are displayed for all three homologues. As could be expected, the larger differences among the investigated anthocyanidins are found for those dihedrals located on the B ring, where the chemical structure differs the most. Indeed, the ϕ and δ_3 angles show almost identical profiles along the series, indicating for all homologues rather large fluctuations centered around 0°. Slight differences arise in the δ_1 and δ_2 distributions, where the 0° conformer shows a not negligible population only for pelargonidin. Concerning the distribution of the dihedrals located on the B ring (δ_4 to δ_6), significant differences arise for delphinidin, whereas the δ_4 profiles of pelargonidin and cyanidin are almost equal. Due to the vicinity of the three OH groups, and to the presence of a third, H-bonded, ethanol molecule, the δ_4 dihedral in delphinidin is displaced from its equilibrium value (180°) by almost 30° and shows larger amplitude motions, which, at variance with the smaller homologues, eventually populate also the ±30° region. Similarly, also δ_5 and δ_6 dihedral profiles show larger amplitude motion with respect of cyanidin's δ_5, again showing maxima (±30°) slightly displaced from the minimum energy value, indicating a less stable situation in the established HB network.

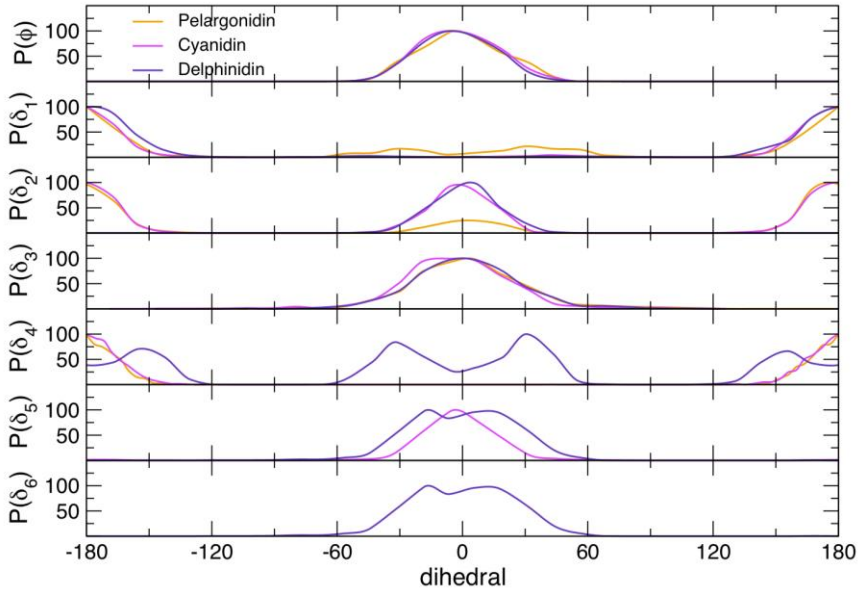

Figure 8 Dihedral distributions (P) achieved for each investigated dye during MD simulations of the dye solvated by ethanol molecules.

The last observation is confirmed by looking at the time evolution of n_{HB}, displayed in Figure 9 together with the distribution of some geometrical quantities relevant in the description the HB network. While for pelargonidin n_{HB} always remains close to its average value, larger fluctuations arise for the other two homologues, delphinidin showing the largest (see also Figure F in the Supporting Information), as the result of a competition established between the intermolecular HBs and the intramolecular ones, formed among the OH groups present on the ring. However, by looking at the last panels of Figure 9, it appears as such competition affects more the aforementioned dihedral motions rather than the HB network features, whose geometrical distributions appear almost unaltered.

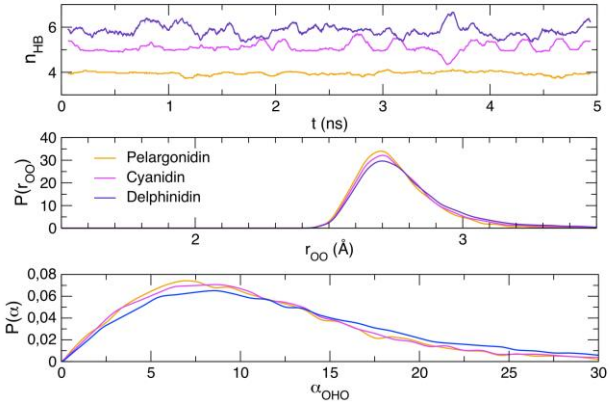

Figure 9 HB network settled around each investigated dye during MD simulations. Top panel: number of HBs (n_{HB}) vs simulation time (t). Middle panel: distribution population of the donor-acceptor distance r_{OO}. Bottom panel: distribution population of the donor-acceptor angle α_{OHO}.

A final deeper insight in the HB network established by ethanol molecules around each dye can be obtained by looking at the atom-atom pair correlation functions, computed between either the hydroxyl dye's Hydrogens (H_n, n=1-6) and ethanol Oxygen atom (O_{solv}, left panels of Figure 10) or between the dye's

Oxygen (O_n, n=1-6) and the ethanol's proton (H_{solv}, right panel). A very similar picture emerges for all dyes, where the ethanol molecules behave as HB acceptors, whereas the dye acts as proton donor: in fact, the g(r) peaks for the H_n-O_{solv} pairs are always found at smaller distances (~1.8 Å) with respect to the O_n-H_{solv} ones. Finally, it could be worth noticing that the lesser stability of the HBs formed by the O_4H_4 group in delphinidin, due to both the steric hindrance of the neighboring OH groups and to the intramolecular HB competition, is also confirmed by this analysis, considering that the $g_{H4\text{-}Osolv}$ peak is less intense and shifted to larger distance with respect to the smaller homologues.

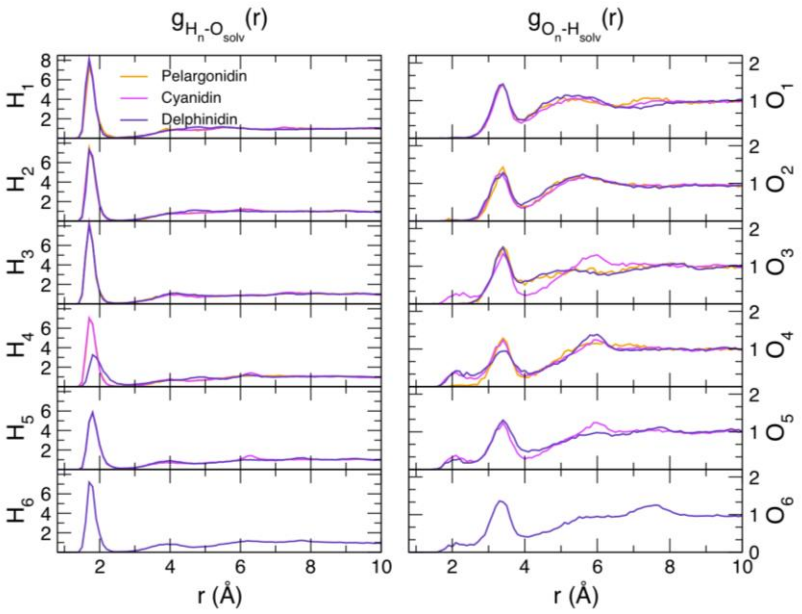

Figure 10 Atom-atom pair correlation functions, $g_{nm}(r)$ between H_n atom of the dye and an ethano Oxygen (n = H_n and m = O_{solv}, left panels) and between O_n atom of the dye and an ethanol Hydrogen (n = O_n and m = H_{solv}, right panels). All functions were computed over the MD trajectories obtained at 300 K and 1 atm for each anthocyanidin.

3.6 Absorption Spectra

Before computing the final absorption spectra for the three investigated anthocyanidins, some preliminary tests were performed on some technical factors that might affect the calculations, as detailed in the SI (see Figure G and H). Based on these results, all calculations have been performed for each homologue on samples of 200 MD frames, convoluting each of the TDDFT stick spectra with a Gaussian function with HWHM of 0.05 eV. A simple average over the resulting set of functions gives the broadened absorption spectrum. To finally compare the simulated spectra with those obtained in the experiment, as displayed in Figure 12, the TDDFT lineshape is corrected by the Δ^{corr} values reported in Table 5 for each homologue. It is important to recall that this shift is not an empirical quantity applied to match the experimental result, but a correction established on the basis of high level ab initio calculations, thus preserving the intrinsic predictive features of the adopted protocol.

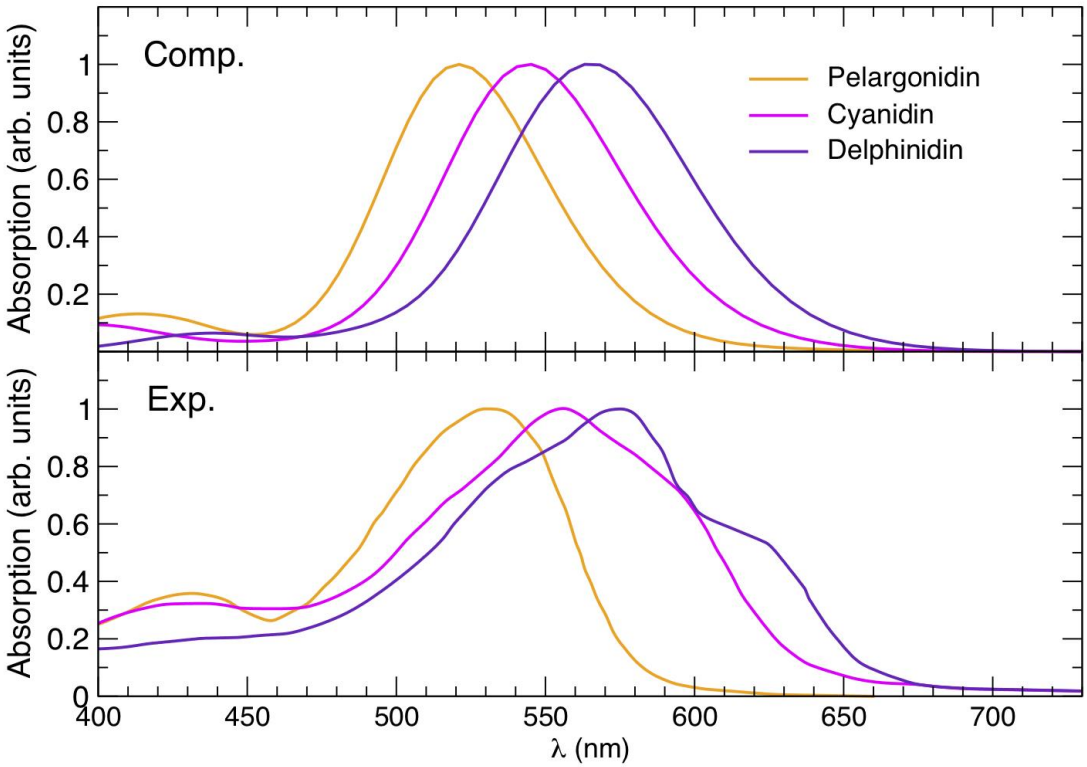

Figure 11 Computed (solid lines) and experimental[30] (dotted/dashed lines) absorption spectra for the three anthocyanidins solvated in ethanol. All computed signals were obtained by averaging the TDDFT vertical energy transitions to the first three excited states, calculated for 200 snapshots extracted from the MD trajectories and convoluted with a 0.05 HWHM Gaussian function.

The agreement of the computed results with the experiment is good. Despite the scale in wavelength, which somehow visually amplifies the differences, the position of the computed bands well matches the experiments. The red shift of the most intense band observed on going from the smallest to the largest homologues is well reproduced by the computations. This also holds for the less intense band at lower wavelength, although the computed intensity is slightly underestimated. A further minor difference is in the computed line-shape profile, which, for every species, has a smaller width with respect to the measured one. For this aspect, it should be taken into account that the band profile is obtained by MD simulations and therefore the effect of nuclear vibrations is then included classically. As discussed elsewhere,[52] a consequence of neglecting quantum effects of the solute vibrations is indeed a significant underestimation of the homogeneous broadening. Yet, in the present work, the former artifact is partially covered by the inhomogeneous component of the broadening, due to the coupling with the solvent which makes the width difference between computed and experimental signals much less evident.

4. Conclusions

In light of the importance that organic dyes of natural origin, occurring in many foods and beverages, may have in several fields as biology, nutrition and enology, as well as in the construction of low-cost and low-environmental impact DSSC, we have reported on an accurate computational study of the absorption spectra

in the visible for three representative members of the class of antocyanidins in ethanol.

We have followed a MD/QM computational route where TDDFT spectra are computed for a wide series of molecular conformations, randomly selected as snapshots from MD in solution, and averaged to obtain the spectrum to be compared with available experimental data.

The choice of the most suitable functional was driven by the comparison, for the equilibrium molecular geometry, of TDDFT spectra with those obtained by an accurate post-HF method such as CI-MRPT2. Despite B3LYP results are closer to CI-MRPT2 ones for planar configurations, we have chosen CAM-B3LYP since it was found to be more well-behaved with the torsion around the angle ϕ. The energy position of the TDFT excitation energies was shifted in order to match those obtained by CI-MRPT2. This has the advantage of making the computational results independent from the experimental measurements, allowing an unbiased comparison between theory and experiments.

For each of the three species, which differ for the number of OH groups in the ring B, in ethanol solution, we have calibrated a system-specific FF to be used in MD simulations. This allows us to take into account at classical level, both intramolecular and solute solvent nuclear effects specifically for each species.

CAM-B3LYP spectra for each molecular configuration pertaining to the 200 selected snapshots of MD, replacing solvent molecules with point charges, were then computed and compared to the experimental data available. The overall result, taking advantage from the specificity of the FF, is indeed very good, both in terms of position, intensity and lineshape profile of the bands in the visible. In addition, MD simulations, although at a classical level, allow the analysis of the solvation pattern for the three different anthocyanidins.

We are confident that the accuracy of the present approach and the possibility to extend its application to larger and more complex systems paves the way to the description of a full organic DSSC.

Aknowledgements

The research leading to these results has received funding from the Italian Ministry of Instruction, University and Research (MIUR), through PRIN 2010-11, 2010PFLRJR (PROxi), and 2010FM738P. Dr. Javier Cerezo is gratefully acknowledged by GP for the many useful discussions.

References

1. O'Regan B, Grätzel M (1991) A low-cost, high-efficiency solar cell based on dye-sensitized colloidal TiO2 films. Nature 353:737.

2. Monat JE, Rodriguez JH, McCusker JK (2002) Ground- and Excited-State Electronic Structures of the Solar Cell Sensitizer Bis(4,4‘-dicarboxylato-2,2‘-bipyridine)bis(isothiocyanato)ruthenium(II). J Phys Chem A 106 7399.

3. De Angelis F, Fantacci S, Selloni A, Grätzel M, Nazeeruddin MK (2007) Influence of the sensitizer adsorption mode on the open-circuit potential of dye-sensitized solar cells. Nano Lett 7:3189-3195.

4. Graetzel M (2009) Recent Advances in Sensitized Mesoscopic Solar Cells. Acc Chem Res 42:1788-1798.
5. De Angelis F, Fantacci S, Selloni A, Nazeeruddin MK, Grätzel M (2011) First-Principles Modeling of the Adsorption Geometry and Electronic Structure of Ru(II) Dyes on Extended TiO2 Substrates for Dye-Sensitized Solar Cell Applications. J Phys Chem C 114:6054-6061.
6. Bignozzi CA, Argazzi R, Boaretto R, Busatto E, Carli S, Ronconi F, Caramori S (2013) The role of transition metal complexes in dye sensitized solar devices. Coord Chem Rev 257:1472-1492.
7. Kuang D, Comte P, Zakeeruddin SM, Hagberg DlP, Karlsson KM, Sun L, Nazeeruddin MK, Grätzel M (2011) Stable dye-sensitized solar cells based on organic chromophores and ionic liquid electrolyte. Sol Energy 85:1189.
8. Odobel F, Le Pleux L, Pellegrin Y, Blart E (2010) New photovoltaic devices based on the sensitization of p-type semiconductors: challenges and opportunities. Acc Chem Res 43:1063-1071.
9. Pastore M, Fantacci S, De Angelis F (2010) Ab Initio Determination of Ground and Excited State Oxidation Potentials of Organic Chromophores for Dye-Sensitized Solar Cells. J Phys Chem C 114:22742.
10. Plannels M, Pellejà L, Clifford JN, Pastore M, De Angelis F, López N, Marder S, Palomares E (2011) Energy levels, charge injection, charge recombination and dye regeneration dynamics for donor–acceptor π-conjugated organic dyes in mesoscopic TiO2 sensitized solar cells. Energy Environ Sci 4:1820.
11. Hagfeldt A, Boschloo G, Sun L, Kloo L, Pettersson H (2010) Dye-Sensitized Solar Cells. Chem Rev 110:6595.
12. Li L-L, Diau EW-G (2013) Porphyrin-sensitized solar cells. Chem Soc Rev 42:291-304.
13. Jinchu I, Sreekala CO, Sreelatha KS (2014) Dye Sensitized Solar Cell using Natural Dyes as Chromophores - Review. In: Pandikumar A, Jothilakshmi R (eds), vol 771. Materials Science Forum. TRANS TECH PUBLICATIONS LTD, LAUBLSRUTISTR 24, CH-8717 STAFA-ZURICH, SWITZERLAND, pp 39-51. doi:10.4028/www.scientific.net/MSF.771.39
14. Narayan MR (2012) Review: Dye sensitized solar cells based on natural photosensitizers. Renev Sust Energ Rev 16:208-215.
15. Senthil TS, Muthukumarasamy N, Velauthapillai D, Agilan S, Thambidurai M, Balasundaraprabhu R (2011) Natural dye (cyanidin 3-O-glucoside) sensitized nanocrystalline TiO2 solar cell fabricated using liquid electrolyte/quasi-solid-state polymer electrolyte. Renew Energ 36:2484-2488.
16. Castañeda-Ovando A, Pacheco-Hernández MdL, Páez-Hernández ME, Rodríguez JA, Galán-Vidal CA (2009) Chemical studies of anthocyanins: A review. Food Chem 113:859-871.
17. Falcone Ferreyra ML, Rius SP, Casati P (2012) Flavonoids: biosynthesis, biological functions, and biotechnological applications. Front Plant Sci 3:222-222.
18. Halbwirth H (2010) The creation and physiological relevance of divergent hydroxylation patterns in the flavonoid pathway. Int J Mol Sci 11:595-621.
19. Rustioni L, Di Meo F, Guillaume M, Failla O, Trouillas P (2013) Tuning color variation in grape anthocyanins at the molecular scale. Food Chem 141:4349-4357.
20. Trouillas P, Di Meo F, Gierschner J, Linares M, Sancho-García JC, Otyepka M (2015) Optical properties of wine pigments: theoretical guidelines with new methodological perspectives. Tetrahedron 71:3079-3088.
21. Shahid M, Shahid ul I, Mohammad F (2013) Recent advancements in natural dye applications: a review. J Clean Prod 53:310-331.
22. Calogero G, Marco GD (2008) Red Sicilian orange and purple eggplant fruits as natural sensitizers for dye-sensitized solar cells. Sol Energy Mater Sol Cells 92:1341-1346.
23. Gómez-Ortíz NM, Vázquez-Maldonado IA, Pérez-Espadas AR, Mena-Rejón GJ, Azamar-Barrios JA, Oskam G (2010) Dye-sensitized solar cells with natural dyes extracted from achiote seeds. Sol Energy Mater Sol Cells 94:40-44.

24. Ekanayake P, Kooh MRR, Kumara NTRN, Lim A, Petra MI, Voo NY, Lim CM (2013) Combined experimental and DFT–TDDFT study of photo-active constituents of Canarium odontophyllum for DSSC application. Chem Phys Lett 585:121-127.
25. Ludin NA, Mahmoud AMA-A, Mohamad AB, Kadhum AAH, Sopian K, Karim NSA (2014) Review on the development of natural dye photosensitizer for dye-sensitized solar cells. Renev Sust Energ Rev 31:386-396.
26. Goto T, Kondo T (1991) Struktur und molekulare Stapelung von Anthocyanen — Variation der Blutenfarben. Angew Chem 103:17--33.
27. Di Meo F, Sancho Garcia JC, Dangles O, Trouillas P (2012) Highlights on Anthocyanin Pigmentation and Copigmentation: A Matter of Flavonoid π-Stacking Complexation To Be Described by DFT-D. J Chem Theor Comput 8:2034-2043.
28. Brouillard R, Mazza G, Saad Z, Albrecht-Gary AM, Cheminat A (1989) The co-pigmentation reaction of anthocyanins: a microprobe for the structural study of aqueous solutions. J Am Chem Soc 111:2604--2610.
29. Harborne JB (1958) Spectral methods of characterizing anthocyanins. The Biochemical journal 70:22-28.
30. Dai Q, Rabani J (2002) Photosensitization of nanocrystalline TiO 2 films by anthocyanin dyes. J Photochern Photobiol A 148:17-24.
31. Liu Z (2008) Theoretical studies of natural pigments relevant to dye-sensitized solar cells. J Mol Struct: THEOCHEM 862:44-48.
32. Calzolari A, Varsano D, Ruini A, Catellani A, Tel-Vered R, Yildiz HB, Ovits O, Willner I (2009) Optoelectronic Properties of Natural Cyanin Dyes. J Phys Chem A 113:8801-8810.
33. Ge X, Calzolari A, Baroni S (2015) Optical properties of anthocyanins in the gas phase. Chem Phys Lett 618:24-29.
34. Anouar EH, Gierschner J, Duroux J-L, Trouillas P (2012) UV/Visible spectra of natural polyphenols: A time-dependent density functional theory study. Food Chem 131:79-89.
35. Ge X, Timrov I, Binnie S, Biancardi A, Calzolari A, Baroni S (2015) Accurate and inexpensive prediction of the color optical properties of anthocyanins in solution. J Phys Chem A 119:3816-3822.
36. Soto-Rojo R, Baldenebro-López J, Flores-Holguín N, Glossman-Mitnik D (2014) Comparison of several protocols for the computational prediction of the maximum absorption wavelength of chrysanthemin. J Mol Model 20:2378-2378.
37. Millot M, Di Meo F, Tomasi S, Boustie J, Trouillas P (2012) Photoprotective capacities of lichen metabolites: a joint theoretical and experimental study. J Photochem Photobiol B 111:17-26.
38. Malcıoğlu OB, Calzolari A, Gebauer R, Varsano D, Baroni S (2011) Dielectric and thermal effects on the optical properties of natural dyes: a case study on solvated cyanin. J Am Chem Soc 133:15425--15433.
39. Sakata K, Saito N, Honda T (2006) Ab initio study of molecular structures and excited states in anthocyanidins. Tetrahedron 62:3721-3731.
40. Barone V, Ferretti A, Pino I (2012) Absorption spectra of natural pigments as sensitizers in solar cells by TD-DFT and MRPT2: protonated cyanidin. Phys Chem Chem Phys:16130-16137.
41. Avila Ferrer FJ, Cerezo J, Stendardo E, Improta R, Santoro F (2013) Insights for an Accurate Comparison of Computational Data to Experimental Absorption and Emission Spectra: Beyond the Vertical Transition Approximation. J Chem Theory Comput 9:2072--2082.
42. Muniz-Miranda F, Pedone A, Battistelli G, Montalti M, Bloino J, Barone V (2015) Benchmarking TD-DFT against Vibrationally Resolved Absorption Spectra at Room Temperature: 7-Aminocoumarins as Test Cases. J Chem Theory Comput 11:5371-5384.
43. Charaf-Eddin A, Cauchy T, Felpin Fc-X, Jacquemin D (2014) Vibronic spectra of organic electronic chromophores. RSC Adv 4:55466--55472.
44. Lopez GV, Chang C-H, Johnson PM, Hall GE, Sears TJ, Markiewicz B, Milan M, Teslja A (2012) What is the best DFT functional for vibronic calculations? A comparison of the calculated

vibronic structure of the S1-S0 transition of phenylacetylene with cavity ringdown band intensities. J Phys Chem A 116:6750--6758.
45. Jacquemin D, Brémond E, Planchat Al, Ciofini I, Adamo C (2011) TD-DFT Vibronic Couplings in Anthraquinones: From Basis Set and Functional Benchmarks to Applications for Industrial Dyes. J Chem Theory Comput 7:1882--1892.
46. Dierksen M, Grimme S (2004) The Vibronic Structure of Electronic Absorption Spectra of Large Molecules: A Time-Dependent Density Functional Study on the Influence of "Exact" Hartree−Fock Exchange. J Phys Chem A 108:10225--10237.
47. Cacelli I, Ferretti A, Prampolini G (2014) Perturbative Multireference Configuration Interaction (CI-MRPT2) Calculations in a Focused Dynamical Approach: A Computational Study of Solvatochromism in Pyrimidine. J Phys Chem A.
48. De Mitri N, Monti S, Prampolini G, Barone V (2013) Absorption and Emission Spectra of a Flexible Dye in Solution: A Computational Time-Dependent Approach. J Chem Theory Comput 9:4507--4516.
49. De Mitri N, Prampolini G, Monti S, Barone V (2014) Structural, dynamic and photophysical properties of a fluorescent dye incorporated in an amorphous hydrophobic polymer bundle. Phys Chem Chem Phys 16:16573--16587.
50. Pedone A, Prampolini G, Monti S, Barone V (2011) Realistic Modeling of Fluorescent Dye-Doped Silica Nanoparticles: A Step Toward the Understanding of their Enhanced Photophysical Properties. Chem Mater 23:5016--5023.
51. Marenich AV, Cramer CJ, Truhlar DG (2015) Electronic absorption spectra and solvatochromic shifts by the vertical excitation model: Solvated clusters and molecular dynamics sampling. J Phys Chem B 119:958--967.
52. Cerezo J, Prampolini G, Santoro F (2015) Comparing Classical Approaches with Empirical or Quantum-Mechanically Derived Force-Fields for the Simulations of Electronic Lineshapes: Apllication to Coumarin Dyes. Theor Chem Acc submitted.
53. Cacelli I, Prampolini G (2007) Parametrization and validation of intramolecular force fields derived from DFT calculations. J Chem Theor Comput 3:1803-1817.
54. Barone V, Cacelli I, De Mitri N, Licari D, Monti S, Prampolini G (2013) JOYCE and ULYSSES: integrated and user-friendly tools for the parameterization of intramolecular force fields from quantum mechanical data. Phys Chem Chem Phys 15:3736-3751.
55. Cacelli I, Ferretti A, Prampolini G (2015) Perturbative Multireference Configuration Interaction (CI-MRPT2) Calculations in a Focused Dynamical Approach: A Computational Study of Solvatochromism in Pyrimidine. J Phys Chem A 119:5250-5259.
56. Cacelli I, Ferretti A, Prampolini G, Barone V (2015) BALOO: A Fast and Versatile Code for Accurate Multireference Variational/Perturbative Calculations. J Chem Theor Comput 11:2024-2035.
57. Laurent AlD, Jacquemin D (2013) TD-DFT benchmarks: A review. Int J Quantum Chem 113:2019--2039.
58. Jacquemin D, Bahers TL, Adamo C, Ciofini I (2012) What is the best atomic charge model to describe through-space charge-transfer excitations? Phys Chem Chem Phys 14:5383--5388.
59. Tomasi J, Mennucci B, Cammi R (2005) Quantum mechanical continuum solvation models. Chem Rev 105:2999-3093.
60. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery Jr. JA, Peralta JE, Ogliaro F, Bearpark MJ, Heyd J, Brothers EN, Kudin KN, Staroverov VN, Kobayashi R, Normand J, Raghavachari K, Rendell AP, Burant JC, Iyengar SS, Tomasi J, Cossi M, Rega N, Millam NJ, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RL, Morokuma K, Zakrzewski

VG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels AD, Farkas Ö, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ (2009) Gaussian 09. Gaussian, Inc., Wallingford, CT, USA
61. Gordon MS, Schmidt MW (2005) GAMESS. Theory and Applications of Computational Chemistry: the first forty years. Elsevier, Amsterdam
62. Cimiraglia R (1984) Second Order Perturbation Correction to CI Energies by Use of Diagrammatic Techniques: An Improvement to the Cipsi Algorithm. J Chem Phys 83:174.
63. Jorgensen WL, Maxwell DS, Tirado-rives J (1996) Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. J Am Chem Soc 7863:11225--11236.
64. Marenich AV, Jerome SV, Cramer CJ, Truhlar DG (2012) Charge Model 5: An Extension of Hirshfeld Population Analysis for the Accurate Description of Molecular Interactions in Gaseous and Condensed Phases. Journal of Chemical Theory and Computation 8:527--541.
65. Aaqvist J (1990) Ion-Water Interaction Potentials Derived from Free Energy Perturbation Simulations. J Phys Chem 94:8021-8024.
66. Caleman C, van Maaren PJ, Hong M, Hub JS, Costa LT, van der Spoel D (2012) Force Field Benchmark of Organic Liquids: Density, Enthalpy of Vaporization, Heat Capacities, Surface Tension, Isothermal Compressibility, Volumetric Expansion Coefficient, and Dielectric Constant. J Chem Theory Comput 8:61--74.
67. van der Spoel D, Lindahl E, Hess B, van Buuren AR, Apol E, Meulenhoff P, Tieleman D, Sijbers A, Feenstra K, van Drunen R, Berendsen H (2010) GROMACS4.5.
68. Bussi G, Donadio D, Parrinello M (2007) Canonical sampling through velocity rescaling. The Journal of chemical physics 126:014101.
69. Parrinello M (1981) Polymorphic transitions in single crystals: A new molecular dynamics method. Journal of Applied Physics 52:7182.