Détermination d'un protocole de calcul pour la prédiction de spectres UV-vis de molécules en solution

Bernardino Tirri

PSL UNIVERSITÉ PARIS

THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL

Préparée à Chimie ParisTech

Determination of a computational protocol for the prediction of UV-vis spectra of molecules in solution

Soutenue par
Bernardino Tirri
Le 4 Mars 2022

Ecole doctorale n° 388
Chimie Physique et Chimie Analytique de Paris-Centre

Spécialité
Chimie physique

Composition du jury :

Esmail, ALIKHANI Professeur, Sorbonne Université	*Président*
Dorothée, BERTHOMIEU Directeur de recherche, Institut Charles Gerhard Montpelier	*Rapporteur*
Maria Cristina, MENZIANI Professor, Université de Modène et de Reggio d'Émilie	*Rapporteur*
Malgorzata, BICZYSKO Professor, Université de Shanghai	*Examinateur*
Carlo, ADAMO Professeur, Chimie ParisTech	*Directeur de thèse*

Chimie Paris ParisTech | PSL

Table of Contents

Chapter 1

1. Introduction and thesis framework

Dyes are attractive for their tuneable electronic properties, allowing broad applications ranging from the design of cutting-edge technologies to textile industries. Depending upon the production sources, dyes can be classified as *natural* or *synthetic* [1]. Being derived from plant and insects, natural dyes are perceived as harmless for the environment, but they are expensive and can be toxic. On the other hand, several synthetic dyes are derived from petrochemical compounds [2] and can be produced in large quantity. However, their impact on both the human beings and the environment cannot be neglected. An open issue exists: how to make the right choice?

The industry has recently moved to reduce the environmental impact of the dye productions especially through the sourcing of renewable raw materials. Developing greener dyes is a timely endeavour and could have a massive impact especially on the textile dye industry. Switching from petrochemical synthesis to biotechnological one, the production could result in a 100-fold reduction of energy use and allows to avoid 100 kilograms of heavy oils and 800 litres of water per kilogram of dye [3]. Compared with regular petrochemical pigments, bio-based dyes lower the carbon footprint by a 9-fold. Applied to the million-ton-scale dye industry, this transition has the same potential in reducing greenhouse gases than cutting the yearly emissions of a million vehicles. A possible green strategy consists in employing enzymatic process [4]. Nowadays, several microorganisms are known to make coloured compounds ranging from orange (*blakeslea trispora*) to violet (*Chromobaterium violaceum*), and red (*Serratia marcesccens*) [5]. This is the goal of the PILI company. Here, bio-sourced compounds are modified by using green chemistry processes for reaching a diversity in the colour palette. In this regard, categorization schemes are useful to distinguish colours [6]. Although used for different purpose [7], these schemes are strongly related to the electronic absorption spectrum. In a nutshell, when a molecule absorbs in the range of visible radiations, it appears coloured and its colour is due to the resultant of the transmitted radiations. The recorded transmitted spectrum is linked to the absorption one, thus the "visible" and the complementary colours are also linked between them. Such colours have to be characterized, because human beings have different colour perceptions. To do it, the Euclidean distances between the colours, the lightness, the chroma (i.e., the amount of saturation) and the

hue angle (i.e., the colour tone) constitute the key elements for distinguishing colours and allow the dyes-industries to classify chromophores.

Taking into account the environmental impact, dyes industries have to broaden their spectrum of interests making more efficient the use of their research, including the process optimization. To speed-up the industrial chain, one can investigate both the photophysical (the physical process induced by light) and photochemical (light induce chemical transformations) process in order to design new chromophores having the desired colour. Upon light-induced processes, molecules leave spectroscopic signatures that are related to their molecular electronic structure. Within this framework, methods aimed at solving the electronic structure problem are useful for both understanding and improving the optical properties of dyes.

Over the years, first-principles methods have been shown as valid tools, and not surprisingly, they are commonly employed for both molecular screening and in silico-design of new chromophores of industrial interest. Among the methods, density functional theory (DFT) and its time dependent extension (TD-DFT) are currently used to investigate the ground and excited state properties of such systems. In the last decades, TD-DFT has extensively demonstrated its capacity to provide excellent solutions in regards to the excited states for several classes of commercial organic compounds of industrial purpose, including anthraquinones and indigos [8,9,10].

Besides the level of theory employed for obtaining the absorption energies, the electronic spectrum is affected by two other factors: the environment and the 3D rearrangement of the chromophore. An accurate modelling of these factors allows the prediction of the Uv-vis spectra with excellent accuracy. Concerning the chemical surrounding, solvent inclusion results to be decisive for the electronic transition energies, especially when the polar solvent is accounted for. The solvent has been modelled in conjunction with DFT simulations by means of a continuum model [11]. Finally, an electronic transition is modulated by a slight change the nuclear configuration form the ground state to the excited state. More precisely, the coupling between the electronic and nuclear degree of freedom (the vibrational coupling) deeply influences the band broadening. An approach to compute the vibronic fine structure consists in using the Franck-Condon (FC) approximation which leads to large computational cost. Despite this drawback, computational studies have shown that the vibronic fine structure cannot be neglected in the quantitative prediction of colours [12], however, as the time-consuming steps cannot be ignored, new computational protocol have to be developed.

While for *rigid chromophores* (i.e., systems characterized by tiny molecular flexibility) the vibrational progression can accurately be predicted using the FC approximation, *flexible chromophores*

are more intricate to model. For these systems, the ground and excited state geometries are quite different, and potential energies surfaces can present several local minima close to each other which prevent the use of the FC approximation and, consequentially, the simulation of their UV-vis spectra represents a current challenge from computational point of view. To solve such intricate problem two possible approaches, that rotate around the phase space sampling, have been developed. *Molecular Dynamic* (MD) simulations, either classical or at the ab initio level [13], are routinely employed for sampling the phase space, here, on the trajectory the absorption spectrum is obtained by extracting on-the-fly some molecular configurations. In the end, the final spectrum is obtained as an average over all computed spectra. Since MD simulations run for too short timespans (generally in the nanosecond order), a small area in the phase space is explored. Consequently, properties are not correctly reproduced along the dynamical trajectory. Despite this drawback, MD simulation are widely applied. Another possible strategy for sampling the phase space is based on the random sampling. This is the idea behind the *Monte Carlo* method, that, allows a huge phase space to be explored [14]. As in MD simulations, inconveniences exist also for the MC method. For instance, the generation of a molecular structure could be traumatic when large perturbative steps are accounted, causing either atom clashes or (aromatic) structures can become too strained. Despite the problems in generating possible unphysical structures, MC simulations require less time in respect to MD simulations. The intention is to overcome the combined challenges of accuracy and computing-time, as such, it is essential to build-up simple methods able to easily reproduce the absorption spectrum for molecules characterized by a certain degree of flexibility.

In this general context, the PILI company is interested in approaches to obtain the molecular spectroscopic signatures which span from colour prediction of metal-free organic dyes to the development of new computational protocols to model the photoreaction mechanism. My project is inserted in this framework. The aim is to provide, and to test, new computational tools to be used in the simulation of the dyes and rationalize the structure/property relationship involving both the ground and the excited state potential energies (figure 1.1).

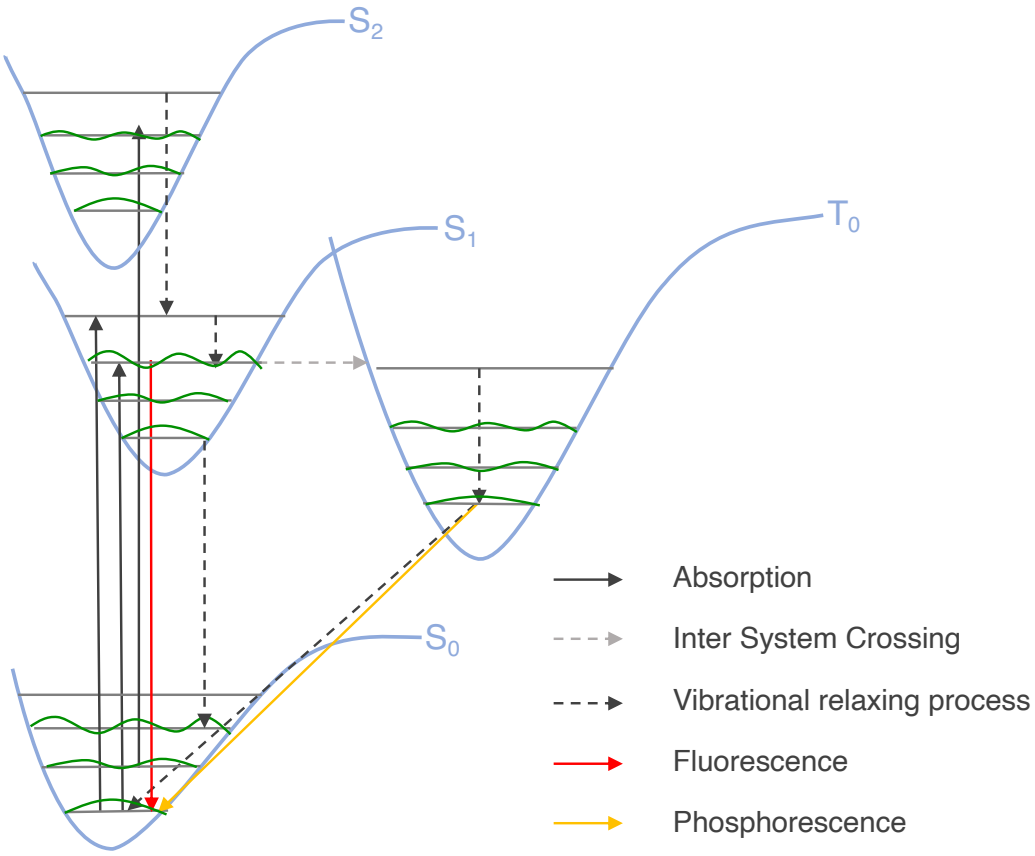


Figure 1.1 Jablonski diagram: possible photophysical phenomena.

During my PhD more than one-hundred organic compounds were studied, but due to confidential disclosure agreement only some of them will be presented.

The manuscript is structured as follows: after this brief introduction, Chapter 2 reviews the theoretical tools needed to understand the approaches used in this thesis. Herein, different methods will be presented. We start by presenting DFT and its time dependent extension. After that, different approaches will be presented to model the absorption spectra. The last part focuses on the molecular mechanic and on the Monte Carlo method.

In Chapter 3, a new computational protocol has been developed for a class of compounds of interest for the PILI company. With the aim to set-up a new computational procedure a set of substituted anthraquinones and indigos, characterized by maximum light absorption that spans from 408 nm to 526 nm, have been tested. In order to obtain their electronic spectra, we adopt two strategies. The first accounts vibronic contributions. Here, absorption bands have been simulated by using two computational models known as time-independent and time dependent approaches. Unlike the first strategy, the second one just accounts of an empirical function in

order to mimic band shapes. Results will be discussed in conjunction with the CIE-Lab colour space [6]. This work will be the focus of a future paper.

In Chapter 4, we present a protocol tailored for merocyanine dyes in their monomeric and dimeric form. In this chapter, the relationship between the structural and electronic features on merocyanine dyes are investigated by combining Monte Carlo simulation with DFT. In addition, in order to give a clear picture on the estimation of the vertical excitation energies, results obtained with the multi-scaling approach are compared to those obtained at the static level, as well as the experimental ones. The results presented in this chapter have been published as research paper [15].

Chapter 5 focuses on a TD-DFT procedure used to investigate the photochemical process involved in the hydrogen abstraction mechanism. In order to study the reactivity, two photoactive anthraquinones, namely AcAQ and TfAQ have been tested. Using the selected compounds, we develop a methodology which aims to track the possible crossing between excited states along a linear synchronous transit (LST) path. Within the method, we integrate the information deriving from the D_{CT}, Π indexes and the spin-orbit coupling. This study is currently in progress, and it will be the focus of a future paper.

Finally, chapter 6 presents general conclusions and perspectives for the modelling of electronic absorption spectra of organic dyes in solution.

Chapter 2

2. Theoretical Background

The molecular behaviour depends intimately on the electronic structure which governs the electronic properties. Here, methods able to solve the electronic structure problem allow us both to understand and to predict molecular properties. Usually, they are classified in three families: non-empirical, semi-empirical and empirical. In the framework of non-empirical approaches, the available theoretical methods are called *ab-initio* when the solutions are generated without reference to experimental data, while semi-empirical methods involve parameters to yield a result in accord with observations. Finally, methods are defined as empirical if they rely on atom-based parameters. Among the *ab-initio* methods, the density functional theory (DFT) represents the most popular quantum mechanical (QM) tool used for the investigation of electronic properties, due to the high accuracy-to-computational cost ratio [16]. Although DFT can be considered as the reference for predicting the electronic and optical properties, it has a computational time that scales with the third power of the number of orbitals making longer simulation unrealizable specially for large systems [17]. A natural alternative consists in using an empirical force field for describing the interaction energy between atoms and molecules required for distorting a molecule [18]. The so-called Molecular Mechanics (MM), thereby approximating the QM surface with a classical mechanic approach. This chapter aims to provide the fundamental theoretical concepts used in this thesis. We start by introducing the key concepts behind the DFT.

2.1 Solving the Schrödinger equation

The electronic properties of a given molecular system can be determined by solving the time-independent Schrödinger equation, that in simplified form is written as:

$$\hat{H}\psi(R_M, r_n) = E\psi(R_M, r_n) \tag{2.1}$$

Here $\hat{H}$ is the Hamiltonian, E is the total energy and ψ is the wave-function, depending on the spatial coordinates (R and r) of the *n* electrons and *M* nuclei. The analytical solution of the Schrödinger equation is only possible for Hydrogenoid systems but it can be solved for the multi-electron systems applying some approximations. The first step to solve the Schrödinger equation is the separation of electronic and nuclear degrees of freedom by invoking the Born-Oppenheimer (BO) approximation [19]. Briefly, due to the fact that the lighter mass electrons move much faster than nuclei, we can solve only the electronic Schrödinger equation for a given nuclear configuration ruling out the nuclear wave function. Through the BO approximation, nuclei can be treated classically and the energy of the system for each configuration of the nuclei determines the *potential energy surface* (PES) [20]. Then an electronic Schrödinger equation can be defined as:

$$\hat{H}_e \psi_e(\{R_M\}, r_n) = E_e \psi_e(\{R_M\}, r_n) \tag{2.2}$$

Where $\hat{H}_e$ is the electronic Hamiltonian and ψ_e is the electronic wavefunction. From equation 2.2 it is possible to see that the electronic wave function depends parametrically on nuclear coordinates ($\{R_M\}$). The total energy for a fixed nuclear configuration including the nuclei interaction term is:

$$E_{total} = E_e + \sum_{K=1}^{M} \sum_{P<K}^{M} \frac{Z_K Z_P}{|R_K - R_P|} \tag{2.3}$$

Where Z is nuclear charge.

2.1.2 Hartree–Fock theory:

In order to understand the fundamental concepts of all wave function methods, let us start by considering the electronic Hamiltonian:

$$\hat{H}_e = \underbrace{-\sum_{i=1}^{N} \frac{1}{2} \nabla^2}_{\hat{T}_e} \underbrace{- \sum_{i=1}^{N} \sum_{A=1}^{M} \frac{Z}{r_{iA}}}_{\hat{V}_{ne}} + \underbrace{\sum_{i=1}^{N} \sum_{j>i}^{N} \frac{1}{r_{ij}}}_{\hat{V}_{ee}} \tag{2.4}$$

Where the terms $\hat{T}_e$, $\hat{V}_{ee}$ and $\hat{V}_{ne}$ are respectively the electron kinetic energy, the electron-electron interaction, and nuclei- electron interaction operators. Focusing on the electronic structure problem, the term that makes difficult to solve the Schrödinger equation for multi-electron systems is the electron-electron interaction term. The first approximation to this problem was to neglect this interaction (an approximation known as independent-particle model or Hartree model).
Since the electrons are fermions, they obey the Pauli exclusion principle which imposes the antisymmetry constraint on the wave function on the solution. Slater developed the multi-electronic wave function as determinant of monoelectronic wave functions, in order to apply this constraint [21].

$$\psi(\chi_1, \dots, \chi_n) = \frac{1}{\sqrt{N!}} \begin{bmatrix} \psi_1(\chi_1) & \cdots & \psi_N(\chi_1) \\ \vdots & \ddots & \vdots \\ \psi_1(\chi_n) & \cdots & \psi_N(\chi_n) \end{bmatrix} \tag{2.10}$$

Where the ψ_N term is the Nth *spin orbital*. These orbitals are written as product between the space and spin functions:

$$\chi_n(r_i) = \varphi_n(r_n)\omega_n(\sigma_n) \tag{2.11}$$

The $\omega_n(\sigma_n)$ function describes the intrinsic spin momentum of an electron, the spin function.
Further, the interaction between each electron and the other *n*-1 electrons is then recast into the Hamiltonian. In other words, each electron moves in the average field generated by the all other *n*-1 electrons. This is the core of Hartree–Fock (HF) theory [20] that gave birth to the Molecular Orbital (MO) family of *first principles* approaches. In the HF, the energy in shorthand form is written as:

$$E_{HF} = \sum_{i=1}^{N} \hat{h}_i + \sum_{i=1}^{N} \sum_{j=1}^{N} \left(\hat{J}_{ij} - \hat{K}_{ij}\right) + \hat{V}_{nn} \tag{2.12}$$

Where $\hat{h}_i$ is the monoelectronic Hartree operator, $\hat{J}$ is the Coulomb operator and $\hat{K}$ exchange operator. Within the HF theory no-self interaction error occurs, indeed, when $i = j$ it is easy to see that the Coulomb "self-interaction" term, $\hat{J}_{ii}$, is compensated by the exchange term, $\hat{K}_{ii}$ [22].

One important consideration has to be made. The HF theory is able to recover about 99% of the "true" energy, however, atomization energies and other that physico-chemical properties require more of 99% [22]. This blemish is related to the specific $\hat{V}_{ee}$ term.

In 1958, Löwdin defined the correlation energy as the difference between the true Energy (E) and the HF solution is due to the correlation energy (E_{corr}) [23]:

$$E_{corr} = E - E_{HF} \tag{2.13}$$

Conceptually is useful separate the correlation energy in *"dynamic"* and *"static"* part. Since the HF theory is derived by employing the mean-field approximation (assuming the independent electron motion), it underestimates the electron distances and thus increase the electron repulsion energies. The correlated motion is called dynamic correlation, and it occurs between electrons in the same orbital as well as between electrons in the molecule.

On the other hand, the static correlation (also known as *"no-dynamic"* correlation) occurs when occupied and unoccupied orbitals are (nearly)-degenerate. In these systems the multi-determinantal nature of the wave function cannot be ignored. The HF theory fails, since is a mono-determinant method. Over the years, several MO approaches have been built in order to get the correlation energy via perturbative (for example the Møller-Plesset second order correction) schemes or by employing variational schemes, for instance configuration interaction (CI)[20]. The accuracy of these methods (so called post-HF) is comparable with experiments, for instance in thermochemistry the errors are below *1 kcal/mol* ("*chemical accuracy"*). However, they are computationally demanding and can be used just for small systems.

2.2 Density Functional Theory:

Another theoretical framework for solving the electronic Schrödinger equation is based on a method that takes advantage of the electronic density: the Density Functional Theory (DFT) [24,25]. The main advantage of DFT with respect to the MO approaches is due to the fact that the multi-particle problem is shifted from a problem in 4N variable (three space and one spin coordinates for each of N electrons) to a problem in three variables (i.e., the electronic density $\rho(r) = N \int d^3r_2 \dots \int d^3r_N \, |\psi(r, r_2, .., r_N)|^2$). This brings a great advantage in terms of computing time, at list in principles, allowing us to study large systems also in complex environments with relatively short time with respect to the wave-function based methods [26]. Thus, among the *first-principles* methods, DFT is the most widely used theoretical tool for solving

the electronic structure problems that range from periodic systems to molecules in complex environments. This theory is based on the two Hohenberg-Kohn theorems.

The first theorem proves that the electronic density determines all the properties of a system in its electronic ground state. The first theorem proof occurs through a "*reductio ad absurdum*" and shows that the ground state electronic density univocally determines the electronic Hamiltonian. In other words, it proves the uniqueness and the existence of a functional relationship between the electronic density and the ground state electronic energy.

The second theorem proves the existence of a functional relationship between the $\rho(r)$ and the ground state electronic energy which obeys the variational principle. This means that for a given guessed electronic density ($\rho_{trial}(r)$) is always higher with respect to the true ground state electronic density $\rho_0(r)$. According to the first Hohenberg-Kohn theorem, the total energy $E[\rho(r)]$ can be expressed as a functional of the density:

$$E[\rho(r)] = T[\rho] + V_{ee}[\rho] + V_{ne}[\rho] = F[\rho(r)] + \int \rho(r)\, v(r) dr \tag{2.14}$$

In equation 2.14, $T[\rho]$ is the kinetic energy functional $V_{ee}[\rho]$, the electron-electron interaction term and $V_{ne}[\rho]$ is the nuclei-electron interaction in an unperturbed system (i.e., without the external potential). Focusing on the second term of equation 2.14, $T[\rho]$ and $V_{ee}[\rho]$ functionals depend only on electrons, thus the two terms can be defined as *universal*. In other words, these two functionals depend on only electrons and they are the same for any multi electron system. With $F[\rho(r)]$ the universal functional of $\rho(r)$:

$$F[\rho(r)] = T[\rho] + V_{ee}[\rho] \tag{2.15}$$

For the variational theorem:

$$E = \min_{\rho_{trial} \to \rho_0} E[\rho_{trial}(r)] \tag{2.16}$$

Then:

$$F[\rho_0(r)] + \int \rho_0(r)\, v(r) dr \leq F[\rho_{trial}(r)] + \int \rho_{trial}(r)\, v(r) dr \tag{2.17}$$

$$E[\rho_0(r)] \leq E[\rho_{trial}(r)] \tag{2.18}$$

Hence, this theorem allows the optimization of the electronic density by minimizing the energy functional $E[\rho_{trial}(r)]$. Unfortunately, the exact functional form for DFT is unknown and several approximations must be considered.

2.2.1 Kohn-Sham DFT

In DFT formalism the energy functional is composed by three terms: the kinetic energy functional, the electron-electron interaction term and the external potential t erm. Historically the first problem has been the accurate formulation of the kinetic energy term. In 1927, Thomas and Fermi proposed a model called *uniform electron gas* for solving the kinetic energy formulation. After that a more robust approach has been developed leading to the birth of modern DFT, the so-called Kohn and Sham DFT (KS-DFT), where the density take form: $\rho(r) = \sum_{i=1}^{N} |\psi_i(r)|^2$. In this approach the kinetic energy of the non-interacting system (T_{ns}) is described by mono-electronic functions ψ_i.

$$T_{ns}[\rho] = \sum_{i=1}^{N} \langle \psi_i \left| -\frac{1}{2}\nabla^2 \right| \psi_i \rangle \tag{2.19}$$

The subscript ns means non-interacting system. The electron-electron interaction term, the second term of equation 2.19, consists of a classical electrostatic term plus a non-classical one ($W[\rho]$):

$$V_{ee}[\rho] = \sum_{i=1}^{N} \langle \psi_i \left| \frac{1}{2} \int \frac{\rho(r')}{|r_i - r'|} \right| \psi_i \rangle + W[\rho] \tag{2.20}$$

And the external potential is:

$$V_{ne}[\rho] = -\sum_{a}^{M} \sum_{i}^{N} \langle \psi_i \left| \frac{1}{2} \int \frac{Z_a}{|r_i - R_a|} \right| \psi_i \rangle \tag{2.21}$$

Then in the KS-DFT the energy functional ($\mathrm{E}[\rho]$) of interacting particles is written as:

$$\mathrm{E}[\rho] = T_{ns}[\rho] + (\mathrm{T}[\rho] - T_{ns}[\rho]) + V_{ee}[\rho] + (W[\rho] - V_{ee}[\rho]) + V_{ne}[\rho] \tag{2.22}$$

The differences terms in equation 2.22 are included in a non-classical term called exchange and correlational (XC) functional, $E_{xc}[\rho] = (\mathrm{T}[\rho] - T_{ns}[\rho]) + (W[\rho] - V_{ee}[\rho])$. In KS-DFT the $E_{xc}[\rho]$ is only unknown term, therefore, approximations are done made on it. XC functionals are discussed in chapter 2.8. The electronic Hamiltonian is written as monoelectronic KS Hamiltonian:

$$h_i^{KS} = -\frac{1}{2}\nabla_i^2 + \frac{1}{2}\int \frac{\rho(r')}{|r_i - r'|} - \frac{1}{2}\int \frac{Z_a}{|r_i - R_a|} + V_{xc} \tag{2.23}$$

Thus, as the HF method, the eigenvalue problem to be solved is:

$$h_i^{KS}\chi_i(r_i) = \varepsilon_i\chi_i(r_i) \tag{2.24}$$

As the HF the KS-DFT is a mono-determinant method. But, unlike the HF, the KS-DFT includes the exact density of non-interacting system. This approach maps the fully interacting system into a non-interacting system that leads with the same ground-state density. However, the KS-DFT formulation would yields the exact ground state electronic energy if the exact form of E_{xc} were known.

It is worth noting that KS-DFT shows the link between the ground state electronic density and the associated wave function. A practical solution for resolving the dependence of the KS-orbitals on the electronic density itself consist in a variationally solution of the KS equations employing an iterative approach. The solution therefore it is self-consistent and then the method is called Self-Consistent-Field (SCF).

2.2.2 Self-Consistent-Field (SCF) procedure

This procedure, Figure 2.1, starts by obtaining an initial guess density ρ_0, by expressing the electronic density in MO. These orbitals are described by using a linear combination of basis functions (section 2.2.3). After, the effective potential (V_{eff}) is computed and inserted into the KS equation, the N eigenvalue KS equations are resolved through diagonalization. The total electronic energy is computed. After this first step, the "new" electronic density is used as "new electronic

guess". The same procedure is replicated up to reach the convergence by fixing a predefined threshold value (ε_{toll}).

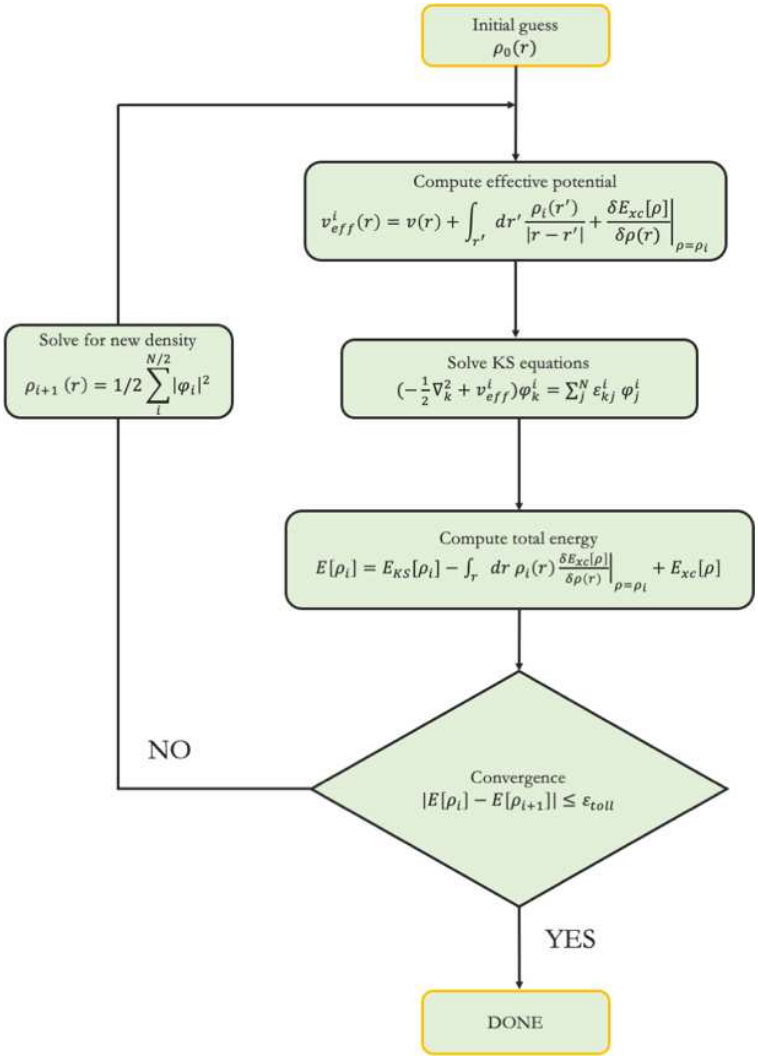


Figure 2.1 Self-Consistent-Field flow chart.

Since the density is expressed in terms of MOs, it is foreseeable that the MO description together with the type of XC functional used have a relevant effect on the level of accuracy. In the next section, we introduce an approximation inherent to all first-principles methods or rather to the basis functions.

2.2.3 Linear combination of atomic orbitals

One approximation in the KS-DFT is referred to the basis set expression [26]. This set of variable functions is used to define where the wave function extends. To do it, MOs are built as linear combination of atomic orbitals (LCAOs) of monoelectronic wavefunctions that have generally the form of atom-centred functions.

$$\varphi_i(r) = \sum_{q=1}^{K} c_{iq}\, \chi_q(r) \tag{2.25}$$

Where φ_i is the molecular orbital, χ_q is the atomic orbital and c_{iq} are the expansion coefficients. Generally, these basis functions can be classified into two types: the Slater (χ_q^{STO}) and the Gaussian-type orbitals χ_q^{GTO}.

The Slater-type orbital (STO) possess an exponential form:

$$\chi_q^{STO}(x,y,z) = N_{a_q}\left(x-x_q\right)^{a_{q,x}}\left(y-y_q\right)^{a_{q,y}}\left(z-z_q\right)^{a_{q,z}} e^{-\alpha_q|r-r_q|} \tag{2.26}$$

Where N_{a_q} is the normalization constant, $a_{q,}$ is the orbital angular moment, r_q is its center and α_q is the exponent. The advantage of the STO is due to the fact that the exponential dependence reproduces the orbital of hydrogen atom, then they are able to capture both the cusp and the exponential decay. However, the computation of the two electron integrals at different centres cannot be performed analytically.

On the other hand, Gaussian-type orbitals (GTOs) are simple to integrate, they are commonly employed.

$$\chi_q^{GTO}(x,y,z) = N_{a_q}\left(x-x_q\right)^{a_{q,x}}\left(y-y_q\right)^{a_{q,y}}\left(z-z_q\right)^{a_{q,z}} e^{-\alpha_q|r-r_q|^2} \tag{2.27}$$

Focusing on equation 2.27, it is easy to note that the quadratic dependence in the exponential term ($|r-r_q|^2$) makes the GTOs inferior with respect to the STOs in two respects. First, unlike the STO which at the nucleus has a cusp, a GTO does not. Second, the GTO with respect to the STO, decays rapidly far from the nucleus. Due to these limitations, a linear combination of *primitive* GTO is used for building contracted Gaussian-type orbitals:

$$\chi_q^{CGTO}(x,y,z) = \sum_{k=1}^{N_g} d_{q,k}\, \chi_q^{GTO}(x,y,z) \tag{2.28}$$

Where the contraction coefficient $d_{q,k}$ controls the shape of the combination of GTOs.

2.2.4 The word of Exchange Correlational functionals

As already mentioned, although the DFT ideally should be able to predict the ground state properties, some approximation must be adopted. This is due to the fact that the main problem in the modern KS-DFT revolves around the exact form of the exchange and correlational (XC) functional. Indeed, today the exact form is unknown and over the years many approximations onto the exchange and correlational term have been proposed. Up to now, different exchange correlational functionals exists and they are categorized according to the so-called Pardew's Ladder [27]. Here the exchange and correlational functionals are sorted on the basis of their complexity and are classified as rungs of this ladder, where the "*Heaven*" of the *chemical accuracy* corresponds with high level of accuracy. Hence increasing the complexity of the XC functionals, the accuracy increase. As results, they are generally classified and discussed in basis of their complexity. We will begin with the simplest approach and, thus, with the local density approximation.

2.2.5 Local Density Approximation

Based on the DFT solution for an hypothetical uniform electron gas [28], the Local Density Approximation (LDA) represents the simplest approach in the XC functional landscape. Approximating the real density with an uniform electron gas means that the electrons moves in a positive background generated by the nuclei. In other words, the density is constant anywhere in the space and functional belonging to this class are called local functionals. Then, not surprisingly, LDA based-functionals are able to describe, with certain accuracy, systems characterized by slowly varying electronic density (i.e., metals).
In LDA the exchange and correlation term is written as:

$$\mathrm{E}_{xc}^{LDA}[\rho] = \int dr\, \rho(r)\epsilon_{xc}\,[\rho(r)] \tag{2.29}$$

Where ϵ_{xc} is the energy density. Here the exchange and correlation contributions are generally treated separately. It is important to note that in this variant the LDA fails in the treatment of spin

polarized systems (SPSs). To describe SPSs, for instance open-shell ones, LDA is generally used in a variant, namely *Local Spin Density Approximation* (LSDA). Conventionally, the exchange and correlation contribution are assumed to be separable.

$$E_{xc}[\rho^{\alpha}, \rho^{\beta}] = E_x[\rho^{\alpha}, \rho^{\beta}] + E_c[\rho^{\alpha}, \rho^{\beta}] \tag{2.30}$$

Here the spin density ρ^{α} and ρ^{β} are separately optimized allowing the of study the SPSs. However, several electronic systems are characterized by a non-uniform electron density, including all chromophores studied in this thesis. Hence, LDA functional fails in describing such systems and, thus, another class of functional (*semi-local functional*) must be introduced.

2.2.6 Generalized Gradient Approximation

The generalized gradient approximation (GGA) introduces the first derivative of the electronic density in the XC functionals [29]. The knowledge of the electronic density gradient allows the study of how non-uniform electron density varies, thus, the XC functionals belonging to this class are called local. Functionals obtained with this method outperform the prediction of molecular properties obtained at LDA level. Specifically, a GGA functional is obtained by adding the gradient of the electronic density to LDA functional:

$$\mathrm{E}_{xc}^{GGA}[\rho(r)] = \mathrm{E}_{xc}^{LDA}[\rho(r)] + \Delta\epsilon_{xc}\left(\frac{|\nabla\rho(r)|}{\rho^{\frac{4}{3}}(r)}\right) \tag{2.31}$$

Since XC functionals are translated in to analytical form and the exchange and correlation terms are conventionally separable, it is possible to express them mathematically:

$$E_x[\rho(r)] = \int \rho(r)\epsilon_x^{unif}(\rho)F_x(s) \tag{2.32}$$

$$E_c[\rho(r)] = \int \rho(r)\epsilon_c^{unif}(r_s, \zeta)F_c(r_s, \zeta, t) \tag{2.33}$$

Where the total density is $\rho(r)$, $\epsilon_x^{unif}(\rho)$ and ϵ_c^{unif} are respectively the exchange and the correlation energy densities obtained with the uniform electron gas approximation (i.e., within LDA approximation). In equation (2.33) the r_s factor is called local Seitz radius ($r_s =$

$[(4\pi/3)\rho(r)]^{\frac{1}{3}}$), ζ is the spin polarization and it is defined as $\zeta = (\rho^{\alpha} - \rho^{\beta})/\rho$ and t is another dimensionless density gradient. The exchange part of the GGA functionals is expressed using the enhancement factor $F_x(s)$ which is dependent on another dimensionless gradient density ($s = |\nabla\rho(r)|/2k_f\rho(r)$, where k_f is the local Fermi wave factor). An example of non-empirical exchange GGA functional has been derived by Perdew, Burke and Ernzerhof (PBE) in 1996 [30].

$$F_x(s) = 1 + k - \frac{k}{1 + \frac{\mu}{k}s^2} \tag{2.34}$$

In PBE exchange functional, the $F_x(s)$ parameters (k = 0.804 and μ =0.220) are obtained from some known physical constraint [31]. As reported for the GGA functional, the logical way for improving the functionals consists in introducing high order derivatives. Thus, next rung on the Jacobs Ladder collects XC functionals called meta-GGA [32], which incorporate other semi-local ingredients such as the Laplacian in the XC functional form, an example is the kinetic energy density ($\tau = (1/2\sum_i|\nabla\varphi(\mathrm{r})|^2)$).

2.2.7 Adiabatic connection: Hybrid GGA functionals

Since the exchange is treated approximatively, LDA or GGA fail in removing the unphysical self-repulsion contained inside the classical V_{ee} term. From this, the self-interaction error (SIE) originates. Calculations performed with these methods over-delocalize the electronic structure in order to minimize this unphysical self-interaction. As a result, the GGA does not correctly describe the molecular electronic structure. To overcome this drawback, one can think to use a theory without the SIE, as the HF, to remove the error contained into the Coulomb integrals (see equation 2.12). The strategy to combine the HF exchange with the DFT one consists in using the *adiabatic connection* model [28]. As discussed in section 2.2.1, the KS approach is able to link the non-interacting system with the interacting one. These two systems can be seen as two limiting cases linked by a λ parameter. This parameter, that ranges from zero to one, allows the switching-off/on of electron-electron interaction. This is the idea behind the adiabatic connection method that allows two boundary cases to be linked. Within the adiabatic connection model the XC energy read:

$$E_{xc}[\rho] = \int_0^1 d\lambda \langle \psi(\lambda) | V_{xc} | \psi(\lambda) \rangle \tag{2.35}$$

A pictorial representation of the integral (2.35) can be helpful for its evaluations. Figure 2.2 depicts the exchange-correlation energy as sum of two areas. For λ equal zero, the electron-electron interaction is switched-off, and, thus, the area of the upper rectangle (marked as ***A***) is defined just by the exchange contribution. When λ is equal to one the electron-electron energy is switched-on (i.e., interacting system). The remaining energy is represented by the area under the curve inside the rectangle ***B***. Thus, the lower area under the curved line is a fraction α of the rectangle ***B***. The α value is unknown, but we can consider it as a parameter to optimize. Hence, the total area ($E_{xc}[\rho]$) can be approximated as follows:

$$E_{xc}[\rho] = E_x^{HF} + \alpha(E_{xc}^{DFT}[\rho] - E_x^{HF}) \tag{2.36}$$

Conventionally, the p parameter (defined as $\alpha = 1 - p$) instead of α is used, and the exchange and correlation energy takes the form:

$$E_{xc}[\rho] = pE_x^{HF} + (1-p)(E_{xc}^{DFT}[\rho]) \tag{2.35}$$

This equation represents the link between the interacting and non-interacting systems. Functionals that contain the exact exchange are called *hybrid functionals*. Over the years, empirical and non-empirical hybrid XC functional have been proposed. For instance, an empirical XC functional popular in the computational community, is the B3LYP [28, 33]:

$$\begin{aligned} E_{xc}^{B3LYP}[\rho] = (1-a)E_x^{LSDA}[\rho] + aE_x^{HF} + b\Delta E_x^{B88}[\rho] + (1-c)E_c^{LSDA}[\rho] \\ + cE_c^{LYP}[\rho] \end{aligned} \tag{2.36}$$

This hybrid functional uses 3 parameters ($a = 0.20$, $b = 0.72$ and $c = 0.81$) obtained by fitting experimental data. In 1999, a non-empirical functional has been developed by Adamo and Barone [31].

$$E_{xc}^{PBE0}[\rho] = E_{xc}^{PBE}[\rho] + a(E_x^{HF} - E_x^{PBE}[\rho]) \tag{2.37}$$

This hybrid GGA functional, known as PBE0, is constructed with the 25% of exact exchange ($a = 0.25$), while the XC part used for constructing the non-interacting system is the PBE functional. This hybrid GGA functional has been extensively used in this thesis.

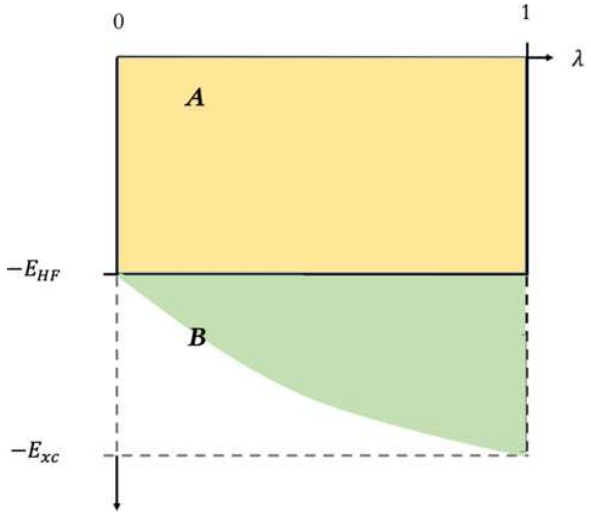


Figure 2.2 Graphical representation used for evaluating the XC energy (equation 2.35). The integral value is the sum of the orange and green areas.

2.2.8 Range Separated Hybrid functionals

Although the hybrid GGA functionals outperform predictions for molecular thermochemistry given by LDA and standard GGA functionals[34], at long range (LR) the asymptote of their X potential become a/r instead of the correct behavior $1/r$ [35]. This yields an underestimation of some molecular properties, for instance: ionization potentials, polarizabilities of π-conjugated chains, and the energy of intra-inter molecular charge transfer (CT) excitations. To overcome the poor description of these molecular properties, hybrid functionals based on the range separated scheme have been introduced.

Range separated hybrid (RSH) functionals are a class of hybrid functionals that improve upon the idea of the hybrid functional by adding a "layer of complexity" on the XC form. In RSH the Coulomb operator splits in to short- and long-range components.

$$\frac{1}{r} = \frac{1 - \mathrm{erf}\,(\omega r)}{r} + \frac{\mathrm{erf}\,(\omega r)}{r} \tag{2.38}$$

The first term of 2.38 accounts for the short-range (SR) interaction, while the second one the LR interaction. The error function (erf) links the SR and LR of the Coulomb operator under the

governance of the ω parameter. In shorthand form, the XC energy for RSH functional is written as:

$$E_{xc}^{RSH}[\rho] = E_x^{SR-DFT}[\rho] + E_x^{LR-HF} + E_c^{DFT}[\rho] \tag{2.39}$$

When a large fraction of HF exchange (100% HF) is incorporated in the LR part the correct asymptotic behaviour is well reproduced. In this thesis, the RSH functional ωB97X-D [36] has been used. This XC functional uses 100% HF exchange for the LR part and 22.2% for the SR one. The additive dispersion term -D- is added to functional ωB97X form, allowing to better describe non-covalent interactions.

2.3 How to study the excited state?

To understand how to study the excited state, in this chapter the DFT is extended at time domain (TD). This method is used for describing the response of a system in its ground-state caused by an external time-dependent perturbation, for instance when the light interacts with matter.

2.3.1 Time Dependent Schrödinger equation

When a system, composed by *N* interacting electrons, evolves in a time-dependent field the Hamiltonian varies in time and the respective Schrödinger equation is:

$$\widehat{\mathcal{H}}_{el}(r,t)\psi_{el}(r,t) = i\frac{\partial}{\partial t}\psi_{el}(r,t) \tag{2.40}$$

This equation allows us to compute the ψ_{el} at time t if the state of the system at the initial time t_0 (i.e., $\psi_{el}(r,t_0)$) is known, hence, one may simply use the ground state one. The time dependent Hamiltonian is composed of three parts:

$$\widehat{\mathcal{H}}_{el}(t) = \hat{T}_e(r) + \hat{V}_{ee}(r,t) + v_{ext}(r,t) \tag{2.41}$$

Here, the first and second terms are respectively the kinetic energy and the electron-electron interaction, while the last one is the external potential. Since we are interested in describing the

absorption of incoming electron-magnetic waves by molecules, one can treat the electric field of the incoming light as a perturbation. For a given fixed nuclei configuration the external potential writes:

$$v_{ext}(r,t) = -\sum_{A=1}^{N_{ATOM}} \frac{Z_A}{|R_A - r|} + E(t)\boldsymbol{\alpha} r \tag{2.42}$$

Where $\boldsymbol{\alpha}$ is the polarization vector associated with the time dependent electric field $E(t)$. From equation 2.42 one may see that the time dependency is explicitly included only inside the second term. Hence, the first term is the usual interaction, while the second one represents the perturbation term.

2.3.2 Time Dependent Density Functional Theory

As previously discussed, the DFT is based on the two Hohenberg-Kohn theorems which assert that all property of a given system can be obtained, in principles, from the ground state electronic density. However, absorption energies have to be found, thus we are interested in obtaining the ground state and the distinctive features deriving by its perturbation. The approach for computing the vertical excitation energy consists in extending the DFT to the time domain. Throughout an excitation process from the ground state to an excited state, charge fluctuations occur and thus states mixing between them. These dynamical effects are captured by TD-DFT. These effects are included into the exchange correlation kernel (f_{xc}) which plays the same role of the XC potential in the ground state theory. In the following we introduce the theoretical basis of the TD-DFT starting from the Runge-Gross theorem.

2.3.3 Runge-Gross theorem

The TD-DFT finds its theoretical foundation in Runge-Gross theorem [37]. Briefly, the theorem proves a one-to-one correspondence between the $v_{ext}(r,t)$ and the time dependent electron density $\rho(r,t)$ for a system evolving from a fixed initial state ($\psi(t_0)$). In other words, for a given system composed by N electrons, two densities $\rho(r,t)$ and $\rho'(r,t)$ evolving from the same initial state (i.e., $\psi_0 = \psi(t_0)$) under two external potentials $v_{ext}(r,t)$ and $v'_{ext}(r,t)$ are different if the potential differ by more than one purely time dependent function (i.e., $v_{ext}(r,t) -$

$v'_{ext}(r,t) \neq c(t)$). Thus, like the first Hohenberg-Kohn theorem, Runge-Gross theorem proves the reverse mapping.

2.3.4 Time Dependent Kohn-Sham approach

The Kohn-Sham scheme at TD-DFT formulation is similar to the ground state formalism. The Runge-Gross theorem, is valid for both non-interacting and interacting systems. Within the time dependent framework, the density is expressed as:

$$\rho(r,t) = \sum_{i=1}^{N} |\psi_i(r,t)|^2 \tag{2.43}$$

The time dependent Kohn-Sham equations possess the form:

$$\left[-\frac{1}{2}\nabla^2 + v_s(r,t)\right]\psi_i = i\frac{\partial}{\partial t}\psi_i(r,t) \tag{2.44}$$

The effective potential is written as a sum of three terms:

$$v_{eff}(r,t) = v_{ext}(r,t) + v_H(r,t) + v_{xc}(r,t) \tag{2.45}$$

Where, $v_H(r,t)$ is the time dependent Hartree potential and $v_{xc}(r,t)$ is the exchange correlation potential. Since one is interested in describing photophysical processes, it is suitable to write the external potential as:

$$v_{ext}(r,t) = v_0(r) + \vartheta(t-t_0)v_1(r,t) \tag{2.46}$$

Where, ϑ is the Heaviside function. It is straightforward to see that the form of Equation 2.46 allows the *"switching-on"* or *"switching-off"* of the time dependency. Indeed, when $t > t_0$ the time dependency turns on, otherwise if $t \leq t_0$ the external potential is equal the external potential in the time independent formalism.

3.3.5 Linear Response Theory

In absorption spectroscopy, one is interested in studying the response of a chemical system in its ground state to a weak electromagnetic field. This means that if the perturbation is small enough the XC potential can be computed around the reference state (i.e., the ground sate). In this case, the linear response theory is considered. For a small perturbation, at time t_0 the system response can be written as a Taylor series:

$$\rho(r,t) - \rho_0(r,t) = \rho_1(r,t) + \rho_2(r,t) + \cdots + \rho_i(r,t) \tag{2.47}$$

$\rho_0(r,t)$ is the ground state density, the densities subscript (i) indicate the perturbational order (i.e., $i = 1,2$ indicate respectively the first and second perturbational order). For a weak perturbation the first order is greater than the higher orders, that may thus be neglected. In addition, a one-to-one correspondence exists between the time density and the time dependent external potentials, in agreement with the Runge-Gross theorem. As the density changes in time, the potential will change too. With no-dependence on the initial state, one may write the time-dependent density as a functional of the potential (i.e., $\rho(r,t) = \rho[v](r,t)$). Linear response corresponds to the lowest-order response formalism ($\rho_1(r,t)$), while quadratic or cubic formalisms treat the second- and third- order terms. In the context of the linear response, the first order change in density can be expressed as:

$$\rho_1(r,t) = \int\int dr'dt'\chi(r,r',t-t')\,\delta v_1(r',t') \tag{2.48}$$

Where v_1 is the first order time dependent potential, while χ is the density-density response function. As in the exchange and correlational functional, the density response function is unknown, but we can express the formally exact density changes by using the Kohn-Sham scheme.

$$\rho_1(r,t) = \int\int dr'dt\chi_{KS}(r,r',t-t')\,\delta v_{eff}(r',t') \tag{2.49}$$

The density-density function for non-interacting systems is:

$$\chi_{KS}(r,r',t-t') = \left.\frac{\delta\rho[v_{eff}](r,t)}{\delta v_{eff}(r',t')}\right|_{v_{eff}[\rho_0](\boldsymbol{r})} \tag{2.50}$$

The effective potential is composed by the external potential, the Hartree and the exchange-correlation contributions. Thus, the functional derivative of third term of equation 2.45 with respect to the ground state density (ρ_0), is the central quantity in TD-DFT or rather the exchange and correlation kernel:

$$f_{xc}[\rho_0](r,r',t-t') = \left.\frac{\delta v_{xc}[\rho](r,t)}{\rho(r',t')}\right|_{\rho=\rho_0} \tag{2.51}$$

Combining the privious results:

$$\rho_1(r,t) = \int dt' \int dr' \chi_{KS}(r,r',t-t') \Bigg[v_{ext}(r',t') + \int dt'' \int dr'' \left(\frac{\delta(t-t')}{|r'-r''|} + f_{xc}(r',r'',t-t') \right) \rho_1(r'',t'') \Bigg] \tag{2.52}$$

This equation shows the dependency of the linearized potential (i.e., the term inside the parenthesis of equation 2.52) on the first order change in density. Since we are interested in computing the vertical excitation energy it is convenient to express the density response in terms of frequencies using Fourier transform.

$$\rho_1(r,\omega) = \int \int dr' dt' \chi_{KS}(r,r',\omega) \Bigg[v_{ext}(r',\omega) + \int \int dr'' \left(\frac{1}{|r'-r''|} + f_{xc}(r',r'',\omega) \right) \rho_1(r'',\omega) \Bigg] \tag{2.53}$$

In the frequency domain the density function becomes:

$$\chi_{KS}(r,r',\omega) = \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} (f_j - f_k) \frac{\psi_j^* \psi_k \psi_j \psi_k^*(r)}{\omega - \varepsilon_k + \varepsilon_j + i\eta} \tag{2.54}$$

The indices j and k correspond respectively to the occupied and virtual orbitals computed at ground state, the f terms are the occupation numbers and η is a positive infinitesimal. The difference $\varepsilon_j - \varepsilon_k$ represents the energy gap between the occupied and unoccupied orbitals. It should be noted that in the density function equation, transition energies occur at energies that are equal to the energy gap between the occupied and unoccupied orbitals of the non-interacting system and, thus, the excitation energies for "real" system are different.

2.3.6 Excitation energies in TD-DFT

When a system in its ground state is perturbed by an external field, two cases can occur. First the frequencies of the perturbation do not match with the energy difference between the ground and excite states, thus the density infinitesimally change. The second case, frequencies match the energy gaps, the change in density is significant and possess a pole at this ω. Thus the "true" excitation energies are obtained from these poles of the response function. For obtaining the excitation energies, an eigenvalue problem has to be solved. The excitation energies (Ω) are computed by using the well-known Casida equation [38]:

$$\begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{B} & \boldsymbol{A} \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} = \Omega \begin{pmatrix} \mathbf{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{-1} \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} \tag{2.55}$$

Where **A** and **B** elements matrix are:

$$A_{ia\sigma,jb\sigma'}(\Omega) = \delta_{ij}\delta_{ab}\delta_{\sigma\sigma'}\omega_{jb\sigma'} + \underbrace{(ia|f_{Hxc}(\Omega)|jb)}_{K_{ia\sigma,jb\sigma'}(\Omega)} \tag{2.56}$$

$$B_{ia\sigma,jb\sigma'}(\Omega) = K_{ia\sigma,jb\sigma'}(\Omega) \tag{2.57}$$

$K_{ia\sigma,jb\sigma'}$ is a two-electron integral over the sum of the Hartree kernel (f_H) and the exchange correlation one (f_{xc}), thus f_{Hxc}. Indexes i and j identify the occupied spin-orbitals while a and b the unoccupied ones, σ and σ' refer to the spin-orbital. The eigenvectors X and Y providing information on transition moments. Although the Casida equation is solved for a finite number of states, TD-DFT benchmarks have shown that the accuracy ranges from 0.15 eV to 0.30 eV [39,40]. In the case of anthraquinones and indigos, the prediction of the maximum absorption energies have been shown small deviations with respect to the experimental results [10, 41].

Equation 2.55 considers only transitions from occupied to unoccupied orbitals, however, in order to obtain the "exact" excitation energy the exact f_{xc} should be known and the equation 2.55 should be solved for all transitions.

2.4 Modelling the environment

The presence of the solvent induces a change in the electron density that affects all molecular properties. From a computational point of view, approaches for estimating the solvent effect are two. The first one is called explicit model. As its name suggests, solvent molecules are explicitly treated. The main advantage of this approach is that it allows the complete description of the specific interactions (i.e., ion pairing, hydrogen bonds, π-interactions etc.). On the other hand, due to the explicit treatment, the method remains computationally demanding.

The second approach is the continuum polarizable model (PCM) [42]. In this model, the environment is treated as a structureless continuum characterized by a dielectric constant. Briefly, an empty cavity is constructed to host the solute that is described by its charge distribution obtained at QM level.

When an absorption spectrum is modelled, two solvent schemes can be applied: equilibrium and nonequilibrium. In the equilibrium case is allowed the full solvent relaxation (i.e., electrons and nuclei). In the non-equilibrium model, on the other hand, just the electrons of the surrounding have time to adapt at the alternate configuration of the chromophore in the excited state. Concerning the model used in this thesis, we choose the non-equilibrium PCM, since absorption process involves short times.

2.5 From the electronic transition to the spectrum

As shown in the previous chapters, the excitation energies between the excited and ground states can be computed in the framework of the linear response theory. A purely electronic transition corresponds to a vertical excitation from the ground electronic state (S_0) to an excited electronic state (S_i) without any structural rearrangement in going to the excited state, Figure 2.3. For an absorption process the vertical excitation starts from the equilibrium geometry at S_0 to finish in a point on the excited PES (S_i). Contrarily, the de-excitation process corresponds to a vertical de-excitation from the relaxed excited state to a point in the ground state PES.

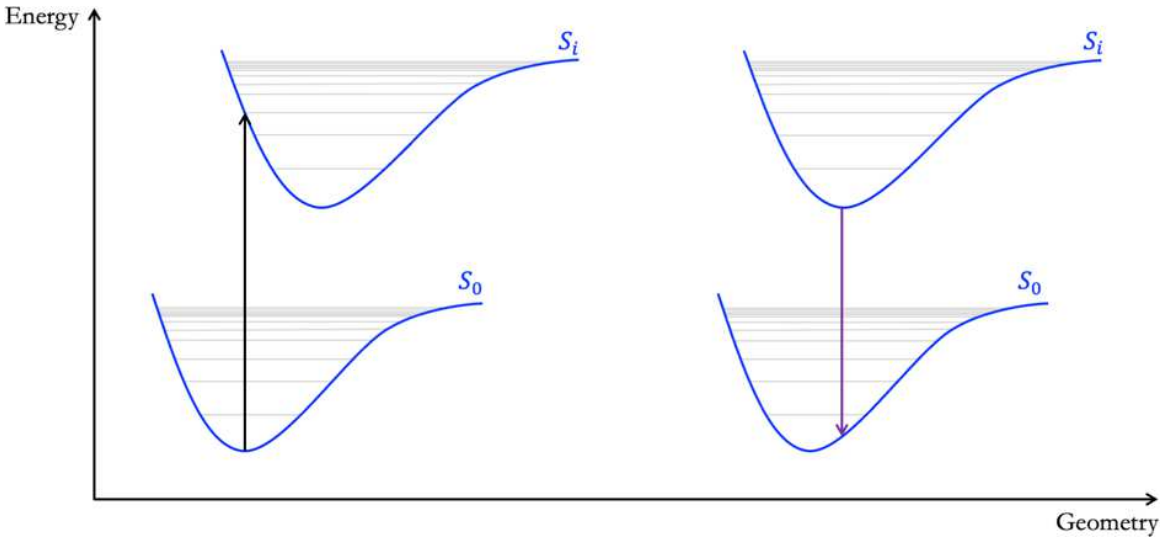


Figure 2.3. Graphical illustration of the vertical excitation (solid black line) and de-excitation (solid green violet) processes.

Hence, directly from a TD-DFT computation a stick spectrum is obtained, but the experimental spectrum is characterized by a well-defined band shape and, thus, a mismatch occurs. Before discussing how to correctly model the spectra, it is essential to link the computed quantity with the experimental part.

2.5.1 Link between theory and experiment

The oscillator strength, the Einstein coefficients or the transition dipole moment are commonly used parameters for describing molecular or atomic optical transitions. In an absorption process, the rate of the $i \rightarrow j$ transition is proportional to the Einstein absorption coefficient $B_{i\rightarrow j}$:

$$B_{i\rightarrow j} = \frac{2\pi}{3\hbar^2}\left|\langle\psi_i|\mu|\psi_j\rangle\right|^2 \tag{2.58}$$

Where ψ_i and ψ_j are the wave functions of the ground and excited state, μ is the dipole moment operator. The term $\langle\psi_i|\mu|\psi_j\rangle$ is the transition moment μ_{ij} that characterizes the transition. The transition strength is commonly used to describe the probability of a transition occurring. Going into the details, matter consists in a set of charges that can interact with the oscillating external electromagnetic field. Hence, absorption takes place when the electric dipole oscillators set into a motion, thus, this occurs when the oscillating external electromagnetic field

possesses frequencies near the matter "characteristic" frequencies ω. The intensity of the interaction is known as oscillator strength ($f_{i\to j}$) which is linked with the Einstein absorption coefficient by the following equation:

$$f_{i\to j} = \frac{\omega m_e h c^2}{\pi e^2} B_{i\to j} \tag{2.59}$$

Where h is the Plank constant, c is the speed of light, m_e is the electron mass and e is the electron charge. For practical purposes, the absorption coefficient $\alpha(\omega)$ can be used for describing the absorption of a light beam passing across a medium. For a gas in standard condition (i.e., 273.15 K and one atmosphere, here and after SPT) absorption coefficient and $f_{i\to j}$ are linked by:

$$\int_{\omega} \alpha(\omega)\ d\omega = N_k \frac{\pi e^2}{c^2 m_e} f_{i\to j} \tag{2.60}$$

Where N_k is the number of atoms per unit of volume. For a beam of light propagating along the l direction, the absorption coefficient is defined as:

$$-\alpha(\omega) = \frac{dI(\omega)}{I(\omega)} \frac{1}{dl} \tag{2.61}$$

From equation 2.61 one may see that the attenuation of the luminous intensity, $dI(\omega)$, is proportional to: the intensity ($I(\omega)$), the thickness of the medium passed through (dl) and the absorption coefficient.

Experimentally, the Lambert-Beer law is used to conduct quantitatively optical measurements, and it is simply written as:

$$ln\left(\frac{I}{I_0}\right) = -\varepsilon(\omega) c_0 l \tag{2.62}$$

Where $\varepsilon(\omega)$ is the absorption molar coefficient, and c_0 is the concentration. In SPT condition for an atomic gas the absorption coefficient is: $\alpha(\omega) = \varepsilon(\omega) c_g \ln(10)$. Then, the link between the theory and experiment is obtained as:

$$f_{i\to j} = \frac{c^2 m_e}{N_k \pi e^2} \int_\omega \alpha(\omega)\, d\omega = 2.303 \frac{c^2 m_e c_g}{N_k \pi e^2} \int_\omega \varepsilon(\omega)\, d\omega \tag{2.63}$$

2.5.2 Modelling electronic spectrum

The modelling of either absorption or emission electronic spectra requires the computation of two factors, the excitation energies and the oscillator strengths for a given transition from the ground to the excited state. Thus, directly from TD-DFT computation spectral lines are obtained. However, absorption or emission bands are never infinitely narrow but they extend over a certain wavelength range. The width of a line is principally caused by three factors: the *natural* broadening, the environment contribution (e.g. solvent) and the vibrational coupling. The first broadening arises from the Heisenberg uncertainty principles ($\Delta E \Delta t \approx h/2$), while the broadening induced by the solvent is related to the fluctuation of the solvent. In the case of the vibrational coupling, the broad band originates from an electronic transition that splits in several peaks due to the vibrational transitions. As described in paragraph 2.4, vertical electronic transitions produce absorption lines. However, an electronic transition is coupled with the vibrational transitions. In absorption, one vibrational transition corresponds to a vertical transition from the vibrational state (S_0) to an excited state one (S_i), Figure 2.4. On the other hand, for an emission process vibrational transition corresponds to a vertical transition from an excited vibrational state to the ground state one.

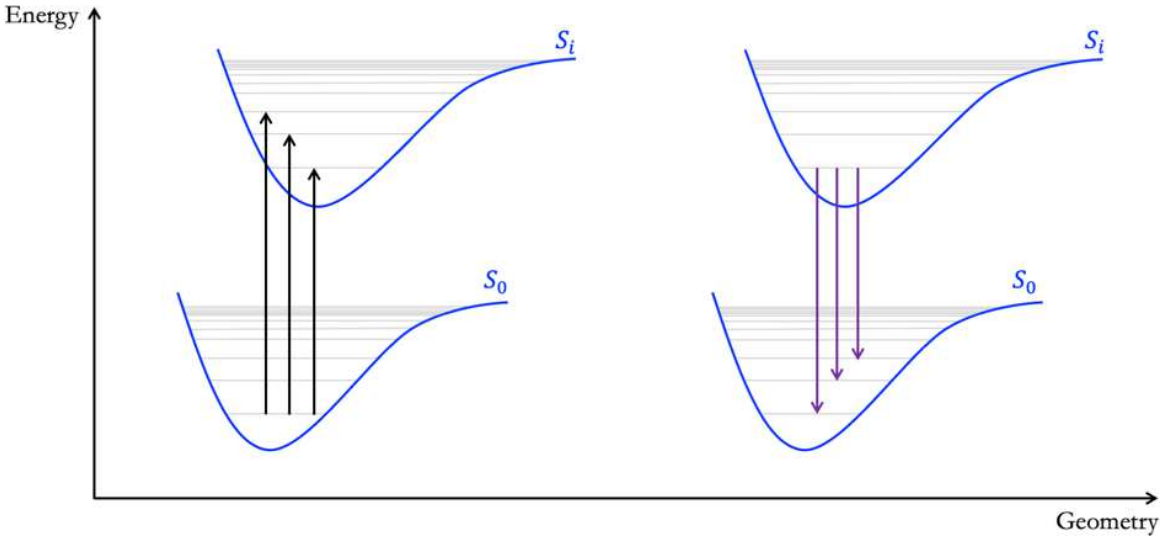


Figure 2.4. Graphical illustration of the vibronic transitions between two singlet states at temperature 0 K. Absorption (solid black line) and emission (solid green violet).

For several chromophores commonly used in dyes industries, including anthraquinones, vibronic coupling largely affects the band structure. In this concern, we analyse the main ingredient necessary for obtaining the vibrationally fine structure starting from the molecular vibrations.

2.5.3 Molecular vibration

A non-linear molecule possesses 3N (with N is the number of atoms) degrees of freedom which split in: 3N-6 vibrations, 3 translational, and 3 rotational degrees of freedom. On the other hand, a linear molecule possesses 3N-5 vibrations, 3 translational, and 2 rotational degrees of freedom. Molecular vibrations can be only described by quantum theory due to the fact that they are quantized motions. In order to get a solid description of the vibrational coupling effect, the quantum treatment of molecular vibrations is discussed in the following.

2.5.4 Nuclear Hamiltonian

Within the Born-Oppenheimer approximation the electron and nuclei movements are uncoupled, allowing the electronic problem to be solved. The nuclear Hamiltonian is composed by two terms:

$$\hat{\mathcal{H}}_n = \hat{T}_n + \hat{V}_n \tag{2.64}$$

The first term on the right side is the kinetic energy operator ($\hat{T}_n$), while the second one is the potential energy operator ($\hat{V}_n$). In terms of normal coordinates, the kinetic energy operator for a nonlinear molecule is:

$$\hat{T}_n = \sum_{k=1}^{3N_n-6} -\frac{1}{2\mu_k}\frac{\partial^2}{\partial q_k^2} \tag{2.65}$$

where μ_k and $q_k$are respectively the reduced mass and the kth normal coordinates. According with the adiabatic approximation for each nuclear configuration the electronic problem is solved, and the nuclear potential energy operator is equal the eigen value of the electronic Hamiltonian:

$$\hat{V}_n = E_e\left(q_1, \dots, q_{3N_n-6}\right) \tag{2.66}$$

2.5.5 Harmonic approximation

In the harmonic approximation, the operator $\hat{V}_n$ is expanded by second order Taylor expansion around the equilibrium geometry. Denoting $\{q_k^{eq}\}$ the collection of the normal coordinates at equilibrium position, the term $\hat{V}$ take the form:

$$\hat{V}\{q^{eq}\} \approx E\{q_k^{eq}\} + J_E\{q_k^{eq}\}(q_k - q_k^{eq}) + \frac{1}{2}(q_k - q_k^{eq})^T H_E\{q_k^{eq}\}(q_k - q_k^{eq}) \tag{2.67}$$

The matrix that collects the first derivate is the Jacobian (J_E) and around the equilibrium geometry it vanishes, because by definition the derivative is equal to zero at stationary points. The matrix H_E is called Hessian, containing the second derivatives of the energy with respect to the normal coordinates. Hence, the nuclear Schrödinger equation becomes:

$$\sum_{k=1}^{3N_n-6}\left[-\frac{1}{2\mu_k}\frac{\partial^2}{\partial q_k^2} + \frac{1}{2}(q_k - q_k^{eq})^T H_E\{q_k^{eq}\}(q_k - q_k^{eq})\right]|\psi_v\rangle = \varepsilon_v|\psi_v\rangle \tag{2.68}$$

Here, $|\psi_v\rangle$ and ε_v are respectively the wave function and the energy of the molecular vibrations. The solutions of this eigenvalue problem are found as product between Hermite polynomials. Within the independent oscillator model, the energy for each harmonic oscillator can be obtained as:

$$\varepsilon_v^{n_k}(k) = \left(n_k + \frac{1}{2}\right)\hbar\omega_k \tag{2.69}$$

Where n_k is the vibrational quantum number and ω_k is the normal mode frequency. Hence, the energy is quantized with differences between neighbouring levels equal to $h\omega_k$. In addition, one can be seen that the energy of the lowest vibrational level (i.e. $n_k = 0$) is not zero leading the so called zero point energy (ZPE)

$$E_0 = \frac{1}{2}\sum_k h\omega_k \tag{2.70}$$

2.5.6 Vibrationally resolved spectra in Time-independent formalism

One route to obtain the vibronic fine structure consists in exploiting a time-independent approach. This approach is based on a sum-over-state formalism. Here, the vibronic spectrum is obtained as the ensemble of all transitions between the vibrational states. All vibrational modes are assumed to be independent from one another. For a radiative transition between two electronic states, the line intensities depend on the square module of the transition dipole moment [43, 44]:

$$\sigma(\omega)_{abs} = \frac{4\pi^2\omega}{3c}\sum_i p_i \sum_f |\langle\psi_i|\mu|\psi_f\rangle|^2 \, \delta(E_f - E_i + \hbar\omega) \tag{2.71}$$

Where $\psi_i\rangle$ and $\psi_f\rangle$ are respectively the initial and final molecular wave functions, p_i is the Boltzmann population of the initial states, δ is the Dirac function, generally replaced by Gaussian or Lorentzian functions used to mimic the experimental broadening. Within the framework of the Born-Oppenheimer approximation, two statements are made [45].
First, the molecular wave functions can be written as a product of two molecular wave functions: the electronic wave function ψ^e and nuclear ψ^n. Second, the dipole moment can be separated into the electronic and nuclear parts ($\mu = \mu_e + \mu_n$).

$$\langle\psi_i|\mu|\psi_f\rangle = \langle\psi_i^e\psi_i^n|\mu_e|\psi_f^e\psi_f^n\rangle + \langle\psi_i^e\psi_i^n|\mu_n|\psi_f^e\psi_f^n\rangle \tag{2.72}$$

Since the electronic wave functions are orthogonal to each other, the second term in equation 2.72 is zero. In addition, since the rotational and translational energies are smaller than the vibrational one, the nuclear wave function is only dependent on the vibrational contribution (ψ^v). Therefore, the electronic transition dipole moment is $\mu_{if} = \langle\psi_i^e|\mu_e|\psi_f^e\rangle$ and can be corrected vibrationally as:

$$\langle\psi_i|\mu|\psi_f\rangle = \langle\psi_i^v|\mu_{if}|\psi_f^v\rangle \tag{2.73}$$

For rigid chromophores, where the ground state equilibrium geometry is close to the excited state one, the Franck-Condon (FC) approximation can be used to solve the vibronic problem. Here, during an electronic transition, the electron jumps between states take place in a short time. In such short time the nuclear positions can be considered frozen around the equilibrium geometry.

Hence the nuclear shifts are small during the electronic transition. To account this, a truncated Taylor expansion can be used.

$$\mu_{if} \approx \mu_{if}(Q_0) + \sum_{k=1}^{N} \frac{\partial \mu_{if}}{\partial Q_k} \tag{2.74}$$

$$\langle \psi_i | \hat{\boldsymbol{\mu}} | \psi_f \rangle \approx \mu_{if}(Q_0) \langle \psi_i | \psi_f \rangle + \sum_{k=1}^{N} \frac{\partial \mu_{if}}{\partial Q_k} \langle \psi_i | Q_k | \psi_f \rangle + .. \tag{2.75}$$

According to equation 2.75, the zeroth-order assumes that the dipole transition moment is constant during the transition and it is called FC approximation. The inclusion of the first-order correction (the so-called Franck-Condon Herzberg-Teller approximation) allows the proper description of the vibronic transitions for systems characterized by weakly-allowed or dipole-forbidden transitions [46, 47].

The computation of the overlap integrals ($\langle \psi_i | \psi_f \rangle$) requires a matrix transformation (Duschinsky transformation) of the normal coordinates.

$$\bar{X} = J\bar{\bar{X}} + K \tag{2.76}$$

In equation 2.76, K is the vector that accounts for the displacement of equilibrium geometry, J is the normal coordinates rotation matrix of the initial state during the transition, $\bar{X}$ and $\bar{\bar{X}}$ are respectively the normal coordinates of the initial and final states. Since K requires the knowledge of the ground and excited state equilibrium geometries, the correct theoretical simulation of the absorption UV-vis spectra requires the knowledge of the PESs involved during a transition. However, for large systems the computation of the PESs is not affordable. To avoid this drawback, the PESs can be built around relevant point defined by the nuclear coordinates through a quadratic Taylor expansion. The harmonic approximation represents the mandatory method at state-of-art level for computing the vibrational fine structure. However, large molecular systems undergo significant structural changes upon a photo-absorption. Hence, harmonic PESs are not superimposed, leading to a lower overlap in the FC region. On the contrary, when the harmonic final-state PES lies in the FC region the PES can reproduced accurately. To deal with it, two computational strategies exist. They are generally categorized in two families: the Adiabatic and the Vertical models [48, 49]. For the first model, the final- and the initial-state are expanded around their respective minimum, leading to higher precision of the band position. On the other hand,

for the vertical method the harmonic final state PES is computed in the FC region and takes into account only the initial state geometry. Hence, the Adiabatic method describes more precisely the vibronic fine structure with respect to the Vertical one. On the other hand, Vertical approach describes more accurately the most intense transition, however, it lacks in the correct description of the band position. In order to both improve the description of the band position and to decrease the computational cost, the vertical model is used in its the gradient variant (VG). Here, the shift of the equilibrium geometry is estimated by the gradient of the excited state energy at the ground state geometry. Logically, one can argue that if no-geometrical changes occur in the final-state PES, the two minima are vertically aligned and thus the approaches are equivalent. Since we are interested in defining both the band position and the band enlargement effects with high level of accuracy, in this thesis two methods have been applied for a set of substitute dyes: the Adiabatic Hessian (AH) and the VG model. This latter has been used to obtain the absorption spectra in conjunction with a time-dependent approach (chapter 2.4.8). An example of vibronic spectra obtained at AH level, it is shown in figure 2.5.

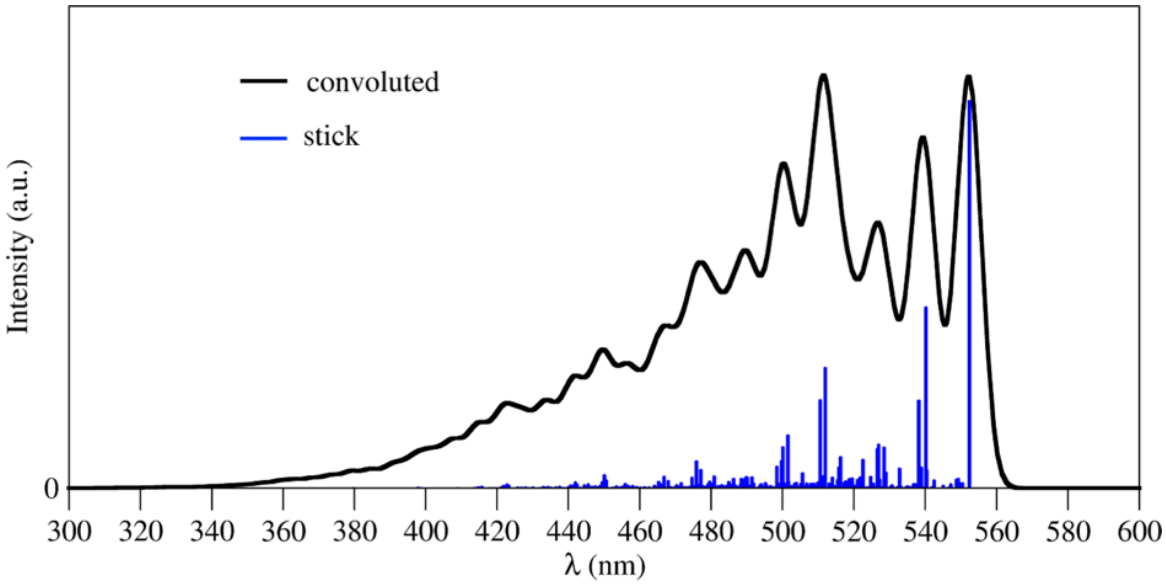


Figure 2.5. Stick and convoluted vibronic spectra of 1,4-Diaminoantharquinone. Data obtained at the ωB97XD/methanol (PCM) level.

2.5.7 Adiabatic Hessian method

In this approach both ground and the excited state PESs are expanded to the second order Taylor formula around their equilibrium geometries [48] (Figure 2.6).

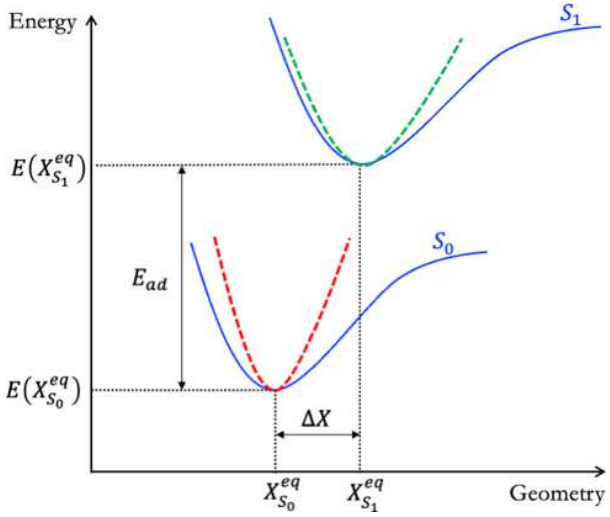


Figure 2.6. Schematic representation of the adiabatic hessian model where solid blue curves correspond to ground- (S_0) and excited- (S_1) singlet electronic states. Inside both PESs, Harmonic PESs (dash lines) are also displayed.

This means that, only frequencies, atomic displacement and J of the normal mode changes are considered. Going into the details, since the gradient is zero at the equilibrium geometry, the harmonic PES is due only given by the Hessian.

$$V(\bar{\bar{X}}) = E_{ad} + \frac{1}{2}\bar{\bar{X}}^T (H^{S_1})^2 \bar{\bar{X}} \tag{2.77}$$

Where H^{S_1} is the Hessian of the excited state, and E_{ad} is the adiabatic energy which corresponds to the difference between the minimum energy of the final state PES with respect to the minimum energy of the initial PES. Using the Duschinsky transformation the expression of $\bar{\bar{X}}$ in function of $\bar{X}$ (or vice versa) can be obtained [50].

2.5.8 Time-Dependent formalism

Although the time-independent model is widely applied for obtaining the vibronic absorption spectra [51, 52], some problems exist. First, since equation 2.71 involves infinite summations, the number of the vibronic transition increases with the molecule size, thus a pre-screening strategy to select the FC integrals that give large contributions to the spectrum is mandatory [43]. Second, several vibrational initial states have to be considered when the temperature is included [53].

A strategy for overcoming the mentioned limitations consist in employing the time dependent formalism. In short, the δ distribution function (see equation 2.71) is expressed as a Fourier transform, thus the sum over state problem is turned in the time dependent problem. In addition, temperature effects are also included [53]. Under the Born-Oppenheimer and Franck-Condon approximations, the averaged absorption cross section $\sigma(E)_n$ writes:

$$\sigma(E)_n \propto f_n \, R_e \int_0^{\infty} dt \, \langle \psi_{00} | \psi_n(t) \rangle \; e^{\frac{i(E+E_{00})t}{\hbar}} e^{-\gamma(t)} \tag{2.78}$$

Unlike equation 2.70, in the time dependent formalism the infinite summation is not present. In equation 2.78, n identify the excited state, E and E_{00} denote respectively the energy of the incident photon and the ground vibrational state energy, f_n is the oscillator strength, R_e is the operator that take the real part, ψ_{00} labels the nuclear wave function of the ground vibrational state on the ground state while $\psi_n(t)$ denotes the time-dependent nuclear wave function on the excited electronic state. The term $e^{-\gamma(t)}$ represents the population decay which includes the Gaussian and the pure lifetime decays.

In this thesis, the computation of the absorption spectra for some organic dyes has been performed using the approach developed by Truhlar and co-worker [54]. The strategy revolves around the locally quadratic model, where both the ground and excited state PESs are locally quadratic, and the transition dipole moment are geometry independent. In the locally quadratic approach, the quantity $\langle \psi_{00} | \psi_n(t) \rangle$ is central [55]. The time-dependent ψ_n evolving under the influence of the excited state vibrational Hamiltonian, H_{ex}, is obtained as the solution of the time-dependent Schrodinger equation:

$$|\psi_n(t)\rangle \, e^{H_{ex}t/\hbar} \, |\psi_{00}\rangle \tag{2.79}$$

Assuming that the excited state PES shares the same frequencies and normal modes as ground state, in the FC region, only the minima are shifted with respect to each other [56]. Within this approximation the nuclear wave functions overlap integral reads:

$$\langle \psi_{00} | \psi_n(t) \rangle = e^{\Sigma_j \left(\frac{\left(\Delta_j^{(n)}\right)^2}{2} \left(1 - e^{-i\omega_j t}\right) - \frac{i\omega_j t}{2} \right) - \frac{i}{\hbar} E_{ad} t} \tag{2.80}$$

Here, $\Delta_j^{(n)}$ denotes the normal coordinates displacement j from the ground sate to the excited state at their equilibrium geometry, E_{ad} identifies the electronic energy difference between the involved electronic states (see figure 5). Further, the vibronic band is approximated expanding by second order Taylor series the exponential term ($e^{-i\omega_j t} \approx 1 - i\omega_j t - 1/2\omega_j^2 t^2$) in equation 2.80:

$$\langle \psi_{00} | \psi_n(t) \rangle = e^{\Sigma_j \left(\frac{i\omega_j t}{2} \left(\Delta_j^{(n)} \right)^2 + 1 \right) - \frac{i}{\hbar} E_{ad} t + \Sigma_j \left(\frac{i\omega_j^2 t^2}{4} \left(\Delta_j^{(n)} \right)^2 \right)} \tag{2.81}$$

Since the exponential term is expanded to the second order, the method is called the locally quadratic second order (LQ2). According to this approximation the absorption cross section is:

$$\sigma(E)_n \approx f_n e^{-\frac{(E - E_{ve})^2}{\hbar^2 \Sigma_j \left(\Delta_j^{(n)} \right)^2 \omega_j^2}} \tag{2.82}$$

Where the E_{ve} term is the vertical excitation energy. From equation 2.82 one can see that the approximation adopted allows the definition of a gaussian vibronic shape, where the denominator identifies the frequency dependent broadenings. On the other hand, expanding to the third order the exponent the LQ3 is derived, leading to the following equation:

$$\sigma(E)_n \approx f_n \int_0^\infty dt \, e^{\Sigma_j \left(\frac{\left(\Delta_j^{(n)} \right)^2 \omega_j^2}{4} \right) t^2} \cos \left(\frac{1}{\hbar} (E - E_{ve}) t + \sum_j \left(\frac{\left(\Delta_j^{(n)} \right)^3 \omega_j^3}{12} \right) t^3 \right) \tag{2.83}$$

The energy difference term is now the argument of the oscillating function. Unlike the Adiabatic hessian model, the approach developed by Thrular and co-workers [54] allows the lowering of the computational time, since the excited state PES is expanded to the second order about the ground state equilibrium geometry according with the VG model.

$$E^{S_1}(\bar{X}) = E_{ve} + (\nabla \mathrm{E}^{S_1})^T \bar{X} + \frac{1}{2} \bar{X}^T (H^{S_0})^2 \bar{X} \tag{2.84}$$

Where ∇E^{S_1} and H^{S_0} are respectively the gradient the Hessian of the excited state at the ground state equilibrium geometry. Key theoretical parameters, including the atomic displacement (ΔX_{ve}) over ground state coordinates, are pictorially represented in Figure 2.7.

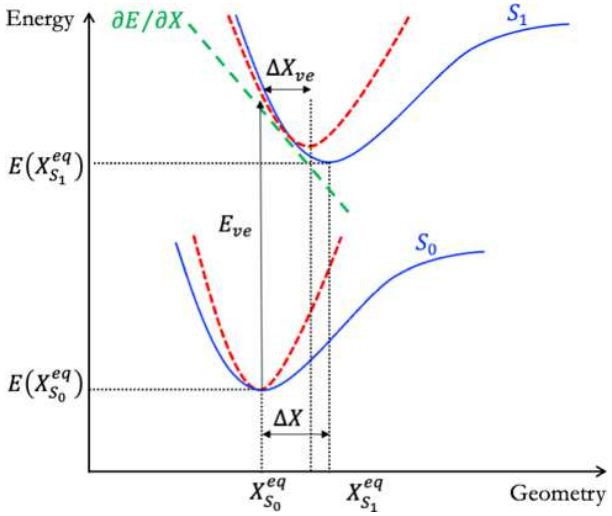


Figure 2.7 The simplified illustration of the vertical model. Solid blue curves indicate the ground- (S_0) and excited- (S_1) singlet electronic states, while the dashed green line is the gradient ∇E^{S_1}. Harmonic PESs are also depicted (red dash lines).

2.5.9 Modelling absorption spectra via convolution functions

Vibronic fine structure can be seen as composed by several vibrational transitions which can be decomposed in single bands [57,58]. Each band is characterized by two parameters: the intensity and the band position. The first parameter is evaluated from equation 2.59. The second one identifies the maximum absorption wavelength (ε_i) of the band. This position represents the transition energy from the vibrational ground state to the vibrational excited state.

As discussed in the previous chapter, absorption bands are characterized by a defined shape that can be obtained by vibronic calculations, however, these computations could require significant computational time. However, alternative methods have been developed to simulate the band shape. A natural approach consists in approximating the absorption profile using the convolution of symmetric functions. This convolution is done via the summation of either Gaussian ($G_i(x)$) or Lorentzian ($L_i(x)$) functions:

$$G_i(x) = N\, e^{-4\ln(2)\left(\frac{x-\varepsilon_i}{\sigma_i}\right)^2} \tag{2.85}$$

$$L_i(x) = N\frac{1}{1+4\left(\frac{x-\varepsilon_i}{\sigma_i}\right)^2} \tag{2.86}$$

Where N is the normalization factor, x is the data point, ε_i is the energy and σ_i is the full width at half maximum (FWHM) of the ith electronic transition. Both functions have "bell"-like shapes, with σ_i that controls their width peak, and on both sides of the peak the tails of the curve quickly fall off on the x-axis. The Lorentzian function has two important features that distinguishes it from the Gaussian one, these differences are highlighted in Figure 2.8. The Lorentzian is a little narrower at its apex, and it extends out further on its sides/edges.

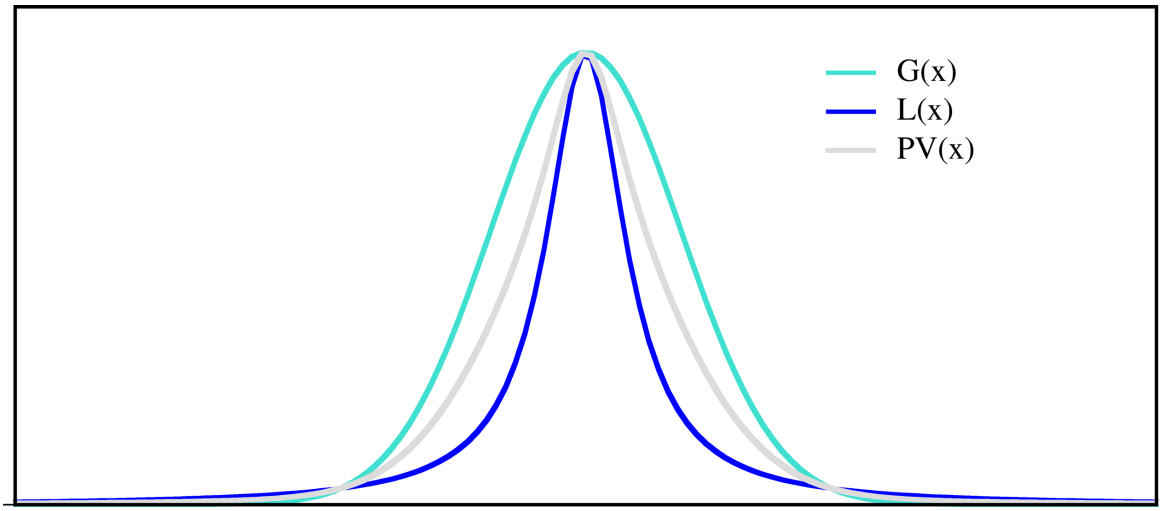


Figure 2.8. Graphical comparison between the Gaussian ($G(x)$), Lorentzian ($L(x)$) and pseudo-Voigt ($PV(x)$) profile are made with σ_i equal 0.5 eV. For $PV\ (x)$ the M parameter is equal 0.5.

One can think to combine these two functions linearly. The resulting function, Figure 2.8, is known as pseudo-Voigt function (PV):

$$PV_i(x) = N\left((1-M)e^{-\left(\frac{x-\varepsilon_i}{\sigma_i}\right)^2} + \frac{M}{1+4\left(\frac{x-\varepsilon_i}{\sigma_i}\right)^2}\right) \tag{2.87}$$

A model based on the Pseudo Voigt function has been used in this thesis and will be discussed in Chapter 3.

2.6 Theory of colour

The relationship between colour and transmission spectra in the context of human perception has been investigated in this thesis. In this chapter, we begin by analysing the human visual system to understand how to construct a theoretical model which connects our colour perception to the spectral properties.

In short, human vision involves two types of photoreceptors found into our retina: rods and cones [59]. The former are responsible for the vision in dark conditions, whereas cones are responsible for vision in bright conditions [59]. More precisely, cones are called S, M, and L and differ between them by their sensibility to short (S), medium (M) and long (L) wavelengths. These cones participate in the process of colour vision, indeed, once light stimulates the cones the resulting signals are processed in the eye and finally are interpreted by the brain. Such intricate process allows us to perceive the colour. However, this spectral feature needs to accurately be described by giving some basic perceptual attributes called: hue, lightness and chroma. The hue allows us to describe the colour tone (here, colours are categorized in a closed ring). The lightness identifies the luminous of a colour in a series of greys which range from black to white. Finally, the chroma is the colourfulness thus the colour saturation.

In addition, once these features have been defined, one should figure out how to quantify the colour perception. This is the aim of the *colourimetry*, which allows the numerical expression of the colour.

2.6.1 Tristimulus values

In the previous part, we pointed out that human vision is mediated by the S, M and L cones which possess different spectral sensitivities. Therefore, a natural consequence is that the colour vision is trichromatic and thus three quantities are sufficient for describing the colour. The colour quantification can be introduced by a method used to specify colour stimuli. The method, that allows the understanding of the additive synthesis of colours, is based on the fact that the colour sensation can be produced by a mixing of three primary stimuli.

Briefly, around 1930, Wright performed experiments in which the combination of the light at 435.8 nm, 546.1 nm and 700 nm the resulting in a colour perception matching with colour produced by a monochromatic light [60]. From these experiments the RGB colour matching functions ($r(\lambda)$, $g(\lambda)$, $b(\lambda)$) have been obtained. However, in practice, the XYZ colour matching function defined by the CIE 1931 are commonly used instead of the RGB ones, since they possess positive values

for all wavelengths. Through this thesis we used these latter which are valid for an observer view of 10 degree.

The colour appearances of a given object (in our case the molecules) under a light source (i.e., the Illuminant), it is derived from the tristimulus values as follow:

$$X = k\int_{380}^{780} S(\lambda)T(\lambda)\bar{x}(\lambda)d\lambda \tag{2.88}$$

$$Y = k\int_{380}^{780} S(\lambda)T(\lambda)\bar{y}(\lambda)d\lambda \tag{2.89}$$

$$Z = k\int_{380}^{780} S(\lambda)T(\lambda)\bar{z}(\lambda)d\lambda \tag{2.90}$$

$$k = \frac{100}{\int S(\lambda)\bar{y}(\lambda)d\lambda} \tag{2.91}$$

Where $S(\lambda)$ is the spectral power distribution of the Illuminant, $T(\lambda)$ is the transmission spectrum of the object and k is a normalising constant.

It should be noted that different source of light (i.e., the Illuminant), entail different colour perception by human beings. Since the CIE recommends to use the Illuminant D65 when the sample is illuminated by the sunlight. Colours reported in this thesis have been obtained using this standard. [61]

In addition, as it can be seen from the previous equations, for a perfectly-reflecting object (i.e., $T(\lambda)$=1) Y is equal to 100. On the other hand, the situation when Y is equal to zero corresponds to a dark object (since it absorbs all light). Thus, the Y value correlates colorimetric calculation with the brightness.

The tristimulus values are linked to the following chromaticity coordinates:

$$x = \frac{X}{X+Y+Z} \tag{2.92}$$

$$y = \frac{Y}{X+Y+Z} \tag{2.93}$$

$$z = \frac{Z}{X+Y+Z} \tag{2.94}$$

Since $x+ y + z$ = 1.00, only two chromaticity coordinates need to be given instead of three [59].

The bidimensional plane constructed by using the x and y chromaticity coordinates (Figure 2.9) is called the CIE 1931 chromaticity diagram and allows us to represent the chromaticity of the light stimulus.

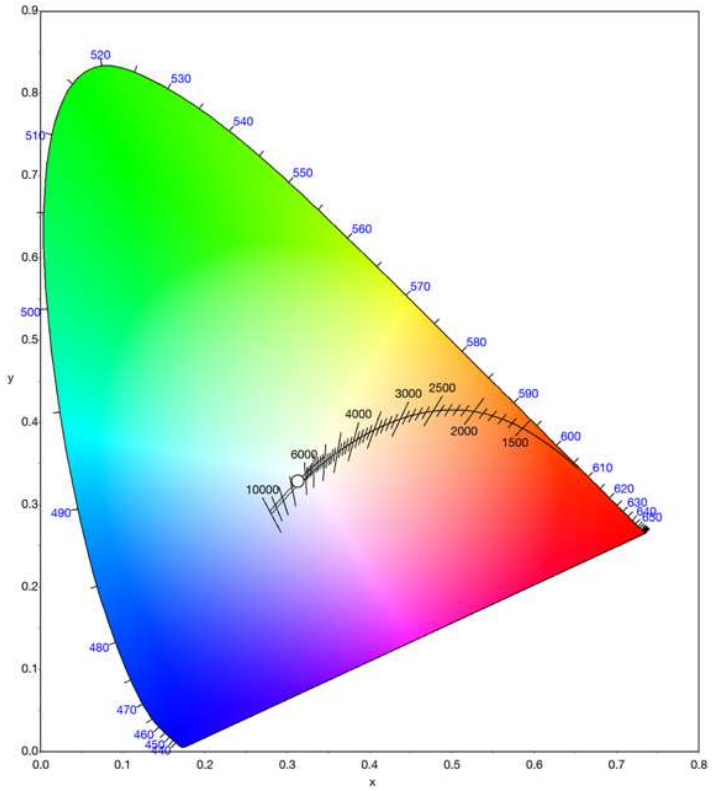


Figure 2.9. CIE 1931 chromaticity diagram. Illuminant D65 (white circle).

In this diagram, the spectral locus is the outer line of this diagram, while the black curved inner line is the blackbody locus showing the colour coordinates of a blackbody radiator at different temperatures.

2.6.2 CIE-Lab colour space

CIE-Lab colour space (Figure 2.10) is the more robust one and it is routinely used for quantitative colour comparation. This is due to the fact that this multidimensional colour space aims to be perceptually uniform. In other words, colours separated by a distance into Lab-model is correlated with the human perception [62].

To get an uniformly spaced value, Y is mathematically translated in other values (L*). Going into the details, differences in lightness are retained into the Y scale. In the resulting scale, L* goes from 0 to 100, and runs in a vertical axis (Figure 2.10).

L* equal to 100 represents a perfect reflecting diffuser (white), while L* equal to zero represents black. Lightness function is represented by the equations:

$$L^* = 116\left(\frac{Y}{Y_n}\right) - 16 \qquad \text{if } \frac{Y}{Y_n} > 0.00856 \tag{2.95}$$

$$L^* = 903.3\left(\frac{Y}{Y_n}\right) \qquad \text{if } \frac{Y}{Y_n} < 0.00856 \tag{2.96}$$

The CIE-Lab involves other two coordinates (a^* and b^*), which can be positive or negative. The chromatic a^* axis extends from red (+ a^*) to green (- a^*), while b^* axis extends from yellow (+ b^*) to blue (- b^*), Figure 2.10. The mathematical transformations are reported below:

$$a^* = 500\left[\left(\frac{X}{X_n}\right)^{1/3} - \left(\frac{Y}{Y_n}\right)^{1/3}\right] \tag{2.97}$$

$$b^* = 200\left[\left(\frac{Y}{Y_n}\right)^{1/3} - \left(\frac{Z}{Z_n}\right)^{1/3}\right] \tag{2.98}$$

The terms X_n, Y_n and Z_n are the nominally white object colour stimulus (i.e., the reference white). As mentioned, the second parameter for defining the colours is the chromacity (C^*), Figure 11. This parameter identifies the amount of saturation of colour, and it is expressed as:

$$C^* = \sqrt[2]{a^{*2} + b^{*2}} \tag{2.99}$$

Lastly, the third colour attributed, the hue metric angle h^* (Figure 2.10) which identifies the colour tone is obtained as [63]

$$h^* = arct(b^*/a^*) \tag{2.100}$$

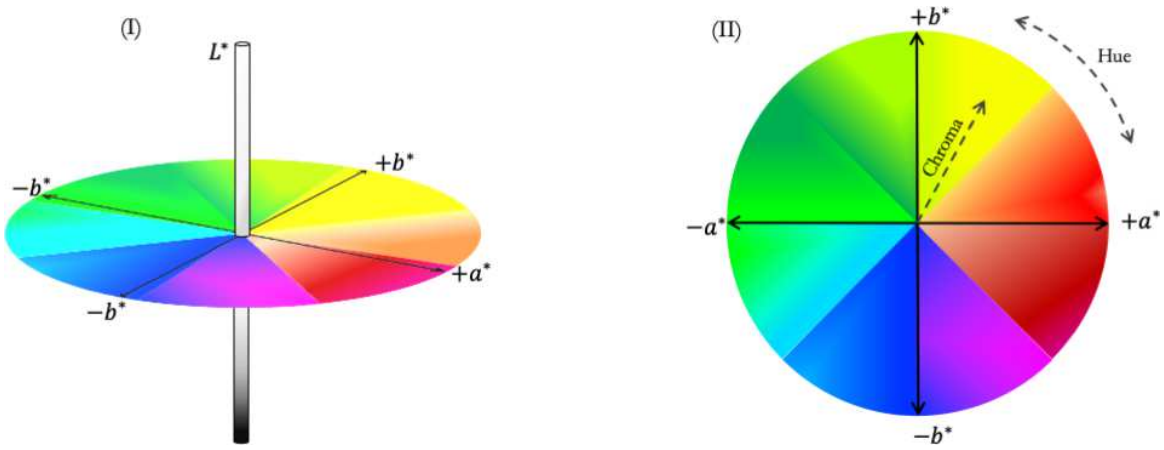


Figure 2.10 Graphical representations of the CIE-Lab colour space (I) chroma (II) and hue angle (II).

From a quantitative point of view the colour evaluation is computed by colour difference within the CIE-Lab colour space, that a predicted colour presents with respect to the reference one (i.e., experimental colour). The parameters involved into the colour differences are ΔL *, Δa * and Δb *. Once obtained ΔL^*, Δa^* and Δb^*, the total difference is established via Euclidean distance (d_{eucl}). Then:

$$d_{eucl} = \sqrt[2]{\Delta L^{*2} + \Delta a^{*2} + \Delta b^{*2}} \tag{2.101}$$

2.7 Molecular Mechanics

Unlike quantum mechanics methods, the Molecular Mechanics (MM) allow us to avoid the explicit electronic description. For instance, the molecular enthalpy can be obtained as sum of bond enthalpies, allowing the computation of the molecular energy starting from single components that define the potential energy, $V^{MM}(r_N)$. Therefore, the energy in MM is obtained by summing some energetic terms, which can express different functionals forms (such as for the $V_{stretching}$) and they include atom-based parameters. They expression is called force field (FF). Among the FF in literature, the AMBER FF is one of the most popular, that has been developed for studying proteins and nuclei acids. During this thesis work, we used a derivation of the AMBER FF namely General AMBER FF (GAFF) that allows the study of a large class of organic compounds.

2.7.1 Potential energy terms

The potential energy function is composed by two terms [64]. One regards the bonded interactions, while the second regards the non-bonded terms.

$$V^{MM}(r_N) = \underbrace{V_{stretching} + V_{bending} + V_{torsional}}_{\text{bonded}} + \underbrace{V_{LJ} + V_{elec}}_{\text{non-bonded}} \tag{2.102}$$

The total potential energy function is dependent on the position (r) of the nuclei (N) with stretching, bending and torsional components (bonded interactions) and the Lennard-Jones potential and electrostatic effects (non-bonded terms). Focusing on equation 2.88, the first three terms write:

$$V_{stret} = \sum_{i}^{N_b} \frac{1}{2} k_{b,i}(b - b_{0,i})^2 \tag{2.103}$$

$$V_{bend} = \sum_{i}^{N_\theta} \frac{1}{2} k_{\theta,i}(\theta - \theta_{0,i})^2 \tag{2.104}$$

$$V_{torsion} = \sum_{i}^{N_{tor}} \sum_{J}^{N_J} \frac{V_{J,i}}{2} \left[1 + \cos(J\omega - \varphi_{J,i})\right] \tag{2.105}$$

In the first two equations, k_b and k_θ are respectively the force constants for the stretching and bending, $b_{0,i}$ and $\theta_{0,i}$ are equilibrium positions. In the torsional term, J represents the multiplicity, V_J is the barrier height, ω is the torsional angle which is defined by a set of four atoms and φ_J is the phase angle.

Regarding the non-bonded interactions the functional forms are reported below:

$$V_{LJ} = \sum_{i,j \neq i}^{N} \left[\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}}\right] \tag{2.106}$$

$$V_{elec} = \sum_{i,j \neq i}^{N} \frac{q_i q_j}{r_{ij}} \tag{2.107}$$

The Lennard-Jones potential V_{LJ} is composed by two terms, a first repulsive term due to the Pauli's exclusion principles while the second one corresponds to an attractive term that is the van der Waals dispersion component. The A_{ij} and B_{ij} terms are parameters obtained via fitting procedures. The electrostatic potential V_{elec} comes from the Coulomb's law and describes the classical interaction between charged atom q.

2.8 Simulation technique: Monte Carlo Method

As mentioned in the introduction, for rigid- or semi-rigid chromophores the vibronically resolved spectra can be computed within the FC region with good accuracy. The situation is instead more complex for flexible chromophores including their aggregates [15]. The correct modelling of their absorption spectra remains a current computational challenge. For these dyes, the ground and excited state geometries are significantly different and, furthermore, they could be characterized by several isoenergetic minima that prevent the use of the FC approximation. In order to overcome this challenge one can think to explore the phase space by using simulation techniques. Here, several methods can be adopted. Among them, the Monte Carlo Metropolis scheme allows the phase space to be efficiently mapped. Within this framework, the following chapter aims at elucidating the fundamental concepts to sample the phase space using the Monte Carlo approach.

2.8.1 Statistical mechanics in a nutshell

Statistical Mechanic (SM) deals with the prediction of properties of a "many-body" system. Here, the system is described in terms of both microscopic constituents and their interactions. The key feature in SM is the partition function Z that provide a bridge between the microscopic and macroscopic worlds.

Since macroscopic properties are the result of large number of molecules characterized by a certain distribution of energies. If the possible energy states can be determined for an individual molecule, SM can be used to compute macroscopic properties.

In this framework, macroscopic quantities are estimated by an efficient sampling of the phase space. This is the essence of the *simulation methods* like Monte Carlo and Molecular Dynamic: to collect a representative sample of the phase space.

2.8.2 Phase space

For a classic system composed by N particles is described in terms of position $\{r_1(t), \dots, r_N(t)\}$ and momentum $\{p_1(t), \dots, p_N(t)\}$. Its time evolution is described by the Newton's equations. In terms of momenta the second Newton's law is:

$$\boldsymbol{F} = m\boldsymbol{a} = m\,\frac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t} = \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} \tag{2.108}$$

The collective description of the states of the system is resumed by a representative point $\boldsymbol{x}$:

$$\begin{aligned} x(t) = \{&q_{x,1}(t), q_{y,1}(t), q_{z,1}(t), \dots, q_{x,N}(t), q_{y,N}(t), q_{z,N}(t), \\ &p_{x,1}(t), p_{y,1}(t), p_{z,1}(t), \dots, p_{x,N}(t), p_{y,N}(t), p_{z,N}(t)\} \end{aligned} \tag{2.109}$$

Where q_i is the ith generalized coordinate. Equation 2.95 defines a parametrical path the is called trajectory, that moves in the *phase space*.

2.8.3 Compute macroscopical observables

Molecular modelling at absolute zero (0K) corresponds to the characterization of the PES. At this temperature it is reasonable to assume that just the fundamental state is populated. For systems that are of interest to SM (i.e., $\mathcal{O}(10^{23})$ particles) at finite temperature, there are several microscopic configurations which lead to the same macroscopic observable.

Originally introduced by Gibbs, the ensemble concept finds its fundamentals on the idea that macroscopic observables are insensitive to precise microscopic details [65]. The ensemble is defined as a collection of systems that share a set of macroscopic properties and evolves under microscopic laws of motion from a different initial condition. Then, heuristically, given an ensemble, macroscopic observables are computed performing averages over the systems into the ensemble:

$$A = \frac{1}{\mathcal{Z}}\sum_{i=1}^{\mathcal{Z}} A(x_i) \equiv \langle A \rangle \tag{2.110}$$

Where A is the macroscopic observable, $A(x_i)$ is the microscopic phase space function and $\mathcal{Z}$ is partition function. This function represents the total number of accessible microscopic states.

2.8.4 Ergodic hypothesis

The main difference between the PES and the phase space is that this latter includes the momentum while PES is a function of the position vector. The average property $\langle A \rangle$ (equation 2.110) can thus be computed following the time evolution (i.e., trajectories) of the system and averaging it over the time, this is the idea behind the *Molecular Dynamics* (MD)simulations [26]. The trajectory is unique in the phase space, because it is dependent on the initial condition. On the other hand, despite their uniqueness, trajectories lead, at long time limit, to the same macroscopical observable, for instance the vibrational spectrum. Taking this into account, an assumption can be done:

$$\langle A \rangle = \lim_{t \to \infty} \frac{1}{t} \int_{t_0}^{t_0+t} A(\tau) d\tau = \lim_{M \to \infty} \frac{1}{M} \sum_{i=1}^{M} A(x_i) \tag{2.111}$$

Eq. 2.111 defines the well-known *ergodic hypothesis* (EH), which assumes that at long time, one can map all points in the phase space, so-that the time-dependent integral is equal to the space-dependent one. The Monte Carlo simulation, as well the MD, is based on the EH.

2.8.5 Monte Carlo method

For system of interest to SM, the number of local minima rapidly grows with the number of atoms. In other words, unlike the static approach that allows the local minima to be searched using optimization algorithms, when a system is characterized by a large number of degrees of freedom a method able to map the phase space is necessary. Hence, starting from the phase space, it is easy to understand that not all (q, p) couple equally contribute to the macroscopic property ($\langle A \rangle$) but with a certain probability $P(q, p)$, so that:

$$\langle A \rangle = \iint dq\, dp\, A(q,p) P(q,p) \tag{2.112}$$

The term $P(q, p)$ is the probability and is expressed as Boltzmann factor:

$$P(q,p) = \frac{e^{-\beta\mathcal{H}(q,p)}}{\iint dqdpe^{-\beta\mathcal{H}(q,p)}} = Z^{-1}e^{-\beta\mathcal{H}(q,p)} \tag{2.113}$$

Where $\mathcal{H}(q,p)$ is the Hamiltonian of the system defined as sum of kinetic $K(q)$ and potential $V(r)$ energy, β is a constant and it is defined by the ratio $1/k_BT$, while the denominator in equation 2.113 is the partition function Z. Since in our case the potential depends on the position, equation 2.112 can be expressed as sum of two integral terms:

$$\langle A \rangle = \int dqA(q)P(q) + \int dpA(p)P(p) \tag{2.114}$$

Now, one can focus on the property independent by the momentum, thus, neglecting the second term in equation 2.114. With this assumption, one can think to explore the phase space by a random sampling, which is the idea behind the *Monte Carlo* method. This means no equations of motion have to be taken into account; thus, the averaged property is directly obtained by a random selection of points in the phase space. At the equilibrium the property is thus computed as:

$$\langle A \rangle = \frac{\int dq\, e^{-\beta E(q)}A(q)}{\int dqe^{-\beta E(q)}} \tag{2.115}$$

2.8.6 Monte Carlo Metropolis procedure

The random sampling via a Monte Carlo method leads to a dramatic drawback. The property slowly converges because several points, which are meaningless to the macroscopic property, are also sampled. For solving this problem, in 1953 Metropolis introduced a strategy which consists of using the "acceptance" probability (P) to select the most relevant points in the phase space [66]. Usually, the Monte Carlo sampling is used within the Metropolis scheme, reason why the method is also called *Monte Carlo Metropolis sampling method*. In this algorithm, new configurations are independent by any previous configuration (Markov chain) but only depend on the previous configuration. In other words, there is no history dependence. In the canonical ensemble (i.e., an ensemble where the number of molecules N, volume V, and temperature T are kept constant

during the simulation, NVT), the Monte Carlo Metropolis scheme follows the steps reported below [67]:

1. Given a structure with nuclear rearrangement q_i, the Monte Carlo Metropolis sampling starts by computing the energy ($E(q_i)$)

2. On the nuclear coordinates q_i a trial movement is applied and, thus, a new configuration is obtained q_j. Subsequently, the energy at this nuclear rearrangement is computed $E(q_j)$.

3. In order to keep points in the phase space the acceptance Monte Carlo Metropolis criterion must be satisfied:

$$P = min\left(1, e^{\beta(E(q_i)-E(q_j))}\right) \tag{2.116}$$

In order to decide what trial movements must be accepted or not, a selection criterion is applied. If $E(q_j)$ is lower in energy than $E(q_i)$ the point is accepted. If $E(q_j)$ is higher in energy than $E(q_i)$, a random point $F$is selected, thus, P is compared with the probability of F. If P is higher than F the structure with nuclear rearrangement q_j is accepted otherwise is rejected. This represents the key factor to escape from the local minima.

As shown in equation 2.113, as well as in equation 2.115, a macroscopic property is dependent on the partition function, thus not surprisingly, the probability is dependent by the ensemble chosen. Because in Monte Carlo Metropolis algorithm, as it is described, is a procedure to find the absolute minimum, we need to introduce a further control on the temperature in order to properly explore the configuration space of the system at given temperature T. This thermostat, we chose, is the Hoover chain [67] one that bring us to a proper ***NVT*** ensemble where the partition function reads:

$$Z(NVT) = \prod_{i=1}^{N} \frac{1}{\Lambda_i^3 N!} \int dq\, e^{-\beta E(q)} \tag{2.103}$$

Where Λ is the De Broglie wavelength ($\Lambda = \sqrt{h^2\beta/2\pi\, m_i}$).

Chapter 3

3. Building simple protocol to predict colors

In this chapter we report our study on two molecular classes: anthraquinone and indigo. Specifically, we tested different computational procedure. On the basis of our results, we selected the best computational protocols able to predict the colors of both anthraquinones and indigos.

3.1 Introduction

Dyeing agents are ubiquitously present in daily life and a significant Research and Development (R&D) effort is being done to produce new substances with specific color-related properties, higher chemical stability, and low environmental impact. Among the best-known coloring agents, anthraquinone (AQ) and indigo (IN) derivatives are extensively used due to their low cost, chemical stability and easy availability in nature [68,69]. They also have a large color tunability thanks to the possibility of functionalization of the parent skeleton with auxochrome groups. Consequently, a large color range palette can be obtained using these families of dyes, thus further increasing their industrial appeal [70,71].

As often the case in well-consolidated fields of Chemistry, new dye molecules are usually obtained by a trial-and-error procedure, based on (simple) empirical models and the large chemical experience acquired over the years. Despite the long-standing knowledge acquired by decades of active research, dye chemistry still requires improvements not only concerning the dye-coloring performances, but also to improve scalability and to lower the environmental impact of the associated industrial processes. For instance, even if innovative biosourced processes to obtain AQ dyes represent a valid alternative with lower environmental impact [72,73], the reactant cost and some possible time-consuming steps in their synthesis are typical drawbacks that cannot be ignored for a final industrial application.

Most of these problems require hard work in laboratory, but rationale property-driven approaches, based on accurate modeling, could speed-up the R&D process reducing the number of candidate molecules to be synthesized and tested. Indeed, modern computational techniques allow for accurate evaluation of a large number of physico-chemical properties, often with accuracy (defined

as "chemical accuracy") comparable to that of the experimental counterpart [74-76]. Further statistics-based approaches, such as Machine Learning or Artificial Intelligence models [77,78], can increase both the accuracy and the number of molecules that can be handled, their robustness requiring also well-defined endpoints and proper descriptors. In the case of (molecular) dyes, their color in solution often represents a discriminating properties. Color can be predicted starting from the absorption spectra [79,80,12], providing that an accurate description of the band shapes over the whole of UV-vis range is given. A large panel of electronic structure methods have been developed over the years to model molecular properties of the excited electronic states, ranging from semi-empirical models to refined post-HF approaches [81-85]. Among others, the so-called Time-Dependent Density Functional Theory (TD-DFT) appeared in the last years as a suitable approach to describe large molecules due to its favorable accuracy/computational time ratio, making it compatible with R&D time scales [49,82, 85-88]. A clear picture of the situation can be obtained by looking at the numerous benchmarks in literature devoted to the assessment of the accuracy of this electronic-structure methods (see for instance refences 44-125), having as pursued accuracy (aka chemical accuracy) a threshold of 0.04 eV for λ_{max}, usually identified as the maximum of the first intense peak [89]. Typical deviations on λ_{max} with respect to experimental data for organic dyes range between 0.1 eV and 0.2 eV, depending on the type of molecule and/or excitation considered (valence, Rydberg, charge transfer) [81,88-91]. Nevertheless, there is no guarantee that functionals providing a high accuracy for the λ_{max} associated to the lowest energy band are also able to provide an accurate description of the associated band shape and, thus, of the perceived color [49-92]

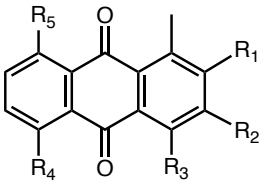
This last task is particularly challenging since a large contribution to the broadening of the UV-vis peaks comes from the coupling between electronic and nuclear degrees of freedom. The vibronic coupling is usually computed using the Franck-Condon (FC) approach and requires the evaluation of structures and (harmonic) vibrational frequencies at both ground and excited electronic states (GS and ES) [44, 93,94]. A part for the requirement of adequate computational resources, FC models can be routinely applied to dyes undergoing to small structural variations upon electronic excitation (defined as rigid or semi-rigid molecules). Technical difficulties (e.g. excited-state geometry optimization) and the relatively-high computational costs have limited the number of benchmarks carried out on the obtained band shape and/or associated color. However, approximated FC approaches, based on simplified Hamiltonians, have been developed and successfully applied, even if their use is not yet generalized [100,101]. Alternative approaches, combining electronic-structure methods with Molecular Dynamics (MD) or Monte Carlo simulations, have been developed for dyes showing larger variations in structures between their

GS and ES (in the case of flexible or floppy molecules) [15, 102-104]. These approaches have the advantage of being based usually on a classical sampling of the ground state potential energy surface (PES), but they indeed require, to obtain an accurate evaluation of the band shape, a relevant number of excited state calculations on the structures extracted from the classical MD trajectories (snapshots).

At the opposite of these computational approaches, a simple and long-standing strategy for deriving UV-vis spectra from quantum chemical calculations consists in convoluting the vertical transitions, computed with a given electronic method (e.g., TD-DFT) with Gaussian or Lorentzian functions [104-106]. Indeed, this type of broadening functions can reproduce both inhomogeneous and homogeneous broadening related to thermodynamics events or lifetime of the lines, respectively. While the electronic-structure calculations provide the position of the band maxima and the corresponding oscillator strengths, the full width at half maximum (FWHM) of the functions used for the convolution is an adjustable parameter, usually fixed in the case of Gaussian functions to 0.2-0.3 eV or determined so to minimize the deviations with the available experimental data. In such a way, broadening factors, such as vibrational fine structure, are just mimicked and the typical band asymmetry due to the vibrational progression is not necessarily reproduced. However, the simplicity of the approach justifies its large applications in literature. It should be remarked, however that other fitting functions, allowing a better simulation of the band asymmetry generated by the vibronic coupling, are currently used to fit electronic spectra at different energies, such as XPS spectra [107]. Their use in UV-vis range is, however, rare.

Here we propose to define the validity of four different approaches for the definition of the band shape and the associated color, starting from TD-DFT calculations. They span from: (I) a refined quantum mechanical model based on the evaluation of the FC coupling to (II) its simplified approach recently proposed by Truhlar and coworkers [54], to two simple convolution procedures based on Gaussian (III) and on a more complex (pseudo-Voigt, IV) broadening functions.

These four models will be tested comparing their results to experimental spectra and then their robustness verified by their application to a selected number of dyes belonging to the indigo (IN) and anthraquinone (AQ) families, as depicted in Figure 3.1.



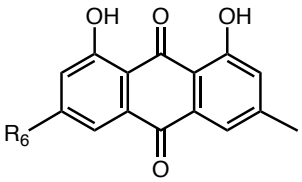
AQ1 $R_1 = H; R_2 = NH_2; R_3 = H; R_4 = H; R_5 = OH$

AQ2 $R_1 = H; R_2 = OH; R_3 = NH_2; R_4 = H; R_5 = OH$

AQ3 $R_1 = H; R_2 = OH; R_3 = NH_2; R_4 = NH_2; R_5 = OH$

AQ4 $R_1 = COOH; R_2 = OH; R_3 = NH_2; R_4 = NH_2; R_5 = OH$

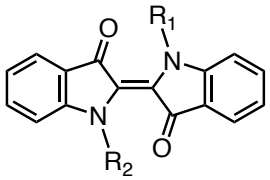
AQ5 $R_1 = H; R_2 = NHC_6H_{12}; R_3 = H; R_4 = H; R_5 = NHC_6H_{12}$



AQ6 $R_6 = OH$

AQ7 $R_6 = H$

AQ8 $R_6 = OCH_3$



IN1 $R_1 = H; R_2 = H$

IN2 $R_1 = CH_3; R_2 = CH_3$

IN3 $R_1 = CH_2CH_2CH_3; R_2 = CH_2CH_2CH_3$

IN4 $R_1 = CO_2tBu; R_2 = CO_2tBu$

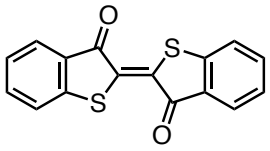
IN5 $R_1 = Bn; R_2 = Bn$

IN6 $R_1 = CH_3; R_2 = p\text{-}CF_3Bn$

IN7 $R_1 = CO_2tBu; R_2 = p\text{-}CNBn$

IN8 $R_1 = CO_2tBu; R_2 = p\text{-}OCH_3Bn$

IN9 $R_1 = CO_2tBu; R_2 = p\text{-}OCH_3Bn$



IN10

Figure 3.1 Sketches of the anthraquinones and indigos investigated in this work.

Beyond the accuracy in absorption spectra maxima and band shape, the property that will be considered as target will be the predicted perceived colour, defined by its coordinated in the CIELAB space as well as the associated chroma and hue angle [62,63]. The two families of dyes, above mentioned, have been considered as representative benchmark.

Beyond their interest for industrial applications, the dyes molecules here studied have been selected since their optical properties have been largely studied and, due to their structural features, the FC approximation is expected to work reasonably [9, 49]. Furthermore, they cover a wide variety of colors, from orange/reddish to blue, allowing the better validations of the models.

Our final aim is to define a computational protocol coupling TD-DFT calculations for vertical excitation energies with a model allowing an accurate and fast prediction of the perceived color of dye in solution.

3.2 Computational details

Albeit all the methods considered in the present section are well defined in Chapter 2 and they are shortly summarized here for sake of completeness and for helping the interest readers.

3.2.1 Underpinning electronic-structure methods

Ground and excited states properties were computed using Density Functional Theory (DFT) and its Time-Dependent (TD-DFT) extension, respectively [102] using the Gaussian software [103]. On the basis of previous works [8-49], we selected two functionals PBE0 [31] and ωB97XD [36] in conjunction with the 6-31++G(2d,2p) basis set. The first functional, PBE0, is known to provide accurate λ_{max} values for a large number of molecule and π- π^* transitions, while the second one, ωB97XD, is more suitable for reproducing vibronic coupling.

Ground- and excited-state geometries have been fully optimized in solvent, using the IEF-PCM model to mimic solvent effects[11]. For consistency with available experimental data, AQ dyes were computed in methanol [83,104]. while different solvents were considered for the IN family: chloroform (for IN1 and IN10) [105, 106], n-dodecane (for IN3) [107], and acetonitrile (for IN2 and for IN4 to IN9) [108]. For both class of dyes, vibrational frequencies were computed at the same level of theory to verify that the identified structures well correspond to true energy minima, in S_0 or S_1 states. TD-DFT calculations provide the vertical excitation energies, the oscillator strengths and, for FC calculation, vibrational states at both ground and excited electronic states. Please note that for all dyes considered a single electronic excitation is occurring in the UV-Vis range and has been considered in the band shape modelling of all systems.

3.2.2 Modeling the shape of UV-vis spectra

As mentioned in the introduction several models of increasing complexity were considered in the present work to model the absorption spectra. In the simplest approach, each transition obtained at TD-DFT level is convoluted with a Gaussian broadening function [109]

$$g_i(x) = \frac{A_0}{\sigma}\sqrt{\frac{4ln2}{\pi}}exp\left(-4ln2\frac{(x-x_{i0})^2}{\sigma^2}\right) \quad (3.1)$$

where x_{i0} is the energy computed at TD-DFT level for the *i* transition, and σ the full width at half maximum (FWHM) of the Gaussian broadening function, here considered to be the same for all electronic transitions. The normalization factor (A_0) is here chosen in order to recover the molar absorptivity expressed in $Mmol^{-1}cm^{-1}$, it is expressed as:

$$A_0 = \frac{N_a e^2}{4 m_e c^2 \epsilon_0 \ln(10)} \tag{3.2}$$

Where N_a is the Avogadro's number, c is the speed of light, e is the charge on the electron, m_e is the electron mass and ϵ_0 is the vacuum permittivity. The full spectrum is then reconstructed as sum of the convoluted bands as:

$$S(x) = \sum_{i=1}^{N} g_i(x) \tag{3.3}$$

This simple (and old model) is widely applied in literature and provides in a fast way the whole adsorption spectrum, having as parameter the FWHM. This latter can be either keep fixed to a given single value (usually between 0.2 and 0.4 eV) or be determined by a fitting procedure when experimental spectra are available. In our case a single transition is contributing to the spectra (N=1 in equ. 3) and the FWHM has been fixed to 0.4 eV and 0.2 eV for AQs and INs respectively. Widely applied in XPS and Raman spectroscopy [110,111], a similar convolution procedure can be defined making use of Pseudo-Voigt (PV) broadening functions. PV functions can be defined either as sum or as product of Gaussian and Lorentzian functions [111-114]. Considering the first approach, the resulting pseudo-Voigt broadening associated to an electronic transition *i* is thus defined as [110]:

$$PV_i(x) = (1 - M) g_i(x) + M \ell_i(x) \tag{3.4}$$

where, M controls the mixing ratio of the Gaussian and Lorenzian functions being expressed respectively as expressed in eq. 3.1 and eq. 3.5, below.

$$\ell_i(x) = \frac{2A_0}{\pi\sigma} \frac{1}{1 + 4\frac{(x - x_0)^2}{\sigma^2}} \tag{3.5}$$

Band asymmetric can be recovered substituting the constant full width at half maximum (FWHM, σ) with a sigmoidal function $\gamma(x)$:

$$\gamma_i(x) = \frac{2\sigma_i}{1 + e^{a(x-\varepsilon_i)}} \tag{3.6}$$

Here, *a* governs the asymmetry and can be considered as characteristic parameter for a given molecular family. Positive *a* values skew the spectrum toward higher wavelength, while negative *a* values allows to skew the spectrum at lower wavelength [110**]**.

For $x = x_0$ the sigmoidal function $\gamma(x)$ reduces to the full width at half maximum, λ_{max}. In this way, each Pseudo-Voigt broadening function is finally expressed as :

$$PV\big(x, \gamma(x)\big) = (1 - M)g'(x) + M\ell'(x) \tag{3.7}$$

With $g'(x)$ and $\ell'(x)$ being defined as :

$$g'(x) = N'_g exp\left(-4ln2\frac{(x - x_{i0})^2}{(\gamma(x))^2}\right) \tag{3.8}$$

$$\ell'(x) = N'_\ell \frac{1}{1 + 4\frac{(x - x_0)^2}{(\gamma(x))^2}} \tag{3.9}$$

The functions $g'(x)$ and $\ell'(x)$ are separately normalized by numerical integration (by the means of the constants N'_g and N'_ℓ) to $\frac{N_A e^2}{4m_e c^2 \epsilon_0 \ln(10)}$ allowing to provide molar absorptivity as described in detail in Appendix A. Three characteristic parameters are needed to the define a PV broadening that are: i) the Gaussian/Lorentzian mixing, M, ii) the a value defining the sigmoidal broadening and iii) the FWHM at $x = x_0$, σ. Two set of values have been used in this work: for the AQ family a=-3000, M=0.9 and σ =0.6 eV; while for the IN one a=-2000, M=0.2 and σ=0.2 eV. These values have been selected on the basis of the results obtained on the AQ1, AQ2, AQ3, IN1 and IN10 molecules.

Finally, to overcome the drawback related to the experimental uncertainties on absorptivity e to compare with computed and experimental spectra, all molar absorptivity were renormalized with

respect to their highest signal allowing to analyze the quality of our prediction in terms of both shape and band position. Gaussian and pseudo-Voigt broadening was performed using an in-house developed code. All characteristic parameters and the mixing Gaussian Lorentzian ratio used in this work are reported in Table 3.1.

Table 3.1. Parameters for Gaussian and pseudo-Voigt functions used to model the band shape for AQ and IN dyes: the dimensionless skewness (a), full width at half maximum in eV (σ) and the mixing parameter (M).

Model	Parameters	Molecules	
		Anthraquinone	Indigo
Pseudo-Voigt	a	-3000	-2000
	σ	0.6	0.2
	M	0.9	0.2
Gaussian	σ	0.4	0.2

A more rigorous way of predicting band shape consists in the calculation of the FC factors governing the vibronic broadening of the absorption band. Some assumptions are usually done, namely that i) the transition dipole moment does not changes with the ground and excited state geometries and that ii) the potential energy surfaces around the GS and the ES behave as quadratic (harmonic approximation). Among the different algorithms implemented to calculate FC factors, the one developed by Barone and co-workers and implemented in the Gaussian program has been used [47].

Here we want just to recall that this method requires ground and excited optimized structures and associated vibrational frequencies, as well as an accurate evaluation of the excitations energies in order to accurately produce absorption band profiles. Previous works [49] have proven that all these properties are not necessarily obtained when using the same DF approximation leading to the derivation of the so-called adiabatic shift model (here after ASM), developed and assessed for rigid and semi-rigid dyes [49]. In this approach the vibrational progression is computed at ωB97XD level and the resulting spectrum is shifted at the λ_{max} obtained at PBE0 [49]. The UV-vis spectrum is then reconstructed by convoluting the obtained transitions between GS and ES vibrational levels using Gaussian functions with FWHM of 0.04 eV to take into account the other physical effects affecting the band broadening (such as the Heisenberg uncertain principle).

Due to the relatively important effort that ES optimization and second derivatives calculations may imply depending on the level of theory used, simplified models allowing to extrapolate excited state minimum geometry and harmonic vibrational structures have been proposed in the litterature. Here we will make use of the approaches developed by Li and Truhlar [54] (LQ2 and

LQ3). Both these models are derived from the time-dependent formulation of the absorption cross sections under the Born Oppenheimer and Condon approximations and the overlap between nuclear wavefunctions is evaluated using the displaced harmonic oscillator assumption. In the case of the LQ2 and LQ3 models the exponential appearing in this overlap integral is simplified using a second and third order Taylor expansion (see Li and Truhlar paper for a detailed derivation [54]). Under all these assumptions in the LQ2 approach, the resulting band for the *n*-excitation can be expressed as:

$$s_n(E) \propto f_n exp\left(-\frac{(E-\Delta E_{0n}^v)^2}{\hbar^2 \sum_j \left[\Delta_j^n\right]^2 \omega_j^2}\right) \tag{3.10}$$

where f_n is the oscillator strength, ΔE_{0n}^v the vertical excitation energy, Δ_j^n represents the displacement of the excited-state geometry with respect to the ground state one, ω_j the ground state vibrational frequencies, the excited state vibrational frequencies being assumed to be the same. Of note the geometry displacement with respect to the ground state one is computed from the excited state energy gradients computed at the GS structure and assuming a harmonic behavior for the ES PES. The LQ3 model is simply an improvement of the LQ2 model implying a third order expansion of the nuclear overlap integrals [54].

In short, the only quantities required by LQ*n* methods are the vertical excitation energies, the ground-state second derivatives -frequencies and normal modes-, and the excited state energy gradients computed at the ground state geometry. All the LQ*n* calculations were carried out with the FCBand package [115].

3.2.3 Color space variables

Once the entire absorption spectrum is obtained, either from experiments or from computations for instance with the approaches illustrated in the previous paragraph, the associated color in the CIELAB space can be computed by numerical integration of the product of the inverse spectrum, a –tabulated- color matching function, here the CIE1964 with 10° field of view, describing the chromatic response of the observer and a –tabulated- function describing the illuminant source, here the D65 as analogously to the procedure illustrated by Beck [79] which is differing only for the assumption of a constant light source in UV-Vis range. This integration has been performed using in-house software.

In such a way, the color of each molecule is identified by its coordinated in the CIELAB color space, a system largely used in industry for color unique definition. In this systems the three color coordinates are: L* (the lightness) and a* and b* defining the mixing of four unique colors of human vision: red, green, blue, and yellow. In particular, a* modulates the change from green to magenta and b* axis links blue to yellow. It is also worth recalling that *L** measures the percent of lightness of the color ranging between zero (corresponding to black) and 100 (for white).
Using this system of color coordinates the Euclidian distance between the points representing the experimental and theoretical predicted colors can be used as simple descriptor to quantify the error associated to the computational protocol applied. This distance (d_{eucl}) is defined as:

$$d_{eucl} = \sqrt{\left(L^*_{theo} - L^*_{exp}\right)^2 + \left(a^*_{theo} - a^*_{exp}\right)^2 + \left(b^*_{theo} - b^*_{exp}\right)^2} \tag{3.11}$$

It should be noted however, that since CIELAB colour space is not perceptually uniform, the Euclidean distance in CIELAB space between two colours defined in eq. 3.11 does not always exactly correspond to the difference in perceived colour. Finally, two additional colour related parameters have been considered and discussed for their relevance for industrial applications, namely hue (or hue angle) and chroma.
The hue, or colour tone, is expressed [63]as:

$$h_{ab} = arctan(b^*/a^*) \tag{3.12}$$

while the chroma, measuring the chromatic intensity (i.e. a low chroma values correspond to pastel color, while high to bright ones) is defined as:

$$C^*_{ab} = \sqrt{a^{*2}+b^{*2}} \tag{3.13}$$

In other words, the hue and the chroma are the angular and radiant component of the color polar representation.

3.3 Results

3.3.1 Defining the reference models for shapes and colours

In order to evaluate to setup and the quality of the selected approaches for the modelling and simulation of band shape and associated colour, the 8 anthraquinones and the 10 indigos sketched in Figure 3.1 have been selected. Experimental spectra, in similar conditions, are either available in literature (all indigos and AQ6, AQ7, AQ8) or have been recorded. Beyond their industrial interest, these molecules span over a wide colour range (from blue to yellow), thanks to the different substitution on their parent skeletons. Since we are focusing on rigid and semi-rigid dyes, green anthraquinones are not present in our data set since they are characterized by the presence of (large) flexible groups, whose modelling requires techniques based Molecular Dynamics or Monte Carlo approaches.

A subset of 5 dyes (AQ1, AQ2, AQ3, IN1, IN10) has been selected for a first screen. Initially, the accuracy in the predicted λ_{max} considering vertical and vibrational resolved spectra using the two DFT functionals above mentioned namely PBE0 and ωB97XD.

Of note for all dyes the UV-Vis spectra is dominated by a single electronic transition which makes tempting to evaluate the accuracy of the computational protocol simply by comparison of the computed λ_{max}. However, AQs are often characterized by asymmetric bands, a clear sign of vibronic coupling, therefore the predicted color is significantly influenced not only by the position of the vertical excitation but also by the band shape associated [49]. A nice example of the impact of vibronic coupling is represented by solvent blue 35 dye [12], a molecule that is predicted yellow by vertical TD-DFT calculation with a simple Gaussian broadening and which band shape (and colour) is reproduced only with the ASM approach [49].

The λ_{max} obtained for our small subset are reported in Table 3.2. In agreement with previous studies on indigos and anthraquinones, PBE0 provides quite accurate results comparing vertical TD-DFT energies with experimental λ_{max} (MAE 0.08 eV) while larger errors are expected and found for the ωB97XD functional (0.36 eV) (see Table 3.2).

Table 3.2 Experimental (λ_{exp}, eV), vertical (λ_{vert}, eV) and vibronic (λ_{vibr}, eV) absorption maxima computed at the PBE0 and ωB97XD level for selected anthraquinones and indigos. In the table are also reported the adiabatic energy difference between the S_0 and S_1 states ($\Delta E_{0\text{-}0}$, eV), and the corresponding ZPE differences (ΔZPE, eV).

		PBE0				ωB97XD			
	λ_{exp}	λ_{vert}	λ_{vibr}	$\Delta E_{0\text{-}0}$	ΔZPE	λ_{vert}	λ_{vibr}	$\Delta E_{0\text{-}0}$	ΔZPE
AQ1	2.76	2.71	2.54	2.40	-0.06	3.21	2.92	2.81	-0.08
AQ2	2.50	2.62	2.26	2.33	-0.07	2.99	2.75	2.63	-0.06
AQ3	2.42	2.36	1.97	2.07	-0.10	2.71	2.32/2.45	2.32	-0.09
IN1	2.05	2.16	1.94	2.00	-0.06	2.40	2.17	2.20	-0.03
IN10	2.25	2.29	2.07	2.12	-0.06	2.49	2.39	2.42	-0.04
MAE		0.076	0.272			0.364	0.098		

Nonetheless, when vibronic calculations are carried out, the ωB97XD functional provides an error comparable to that obtained at PBE0 vertical excitations level (of 0.10 eV with respect to 0.08 eV, see Table 1) and way better than a full vibronic treatment at PBE0 level (0.27 eV). These results, even if limited for a deep statistical analysis, suggest that even the simple description of the λ_{max} associated to these molecules deserve careful analysis.

Keeping these considerations in mind, we move to the analysis of the full UV-vis spectra and the associated color. In Figure 3.2 and 3.3 are reported the simulated spectra for AQ3 and IN10, two molecules taken as representative examples, while the spectra of all the remaining molecules are collected in Appendix A (Figure SI3.1-SI3.14).

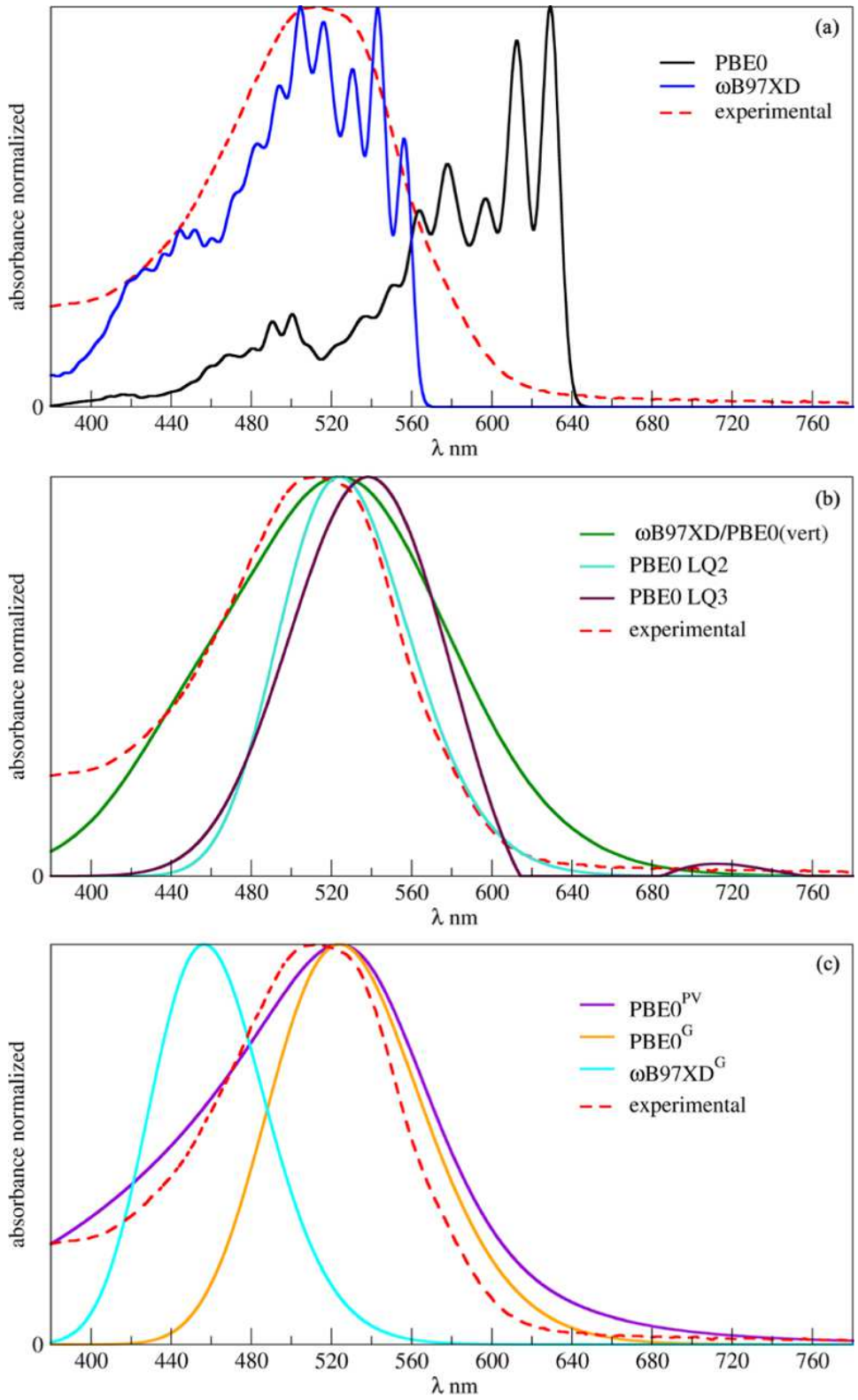


Figure 3.2 Absorption spectra for AQ3, obtained with the different modelling approaches: (a) PBE0^{vib}, ωB97XD^{vib}; (b) ωB97XD/PBE0^{v} LQ2 and LQ2; (c) PBE0^{PV}, PBE0^{G} and ωB97XD^{G}.

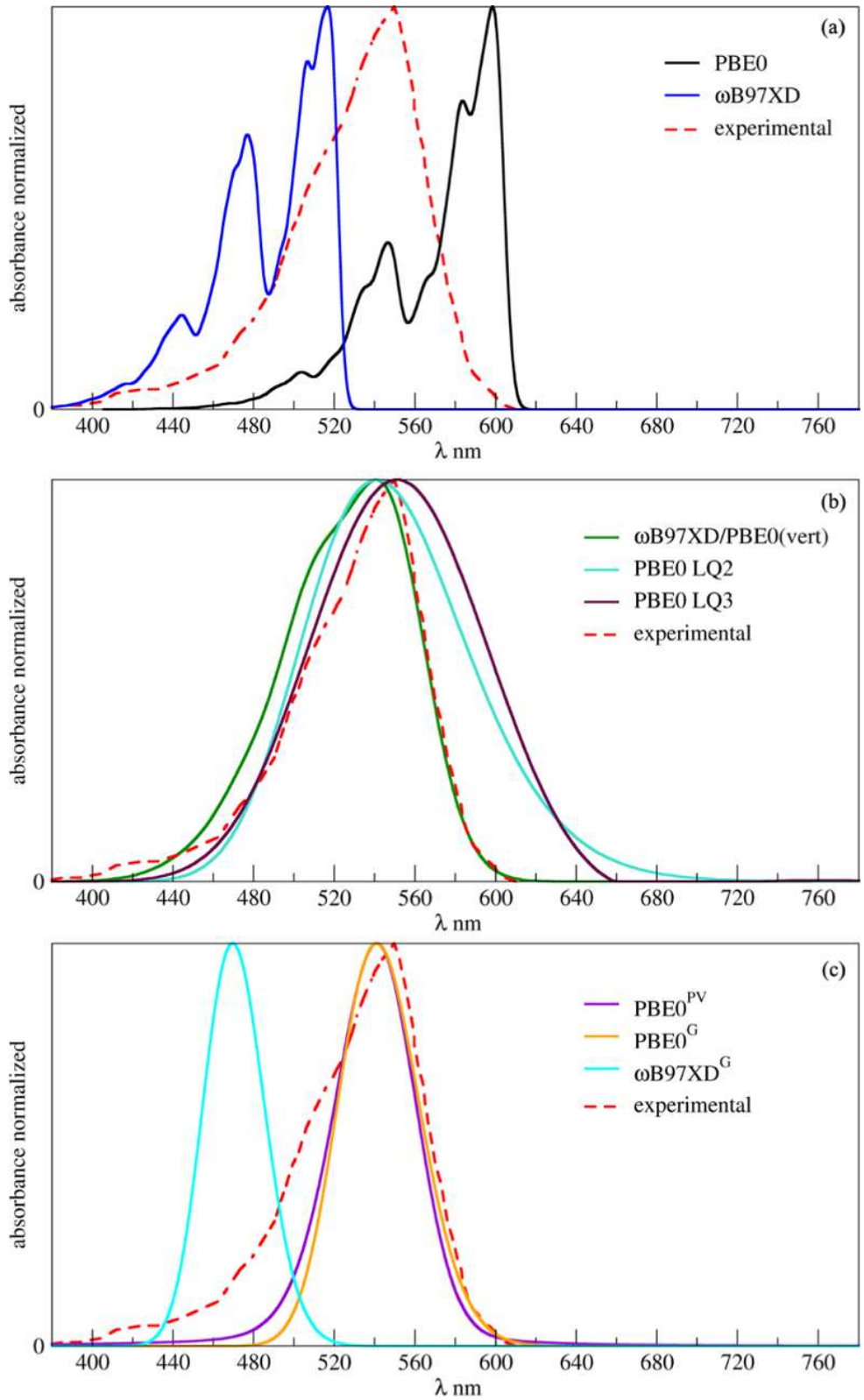


Figure 3.3 Convoluted absorption spectra for IN10, obtained with the different modelling approaches: (a) $PBE0^{vib}$, $\omega B97XD^{vib}$; (b) $\omega B97XD/PBE0^{v}$ LQ2 and LQ2; (c) $PBE0^{PV}$, $PBE0^{G}$ and $\omega B97XD^{G}$.

Full vibronic coupling model at PBE0 and ωB97XD level (hereafter PBE0^{vib} and ωB97XD^{vib}) correctly reproduce the experimental band shape, even if the latter gives a better description of the shoulder at shorter wavelengths (high energy), indicating a more accurate reproduction of the vibrational progression. Nevertheless, if the ωB97XD^{vib} band of AQ3 has a very good overlap with the experimental one, this is not the case for IN1, as well as for others anthraquinones and indigos reported in the Appendix A. Please notice that a these vibronic spectra have been reproduced using a very small FWMH parameter (0.04 eV), thus to evidence the underlying vibrational structure of the UV-vis band.

Following the ASM strategy, we coupled the shape of the ωB97XD ^{vib} with the low deviation of PBE0 for λ_{vert}, we called this approach vertical shift (hereafter ωB97XD/PBE0(vert)), provides an excellent reproduction of the band shapes for all the five molecules (see Figure 3.2, 3.3 and Appendix A Figure SI3.3, SI3.8), thus further supporting its choice as reference method. Approximate methods like LQ2 and LQ3 applied using PBE0 (indicated as PBE0 LQ2 and PBE0 LQ3) are also remarkably accurate, giving a correct simulation of the band asymmetry at higher energy, even if less evident than for the full vibrational approach. Since vertical excitation energies predicted at ωB97X-D level are blue-shifted with respect to the experimental λ_{max}, we did not perform ωB97XD LQ_n calculations. It is also noteworthy that the band maximum is slightly red-shifted for LQ2 with respect to LQ3, so that the band tail at low energy (high wavelengths) is better reproduced (see Figure 3.2 and 3.3). Simple convolution with Gaussian functions (PBE0^{G} and ωB97XD^{G}) leads to an asymmetric band towards higher wavelength as clearly showed in the case of AQ3 (see Figure 3.2).

Since experimental INs bands are narrow than the AQs ones, we decided to use different FWHM values. The FMWHs used in the present case have been set equals to 0.4 eV for anthraquinones and 0.2 eV for indigos, lead to large bands that are well centred on the experiments for PBE0 (hereafter PBE0^{G}) and significantly shifted at lower wavelengths for ωB97XD (hereafter ωB97XD^{G}) as expected from the previously discussed λ_{max}.

The same FMWHs were used for the convolution with the pseudo-Voigt functions, leading to the PBE0^{PV} model. The pseudo-Voigt broadening provides spectra with a better overlap with the experimental one, also in the tail regions (see Figure 3.2). This result is due both to the use of two different functions, that well reproduce the high-energy asymmetry, but also to presence of the a parameter allowing a better small displacement of the band maximum. It is worth to underline that for the PV functions the same set of 3 parameters (M, a and σ) has been considered for each molecular family, that 1 set for anthraquinones and 1 set for indigos. These parameters, reported

in Table 3.1, have been adjusted to decrease the errors on the Euclidean distances between the experimental and computed colours.

In order to obtain a more quantitative analysis of the error associated to the computed band shapes instead of computing a point-by-point distance between the experimental and the simulated spectra, we calculated the predicted and experimental colour, reported in Table 3.3.

Table 3.3 Colors from the experimental and theoretical spectra, evaluated using different computational protocol: PBE0^{vib}, ωB97XD^{vib} and ωB97XD/PBE0(vert): fully vibronic coupling; LQ2 and LQ2: approximate vibronic models; PBE0^{G} and wB97XD^{G}: Gaussian convolution; PBE0^{PV}: pseudo-Voigt convolution.

	exp	PBE0^{vib}	ωB97XD^{vib}	ωB97XD/PBE0 (vert)	LQ2	LQ3	PBE0^{G}	ωB97XD^{G}	PBE0^{PV}
anthraquinones									
AQ1									
AQ2									
AQ3									
indigos									
IN1									
IN10									

In this Table are reported the experimentally perceived color for the 5 considered molecules as well as those obtained by the computational approaches, computed using the above illustrated method. A quick inspection to this table, clearly show that the colors computed the full vibronic approaches PBE0^{vib} and ωB97XD^{vib} predict blue anthraquinones in clear contrast with the yellow-orange experimental data. This result is even more surprising in view of the good reproduction of the band for PBE0^{vib} and it can be ascribed to the sharp decay of the vibronic spectra at higher wavelengths.

The other remarkable difference concerns the two indigos IN1 and IN10, whose colours are badly reproduced by the ωB97XD^{G} model, a results not so surprising in view of the significant shift of the corresponding bands. The relatively good results of this model for the anthraquinones are related to the tail of the Gaussian functions that clearly penetrate in the high-wavelength region thus affecting the computed color (see Figure 3.2).

All the other models give a, at least, qualitative reproduction of the color predicting yellow-orange anthraquinones and blueish indigos (see Table 3.2), in good agreement with the experimental data. In order to discriminate between them a more quantitative analysis has been performed using the CIELab space coordinates, chroma and hue, whose values are reported in Table 3.3.

Table 3.4 CIE colorimetric coordinates (L*,a*,b*), chromas (C) and metric hue angle (h) of selected AQs, derived from experimental and simulated spectra. PBE0^{vib}, ωB97XD^{vib} and ωB97XD/PBE0^{v} : fully vibronic coupling; LQ2 and LQ2: approximate vibronic models; PBE0^{G} and ωB97XD^{G}: Gaussian convolution; PBE0^{PV}: pseudo-Voigt convolution.

	param	exp	PBE0^{vib}	ωB97XD^{vib}	ωB97XD/PBE0 (vert)	LQ2	LQ3	PBE0^{G}	ωB97XD^{G}	PBE0^{PV}
	L^*	89.22	63.21	74.07	89.64	89.78	89.95	92.17	99.83	88.47
	a^*	-18.65	-35.12	87.28	6.59	8.85	11.74	3.23	-4.37	6.81
AQ1	b^*	76.33	-40.01	-174.97	76.54	70.31	69.44	69.40	9.30	74.74
	C	78.57	53.23	195.53	76.82	70.87	70.42	69.47	10.28	75.05
	h	103.73	228.72	296.51	85.08	82.82	80.40	87.34	115.14	84.79
	L^*	74.94	74.21	70.05	82.54	87.73	86.38	87.00	98.80	82.72
	a^*	44.11	-21.93	18.55	27.99	20.7	27.59	21.95	-14.26	21.09
AQ2	b^*	39.44	-58.48	-123.15	55.51	47.50	42.93	47.92	42.00	62.21
	C	59.17	62.45	124.53	62.16	51.82	51.03	52.71	44.36	65.69
	h	41.80	249.44	98.56	63.24	66.45	57.27	65.39	108.76	71.28
	L^*	66.64	76.46	61.65	55.63	69.03	65.43	64.59	92.15	58.60
	a^*	55.51	-12.56	-10.19	51.16	67.22	69.29	66.90	3.30	49.89
AQ3	b^*	6.94	-30.04	-30.32	-13.53	-40.58	-43.46	-42.44	69.34	-1.59
	C	55.94	32.55	31.99	52.92	78.52	81.79	79.23	69.42	49.92
	h	7.13	247.31	251.42	345.19	328.88	327.90	327.61	87.28	358.18

The a* and b* experimental values are generally well reproduced by the theoretical models, even if at a different degree. The ωB97XD/PBE0 (vert) model is particularly accurate for b* value of AQ1 and AQ2, but underestimates the corresponding a* values. The opposite is observed for AQ3, where a* is better reproduced than b*. A similar trend is also found for the PBE0-LQ2 and $PBE0^{PV}$ models, even if further deviations can be observed.

The $PBE0^{vib}$ and $\omega B97XD^{vib}$ predict very low values for b* of all three anthraquinones, corresponding to a clear shift of the color in the blue region, coupled with a further displacement along the a*, toward green for AQ2 and AQ3 and toward red for AQ1. The behavior of $\omega B97XD^{G}$ is instead more erratic, underestimating b* for AQ1 and overestimating it for the other two anthraquinones.

These trends are well summarized in the hue values: overestimated for $PBE0^{vib}$ and $\omega B97XD^{vib}$, all the other methods give lower deviations.

The lightness values are instead correctly reproduced by all the models, due to the normalization of the optical densities, even if some differences can be observed for $PBE0^{vib}$ and $\omega B97XD^{vib}$ (a darker AQ1) or $\omega B97XD^{G}$ (whiter AQ2 and AQ3).

Finally, the chromas are also predicted within a reasonable accuracy, the ωB97XD/PBE0 (vert) being the most accurate approach for all molecules. The two other vibronic models, $PBE0^{vib}$ and $\omega B97XD^{vib}$, significantly underestimate the chromatic intensity for AQ3, thus predicting more pastel colors. The $\omega B97XD^{vib}$ also suggest a very bright AQ1.

Similar trends can be also found for the two indigos, IN1 and IN10, whose color parameters are reported in Table 3.4. Some notable differences can be, however, noticed. All the methods, except $PBE0^{vib}$, and $\omega B97XD^{vib}$, predicts IN1 more red than in experiment. The effect on the perceived color is, however, modest, since the a* values are relative small, the only exception being $\omega B97XD^{G}$ (see Tables 3.3 and 3.5).

Table 3.5. CIE colorimetric coordinates (L*,a*,b*), chromas (C) and metric hue angle (h) of selected indigos, derived from experimental and simulated spectra. PBE0^{vib}, ωB97XD^{vib} and ωB97XD/PBE0(vert) : fully vibronic coupling; LQ2 and LQ2: approximate vibronic models; PBE0^{G} and ωB97XD^{G}: Gaussian convolution; PBE0^{PV}: pseudo-Voigt convolution.

	param	exp	PBE0^{vib}	ωB97XD^{vib}	ωB97XD/PBE0 (vert)	PBE0^{PV}	LQ2	LQ3	PBE0^{G}	ωB97XD^{G}
IN1	L^*	71.10	89.50	7.51	63.72	76.56	61.49	61.54	78.58	84.84
	a^*	-16.69	-23.15	-16.1	38.82	5.64	13.87	8.48	-4.61	49.16
	b^*	-36.21	-11.72	-50.67	-54.58	-38.68	-63.28	-62.00	-36.78	-20.94
	C	39.87	25.95	53.16	66.98	39.09	64.78	62.57	37.07	53.44
	h	245.25	206.85	252.38	305.42	278.29	282.36	277.79	262.86	336.93
IN10	L^*	69.51	80.57	71.74	69.97	70.96	60.51	59.09	78.28	96.10
	a^*	39.14	-0.02	-13.57	62.91	50.31	58.31	54.18	46.71	1.95
	b^*	-41.70	-32.74	-81.42	-33.05	-34.15	-58.21	-58.62	-36.28	33.69
	C	57.20	32.74	82.54	71.06	60.81	82.39	79.82	59.14	33.74
	h	313.18	269.97	260.54	332.29	325.83	315.05	312.74	322.17	86.69

Taken together, all these results indicate that from one side hue, chroma and lightness are less sensible to the choice of the computational protocol, providing at least a qualitative agreement with the values obtained from the experimental data. CIE coordinates are, instead more sensible to the computational level, showing a significant variation in going from a model to another one. More in details, the PBE0^{vib}, and ωB97XD^{vib} and ωB97XD^{G} models are those providing the largest deviations with respect to the experimental values, whereas the ωB97XD/PBE0 (vert) protocol is the most accurate among those considered here. All these results allow also a better analysis, beyond a simple visual discrimination of the computed perceived colors, as those reported in Table 3.3.

3.3.2 Validation on the full set of dyes

The analysis has been then extended to all the other anthraquinones and indigos reported in Figures 3.1. All the simulated colours and CIELab coordinates for the complementary indigos and anthraquinones are provided in the Appendix B and Appendix C (Tables SI3.1-SI3.2 and Tables SI3.14-SI3.15), while their simulated spectra are reported in Appendix A (Figures SI3.1-SI3.13). Generally speaking, these data follow the trends already discussed in the previous paragraph for the 5 protype molecules, therefore they will not discussed in details. However, a point must be underlined. Unfortunately, it was not possible to optimize three indigos at the excited electronic state. Therefore, statistics obtained for the three approaches using a full vibronic treatment, that are ωB97XD/PBE0 (vert), PBE0^{vib} and ωB97XD^{vib}, are limited. If this problem further reduces the molecular space, it also well underlines the limit of this approach. Indeed, a single floppy motion, activated by electronic excitation, could make problematic or, as in the present case, impossible to correctly characterized the bright excited state and will impede the full vibronic treatment. In case like this, as already mentioned, Molecular Dynamics simulations can be considered, even if quantum effects such as Zero Point Energy are difficult to reproduce. In Figures 3.4 to 3.6 are reported the deviations with respect to the experimental values for all the considered anthraquinones and indigos, including the previous 5 molecules already discussed in details. While for Chromas and hue angles are reported the mean of the absolute deviations (MAD), for the three CIELab coordinates, the error, always expressed as MAD, is computed as Euclidian distance, as mentioned in the computational details.

Let us start from this latter indicator. The PBE0^{PV} model is the better performing, a with deviation of about 20 units in the CIELab space coordinates. This result confirms the ability of this model in the band shape, including vibronic asymmetry, and colorimetric indexes at a negligible

computational cost. The number of parameters (4) could justify the good accuracy obtained, since the band shape and the peak position can be tuned in order to improve the agreement with the experimental values. Nevertheless, the considered parameter set is unique for a given family of molecules, anthraquinones or indigos, so that their optimization is not molecule depending, thus enlarging its applicability domain at least to families of molecules.

The PBE0-LQ2 and PBE0-LQ3 models represent then a good compromise between the empirical PV convolution and the full vibronic treatment of the ωB97XD/PBE0 (vert) approach. Indeed, they are only slightly less accurate than the $PBE0^{PV}$ model, without the need of adjusted parameters, and the errors are comparable for the two classes of molecules. The lower error with to respect ωB97XD/PBE0 (vert) is however misleading since in this last case only 2 indigos are considered in the statistics, which is not the case for all the other models. Indeed, ωB97XD/PBE0 (vert) is (slightly) more accurate than LQ2 and LQ3 on the anthraquinones where more systems are considered. Nevertheless, this result points out that the LQ2 and LQ3 allow to circumvent the mentioned numerical problems associate with floppy motions, since they do not require a full characterization of the excited state energy minima. The good performance obtained thus confirm the validity for (semi)rigid molecules and the robustness of these two protocols for simulating band shape in a nonempirical way.

Finally, the $PBE0^{v}$ model also gives reasonable results, but is more accurate on the indigos than on the anthraquinones, while $\omega B97XD^{vib}$ shows very large deviations for both families of molecules.

The trends on the CIELab coordinates are reflected by the errors for the hue angles, even if the $PBE0^{PV}$ gives a deviation on the AQs larger than that provided by the LQ2 and LQ3 models. Concerning the AQs, it worth to stress that errors obtained at $\omega B97XD^{G}$ are lower than those produced by all other methods tested in this work. A careful analysis on the benchmarked AQs (see Appendix B) reveals that the $\omega B97XD^{G}$ approach predicts colorimetric coordinates with proper sign for 7 AQs, thus the hue metric angles are computed on the same experimental circumference arc.

The vibronic approaches then follow in the mentioned order, ωB97XD/PBE0 (vert) being better than $PBE0^{vib}$ and $\omega B97XD^{vib}$, at least for anthraquinones.

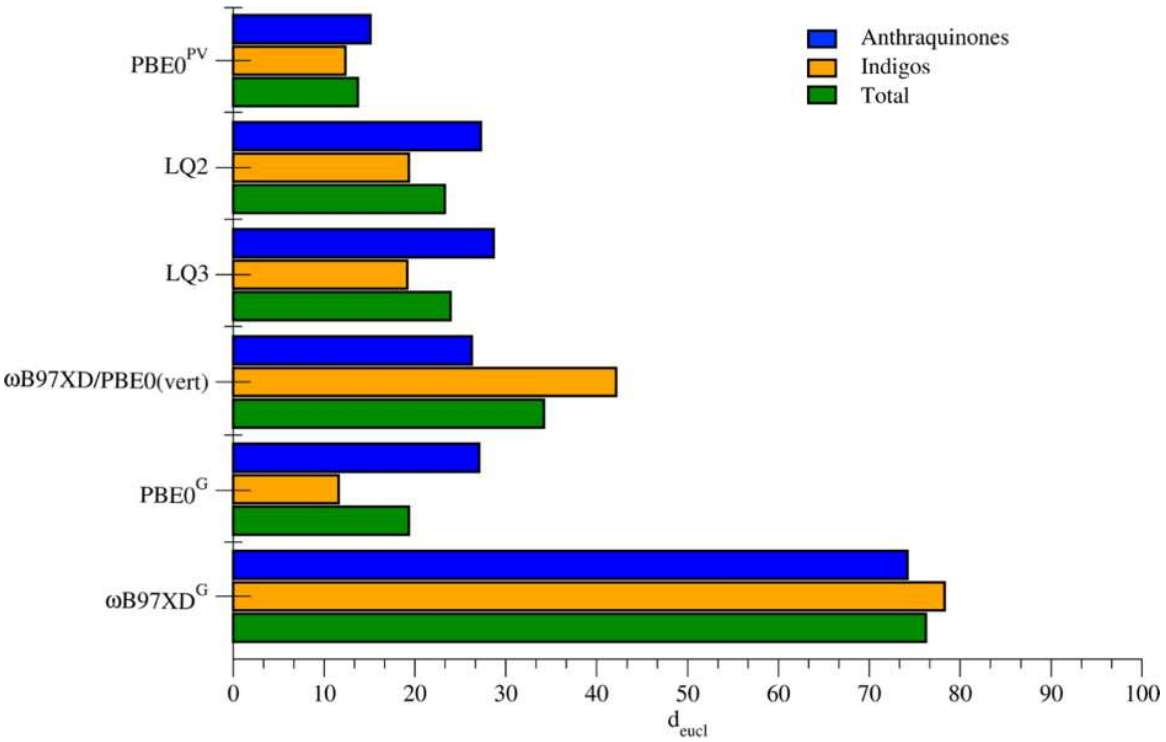


Figure 3.4 Mean Absolute Error on the Euclidian distances for the CIE colorimetric coordinates derived from experimental and theoretical spectra.

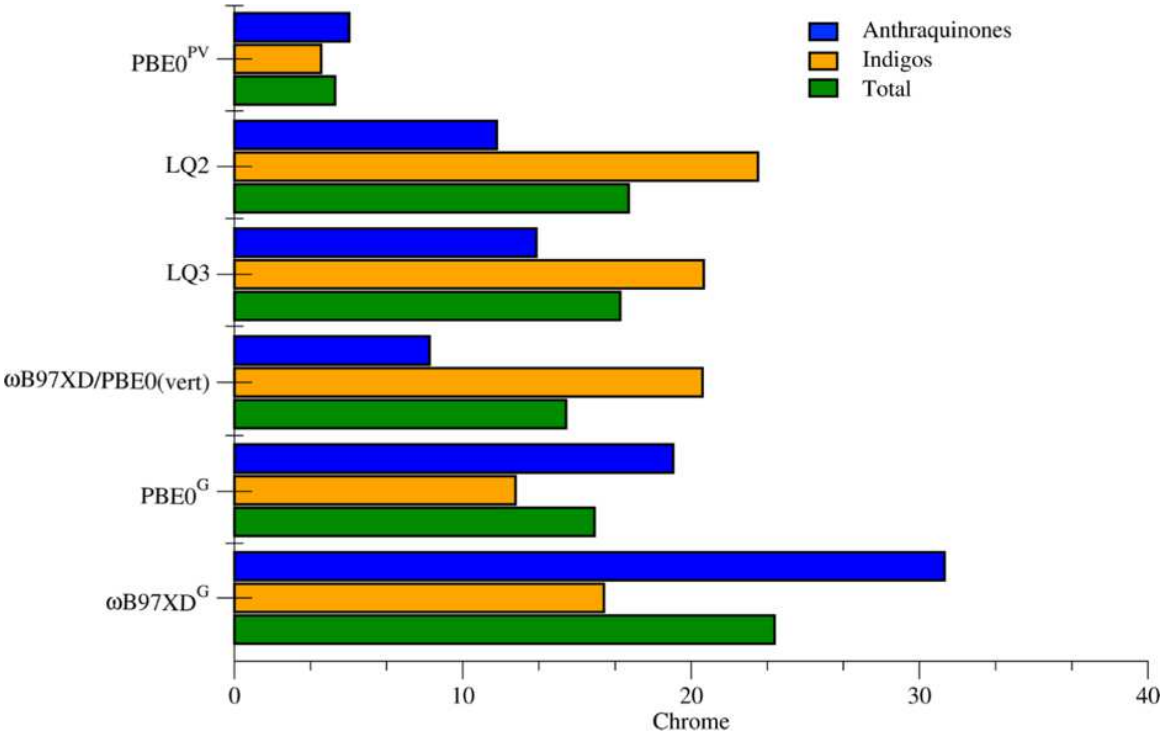


Figure 3.5 Mean Absolute Error on the chromas derived from experimental and theoretical spectra.

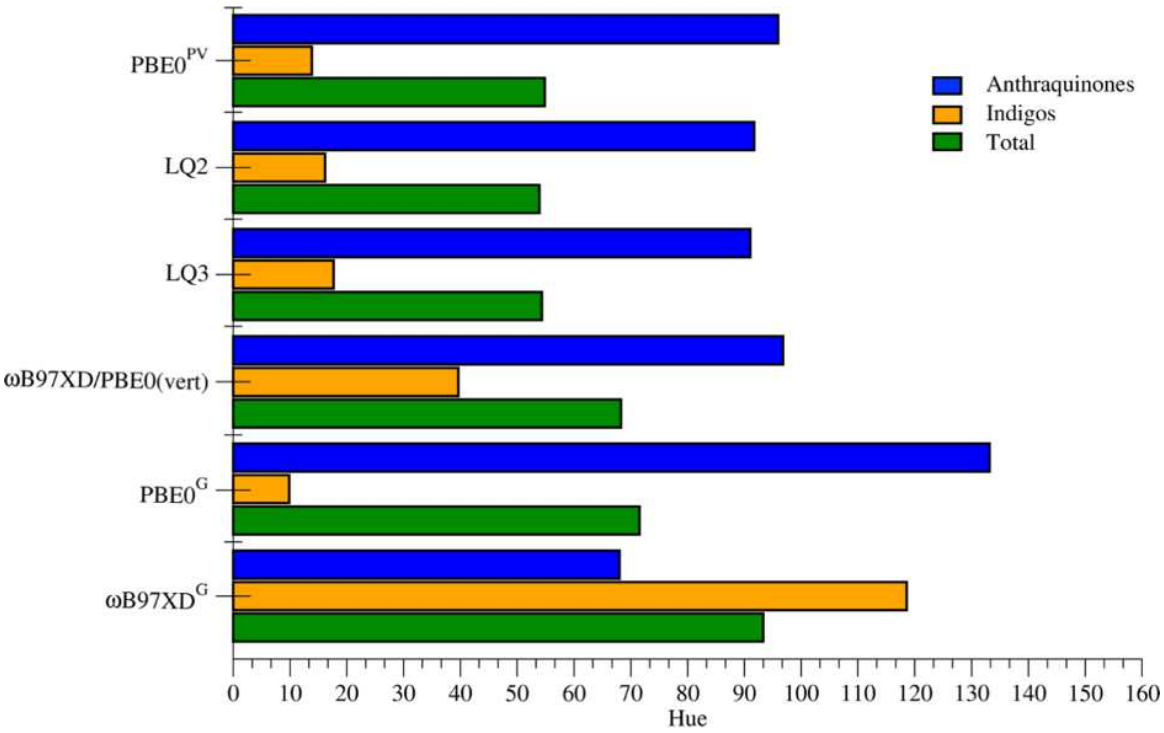


Figure 3.6 Mean Absolute Error on the metric hue angle derived from experimental and theoretical spectra.

3.4 Conclusion

In this work, we have tested different computational methods for predicting the spectra of functionalised AQs and INs. Although the AH approach is considered a reference for predicting the spectra of rigid compounds, it is not reliable for describing floppy molecules. Indeed, our results show that for floppy INs dyes the FC spectrum is not computable. Moreover, the characterisation of the FC minimum leads to large errors for the tested dyes.

Based on our results, we have selected two cost-effective protocols. One considers the LQ*n* method, while in the second the absorption band was modelled with a PV function. Since both the λ_{max} and the band shape play a crucial role in determining the colour of the tested compounds, we performed a quantitative colour analysis in CIELab colour space. Here, both PBE0^{PV} and the LQ*n*-based protocols are able to predict both a* and b* coordinates for the tested dyes with sufficient accuracy. Indeed, Euclidean distances, chromas and hue angles computed at PBE0^{PV} level are similar than those predicted with the LQ*n* protocols. Comparing the results deriving from these two protocols, the chroma shows large differences. Smaller deviations are obtained with the PV model, suggesting that the correct simulation of the electronic spectrum affects the prediction

of the a* and b* colorimetric coordinates. It should be noted, however, that the PV method is characterised by ad-hoc parameters. On the other hand, LQ*n* methods do not require external parameters and are able to reproduce the colorimetric coordinates with similar accuracy for the dyes benchmarked in this work.

Chapter 4

4. Modelling UV-vis spectra using a multiscale approach

In this chapter we report our study on a class of flexible dyes. Specifically, we show, choosing merocynine molecules as an example, that the combination of a MC sampling of the configurational space with TD-DFT is workable compromise to solve the combined challenge of accuracy and computational cost to estimate vertical excitations energies and simulate absorption spectra of molecules and aggregates.

4.1 Introduction

Merocyanines are a class of polar dyes showing a variety of colors, and unique self-assembly behavior arising from their large dipolar character [116]. This class of molecules is characterized by the presence of a donor and an acceptor group connected by an ethylene or polyethylene bridge, Figure 4.1.

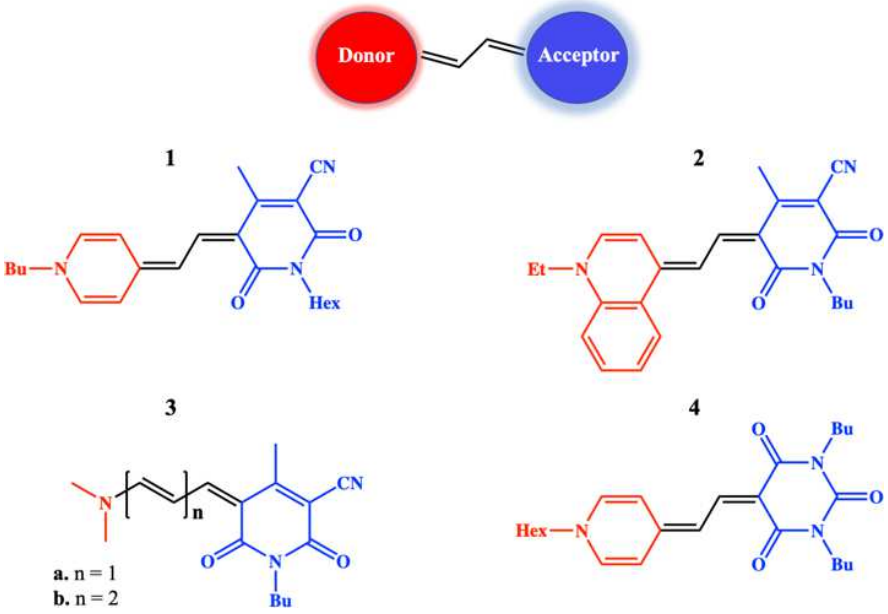


Figure 4.1 Schematic merocyanine representation investigated in this PhD thesis. But, butyl; Hex, hexil; Et, etil.

Thus, the so-called bond-length alternation (BLA) represent a relevant structural parameter of merocyanines because accounts the difference between single and double carbon–carbon bonds of the polyethylene bridge. Large (small) BLAs are an indication of a reduced (increased) delocalization of the π electrons and of a worsen (improved) coupling between the donor-acceptor moieties, thus leading to a blue (red) shift of the spectra. This can be translated in terms of electronic structure, large BLA increases the weight of the zwitterionic (charge-separation) form, while small BLA indicates formally neutral structure [117]. Nonetheless, BLA is by itself a challenging property for computational approaches such as DFT, because its correct reproduction requires the reduction (or even cancellation) of the Self-Interaction Error, which, inducing a large electronic delocalization, leads shorter BLAs [118, 119]. Briefly, over the years, efforts have been made to the theoretical description of the merocyanine dyes using density-based method (DFT and TD-DFT), that allows a reliable prediction of electronic absorption spectra in terms of λ_{MAX} and band shape in different environment [102,120,121]. Going into the details, several computational studies focus on the most suitable exchange-correlational functional to model the absorption spectra of the merocyanine dyes. Because the most intense transition in merocyanine shows a charge-transfer character, standard DFT approach are expected to fail [122]. However, also range separated hybrid functional do not significantly improve the results for this class of chromophore [102, 123, 124]. Actually, it has even been suggested that the insensibility of the computed transition to the choice of the DFT approach is a fingerprint of the merocyanine [125]. In addition, their vibrational fine-structure is a matter of debate. From one side, theoretical and experimental studies suggest that the broadening of the first absorption band is related to the C=C vibration of the polymethine bridge [125-126], on the other side, the complete modelling of their absorption spectra obtained by using the Franck-Condon approach indicate that the origin of this fine structure can be ascribed to a number of different vibrational modes [124, 127-130]. For instance, vibrational levels are considered as being equal for both the ground and excited states (adiabatic approximation) or only one degree of freedom is explicitly considered [129,130]. As a rule of thumb, the vibronic coupling is expected to shift the λ_{max} of about 0.3 eV [129].

Furthermore, simulating the merocyanine behavior in complex environment is complicated, because: these chromophores undergo to strong solvatochromism [131] and the solvent effect cannot be in turn avoided in the modelling [130, 132, 133]. Furthermore, in solution merocyanine form H-type dimer aggregates (anti- parallel or parallel, see figure 4.2) [134, 135] All these issues are strongly entwined and, often, they have been investigated separately.

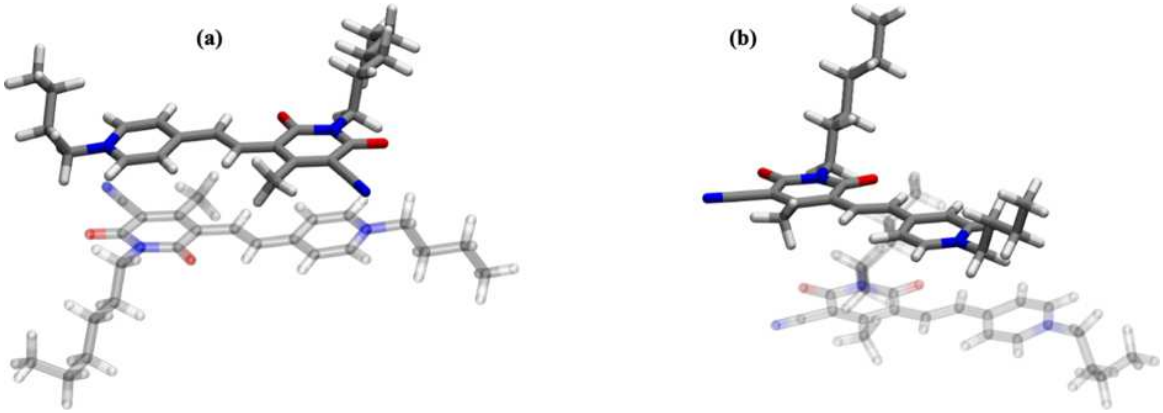


Figure 4.2. representation of the H-type aggregates: (a) antiparallel co-facial π-staking, (b) parallel co-facial π-staking.

Here, focusing on both the monomeric and dimeric forms of the merocyanines **1**-**4** reported in Figure 1, we tried to analyzed the relationships between the structural and electronic features of merocyanines exploring their complex phase space with a stochastic approach relying on Monte Carlo (MC) molecular simulations. The flexible structure of merocyanines adds another level of complexity in the modeling. Indeed, standard methods to obtain the vibrationally resolved optical spectra [49,53,136-137] cannot be applied in this situation due to the presence of several energetically accessible minima existing on the ground and/or excited state potential energy surfaces (PESs) [138-140]. In such cases a single structure can hardly be representative of the real phase space and ensemble averages become mandatory. For this reason, we simulated the absorption spectrum as an average of those vertically computed for a statistically representative average over an ensemble of molecular structures, extracted from the ground state PES. In this respect, the widely applied strategy to build the statistical ensemble is based on Molecular Dynamics (MD) simulations, where the time evolution of the system is achieved by integrating the classical Newton's equations of motion, the corresponding potential energy being computed either at classical or quantum chemistry level of theory [138-142]. Although MD simulations are widely and successfully applied, some concerns still arise. Indeed, when large energy barriers (often of several k_BT) to conformational changes (e.g. bond rotations in flexible molecules, geometrical rearrangements in aggregates) exist, it is easy to be trapped in local -low energy- conformations, leading to a poor quality statistical sampling. Increasing the temperature of the simulation, in conjunction or not with simulated annealing approaches or replicate trajectories starting from different initial conditions are common procedures to improve the space-phase sampling [143]. Driven dynamics (metadynamics, umbrella sampling) are also viable alternatives, requiring the identification of the relevant reaction coordinate(s) [144, 145]. Nevertheless, spectra simulated at high temperature are clearly broadened with respect to the experimental ones, and more replicas

of the trajectory or driven dynamics can be computationally difficult to achieve especially in the case of ab initio molecular dynamics. For these reasons, we relied on an alternative philosophy to escape from local minima on the PES based on random sampling with the Monte Carlo method. In this approach, configurations in the phase space are generated by randomly changing the coordinates of a randomly chosen atom of the system, and accepting the new configuration according to the Metropolis sampling algorithm [146]. The random moves and the possibility to accept higher energy configurations allow MC simulations to avoid being trapped on local minima reaching energetically distinct regions of the PES. Although this approach is very powerful, few examples are reported in literature coupling MC sampling with quantum chemistry calculation for the simulation of UV-Vis absorption spectra of molecules, mainly focused on the organization of the solvent molecules around the chromophore [147-149]. The main aim of the present work is thus to show the potentiality of a combined Monte Carlo/Quantum Chemistry, here Time Dependent Density Functional Theory (TD-DFT), approach to simulate the electronic absorption spectra of the merocyanines family of flexible molecules and their aggregates (here dimers).

4.2 Computational details

Two computational protocols were considered for the calculation of UV-Vis spectra. The first one, referred to as *static*, consists of a full optimization of the molecular structures using the PBE0 functional [31] and the 6-31+G(d,p) basis set [150]. In the case of dimers, dispersive interactions were accounted through the Grimme empirical correction (D3) [151]. Vertical absorptions were then computed at the TD-DFT level, using global (PBE0) and range separated (ωB97XD) [36] hybrid functionals in conjunction with the 6-31+G(d,p) basis set. The first functional, PBE0, is widely used for the computation of UV-Vis spectra involving valence transitions [40], while the second, containing a larger quantity of HF exchange, is more suitable for Charge Transfer (CT) excitations and it was already applied to study model merocyanines [102].

Solvent effects were taken into account using the IEF-PCM model, considering dioxane as solvent as in the experimental protocol [153]. Due to the aprotic nature of this solvent, no specific solute-solvent interactions are expected, so that the IEF-PCM provides accurate results for the electronic transitions [152].

The second approach, hereafter referred to as *MC*, is based on the generation of ensemble configurations for both monomer and dimeric forms of merocyanines **1-4** (Figure 4.1) using MC Metropolis simulations. Our method is schematized in Figure 4.3 and works as it follows:

1. The statistical ensemble is constructed through the Monte Carlo sampling of the ground state PES.
2. Vertical excitation on randomly selected configuration are carried out at TD-DFT level.
3. The electronic absorption spectrum is obtained by a Gaussian broadening the so computes vertical excitations.

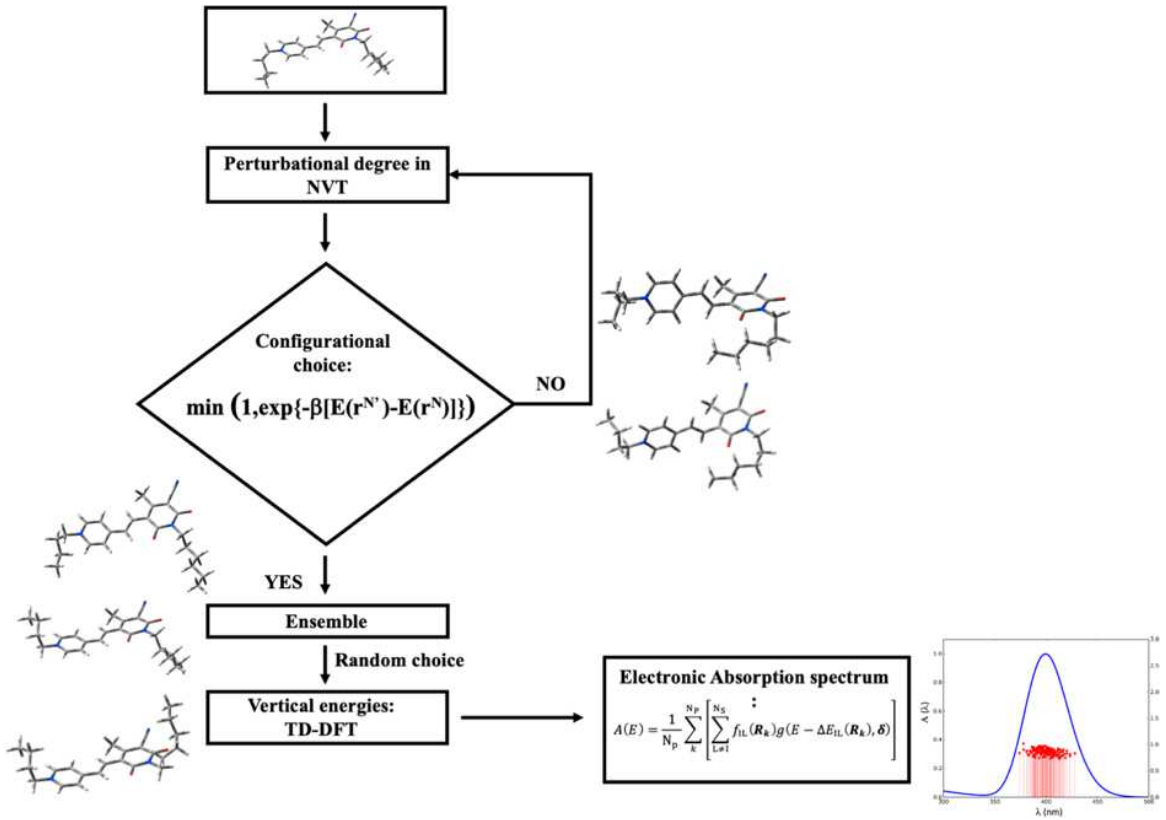


Figure 4.3: schematic representation in flowchart style of the steps involved in the Monte Carlo/DFT protocol.

Going into the details, the energy of the Monte Carlo configurations was computed at molecular mechanics level using standard Generalized Amber Force Field (GAFF) [154]. Production MC simulations were carried out for $1*10^7$ steps at a 300K temperature. Hence, 500 structures were randomly extracted to simulate the UV-vis spectrum and their vertical electronic transitions were computed using the same TD-DFT computational level as for the *static* approach, including solvent effects. Finally, the absorption spectrum ($A(E)$) was simulated as:

$$A(E) = \frac{1}{N_p}\sum_{k}^{N_p}\left[\sum_{I\neq L}^{N_s} f_{IL}(R_k)\, g(E - \Delta E_{IL}(R_k), \delta)\right] \quad (3.14)$$

where vertical excitation energies ($\Delta E_{IL}(R_k)$) between the ground electronic state(I) and excited state (L) and corresponding oscillator strengths ($f_{IL}(R_k)$) were computed for each sampled configuration (R_k), N_s is the number of excited states considered, N_p is the number of sampled points, and g represents a normalized Gaussian function catered on ΔE_{IL} of full width at half maximum (δ) of 0.3 eV. Convergence of the sampling was tested for the monomer **1** molecule by using a sampling of 600 snapshots providing the same computed spectral properties (Figure 4.4). All first-principles calculations were carried out using the Gaussian suite [103], while MC computations were performed using the LAMMPS code [155].

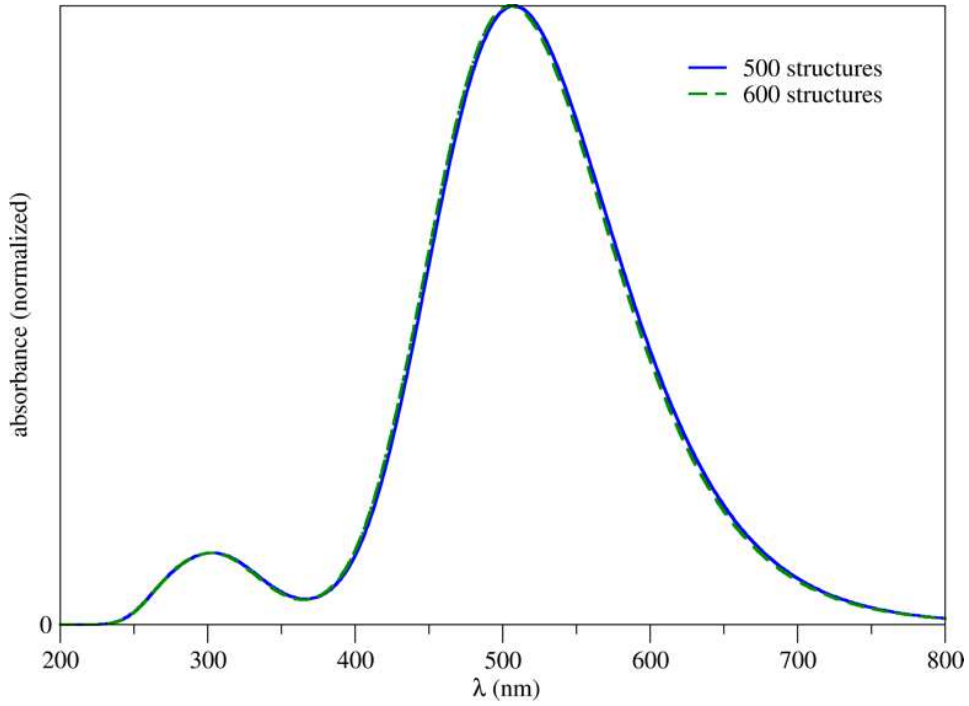


Figure 4.4. Convergence of the sampling was tested for the monomer 1 molecule by using a sampling of 600 providing the same value of the computed spectral properties.

4.3 UV-visible spectra: static versus MC approach

In Table 4.1 are collected thee computed absorption energies maxima (λ_{MAX}) for all monomers and dimers of the merocyanine sketched in Figure 4.1, obtained with the two approaches described above. From literature data it's known that merocyanine dyes undergo vibronic coupling significantly affecting the position of the λ_{MAX} computed by a static approach . The red-shift of the main peak, with respect to the vertical approximation, was estimated to be about 0.3 eV [15]. The comparison of data coming from *static* TD-DFT calculations with the experimental data [153], included in Table 4.1, not surprisingly reveals that the vertical transitions energies computed at the PBE0 level for all the monomers are significantly overestimated (between 0.4 and 0.5 eV). The

computed Mean Absolute Deviations (MADs) are quite large, but essentially in agreement with the errors reported in the literature for similar systems [5, 9, 10]. Furthermore, if the estimated 0-0 correction is subtracted (0.3 eV), the deviations are within the expected range of accuracy for TD-DFT methods (0.1-0.2 eV), always taking into account the above reported caveats. These discrepancies slightly increase if the range-separated ωB97XD functional is considered (+0.5 eV for **1**), in agreement with data in literature [5]. For these reasons, ωB97XD will be not further considered for the other molecules and only PBE0 results will be discussed. Despite these significant deviations between theory and experiments, the experimental relative trends are correctly reproduced by the PBE0 results. Indeed, the outcomes obtained (Table 4.1) show significant blue-shifts going from molecule **1** to **4** (between 0.09 and 0.56 eV), in the same range of the experimental observation obtained by Wortmann and co-workers (between 0.01 and 0.56 eV) [153], with the exception of molecule **2**, for which a red-shift of 0.22 eV is found, in agreement with the experimental trend.

Table 4.1 Computed λ_{MAX} (in eV) using a static and a MC approach. All the electronic transitions were computed at the PBE0/IEF-PCM level of theory, except for the values in parenthesis, which were evaluated using the ωB97XD/IEF-PCM model. Mean Absolute Deviations (MADs in eV) with respect to the experimental values are also reported.

	Monomer					**Dimer**		
	static	MC	static/ dimer[a]	MC /dimer[b]	exp[c]	Static	MC	exp[c]
1	2.60 (2.65)	2.46 (2.49)	2.55	2.45	2.18	2.84 (2.96)	2.66 (2.72)	2.52
2	2.38	2.19	2.34	2.18	1.96	2.54	2.37	2.22
3a	3.16	2.89	3.12	2.89	2.68	3.30	3.04	2.89
3b	2.69	2.49	2.65	2.42	2.19	2.87	2.64	2.42
4	2.87	2.70	2.84	2.68	2.41	2.99	2.81	2.60
MAD[d]	0.46	0.26	0.42	0.24		0.38	0.17	

a) computed as the average value of the two monomers at their static geometry in the dimer;
b) computed as the average value of the two monomers at their MC geometry in the dimer;
c) from reference [153];
d) Mean Absolute Deviation.

Figure 4.5 depicts the topology of the Frontier Molecular Orbitals (FMOs) of the merocyanine monomers. According with the TD-DFT calculations, the most intense absorption corresponds to a $\pi \rightarrow \pi^*$ transition, from the highest occupied molecular orbital (HOMO) to the lowest unoccupied molecular orbital (LUMO). These FMOs are delocalized over the whole core of the

molecules. Hence, any structural deformation modifying the electronic conjugation of the backbone impacts the computed absorption λ_{MAX}. On the other hand, the CT character appears limited.

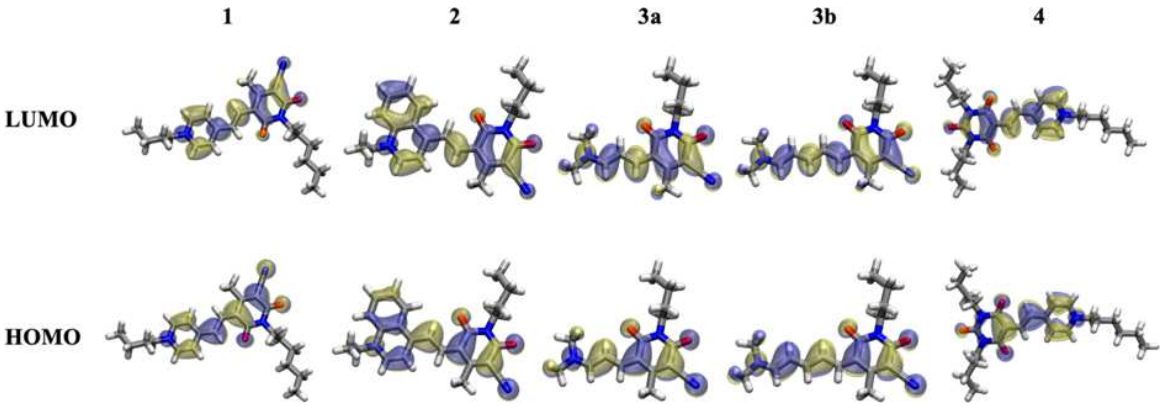


Figure 4.5 HOMO/LUMO illustration for merocyanine monomers investigated in this thesis. Blue region indicates negative isosurface, while yellow ones indicate positive isosurfaces.

The trend found for monomers are kept for the dimers, indeed, the blue-shifts going from merocyanine **1** to merocyanine **4**, in the same range of the experimental observation excepting for the chromophore number **2**. All the static predictions, Table 4.1, refer to co-facial antiparallel dimers (H-type aggregate). This structure represents the global minimum on the ground state PES. As shown for the monomers, a significant overestimation of the λ_{MAX} is found at the PBE0 level. However, these deviations falling within 0.32 eV (dimer **1** and dimer **2**) and 0.45 eV (dimer **3b**), are smaller than the corresponding monomers. As matter of fact the computed MAD for dimeric species is lower than the MAD computed on the monomers. Noticeably, and as experimentally observed, all λ_{MAX} are blue-shifted in going from the monomer to the dimer [153]. Comparing the dimers absorption spectra, the one of dimer **2** is characterized by a red-shift (0.30 eV) of the maximum absorption, with respect to **1**, the same found in the experiments [153]. In contrast, the molecule **3b** is slightly blue shifted at theoretical level (0.03 eV), while it results red-shifted (0.10 eV) with respect to the recorded spectrum [153].

Resuming, *static* TD-DFT calculations on the monomeric and H-type dimeric forms of merocyanines provide a picture in qualitative agreement with the experimental data, even if the obtained errors are quite large, though in line with previous literature data.

A sizable improvement, compared to the *static* approach, is obtained simulating the electronic absorption spectra as an ensemble average over multiple configurations. Indeed, the MAD is halved going from the *static* to the *MC* approach for both monomers and dimers. Going into the details, the MAD is 0.46 eV for the static monomers whereas it becomes 0.26 eV using a MC

sampling. For comparison the vertical transitions are computed at the same PBE0/6-31+G(d,p) level of theory in both the *static* and the MC spectra. As in the case of the *static* approach, the largest error is obtained for merocyanine **3b**, suggesting the subtle effects played by the length of the polymethine bridge and the BLA. The relative shifts computed for the different merocyanines in their monomeric forms also fairly agree with the *static* approach and the experimental results. The character of the brightest electronic transition corresponds to a $\pi \rightarrow \pi^*$ transition where HOMO and LUMO orbitals are delocalized over the whole core of the molecule.

The MC sampling leads to even better results in the case of dimers, for which the MAD decreases from 0.38 eV (*static* approach) to 0.17 eV (*MC* approach). In this case, the capability of MC to account for several configurations accessible in the phase space is crucial to reduce the simulated/experimental error on λ_{MAX}. Compound **3b** represents the trickiest molecule to simulate (error of +0.22 eV), also in its dimeric form. Overall the MC results are in fairly agreement with the experimental spectra both in terms of absolute computed-experimental shifts than in relative shifts between the different merocyanines.

4.4 Structural effects underpinning spectra

In this section we are going to examine how the structural features of merocyanines, and their dimeric species influence the λ_{MAX} and how the two different approaches can describe them. As previously mentioned, the major limitation of the *static* approach relies on the impossibility to consider all the energetically accessible minima existing on both ground and/or excited PESs. On the other hand, the MC simulations though cannot be able to provide the ZPE corrections due to their quantum nature, allow to take into account several energetically accessible conformations, thus supplying the contribution to λ_{MAX} coming from the distortions of flexible geometrical parameters. As matter of facts, several degrees of freedom can drastically influence the electronic properties of this family of chromophores. Relevant parameters, to be considered, are: *i*) the bond length alternation (Figure 4.6 a); *ii*) the relative orientation between the pyridone and the pyridine rings (amino group for **3**), linked by the conjugated chain (defined by the angle θ, (Figure 4.6 b) and *iii*) for the dimer, the relative distance (and orientation) of the two monomers. Small (large) BLAs are indication of increased (reduced) delocalization of the π electrons and of an improved (worsen) coupling between the two moieties, thus leading to a red (blue) shift of the spectra. A similar effect on the electronic structure could be expected by a coplanar rearrangement of the pyridone and the pyridine rings (or amino moiety for **3**), where smaller θ values correspond to smaller transition energies. Short intermolecular distances, represented as the distance between the

centre of mass (CoM) of the two monomers, lead to low energy transitions while an opposite effect is expected for larger inter-dimer distances.

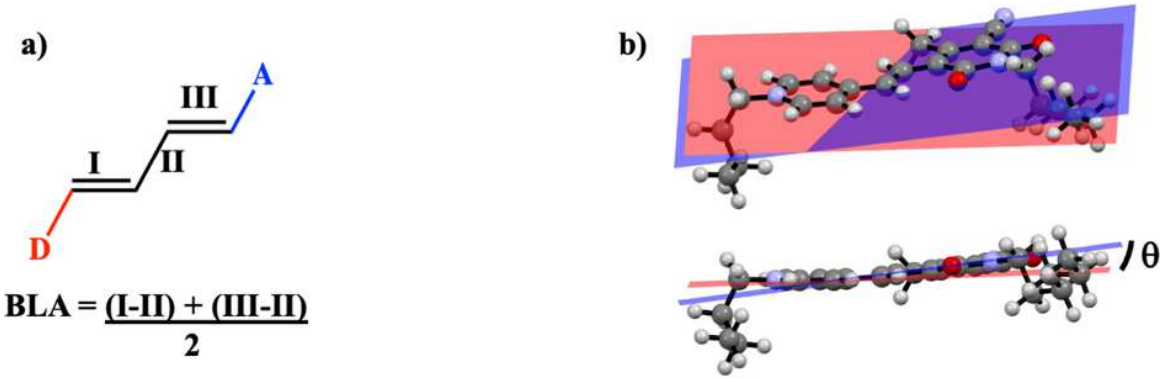


Figure 4.6 Schematic representation of the bond length alternation (a) and planes involved between the donor acceptor moieties (b).

The main geometrical parameters computed for merocyanines **1**-**4** are collected in Table 4.2 and Table 4.3 whereas the structures of the monomers are depicted in Figure 4.7. Going from static to MC calculations, small variations of the angle θ are observed, the largest ones being of about ± 4 degrees (for **2**, **3a** and **3b**), thus suggesting that this parameter, at least for monomers, is not relevant.

Table 4.2 Computed main geometrical parameters corresponding to the static approach or averaged MC simulations for the merocyanine monomers. The energy differences ΔE) between optimized PBE0 and MC structures are also reported. Values in parenthesis refer to structures optimized with the ωB97XD functionals.

	Monomer				
	θ^a (degrees)		BLA^b (Å)		ΔE (kcal/mol)
	static	MC	static	MC	
1	3.0 (1.0)	1.8	0.023 (0.024)	0.007	-3.8
2	8.0	3.8	0.017	0.023	-3.2
3a	0.4	4.0	0.008	0.003	-3.3
3b	5.3	7.6	0.015	0.038	-3.2
4	1.6	1.8	0.012	0.024	-4.3

Notes: a) refer figure 4(b); b) refer figure 4(a)

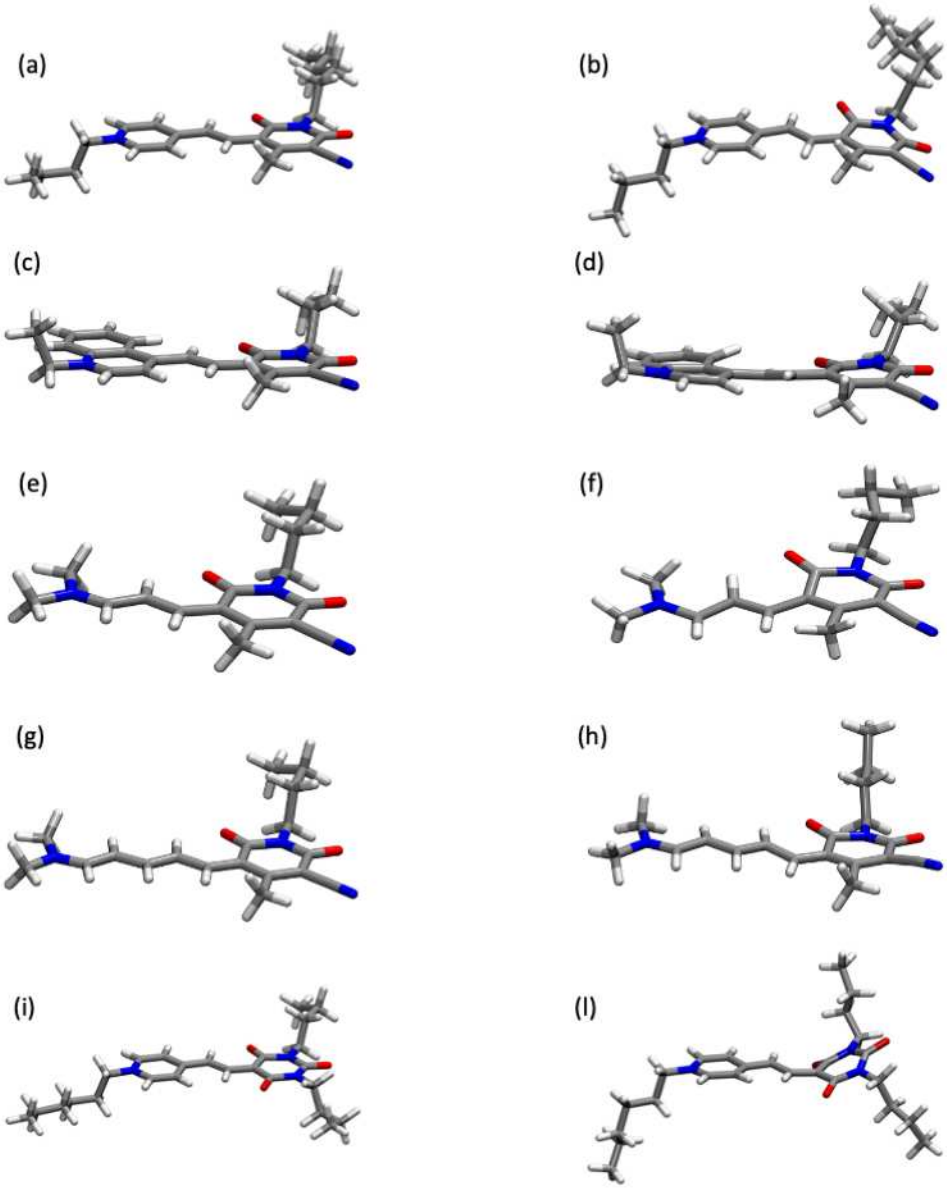


Figure 4.7 Optimized static (left) and Monte Carlo averaged structures (right) for merocyanine 1 (a and b), merocyanine 2 (c and d), merocyanine 3a (e and f), merocyanine 3b (g and h) and merocyanine 4 (i and l). Spectra are reported in Appendix D.

Concerning BLAs, the small lengths of the polymethine chain should prevent large errors at DFT level, generally observed in longer oligomeric chains [156, 157] and, indeed, PBE0 and ωB97XD functionals give very close values for molecule **1**, even if their errors become significantly different when larger systems are considered. Significant variations are, instead, observed in going from static to MC calculations, with different behavior for the considered systems. In particular, the BLAs increase for **2**, **3b** and **4** and decrease for **1** and **3a**, although in some cases the variations are quite small. Overall, a small energy difference (about 3 kcal/mol) between the PBE0 optimized

and the MC average structures is found for all the molecules. This energy difference is not only related to the BLA since the two structures differ also in other geometrical parameters (i.e., lateral chains conformation), as can be inferred from the structures reported in Figure 4.7. More interestingly, it cannot be traced from our results a direct relationship between any single selected structural parameter, including the BLAs or θ angles, and the red-shifts observed in going from static to MC values. Such a shift results from a global deformation rather than a single dominant degree of freedom. In other words, the hypothesis of a dominant effect of the C=C double bond variation on the merocyanine spectra cannot be confirmed by our calculations. Therefore, the net effect of stochastic calculations is the red-shift of the λ_{MAX} providing better agreement with the experimental spectra. It should be also pointed out that the resulting MAD is of the same order of magnitude as the E^{0-0} estimated for similar systems (0.3 eV). [135]

In an attempt to provide a comparison even with the vibrationally resolved absorption spectrum, the structure of the lowest electronic excited state (S_1) of molecule **1** was optimized at the PBE0/6-31G level, Figure 4.8.

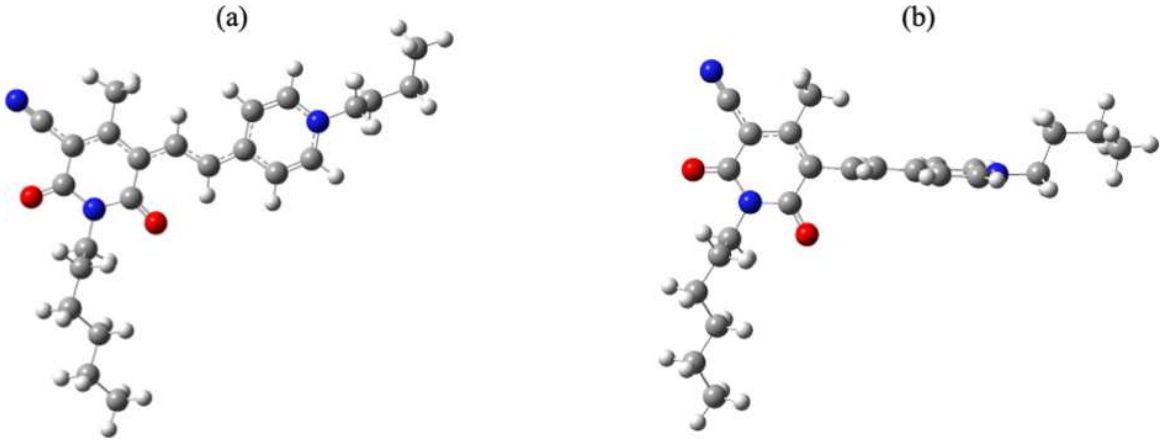


Figure 4.8 PBE0/6-31G/IEF-PCM(1,4-dioxane) Optimized structures of ground (a) and first excited (b) states of merocyanine **1**.

However, because of the flexibility of the molecule, the obtained S_1 minimum energy structure is strongly distorted with respect to the ground state one (θ close to 90°). As a consequence, the FC approximation is no longer valid and it is not possible, at least with usual approaches, to compute the vibronic contribution.

Once defined the global molecular flexibility as the main factor affecting the absorption energies of the merocyanine monomers, it is interesting to establish how these structural features can impact the aggregate formation. All the molecules considered are push-pull systems possessing a very high dipole moment on the ground electronic state (20.3 Debye for **1**). Previous studies have shown that dipole-dipole interactions are indeed the driving force for their self-assembly in dilute

solution leading to dimers showing an antiparallel orientation of the dipole moment of the constituting monomers and thus an overall null dipole moment. Furthermore, the presence of a null dipole moment of dimers avoids their further clustering and thus the formation of larger aggregates [135].

In agreement with the experimental observation, the MC sampling on the merocyanine dimers indicates that most of the stable dimeric forms are co-facial and antiparallel aligned (H aggregates) while only a few of them have a co-facial parallel arrangement of the two units (H dimer). In particular static calculations indicate that the co-facial antiparallel aggregates are most stable than the parallel ones of 18.1 kcal/mol and this is confirmed by the MC calculations showing less than 1% of parallel dimers. Furthermore, due to the relative high difference in energy between the two forms it is clear that it is difficult to go from one to the other arrangement. Due to the statistical sampling which massively favor the H dimers, in agreement with the experimentally observed blue-shift, only these latter will be considered in the following. These dimers appear even more flexible than the isolated monomers, since they undergo large structural modifications, as illustrated by the data reported in Table 4.3 and Figure 4.9.

Table 4.3. Computed main geometrical parameters corresponding to the static approach or averaged MC simulations for the merocyanine dimers. The energy differences (ΔE) between optimized PBE0 and MC structures are also reported. Values in parenthesis refer to structures optimized with the ωB97XD functionals.

	Dimer						
	θ (degrees)		BLA (Å)		center of mass (Å)		ΔE (kcal/mol)
	static	MC	static	MC	static	MC	
1	13.1 (12.7)	4.4	0.042 (0.06)	0.024	5.137 (5.080)	5.450	-7.1
2	44.1	2.6	0.038	0.018	5.610	4.320	-14.2
3a	48.9	2.1	0.011	0.030	4.110	4.360	-6.1
3b	55.8	8.4	0.007	0.026	3.908	4.350	-6.7
4	8.9	3.0	0.033	0.011	4.770	5.632	-9.0

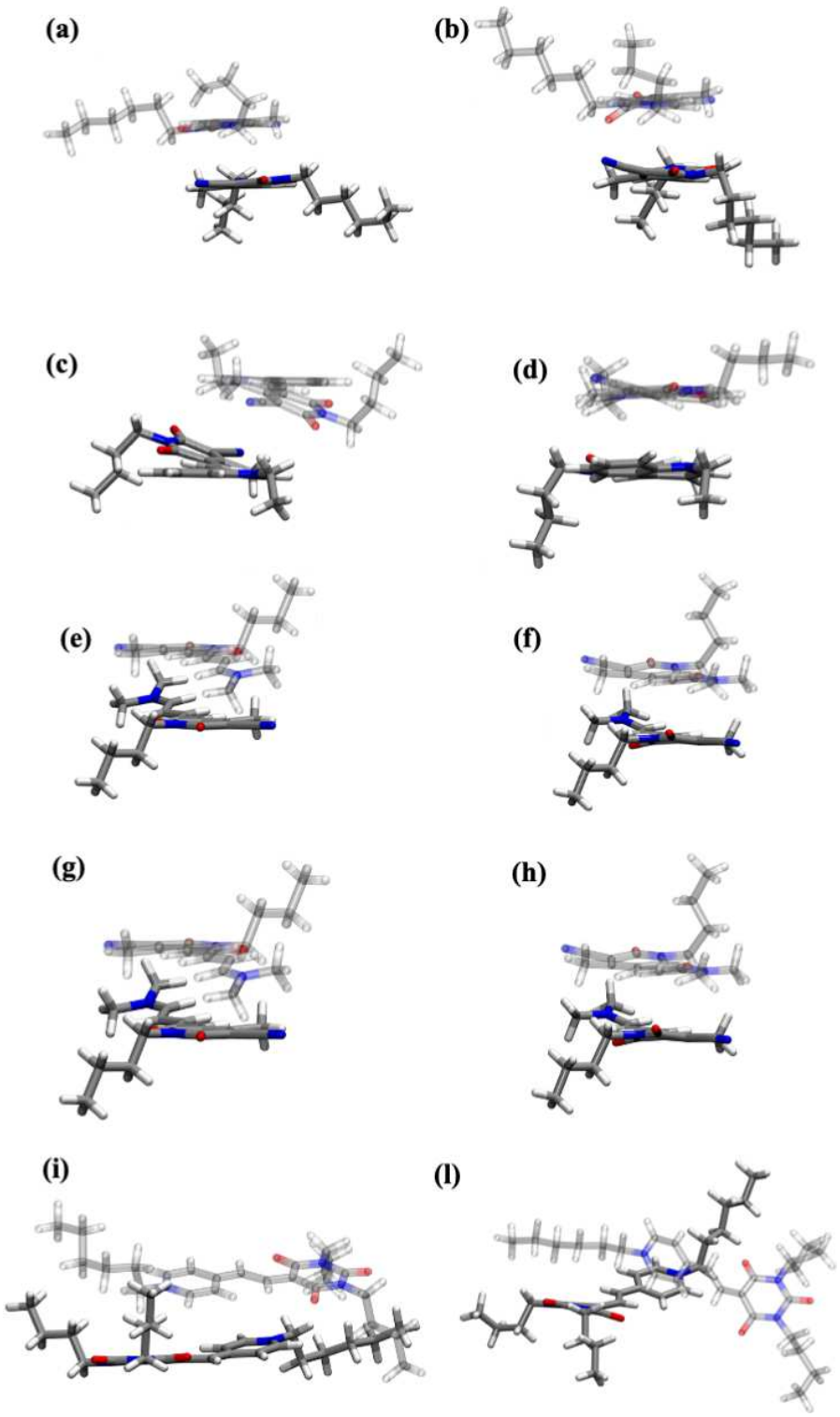


Figure 4.9 optimized static (left) and Monte Carlo averaged structures (right) for dimers merocyanine 1(a and b), merocyanine 2 (c and d), merocyanine 3a (e and f), merocyanine 3b (g and h), and merocyanine 4 (i and l). Spectra are reported in appendix D.

Indeed, significant variations are observed for both θ and BLAs. Starting from the first parameter, the static structures of dimers **2**, **3a** and **3b** are characterized by large torsional angles (θ around

50 degrees), while dimers **1** and **4** still retain an almost planar arrangement of their moieties (θ around 10 degrees). As above-mentioned, their corresponding isolated monomers are all almost planar. The weak non-covalent interactions between the pyridine rings force the planar structure of the monomers in molecules **1** and **4**, while the flexible amino terminal chain in dimers **3** prevents such arrangement. In molecule **2**, instead, the head-tail rearrangement is characterized by a steric repulsion between the ethyl and butyl chains substituting group of the two aromatic rings, which induces a shearing of the two moieties. As a consequence, the leucoline moieties can rotate around a C-C bond of the polymethine chain. These distorted structures break the extended π conjugation and induce a significant increase of the BLAs in the dimers exhibiting the largest values of θ (for molecules **1**, **2**, **4**). Therefore, the zwitterionic form is favored over the neutral one.

The intermolecular distance was, computed as the distance between the centers of mass (CoM) of the two moieties. Shorter distances correspond to a more compact structure, whereas the longer distances arise from a relative shift of the two units to reduce the steric hindrance. Thus, the intermolecular distance strongly depends on the steric repulsion in the packed structures, so that a large panel of values is found, ranging from 3.9 (molecule **3b**, the most compact) to 5.6 Å (molecule **2**). It is interesting to note that the MC structures are relatively higher in energy compared to the optimized ones. The difference in energies between the two sets is just twice than the value observed for the monomer (7-9 kcal/mol, see Table 4.2), the exception being the dimer **2** where very compact rearrangement (CoM = 4.3 Å) lead to significantly higher energies (ΔE = 14.2 kcal/mol).

Going from the monomers to the dimers, all the transitions are blue-shifted, so it seems that the mutual polarization of the two monomers in the aggregate plays the major role. Therefore, we have simulated the electronic spectra of the monomer at the geometry in the corresponding dimer, for both optimized static and MC structures. As it appears from the results collected in Table 4.3, all the transitions are close (within a few thousandths of eV) to those computed for the corresponding isolated monomers, regardless of the approach used to obtain such a structure (static or MC). As for the isolated monomer, it is not possible to draw a direct relationship between the BLA and the θ parameters, the main factor affecting the blue-shifts of the dimers seems to be their formation, accounted by the distance of the CoMs. Nevertheless, there is not a correlation between the blue-shifts in the transitions and the energy of formation of the dimers. Similarly, no correlations were found between the static and MC transitions for the dimers and their corresponding CoMs distances (or their differences). However, the numerical improvement in computed absorption energies appears evident looking at the computed MAD (0.17 eV for the

MC dimers), well below the expected value for such systems. Figure 4.10, where the computed spectra for the dimer **1** are reported, also exemplifies this improved agreement.

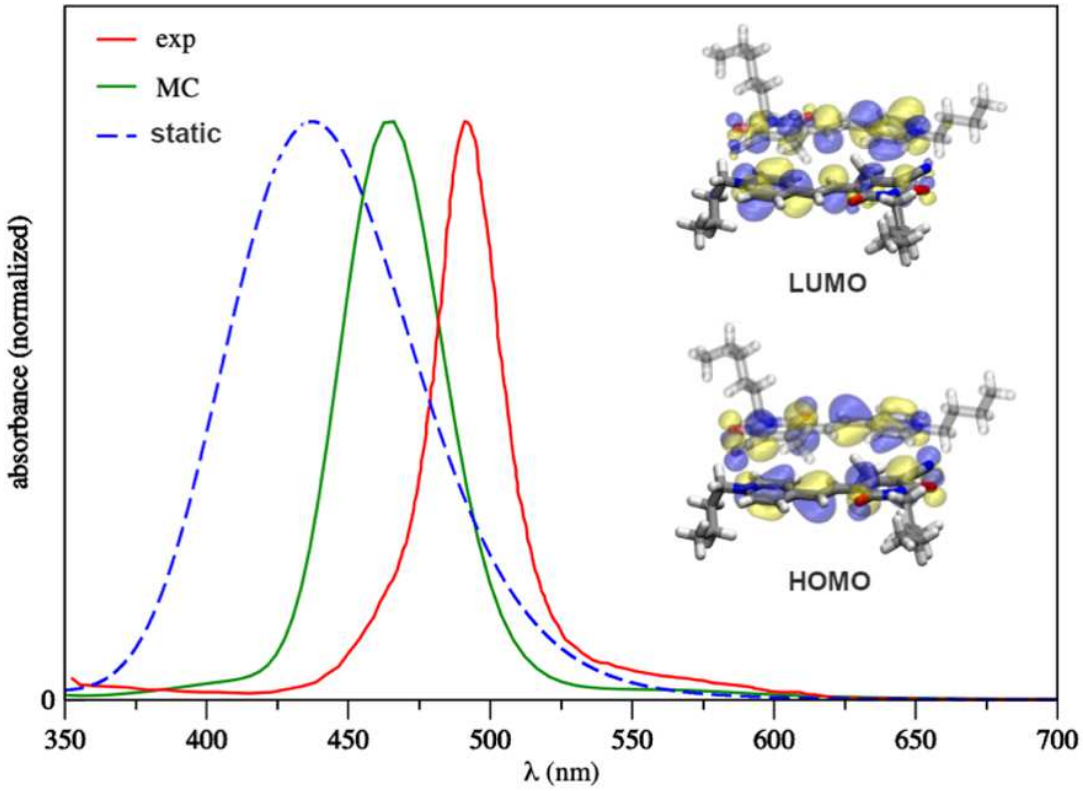


Figure 4.10 Absorption spectra of dimer **1** in 1,4 dioxane computed with the static (blue dashed line) and stochastic (green solid line) approaches. The experimental spectrum is also reported for comparison (red line)

4.5 Conclusion

In conclusion, the overall agreement between MC results and experimental data suggests that the main structural effects on the electronic spectra are captured by the MC approach combined with vertical TD-DFT calculations. Although classical MC simulations do not allow to reconstruct the vibronic progression, they take into account the flexible nature of the molecules and its effect on the λ_{MAX} which is overall shifted leading to a better agreement with the experimental spectra. Our results reveal that despite its simplicity a combined MC/TD-DFT approach is reliable to describe flexible systems that cannot be properly described in a FC static framework. Indeed, in the case of the merocyanines dyes the characterization of the FC minimum leads to an error bar (0.3-0.5 eV) for the spectra of the monomers obtained with standard functionals, including global or range-separated hybrids, in perfect agreement with literature data [136, 158].

Finally, to the best of our knowledge, a detailed FC analysis of merocyanines dimers has never been performed, the only study concerning two units bound by aliphatic chains which significantly damp intermolecular soft modes. This is not surprising, in view of the difficulties that can be expected in the optimization of the dimers on the excited electronic states. In this respect, the MC stochastic sampling here proposed is a reliable alternative to sample the dimer conformational space. Promising applications can be expected in the future, where the MC approach could be used to generate initial conditions for ab initio molecular dynamics.

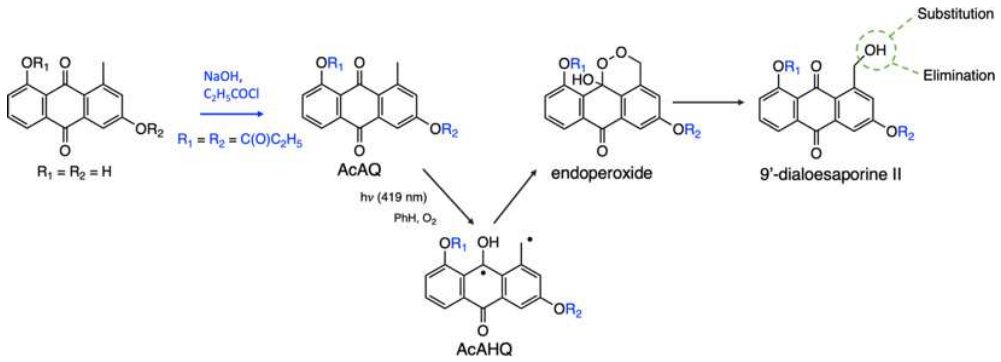
Chapter 5

5. Modelling Hydrogen Abstraction

In this chapter we report our study on the photochemical hydroxylation of two functionalized anthraquinones. Precisely, we describe a novelty scheme which couples the information derived from the density-based indexes with the Sin-Orbit Coupling in order to give a quantitative prediction of possible crossing between excited states.

5.1 Introduction

Photolysis of functionalized anthraquinones is one of the main methods to synthetise an hydroxylation product [159,160]. The photoproduct can be used as a starting point to make new chromophores, since the alcohol moiety can be easily transformed (see Scheme 5.1) [161]. Within this framework, understanding the photoinduced reaction is crucial, being the initial stage to produce new dyes of interest for PILI company. As shown in scheme 5.1, an anthraquinone (AcAQ) readily abstracts the hydrogen from its methyl group upon irradiation at 419 nm (2.95 eV) [161, 162]. Absorption causes the promotion of an electron from the ground state to the excited electronic state, then the excited AcAQ undergoes an intersystem crossing (ISC) to the triplet state, promoting the formation of the 1-4 diradical anthrahydroquinone (AcAHQ). This diradical is able to react with molecular oxygen to yield the endoperoxide. Here, a second intersystem crossing occurs [161]. Continuing the photolysis, the hydroxylation of 9'-dialoesaporine II is produced.



Scheme 5.1 Photochemical hydroxylation reaction investigated to produce the 9'-hydroxylaloesaporine II.

The major limitation of the Scheme 5.1 reaction is the low quantum yield of the 9'-dialoesaporine II products due to the conversion of the diradical intermediate back to the starting ketone, also known as internal dismutation [163]. Thus, the key step in this multistage reaction is the AcAHQ production, whose energetic barrier governs the endoperoxide formation. Fundamental for the AcAHQ production is to study the deactivation pathways from the initially singlet excited bright state to the triplet one. In this concern we modelled the non-radiative decay pathway for two substituted anthraquinones characterized by acetate (Ac) and the triflate (Tf) as protective units, Figure 5.1. In addition, although anthraquinones in their excited triplet states are relatively strong oxidants and can thus promote an intermolecular hydrogen transfer, the hydrogen abstraction can simultaneously occur in the decay mechanism making its modelling a challenging topic from a computational point of view.

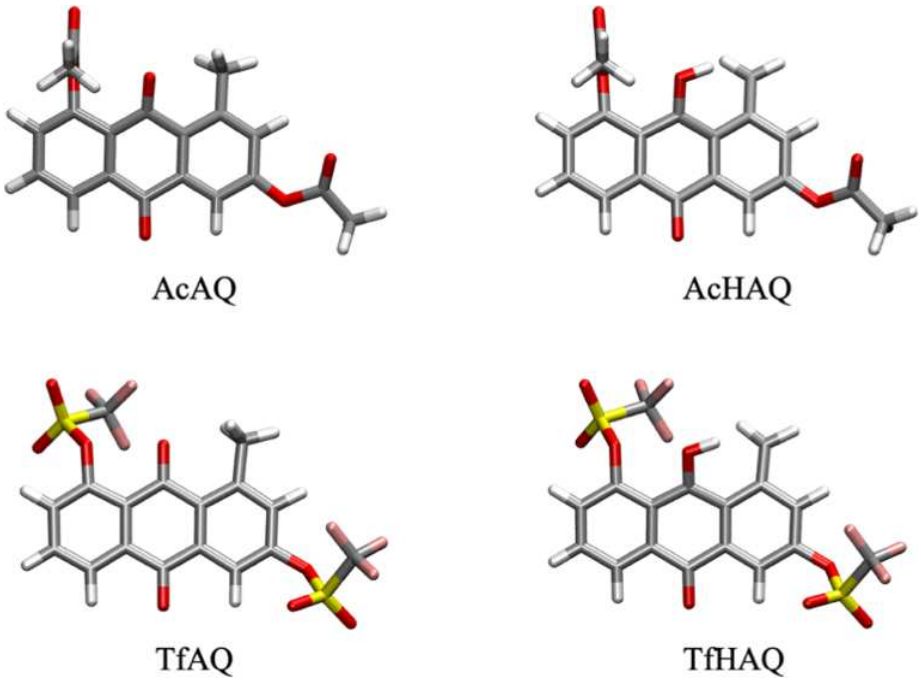


Figure 5.1 Stick representation of two disubstituted anthraquinones/anthrahydroquinone investigated in this work.

5.2 Spectra of AcAQ and TfAQ

Ground state geometries of both AcAQ and TfAQ compounds have been optimized at the PBE0/6-31++G(2d,2p) level in benzene (PCM). Since experimental spectra for the selected compounds are not available in the literature, Figure 5.2 presents only the simulated spectra. Broadening effects were here mimicked by the convolution of Gaussian functions.

The spectrum of AcAQ (orange solid line) is characterized by two intense picks, one is located around 346 nm ($S_0 \rightarrow S_3$) and the other one around 273 nm ($S_0 \rightarrow S_7$). For these two bright transitions the oscillator strengths (f) are respectively 0.134 and 0.154.

The absorption spectrum of TfAQ (dash orange line) is characterized by two intense peaks. The first peak is located around 331 nm ($S_0 \rightarrow S_3$, $f = 0.124$), while the second one is observed at 244 nm ($S_0 \rightarrow S_{10}$, $f = 0.360$). As a result, in comparison with AcAQ, the spectrum of TfAQ is slightly blue shifted.

For both compounds, the first two electronic transitions ($S_0 \rightarrow S_1$ and $S_0 \rightarrow S_2$) are hidden by the predominant one ($S_0 \rightarrow S_3$). Indeed, for AcAQ the f are equal to 0.005 ($S_0 \rightarrow S_1$ at 412 nm) and 0.001 ($S_0 \rightarrow S_2$ at 370 nm), while for TfAQ the f are equal to 0.002 ($S_0 \rightarrow S_1$ at 415 nm) and 0 ($S_0 \rightarrow S_2$ at 374 nm).

Based on these findings, one may deduce that the vertical excitation energy for the lowest "dark" AcAQ state is consistent with the experimental wavelength (419 nm) [161]. However, if we consider irradiation with blue light (around 346 nm), the S_3 state is directly populated and, thus, could affects the photolysis rate of 1-methyl-bi (Ac)/(Tf)-AQ in benzene. This hypothesis is consistent with other experimental data. Indeed, similar photosensitive compounds used for photolysis reaction are irradiated upon the first bright electronic transition [164-165].

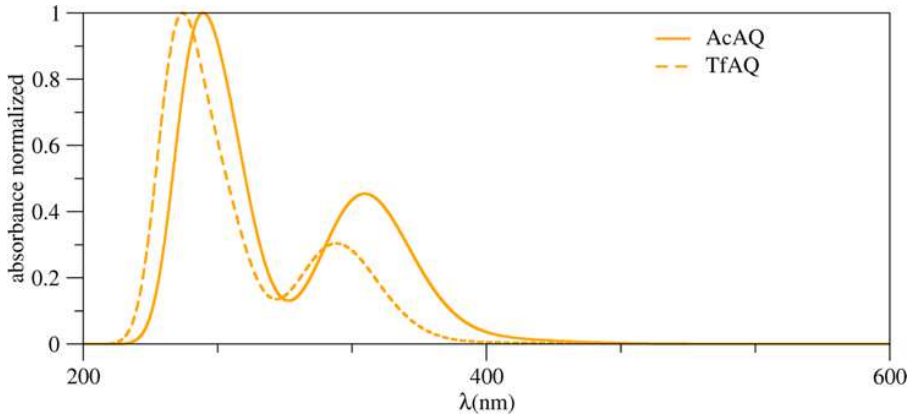


Figure 5.2 Computed electronic absorption spectra of AcAQ (solid orange line) and TfAQ (dash orange line) in benzene. The normalized spectrum has been obtained by convolution of Gaussian functions assigning an arbitrary full weight at half maximum equal to 0.3 eV.

5.3 Modelling Hydrogen abstraction pathway

Our goal is to model the photochemical hydroxylation in the framework of DFT. To derivate a model for the excited-state deactivation and identify the reaction channel, we have used an approach based on the linear synchronous-transit (LST) method. LST consists in constructing

a linear interpolation pathway starting from the optimized structure of reactant and product [166]. Going into the details, in the first reaction step the absorption populates a singlet state with highest oscillator strength, S_k, while the product (AcAHQ or TfHAQ) is in its triplet state (T). Thus, according to the results reported in paragraph 5.2, the LST path has been constructed keeping both AcAQ (TfAQ) and AcAHQ (TfHAQ) structures in their S_3 minximum. Giving that a decay mechanism involves different electronic states with the same or different spin multiplicity the excited state-deactivation has been detected by varying the electronic state but keeping fixed the LST structures obtained from the reference state (S_3). This approach is consistent with the effective modelling of the decay mechanism, since the structural changes are much slower than the electronic decay. Figure 5.3 shows the possible pathways involving the excited states for AcAQ (TfAQ) to produce the corresponding protected anthrahydroquinone.

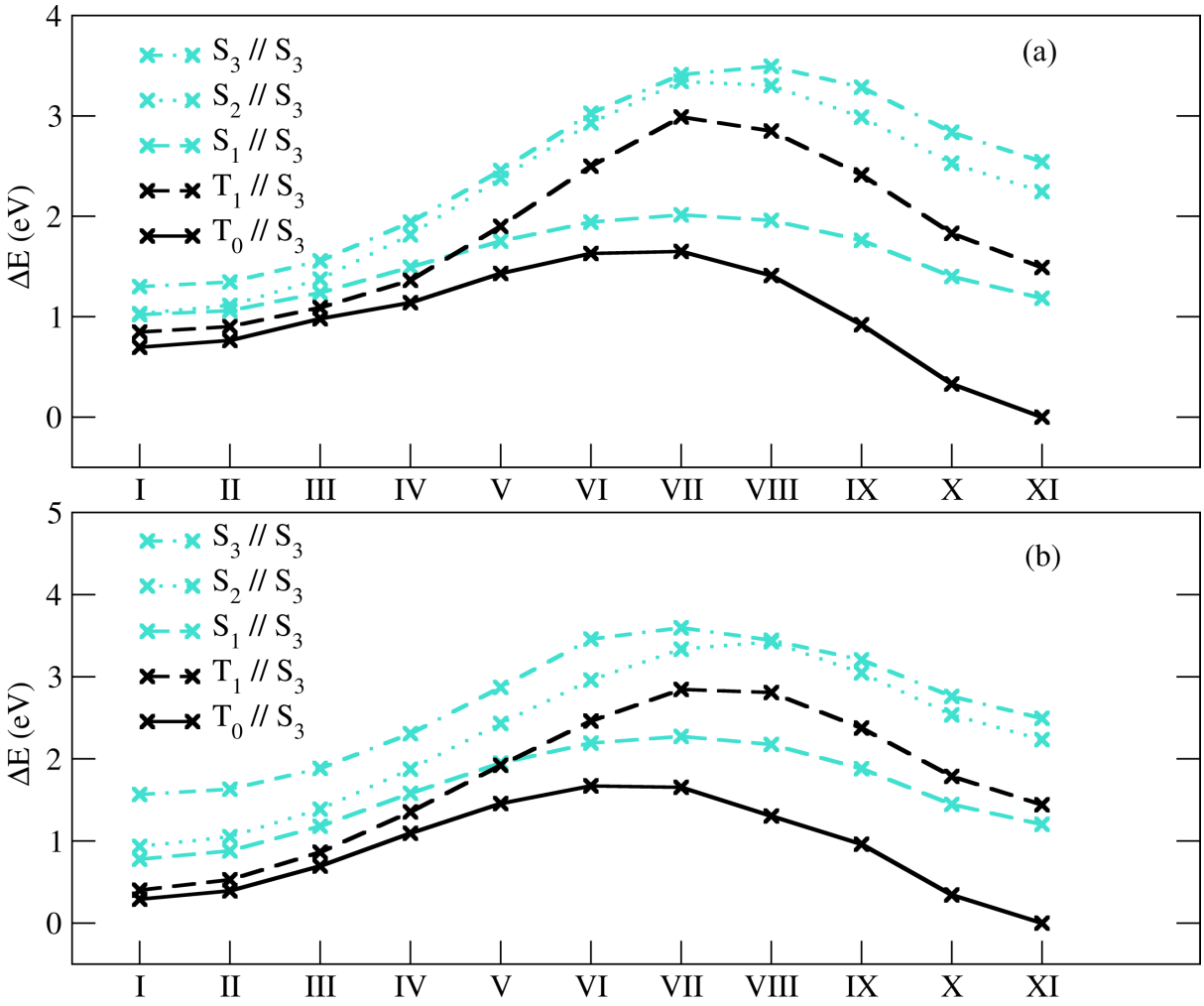


Figure 5.3 Electronic energy variation against the reaction coordinates. Curves identify the state and geometry: Ai//Bj, Ai identifies the electronic state; Bj identifies the optimized geometry for a given electronic state. Level of theory: PBE0/631++G(2d,2p)/PCM (benzene). Curves scaled with reference to T_0//S_3 minimum energy. Hydrogen abstraction energy profiles: upper panel (a) refers to acetate (Ac) substituted AQ, bottom panel (b) refers to triflate (Tf) functionalized AQ.

The energy profiles obtained at the S_3 optimized geometry (turquoise and black curves Figure 5.3) identify the most probable pathway, because both the rate and the intersystem crossing efficiency can be explained by small electronic and nuclear variations that the chromophore needs to undergo during the singlet–triplet transition. Efficient decay pathways emerge in regions where two states are close in energy ($\Delta E_{ij} \rightarrow 0$ eV).

Starting the discussion with the AcAQ, the first bright excited state is computed to be at 3.75 eV and corresponds to the third excited state (S_3). All energy differences are reported in Table 5.1. A possible channel is the $S_3 \rightarrow S_2$ decay, appearing around step number V (see also upper panel of Figure 5.3). For the S_2 state an intersystem crossing with the state T_1 at the optimized S_3 geometry takes place around step number VII.

Table 5.1 Energy differences (eV) between the singlet-singlet, singlet-triplet and triplet-triplet at given LST step, data refers the Ac functionalized AQ.

LST	$\Delta E_{S_3-S_2}$	$\Delta E_{S_2-S_1}$	$\Delta E_{S_1-T_1}$	$\Delta E_{S_3-T_1}$	$\Delta E_{S_2-T_1}$	$\Delta E_{T_1-T_0}$
I	0.28	0.01	0.17	0.46	0.18	0.15
II	0.23	0.05	0.16	0.44	0.21	0.14
III	0.18	0.14	0.14	0.46	0.28	0.12
IV	0.13	0.32	0.13	0.58	0.45	0.22
V	0.08	0.63	-0.15	0.56	0.48	0.46
VI	0.10	0.99	-0.57	0.52	0.42	0.87
VII	0.07	1.33	-0.98	0.41	0.35	1.35
VIII	0.19	1.34	-0.89	0.64	0.45	1.44
IX	0.30	1.23	-0.65	0.87	0.58	1.49
X	0.31	1.13	-0.43	1.00	0.69	1.50
XI	0.30	1.06	-0.31	1.05	0.75	1.49

Like the LST path of AcAQ, the Tf-substituted AQ is characterized by the same intersystem crossings (bottom panel of Figure 5.3) along its hydrogen abstraction path. Table 5.2 summarizes the electronic energy variations between points in the PESs constructed with the LST procedure for the Tf substituted AQ. Around LST step number VI, the first excited triplet state crosses the S_1 curve. A possible channel is the $S_3 \rightarrow S_2$ decay around LST step number VIII ($\Delta E_{S_3-S_2}$ = 0.02 eV). For the S_2 state an intersystem crossing with the state T_1 at the optimized S_3 geometry probably takes place around structure VIII ($\Delta E_{S_2-T_1}$ = 0.62 eV). On the other hand, the energy difference between the S_3 and T_1 around step VIII is computed to be 0.64 eV and thus a possible intersystem crossing occurs, and thus a possible ISC can occur.

Table 5.2 Energy differences (eV) between the singlet-singlet, singlet-triplet and triplet-triplet at given LST step, data refers to Tf functionalized AQ.

LST	$\Delta E_{S_3-S_2}$	$\Delta E_{S_2-S_1}$	$\Delta E_{S_1-T_1}$	$\Delta E_{S_3-T_1}$	$\Delta E_{S_2-T_1}$	$\Delta E_{T_1-T_0}$
I	0.63	0.16	0.38	1.16	0.53	0.11
II	0.58	0.17	0.35	1.10	0.53	0.13
III	0.50	0.21	0.32	1.02	0.52	0.17
IV	0.43	0.29	0.23	0.95	0.52	0.26
V	0.44	0.48	0.03	0.95	0.51	0.47
VI	0.50	0.77	-0.27	1.00	0.50	0.79
VII	0.26	1.06	-0.57	0.75	0.49	1.19
VIII	0.02	1.25	-0.63	0.64	0.62	1.50
IX	0.16	1.17	-0.50	0.83	0.67	1.42
X	0.23	1.09	-0.34	0.97	0.75	1.45
XI	0.26	1.03	-0.24	1.05	0.79	1.44

5.5 Spin Orbit coupling

The Spin–orbit coupling (SOC) plays a crucial role in several chemical and physical phenomena. For instance, photochemical reactions in electronic excited states can be affected by SOC by promoting intersystem crossing (ISC) between states with different spin multiplicity. In our case, the hydrogen abstraction reaction produces the photoproduct AcAHQ (or TfAHQ) at T_1, which reacts with the $^3O_2$ to form the endoperoxide species (Scheme 5.1), therefore the SOC cannot be neglected.

In shorthand, the spin-orbit Hamiltonian is written as the sum of two terms [169]:

$$\mathcal{H}^{SO} = \frac{e^2\hbar}{2m^2c^2}\left[\underbrace{\sum_i \sum_\alpha Z_\alpha \boldsymbol{s}_i \left(\frac{r_{\alpha i}}{r_{\alpha i}^3} \times \boldsymbol{p}_i\right)}_{\text{One-Electron}} - \underbrace{\sum_{i \neq j} \left(\frac{r_{ij}}{r_{ij}^3} \times \boldsymbol{p}_i\right)(\boldsymbol{s}_i + 2\boldsymbol{s}_j)}_{\text{Two-Electron}}\right] \tag{5.1}$$

Where $\boldsymbol{\alpha}$ denotes the nuclei and $\boldsymbol{Z}$ is charge, $\boldsymbol{s}$ is the spin operator, the product between the distance ($\boldsymbol{r}$) and the momentum operator ($\boldsymbol{p}$) is the orbital angular momentum operator ($\boldsymbol{\iota} = \boldsymbol{r} \times \boldsymbol{p}$). The structure of this Hamiltonian shows that the SOC is proportional to the nuclear charge the so called "heavy atoms" effects, indeed, for studying the SOC generally hydrogen atoms

are replaced with bromine or iodine [170]. To avoid the expensive two-electron part, the mean-field (MF) approximation, which is defined through the averaging of the two-electron contributions to the SO coupling matrix elements over the valence electron shell, is almost universally used. [171]. Although replacing the two-electron term with one-body term sounds dramatic, the method is known to provide accurate prediction [172].

Employing the MF approximation, the Spin−orbit matrix elements have been computed in benzene (PCM) by using the DALTON suite [173]. The SOC has been computed according to the equation [174, 175]:

$$SOC_{S_m,T_1} = \sqrt{\sum_n \left|\left\langle \varphi_{S_i} \middle| \hat{\mathcal{H}}_{SO} \middle| \varphi_{T_{j,n}} \right\rangle\right|^2} \tag{5.2}$$

Where the sum is over the three triplet components (i.e., $n = x, y, z$). To compute the SOC, the 6-61++G(d,p) basis set has been used.

Furthermore, since few hybrid GGA functionals are available in DALTON, the B3LYP has been selected for performing SOC calculations. Our choice is motivated by the fact that both PBE0 and B3LYP predict similar excitation energies [39]. As a matter of fact, the decision has been backed up by the energy profiles reported in Figure 5.4.

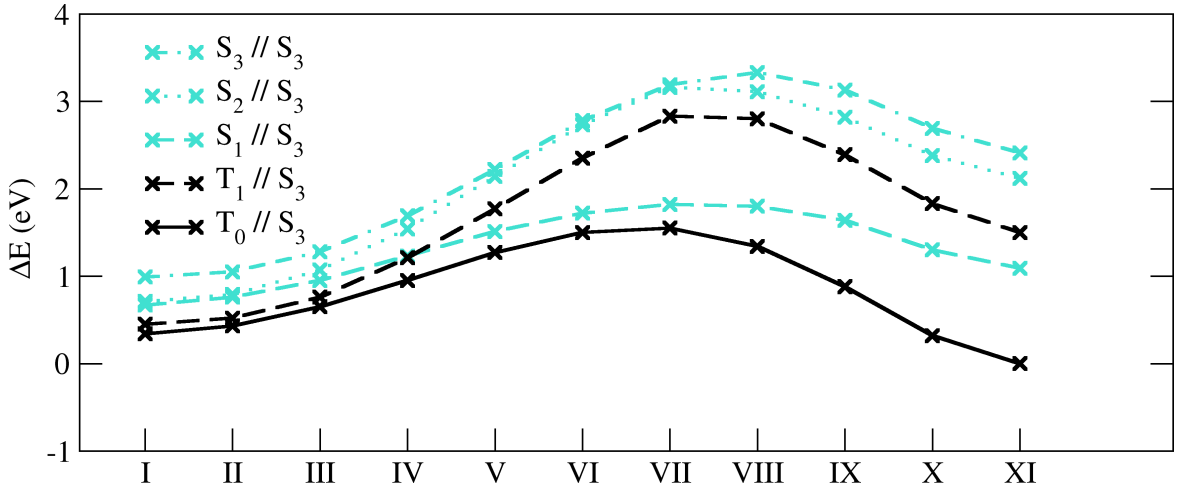


Figure 5.4 AcAQ energy profile. Electronic energy profile computed at B3LYP/6-31++G(d,p)/Benzene(PCM) level on LST structures obtained at PBE0/6-31++G(2d,2p)/Benzene(PCM) level.

From Table 5.3, it is easy to see that the largest SOCs are located in the region spanning from structure VI to structure IX. In details, the $S_3 - T_1$ crossing (151.45 cm^{-1}) corresponds to the global maximum, but considerable contributions are also derived from the $S_2 - T_1$ (29.32 cm^{-1})

and $S_1 - T_1$ (118.84 cm^{-1}) intersystem crossings. Thus, the intersystem conversion from the higher singlet states to the lowest triplet excited state could provide a direct deactivation channel to populate the T_1. On the other hand, lowest SOC values are 0.47 cm^{-1} (structure XI), 5.78 cm^{-1} (structure XI) and 3.33 cm^{-1} (structure IV).

Alike the previous system, for $S_1 - T_1$ the largest SOC (308.65 cm^{-1}) is found around LST step number VIII (see Table 5.3). On the other hand, both $S_2 - T_1$ and $S_3 - T_1$ couplings achieve their global maxima (25.63 cm^{-1}, and 7699.63 cm^{-1}) around step number XI. Since the SOC values for TfAQ are higher than those obtained for AcAQ, it can be identified as a possible candidate for producing the 9'-dialoesaporine II [161].

Table 5.3 SOC value (cm^{-1}) computed between the lowest triplet and singlet states for the investigated systems at B3LYP/6-31++G(d,p)/PCM(Benzene) level.

LST	AcAQ			TfAQ		
	SOC_{S_{1},T_1}	SOC_{S_{2},T_1}	SOC_{S_{3},T_1}	SOC_{S_{1},T_1}	SOC_{S_{2},T_1}	SOC_{S_{3},T_1}
I	2.18	14.96	13.34	0.70	0.47	3.07
II	7.55	15.67	11.87	0.94	3.35	3.86
III	12.07	14.65	6.86	0.99	8.94	3.88
IV	8.41	15.01	3.33	0.44	14.56	4.25
V	3.54	15.37	5.06	2.29	17.90	7.94
VI	2.31	29.32	5.93	6.94	23.48	45.13
VII	11.87	14.25	7.05	17.92	18.98	12.93
VIII	118.84	17.44	30.76	308.65	23.27	38.41
IX	19.33	25.13	151.45	16.66	25.64	7699.64
X	0.60	10.84	14.37	0.27	10.37	15.64
XI	0.47	5.78	14.99	0.23	5.55	15.74

5.6 Spin Density Distributions and Structural considerations

To provide a clear picture of anthraquinone reactivity, the spin density distributions (SSDs) have been utilized as indicators. Moving from structure I to structure XI, SDDs on both photoactive compounds change considerably. However, a general trend can be inferred.

Starting with AcAQ, Figure 5.5 collects all SSDs along its LST path. These distributions are characterized by two distinct areas, one centered on the carbonylic groups and the other one on

the aromatic skeleton. Going from structure I to structure IV the SSD is mainly localized on C=O groups. Continuing the photolysis, a fraction of spin density moves from the C=O group to the CH_2 one, which serves as an anchor site for the $^3O_2$. Unlike the ketone part, acyl groups have negligible spin densities. Since SSDs in $T_1//S_3$ are similar to other known photoactive anthraquinones [176], high photoactivity is detected when the triplet SSD is not localized on the anthracene skeleton [177].

Figure 5.6 depicts the SSDs of TfAQ along its LST path. As in the case of AcAQ, carbonylic stretches and similar localization of the SSDs on the C=O group have been observed. Indeed, also in this case, SSDs are not localized on the protective units. In contrast to AcAQ, on step IX (where the ISC is suspected to be) the spin density on the CH_2 group in γ position with respect to the C=O group is itself in resonance on the right ring of the aromatic skeleton.

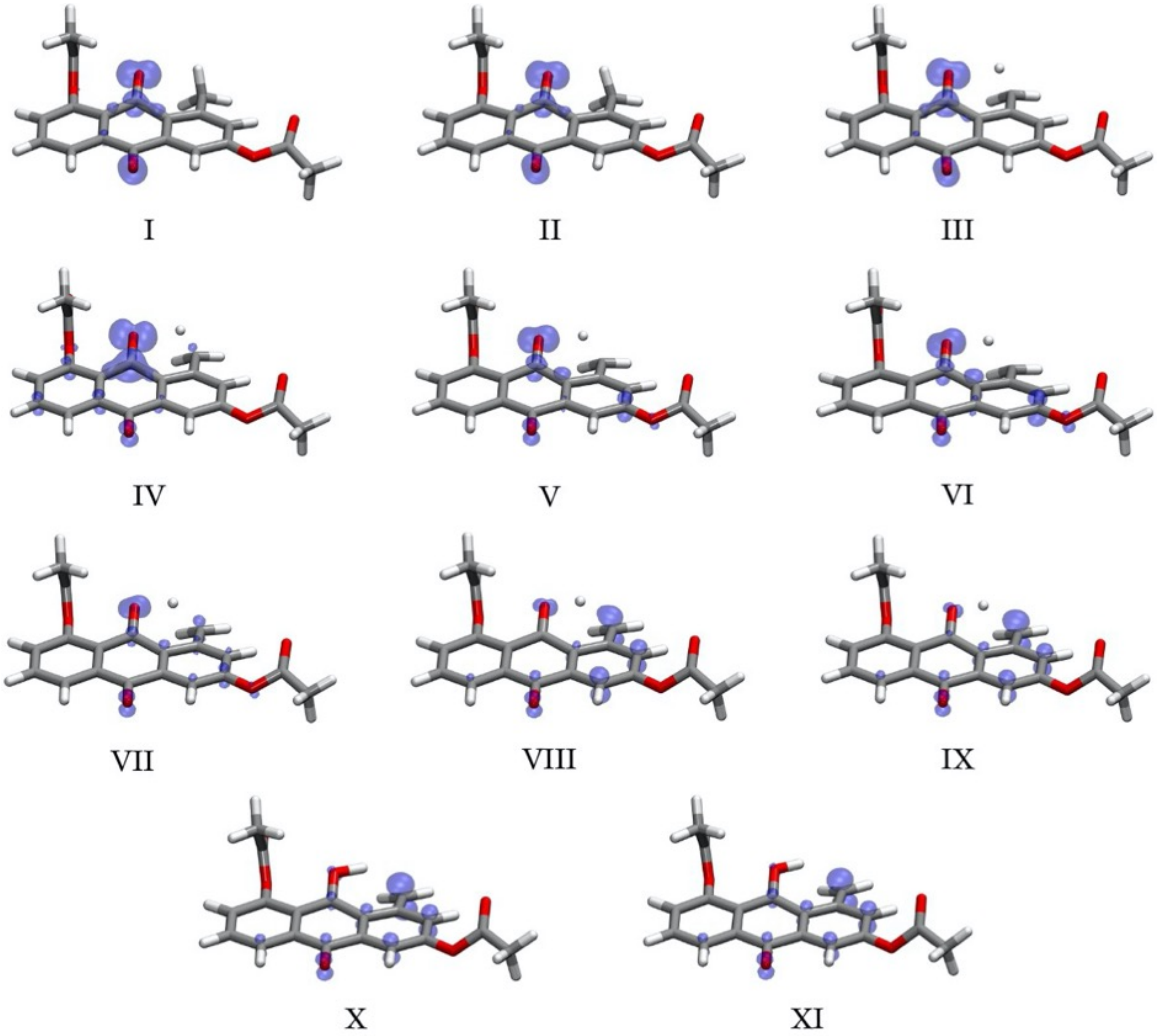


Figure 5.5 Spin density distributions (isovalue 0.02) for all Ac substituted AQ ($T_1//S_3$).

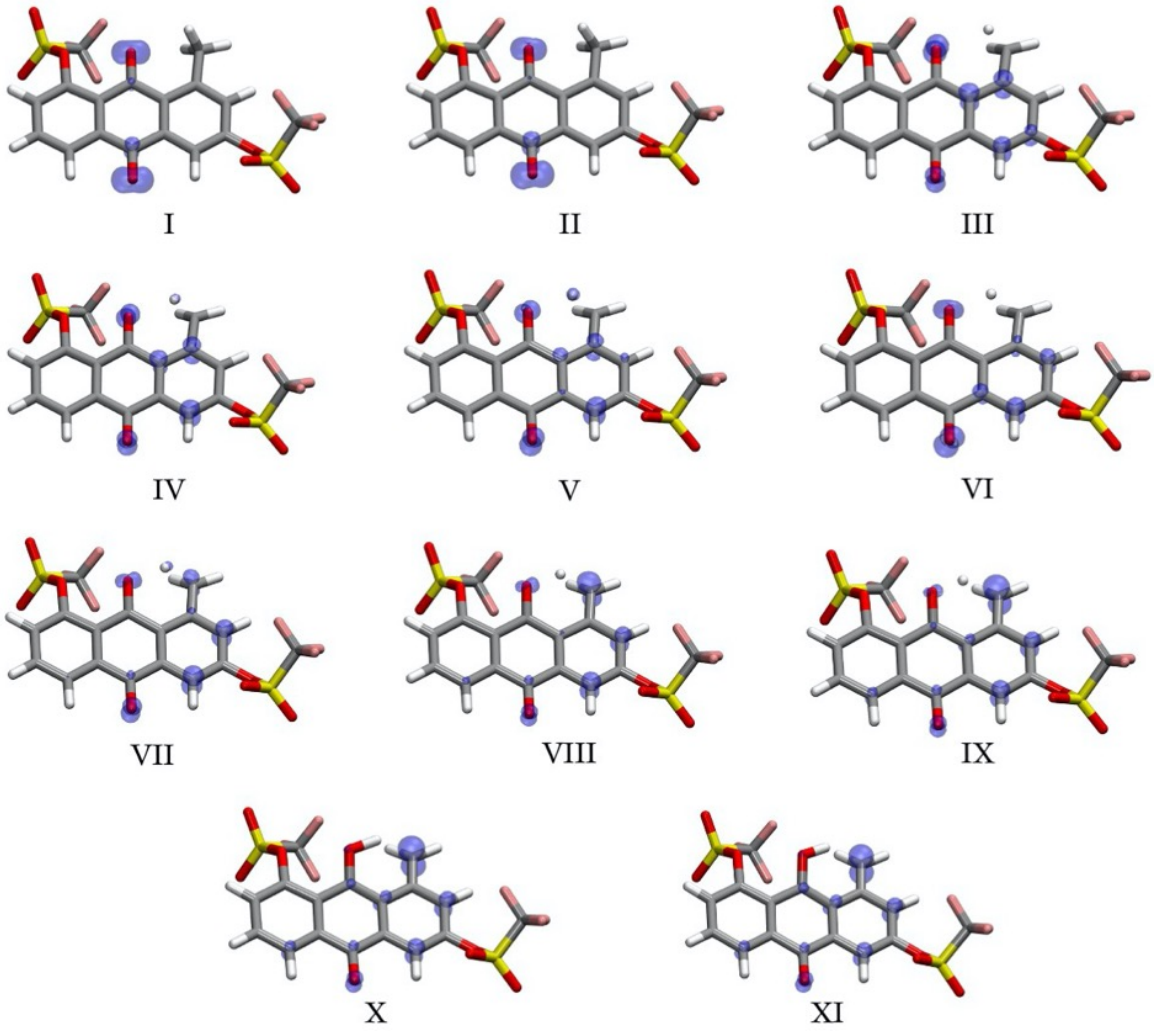


Figure 5.6 Spin density distributions (isovalue 0.02) for all Tf substituted AQ ($T_1//S_3$).

When the C=O bond lengths of the structure I obtained at $S_0//S_0$ and $S_3//S_3$ are compared (Figure 5.7, Tables 5.4 and 5.5), our computation reveals that the bonds length stretches and increases by 0.01 Å in $S_3//S_3$. Furthermore, on the structure I, these variations are quite small and thus can be considered as additional validity proof of the scheme adopted to model the hydrogen abstraction path. When going from structure I to structure XI, the C=O stretch is counterbalanced by a tightening of C-C bond length in the central ring (r_2), allowing a gaining in the sp^2 character for localizing the unpaired electrons. The same trend characterizes the C-C bond length in γ to the carbonyl group.

In particular, at LST step number IX the carbonyl group increases (0.07 Å) in bond length that is found to be equal to that of the bond length obtained of unsubstituted AQ, [177].

Table 5.5 collects the bond length for TfAQ compound. As reported for the first system, no significative geometrical change on the structure I has been observed and, in addition, the same

analysis performed on the first system can be inferred for TfAQ. However, unlike the AcAQ, around LST step number X the carbonyl group increases (0.06 Å) in bond length that is found to be equal to that the bond length obtained for a similar AQ compound [176]. Overall, these results suggest that a possible reaction between $^3O_2$ and the photoactive molecules in their triplet states can take place around the final stage of the LST path, since equilibrium geometries are similar to some known triplet photoactive anthraquinones [176].

Table 5.4 Bond lengths (Å) of AcAQ and TfAQ for the optimized geometry ($S_0//S_0$).

LST	AcAQ($S_0//S_0$)				TfAQ($S_0//S_0$)			
	r_1	r_2	r_3	r_4	r_1	r_2	r_3	r_4
I	1.22	1.48	1.41	1.50	1.21	1.48	1.41	1.50

Table 5.5 Bond lengths in Å for AcAQ and TfAQ along the hydrogen abstraction path.

LST	AcAQ($S_3//S_3$)				TfAQ($S_3//S_3$)			
	r_1	r_2	r_3	r_4	r_1	r_2	r_3	r_4
I	1.23	1.49	1.45	1.49	1.24	1.46	1.42	1.50
II	1.24	1.48	1.45	1.48	1.24	1.46	1.42	1.48
III	1.25	1.48	1.45	1.47	1.25	1.45	1.42	1.47
IV	1.26	1.47	1.44	1.46	1.25	1.45	1.42	1.45
V	1.27	1.47	1.44	1.45	1.26	1.45	1.42	1.44
VI	1.27	1.46	1.44	1.44	1.26	1.44	1.42	1.43
VII	1.28	1.45	1.44	1.43	1.27	1.44	1.42	1.42
VIII	1.29	1.45	1.44	1.42	1.28	1.44	1.43	1.41
IX	1.30	1.44	1.44	1.41	1.29	1.44	1.43	1.41
X	1.30	1.44	1.43	1.40	1.30	1.44	1.43	1.40
XI	1.31	1.43	1.43	1.39	1.30	1.43	1.44	1.40

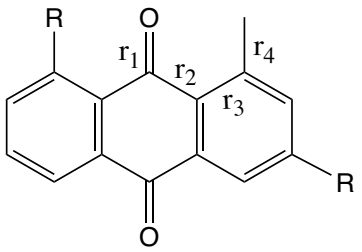


Figure 5.7 Sketch of substituted (R= Ac or Tf) AQ compound considered in this work. Bond lengths are reported in Table 5.6 and table 5.7.

5.4 Density based index to study the reactivity

In order to characterize the nature of the excited states a density-based index, the so-called D_{CT}, has been used [168,178]. Here, the basic idea is to consider the difference in the total electronic density computed for the ground and the excited states and to spot the regions of electron density depletion and electron density increment upon excitation (ρ^+ and ρ^- respectively). The distance between the barycenters associated to these two density distributions corresponds to the average particle-hole distance, that is the D_{CT}. In other terms, large D_{CT} values correspond to large CT distance et vice versa. Therefore, this index provides an estimate of the spatial extent of a given electronic transition, allowing us to monitor the changes in the character of the excited states. In order to characterize the possible crossing taking place at the excited state, D_{CT} values with respect to S_1, S_2 and S_3 states have been computed. The obtained results are plotted in Figure 5.8.

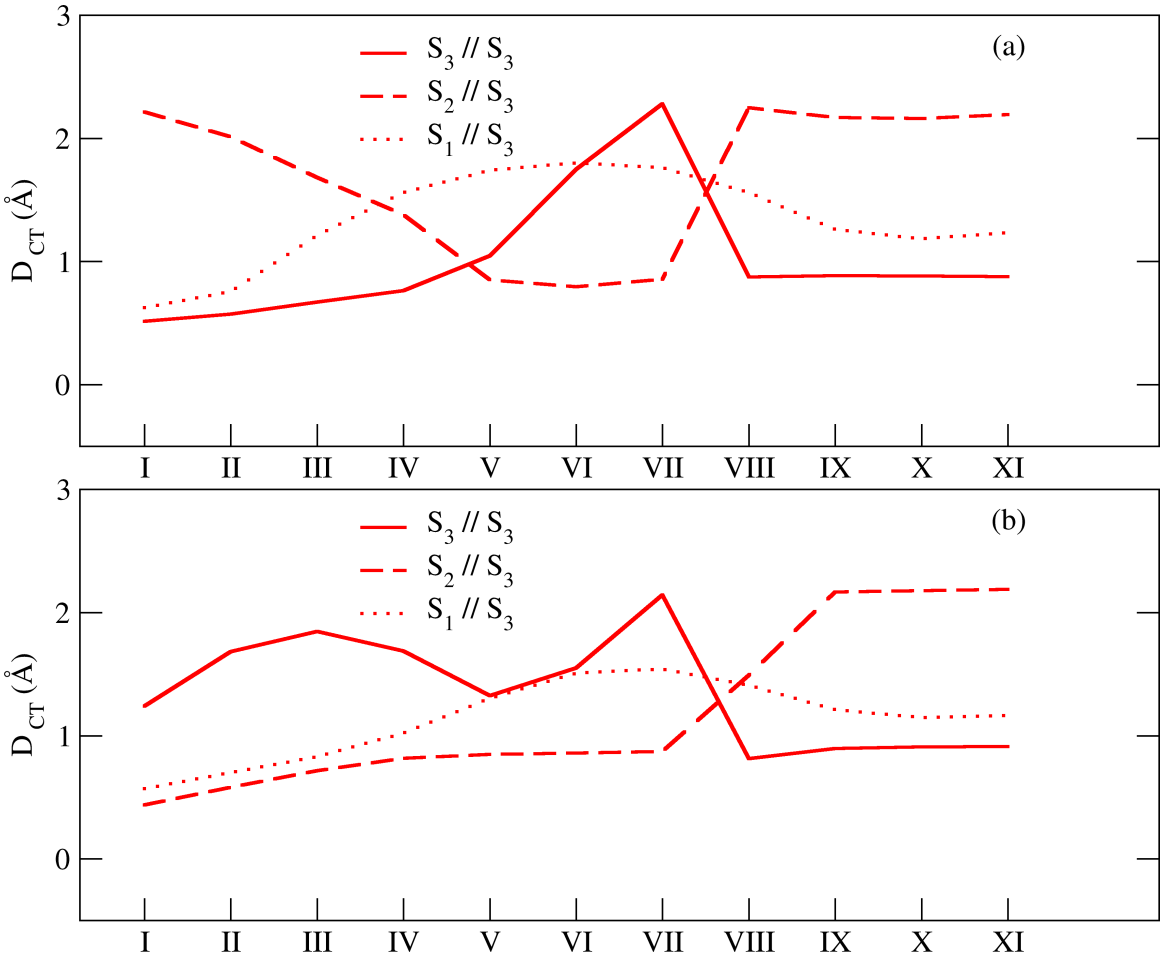


Figure 5.8 D_{CT} evolution along the LST scan. (a) Ac-substituted AQ, (b) Tf-functionalized AQ.

The upper panel of Figure 5.8 reveals that the S_1 and S_2 D_{CT} profile cross around structure IV. At the same LST step, the first excited state corresponds to a dark ($f = 0.04$, see Table 5.6) CT state characterized by a high D_{CT} while the second excited state S_2 is characterized by a lower D_{CT} value and corresponds to a dark locally excited (LE) state.

Continuing the hydrogen abstraction, the difference in D_{CT} between the S_3 and S_2 states tends to zero and the character of these states mixes up around structure number V (1.045 Å and 0.851 Å). At this LST step the second excited state gains more local character, while the third state gains in CT character. Finally, around the region which goes from LST step number VI to step number VIII the nature of the third and second excited state is inverted as reflected by the values of the corresponding D_{CT}.

The bottom panel of Figure 5.8 depicts the D_{CT} variation of Tf-substituted AQ. From structure III to structure V, the CT character of the third excited state decreases. The first excited state, on the other hand, gains in CT nature and corresponds to a dark state (Table 5.7). The D_{CT} values for S_1 and S_3 at step number V are 1.304 Å and 1.325 Å, respectively.

As seen with AcAQ, around structure VIII the third and second excited states cross each other. Here, the higher dark excited state gains LE character, on the other hand, the second excited state gains CT character (0.813 Å and 1.492 Å).

Overall, the computed D_{CT} value for both investigated systems show similar D_{CT} profile, suggesting a possible crossing around structure VIII.

Table 5.6 oscillator strength (f_{s_i}) and D_{CT} (Å) computed values for acetate protecting AQ.

LST	$S_0 \rightarrow S_1$		$S_0 \rightarrow S_2$		$S_0 \rightarrow S_3$	
	f_{s_1}	D_{CT}	f_{s_2}	D_{CT}	f_{s_3}	D_{CT}
I	0.003	0.625	0.078	2.215	0.004	0.513
II	0.007	0.754	0.085	2.013	0.004	0.572
III	0.028	1.210	0.079	1.682	0.001	0.670
IV	0.043	1.561	0.071	1.379	0.003	0.762
V	0.046	1.741	0.050	0.851	0.028	1.045
VI	0.045	1.800	0.011	0.795	0.072	1.749
VII	0.050	1.761	0.002	0.855	0.007	2.280
VIII	0.072	1.560	0.008	2.248	0.001	0.873
IX	0.115	1.262	0.013	2.170	0.001	0.884
X	0.161	1.183	0.035	2.162	0.001	0.882
XI	0.193	1.234	0.062	2.192	0.001	0.877

Table 5.7 Oscillator strength (f) and D_{CT} (Å) computed values for Tf protecting AQ.

LST	$S_0 \rightarrow S_1$		$S_0 \rightarrow S_2$		$S_0 \rightarrow S_3$	
	f_{s_1}	D_{CT}	f_{s_2}	D_{CT}	f_{s_3}	D_{CT}
I	0.000	0.570	0.000	0.439	0.060	1.239
II	0.000	0.700	0.000	0.582	0.118	1.683
III	0.001	0.831	0.001	0.716	0.147	1.847
IV	0.012	1.020	0.004	0.816	0.146	1.687
V	0.035	1.304	0.004	0.849	0.130	1.325
VI	0.055	1.508	0.003	0.860	0.094	1.550
VII	0.073	1.543	0.002	0.873	0.010	2.142
VIII	0.098	1.410	0.010	1.492	0.002	0.813
IX	0.131	1.212	0.022	2.168	0.001	0.897
X	0.164	1.147	0.044	2.177	0.001	0.908
XI	0.187	1.162	0.064	2.188	0.001	0.913

In order to identify qualitatively the potential energy region where a crossing between the states is suspected, the Π index has been used [178,179]. This index couples the D_{CT} information with a charge displacement (q_{CT}) and an energetic term.

$$\Pi = \frac{1}{\Delta E \; q_{CT} \; D_{CT}} \tag{5.3}$$

High values of Π will define the regions corresponding to the highest decay probability. Although the energy gap between the two states will be the leading term for many photochemical reactions implying a crossing of states ($\Delta E \rightarrow 0$), for a given value of the energy gap a decay will be more efficient the more similar the electronic densities of the starting and final states are, i.e. when the product $D_{CT} \cdot q_{CT}$ is small. Otherwise stated, the smaller the electron reorganization between the two states, the more likely the decay will be. As a result, efficient decay pathways are likely to emerge in those regions where two states come close in energy ($\Delta E \rightarrow 0$), or when the electron densities of two states involved are similar ($D_{CT} \cdot q_{CT} \rightarrow 0$), or when both of these criteria are simultaneously verified. Figure 5.9 shows the more probable decays identified by the Π index on all LST steps.

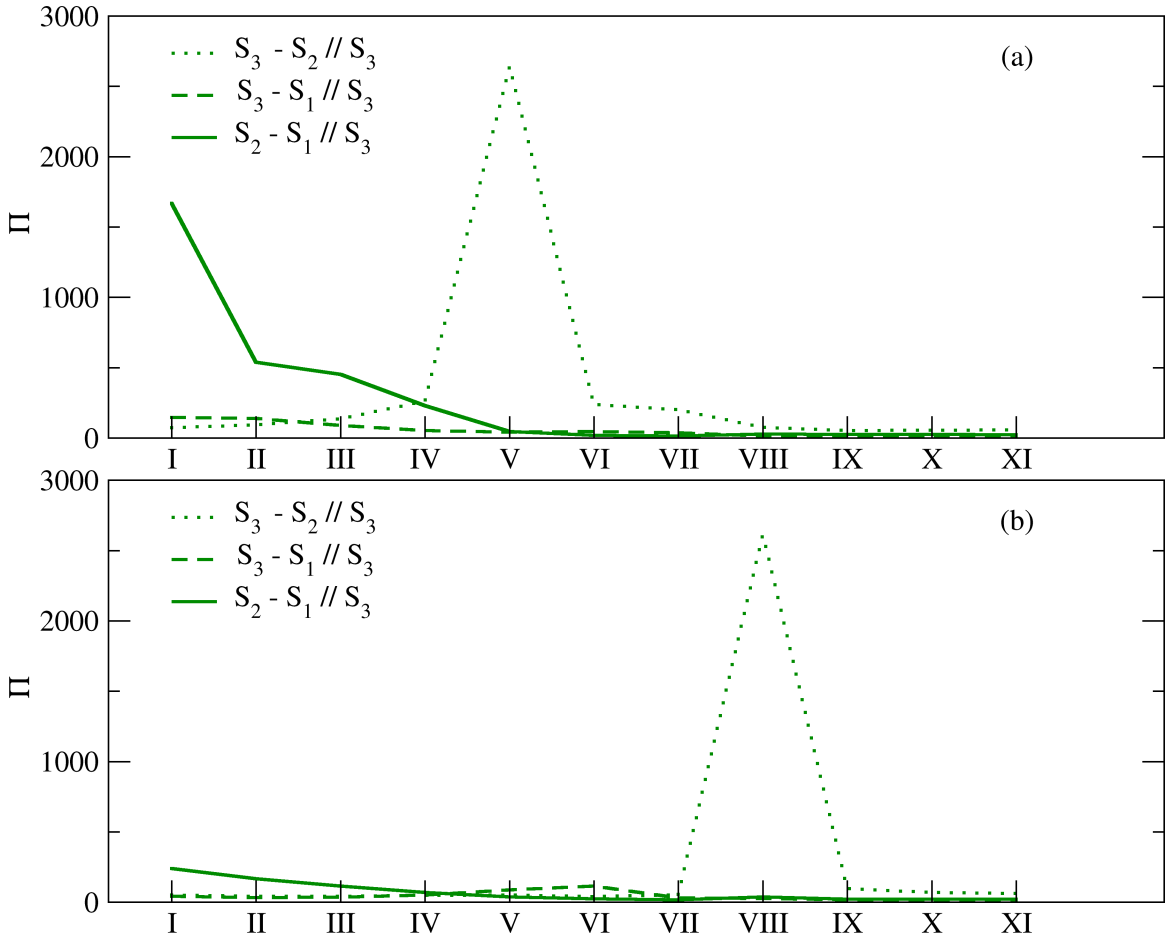


Figure 5.9 Evolution of Π index along the reaction coordinate. upper panel (a) refers to acetate (Ac) substituted AQ, bottommost panel (b) refers to triflate (Tf) functionalized AQ.

As mentioned above, the Π index allows us to identify regions of the excited state potential energy surfaces where decay pathways are more likely to occur.

As can be seen in Figure 5.9, for the AcAQ the global maximum is localized on LST step number V and refers to the S_3-S_2 transition. The oscillator strength of S_3 (f_{S_3} = 0.028) indicates low absorption, suggesting the presence of non-radiative channel which allows the electronic population transfer from S_3 to S_2 (f_{S_2} = 0.050). This decay is essentially energy driven ($\Delta E_{S_3-S_2}$ = 0.08 eV). Continuing the photolysis, no crossing between singlet states have been found. For TfAQ the Π profile reaches its global maximum around structure VIII where the oscillator strength variation is computed to be 0.0088. As observed for AcAQ, the emissive decay is energy driven ($\Delta E_{S_3-S_2}$ = 0.018 eV), Figure 5.3 (b). The photochemical path derived from the Π profile is coherent with the electronic energy variation, the D_{CT} and the Δq_{CT} profile (Figure 5.10).

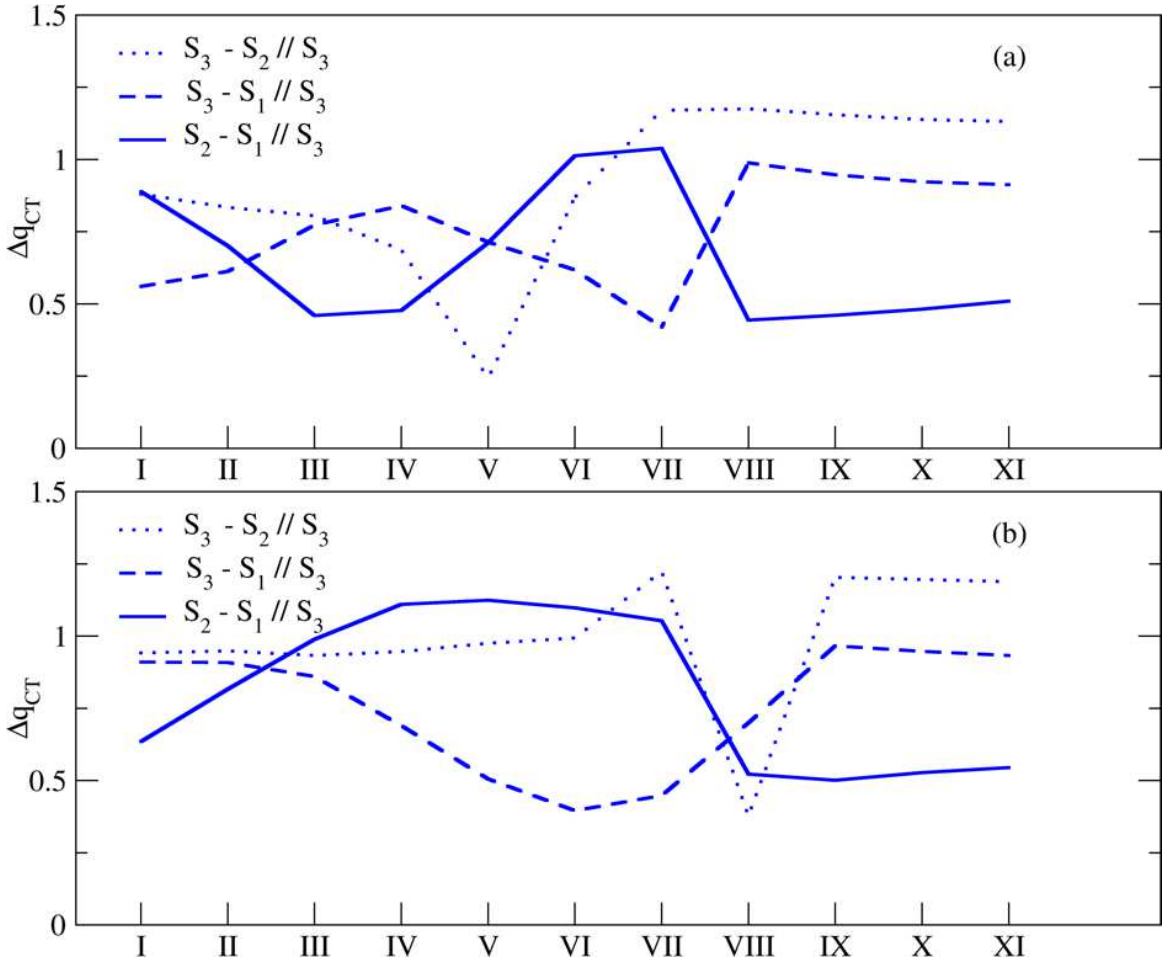


Figure 5.10 Δq_{CT} evolution along the LST scan for (a) Ac-substituted AQ and (b) Tf-functionalized AQ.

5.6 Conclusion

Accounting for the experimental evidence on 1-methylanthraquinone [181] which demonstrated that the photochemical migration of hydrogen can occur in both singlet and triplet excited states, the proposed mechanism accounts for all possible non-radiative channel decays and has shown that the hydrogen abstraction proceeds without spin-conservation leading to a triplet biradical. Contrarily the hypothesis that the protected AQ actively undergoes photoreduction by abstracting hydrogen atom in the T_1 state [162], our computations show that the direct intersystem crossing is unfavourable, indeed, the computed SOC_{S_3,T_1} on structures I are 13.34 cm^{-1} and 3.06 cm^{-1} for AcAQ and TfAQ, respectively. In addition, according to our model, hydrogen abstractions could not occur in the first stage, because the energy barriers computed on the T_1 surface for AcAQ and TfAQ are $\Delta E_{T_1(I)-T_1(VII)}$= 2.15 eV and $\Delta E_{T_1(I)-T_1(VII)}$= 2.91 eV. Further, both S_3-S_2 and S_3-S_1 Π

curves possess minimum values around the first photoreaction stage, thus these decays cannot occur.

Concerning the Ac protecting AQ, the S_3-S_2 Π curve reaches its global maximum around structure V. Continuing the hydrogen abstraction an intersystem crossing is observed around structure VI and the energy barrier computed on the S_2 state results to be $\Delta E_{S_2(VII)-S_2(VI)}$= 0.41 eV. On the other hand, for TfAQ system a more reasonable decay channel takes place around step number VIII involving the S_3 and S_2 states, at IX step a possible intersystem crossing between S_2 and T_1 appears, and the computed energy barrier is $\Delta E_{S_2(VIII)-S_2(IX)}$= 0.37 eV.

Overall, all steps have $\Delta E_{S_2-T_1}$gap lower than one the necessary to activate the oxygen molecule from its triplet ground state to singlet excited state 0.97 eV [180, 162]. Thus, in agreement with the experimental data [162], the two tested photoactive molecules react with the $^3O_2$. This observation is coherent with the experimental results inferred by Elkazaz and co-workers [161], as matter of the fact that a $^1O_2$ quencher had no effect on the rate of photocleavage leading the endoperoxide formation [161].

Another important point is that, although the intramolecular hydrogen migration for the investigated compounds possesses small energy barriers, they vary depending on the functional group used, indeed, the electron-withdrawing groups installed in γ with respect to the carbonylic moiety positively impact on the diradical formation [182]. Overall, our results allow us both to identify a non-radiative decay channel connecting S_3 to T_1 via S_2, and to demonstrate the existence of a narrow channel which leads to an apparent barrierless mechanism.

Chapter 6

6. Conclusions and perspectives

The purpose of this thesis was to study the photophysical and photochemical phenomena of organic chromophores in solution. To this end, we decided to investigate and to model the light induced phenomena in the framework of TD-DFT. With this tool we set-up three cost-effective computational procedures for three classes of chromophores. The first two molecular classes, anthraquinones and indigos, are currently used in the textile industry, since they show a variety of colors. The third class, merocyanines, is routinely employed in different fields ranging from industrial paints to sun filters.

Our first study focused on the spectroscopic signatures of derived anthraquinones and derived indigoes. The former has two keto groups installed on the central ring of the anthracene skeleton and functional groups which replaces hydrogen atoms of the aromatic core, whereas the latter chromophores have two indole (or thioindole)-derived structures linked by a double C=C bond and functional groups which replace the hydrogen atom of the amino group. In order to predict the color of these natural dyes, we performed a spectroscopic investigation at DFT/PCM/TD-DFT level which provided Uv-Vis spectra. Here, we tested two XC functional: the PBE0 and ωB97XD. Our computations have shown that the PBE0 functional reproduces better the experimental absorption peaks than ωB97XD, hence the PBE0 functional has been selected for obtaining the vertical excitation energy throughout this work. To account the band shape, we used two methods. The first is based on a convolution of Gaussian or of Pseudo-Voigt functions, whereas the second method accounts the vibrational fine structure. Concerning this latter, vibrational progressions have been obtained by using three approaches: Adiabatic Hessian, the LQ2 e LQ3. Only for the Adiabatic Hessian method, we decided to shift the ωB97XD band at the PBE0 vertical absorption wavelength, since the vibronic band shape is better described by ωB97XD functional than the PBE0 one. We called the resulting approach ωB97XD /PBE0(vert). The Euclidean distances, computed between the experimental and computed colors in the CIE-Lab colors space, detected lower errors when LQ2 and LQ3 approaches are used instead of

ωB97XD /PBE0(vert). Interestingly, the results showed that the application of the pseudo-Voigt and LQ2 and LQ3 approaches allows the accurate replication of the absorption spectrum. Overall, errors over the Euclidian distances between the computed and the experimental colors are small for the pseudo-Voigt, the LQ2 and LQ3 methods.
Our analysis allowed us to set-up two low-cost procedures for obtaining the absorption spectra and the spectral features of organic dyes in solution. Although the vibronic methods evaluated were successful in predicting spectral features, only the Adiabatic Hessian, LQ2, and LQ3 were tested. As a result, other methods such as the Vertical Hessian and Vertical Gradient models should be tested are indeed the perspectives work of this project. Furthermore, the developed protocols provide the molecular finger print with sufficient accuracy, it should be stressed that we benchmarked 19 dyes that have maximum absorption wavelength ranging from 408 nm to 526 nm, implying that further investigations are needed to validate our approach in the complete UV-vis range.

The second study focused on five polymethine dyes in both their monomeric and dimeric forms. These chromophores, called merocyanines, are typical flexible dyes and they constitute the third molecular class studied in this thesis. In order to consider the impact of the possible molecular rearrangements on the absorption property two computational approaches have been selected. The first strategy refers to a static approach, where, vertical TD-DFT calculation were performed on just one optimized ground state structure. The second strategy corresponds to a combination of molecular mechanics and a DFT procedure. More precisely, we combined a *MC* sampling of the configurational space with TD-DFT to estimate vertical excitations energies for computing the absorption spectra. Interestingly, when the vertical excitation energies are obtained by static approach, the computed MAD are larger than those obtained at MC/DFT level. To better investigate this, an analysis of the relevant geometrical parameters have been performed on the optimized ground state structures and averaged MC ones. Overall, our results showed that the observed shifts are due to small deformations of the molecular geometry, thus the combination of MC sampling with quantum mechanics calculations provides a workable compromise to solve the combined challenge of accuracy and time-consuming problem. Indeed, the simulated absorption spectra agree with the experimental ones, suggesting the general reliability of the method. In the future, a possible strategy can be developed taking into account the averaged configurations. Indeed, they can be used as staring point in a molecular dynamics simulation and the spectra can be extracted on-the-fly on the molecular trajectory.

The last work has been focused on the reactivity in the excited state. Here, two substituted anthraquinones have been selected as prototypical compounds for obtaining hydroxylation products. In order to model the photo reaction mechanism, the LST approach has been used. On this path, we integrated the results obtained from the Π index with the SOC values, in order to provide a clear picture of the possible decay pathways. We applied the index Π to estimate the crossing probabilities between singlet excited state of both AcAQ and TfAQ molecules. We identified one relevant channel which occurs with large probability and corresponds to the S_3 and S_2 crossing. On the other hand, the spin orbit coupling allowed us to infer a coherent picture of where the intersystem crossing occurs in this photochemical process. Overall, our computations have shown that the intersystem crossing can occur around the final stage of the LST path, leading to an apparent barrier less mechanism.
In conclusion, although the method provides a good description for tracking possible intersystem crossings, a further investigation can be performed by including more than one excited triplet state.

To conclude, computational protocols developed in this project have been proven to efficiently model the complex phenomena that involve the excited states and are able to efficiently support the PILI company in the design of new dyes to reduce the environmental impact.

7. Appendix

7.1 Appendix A

As detailed in the main text, the Pseudo-Voigt broadening function used in our work is expressed as:

$$PV\big(x,\gamma(x)\big) = (1-M)g'(x) + M\ell'(x) \quad \text{(EQSI.1)}$$

With $g'(x)$ and $\ell'(x)$ being defined as:

$$g'(x) = N'_g exp\left(-4ln2\,\frac{(x-x_{i0})^2}{(\gamma(x))^2}\right) \quad \text{(EQSI.2)}$$

$$\ell'(x) = N'_\ell \frac{1}{1+4\,\frac{(x-x_0)^2}{(\gamma(x))^2}} \quad \text{(EQSI.3)}$$

The functions $g'(x)$ and $\ell'(x)$ are separately normalized to $\frac{N_A e^2 f}{4m_e c^2 \epsilon_0 \ln(10)}$ by imposing that:

$$N'_g = \left(\frac{N_A e^2 f}{4m_e c^2 \epsilon_0 \ln(10)}\right) \Big/ \left(\int_{-\infty}^{\infty} exp\left(-4ln2\,\frac{(x-x_{i0})^2}{(\gamma(x))^2}\right) dx\right) \quad \text{(EQSI.4)}$$

$$N'_\ell = \left(\frac{N_A e^2 f}{4m_e c^2 \epsilon_0 \ln(10)}\right) \Big/ \left(\int_{-\infty}^{\infty} \frac{1}{1+4\,\frac{(x-x_0)^2}{(\gamma(x))^2}}\, dx\right) \quad \text{(EQSI.4)}$$

The two integrals appearing in EQSI.4 and EQSI.5 have been computed numerically. Using such normalization factors, the molar absorptivity ($\varepsilon(x)$) expressed in L.mol^{-1}cm^{-1} is indeed directly obtained.

Figure:

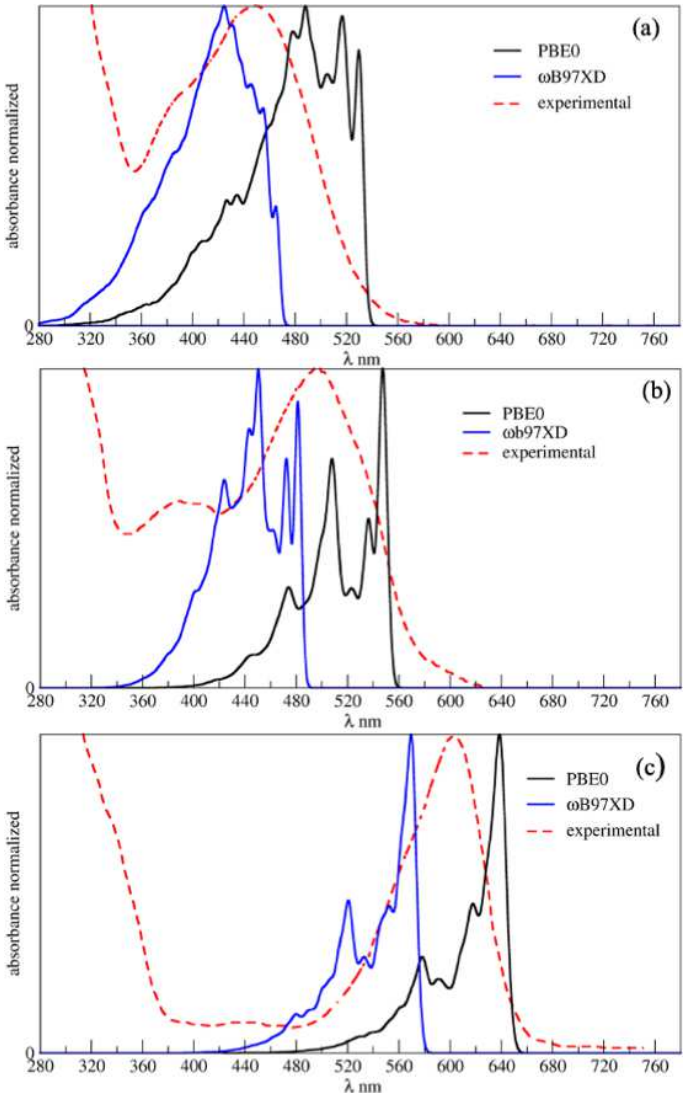


Figure SI3.1 Experimental and FC spectra for: (a) AQ1, (b) AQ2, (c) IN1.

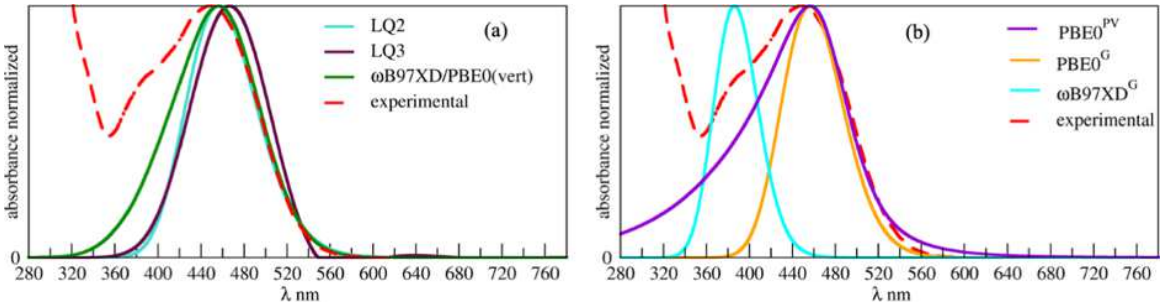


Figure SI.2. Experimental and computed spectra for AQ1 molecule. Panel (a): ωB97XD/PBE0(vert), LQ2, LQ3. Panel right (b): $PBE0^{PV}$, $PBE0^{G}$, $\omega B97XD^{G}$ protocols.

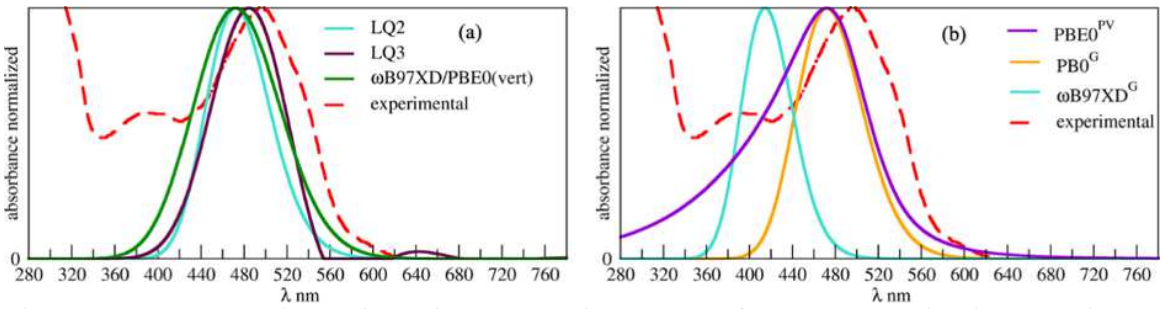


Figure SI.3.2 Experimental and computed spectra for AQ2 molecule. Panel (a): ωB97XD/PBE0(vert), LQ2, LQ3. Panel (b): $PBE0^{PV}$, $PBE0^{G}$, $\omega B97XD^{G}$ protocols.

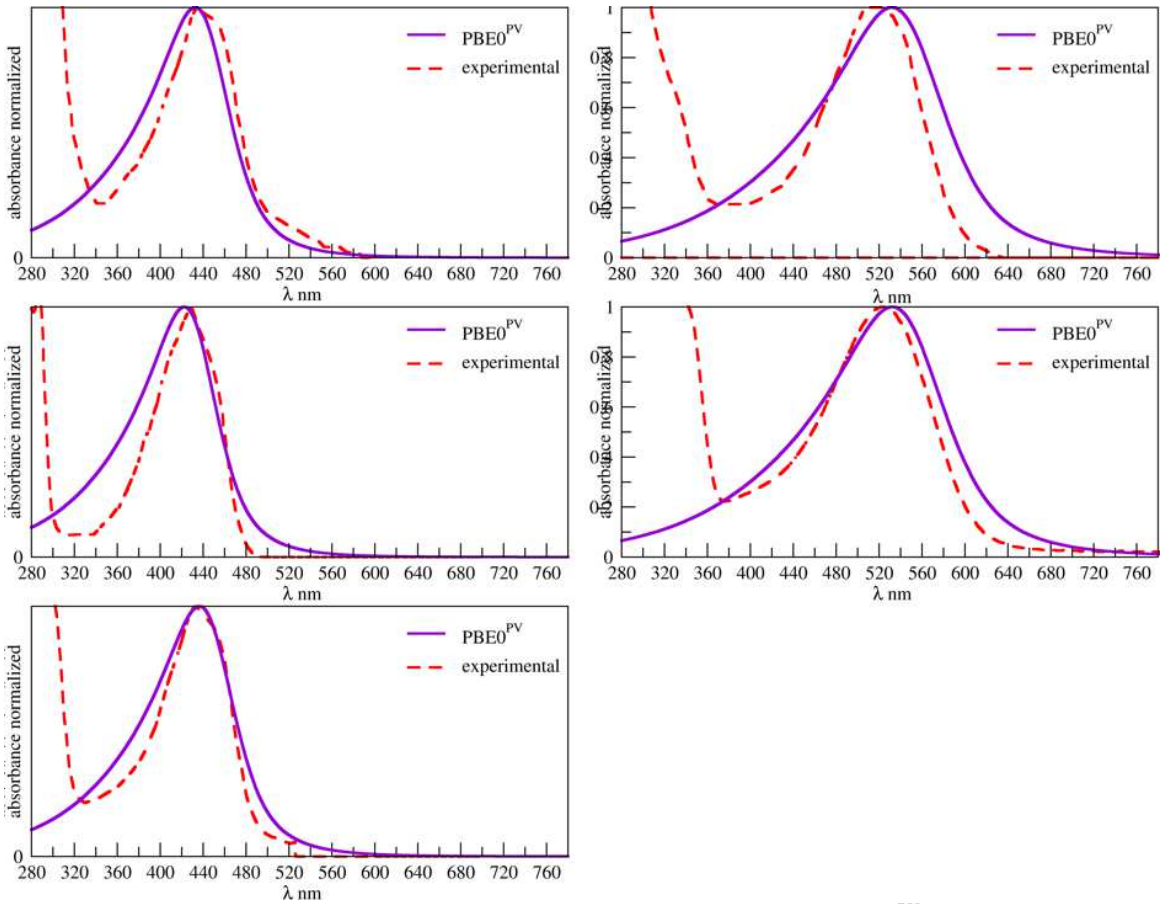


Figure SI3.3 Experimental and computed spectra with the protocol $PBE0^{PV}$; (a) AQ6, (b) AQ7, (c) AQ8, (d) AQ4 and (e) AQ5.

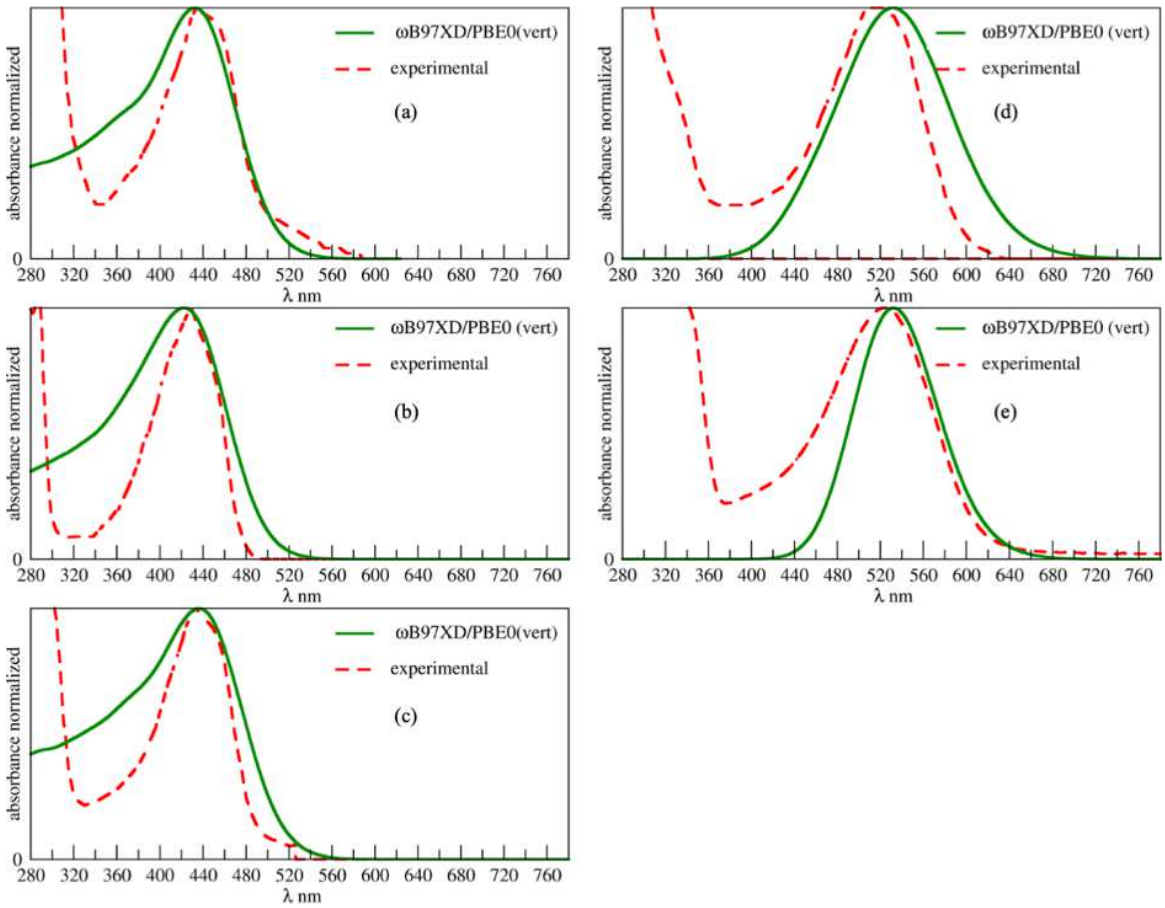


Figure SI3.4 Experimental and computed spectra with the protocol ωB97XD/PBE0(vert): (a) AQ6, (b) AQ7, (c) AQ8, (d) AQ4 and (e) AQ5.

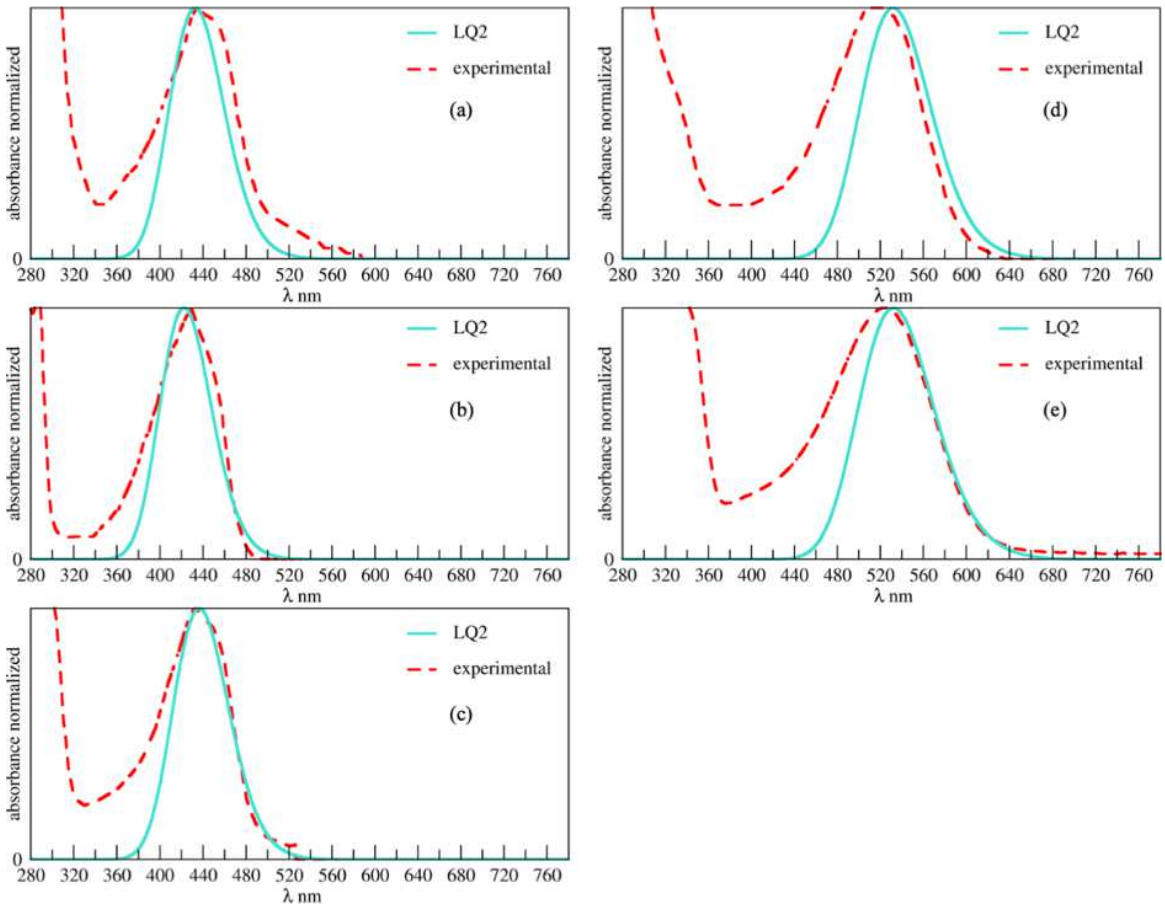


Figure SI3.5 Experimental and computed spectra with the protocol LQ2: (a) AQ6, (b) AQ7, (c) AQ8, (d) AQ4 and (e) AQ5.

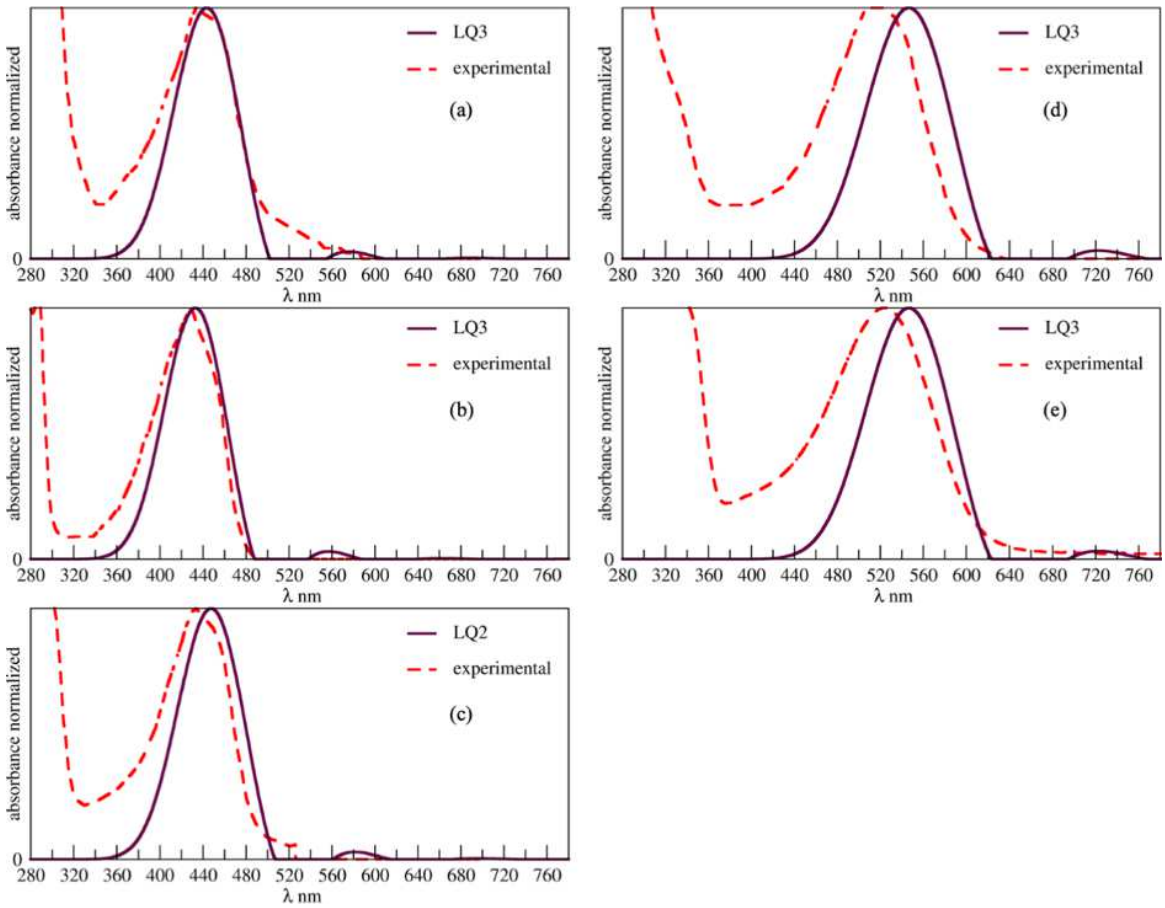


Figure SI3.6 Experimental and computed spectra with the protocol LQ3: (a) AQ6, (b) AQ7, (c) AQ8, (d) AQ4 and (e) AQ5.

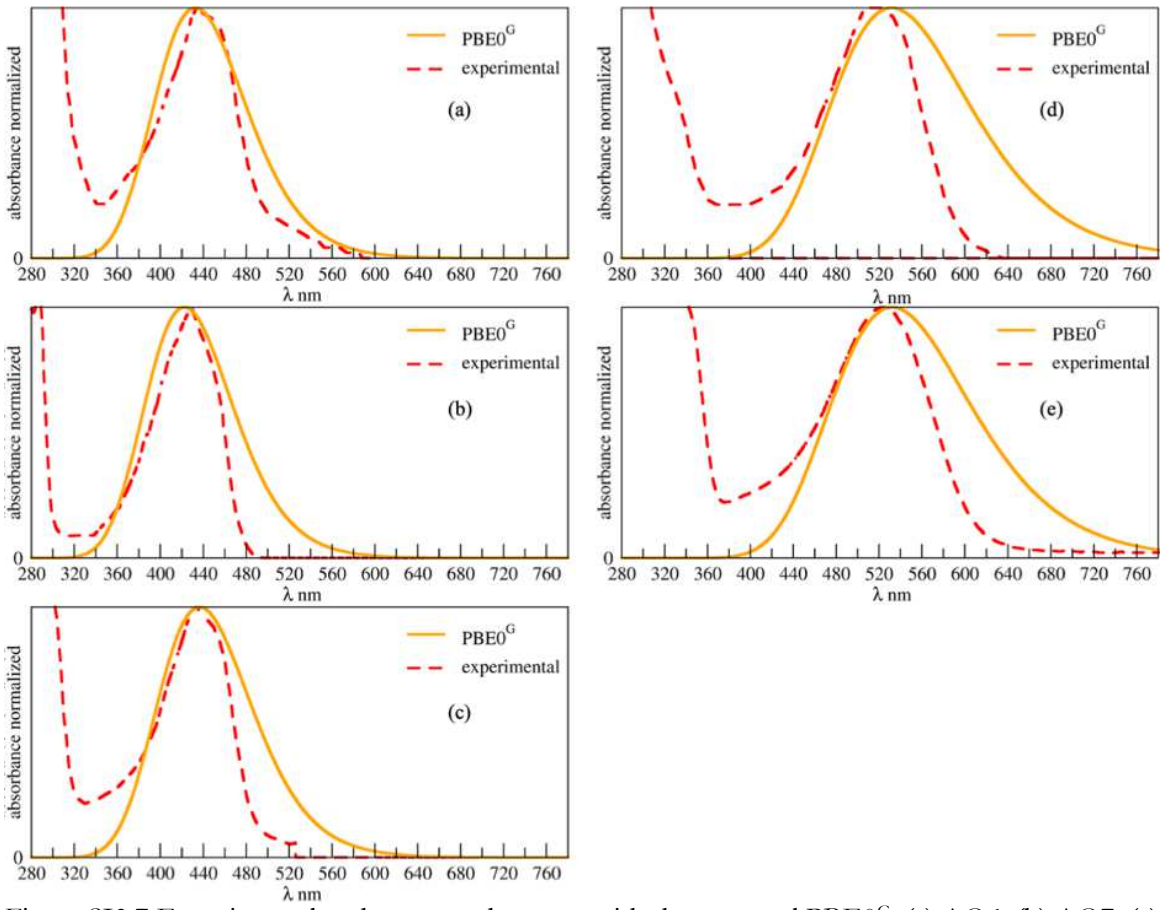


Figure SI3.7 Experimental and computed spectra with the protocol PBE0^{G}: (a) AQ6, (b) AQ7, (c) AQ8, (d) AQ4 and (e) AQ5.

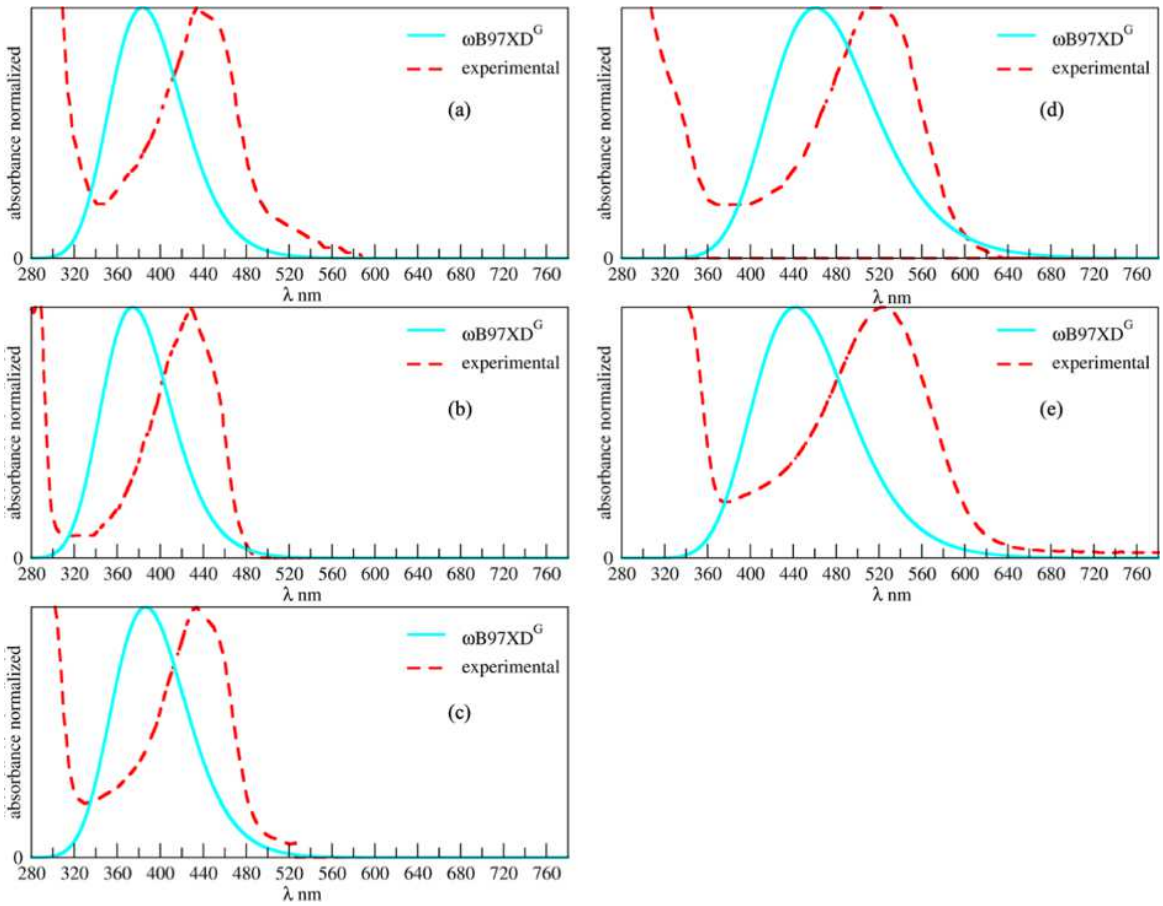


Figure SI3.8 Experimental and computed spectra with the protocol ωB97XD^{G}; (a) AQ6, (b) AQ7, (c) AQ8, (d) AQ4 and (e) AQ5.

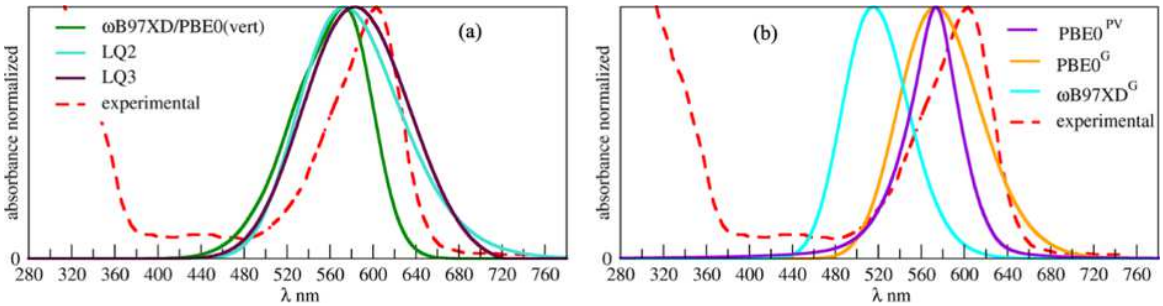


Figure SI3.9 Experimental and computed spectra for IN1 molecule. Panel left side: (a) ωB97XD/PBE0(vert), LQ2, LQ3. Panel right side: (b) PBE0^{PV}, PBE0^{G}, ωB97XD^{G} protocols.

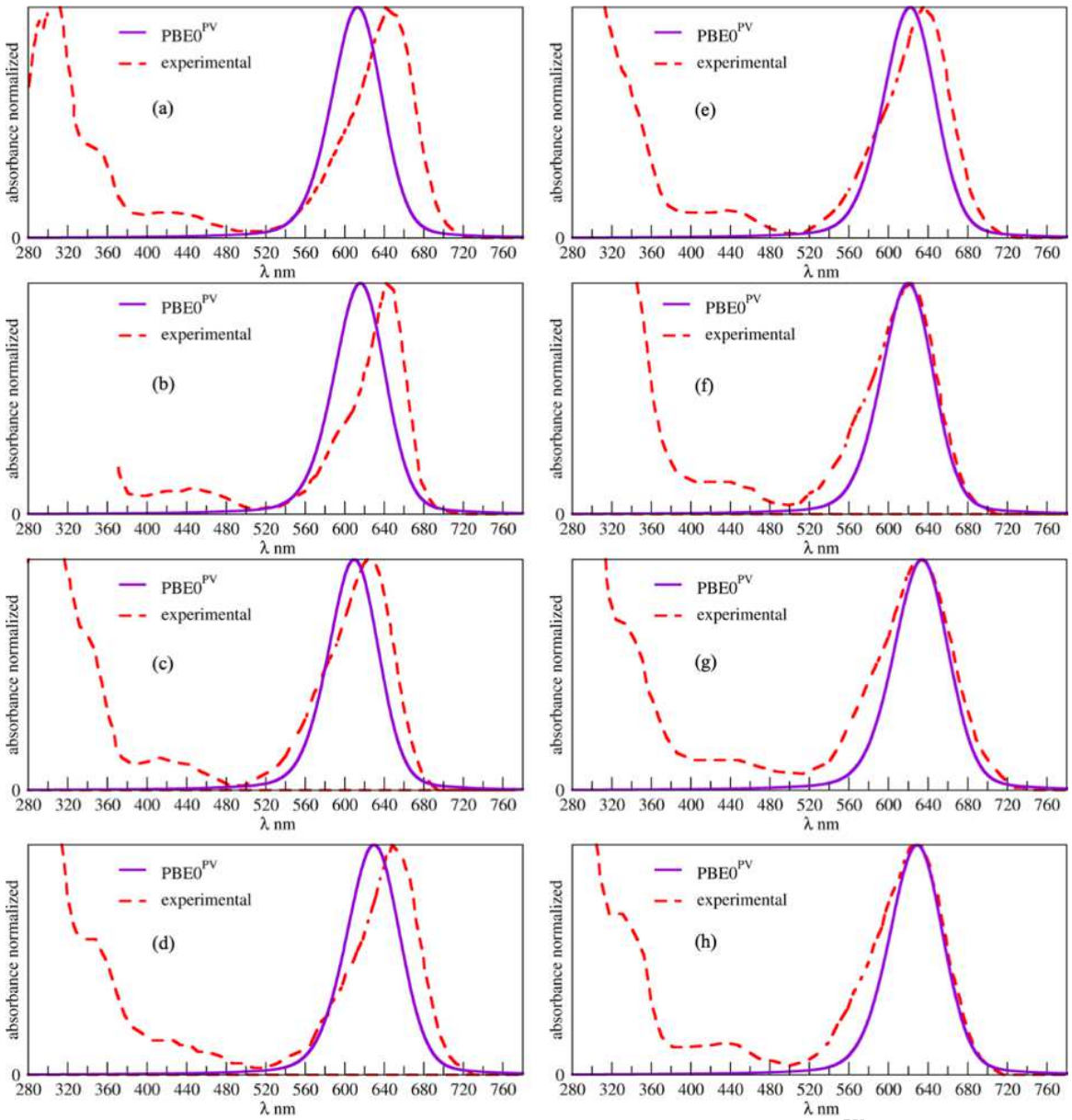


Figure SI3.10 Experimental and computed spectra with the protocol PBE0^{PV}; (a) IN2, (b) IN3, (c)IN04,(d) IN05 and (e) IN06, (f) IN07, (g) IN08, (h)IN09.

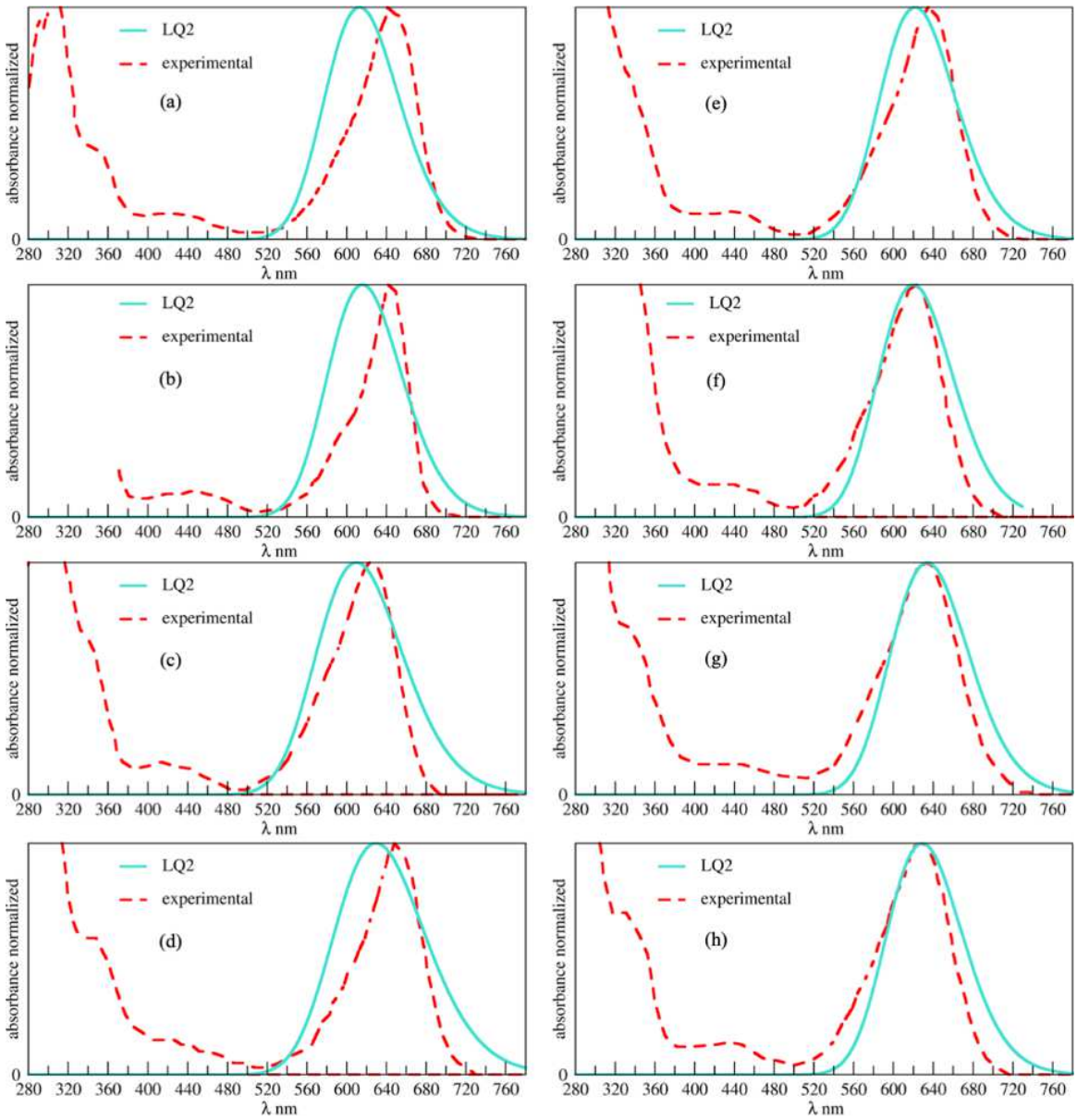


Figure SI3.11 Experimental and computed spectra with the protocol LQ2; (a) IN2, (b) IN3, (c)IN04,(d) IN05 and (e) IN06, (f) IN07, (g) IN08, (h)IN09.

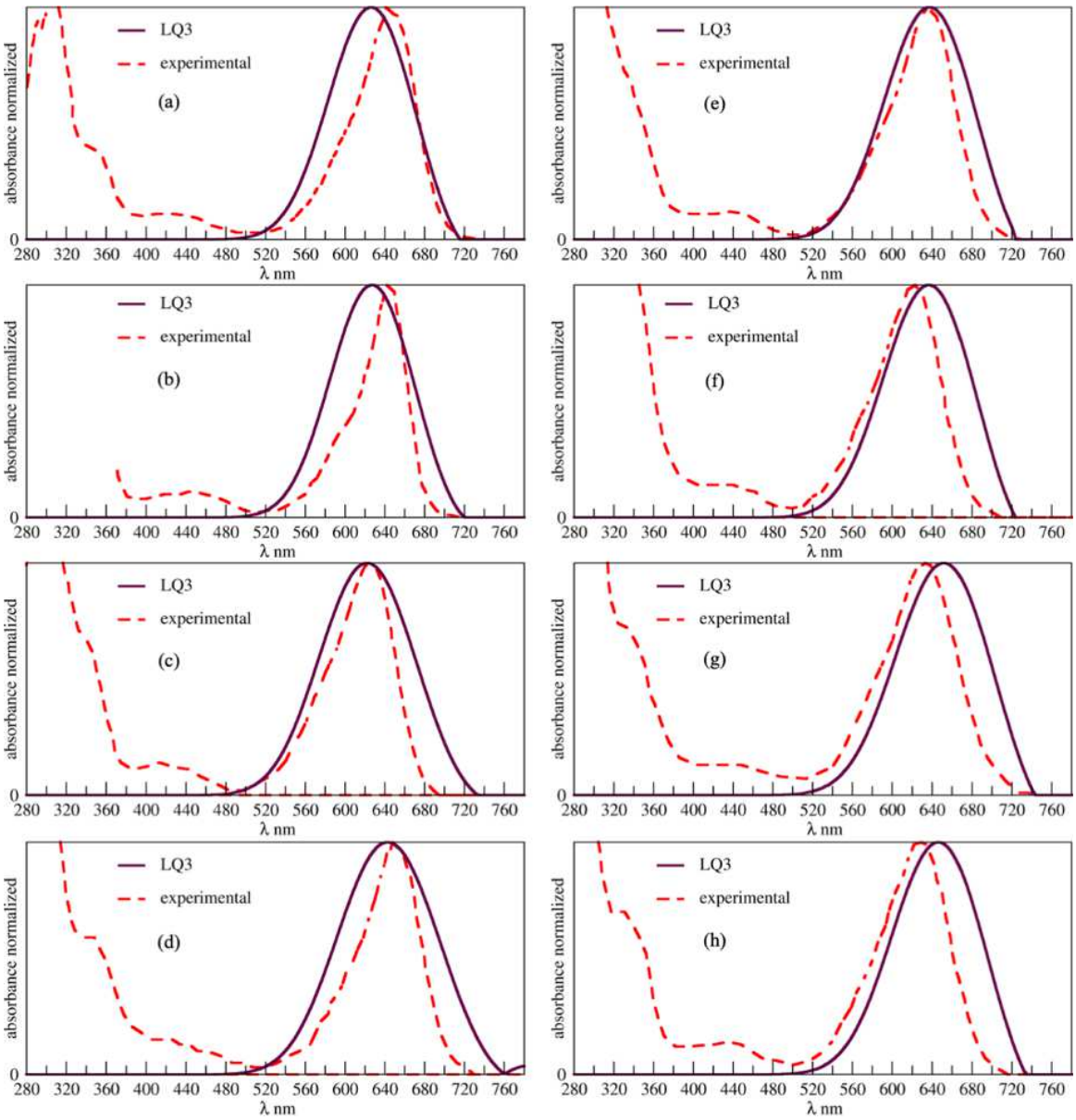


Figure SI3.12 Experimental and computed spectra with the protocol LQ3; (a) IN2, (b) IN3, (c)IN04,(d) IN05 and (e) IN06, (f) IN07, (g) IN08, (h)IN09.

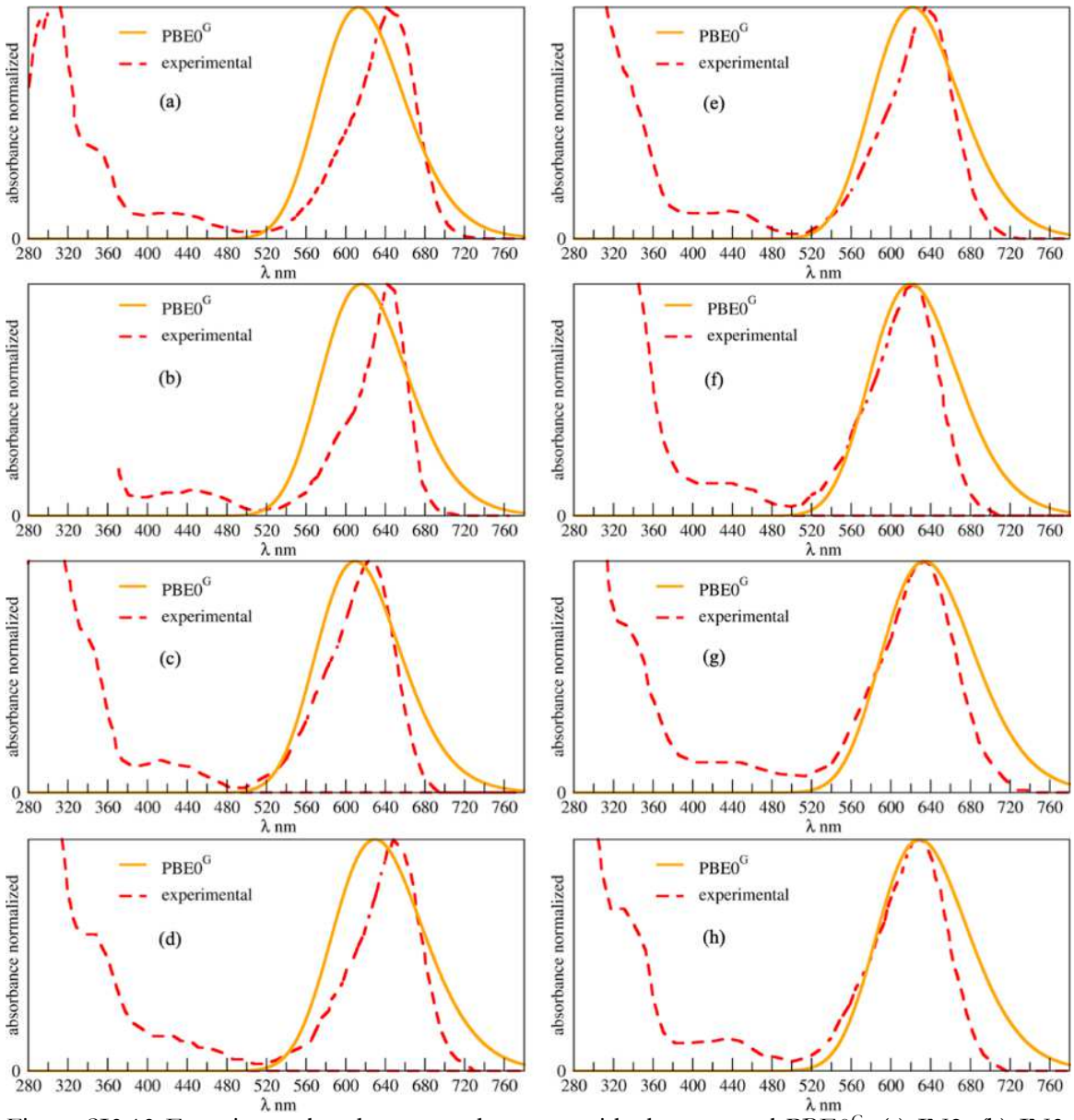


Figure SI3.13 Experimental and computed spectra with the protocol PBE0^{G}; (a) IN2, (b) IN3, (c)IN04,(d) IN05 and (e) IN06, (f) IN07, (g) IN08, (h)IN09.

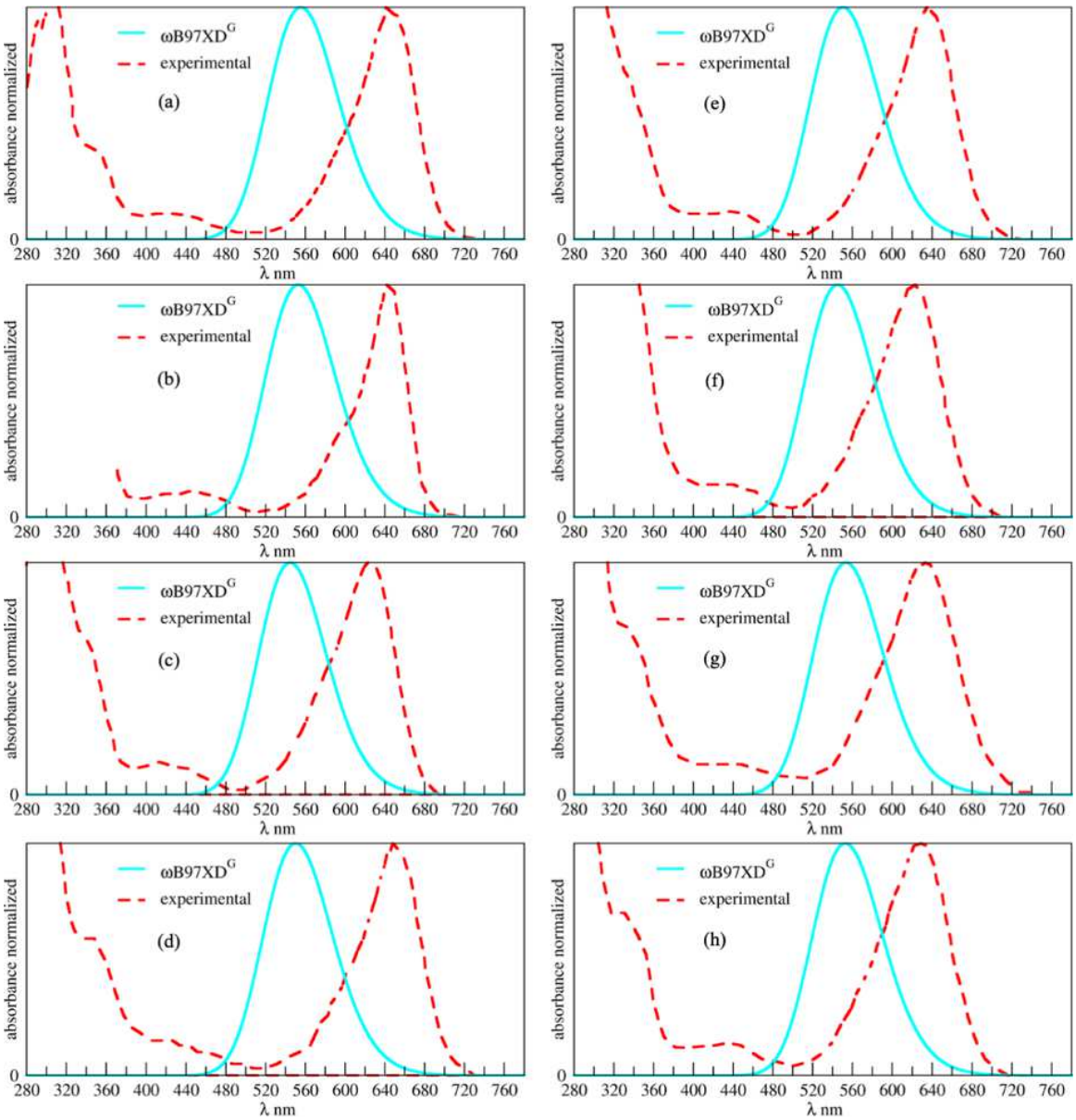


Figure SI3.14 Experimental and computed spectra with the protocol ωB97XD^{G}; (a) IN2, (b) IN3, (c)IN04,(d) IN05 and (e) IN06, (f) IN07, (g) IN08, (h)IN09.

7.2 Appendix B

Table SI3.1. CIE colorimetric coordinates (L*, a*, b*) parameters. Euclidian distance between the computed and experimental color (d_{eucl}), metric hue angle (h) and chromas (C) of selected anthraquinones, derived from experimental and simulated spectra. ωB97XD/PBE0(vert): fully vibronic coupling. LQ2 and LQ3: approximate vibronic models. $PBE0^{G}$. $\omega B97XD^{G}$: Gaussian convolution function; $PBE0^{PV}$: pseudo-Voigt convolution.

	param	Exp.	ωB97XD/PBE0(vert)	$PBE0^{PV}$	LQ2	LQ3	$PBE0^{G}$	$\omega B97XD^{G}$
AQ4	L^*	67.00	55.04	67.16	67.16	63.58	62.18	90.85
	a^*	60.46	54.90	64.99	64.99	63.73	64.76	8.19
	b^*	0.30	-32.41	-47.88	-47.88	-50.80	-50.98	64.69
	C	60.46	63.75	80.72	80.72	81.50	82.42	65.20
	h	0.28	329.44	323.62	323.62	321.44	321.79	82.79
	d_{eucl}		35.27	48.38	48.38	51.31	51.68	86.30
AQ5	L^*	61.59	62.61	65.26	65.26	61.21	62.08	95.50
	a^*	56.67	65.42	65.05	65.05	61.56	64.57	-9.15
	b^*	-6.33	-47.82	-49.59	-49.59	-52.97	-51.35	73.11
	C	57.02	81.03	81.79	81.79	81.21	82.50	73.68
	h	353.63	323.83	322.68	322.68	319.29	321.51	97.14
	d_{eucl}		42.42	44.22	44.22	46.90	45.72	108.60
AQ6	L^*	92.52	94.75	97.10	97.10	96.51	97.01	99.86
	a^*	-3.16	-24.72	-13.91	-13.91	-16.08	-13.67	-3.59
	b^*	71.45	72.54	65.81	65.81	74.68	66.24	7.60
	C	71.52	76.63	67.27	67.27	76.39	67.64	8.40
	h	92.53	108.82	101.94	101.94	102.15	-78.34	115.29
	d_{eucl}		21.70	12.98	12.98	13.91	12.56	64.27
AQ7	L^*	98.59	95.88	98.28	98.28	97.58	98.17	99.94
	a^*	-18.02	-47.30	-15.35	-15.35	-17.79	-15.21	-1.54
	b^*	56.85	62.72	52.32	52.32	62.37	53.45	3.37
	C	59.64	78.55	54.52	54.52	64.86	55.58	3.71
	h	107.59	127.03	106.35	106.35	105.92	105.89	114.62
	d_{eucl}		29.99	5.27	5.27	5.62	4.42	55.97

AQ8	L^*	95.79	93.62	96.60	96.60	96.02	96.44	99.82
	a^*	-10.49	-13.92	-12.62	-12.62	-14.38	-12.16	-4.46
	b^*	69.88	76.73	69.27	69.27	76.72	69.85	9.52
	C	70.66	77.98	70.41	70.41	78.06	70.90	10.52
	h	98.53	100.28	100.32	100.32	100.62	-80.13	115.12
	d_{eucl}		7.96	2.36	2.36	7.88	1.79	60.79

Table SI3.2. CIE colorimetric coordinates (L*, a*, b*) parameters. Euclidian distance between the computed and experimental color (d_{eucl}), metric hue angle (h) and chromas (C) of selected indigos, derived from experimental and simulated spectra. LQ2 and LQ3: approximate vibronic models. PBE0$^{G.}$ ωB97XD^{G}: Gaussian convolution function; PBE0^{PV}: pseudo-Voigt convolution.

	param	exp	PBE0^{PV}	LQ2	LQ3	PBE0^{G}	ωB97XD^{G}
IN2	L^*	86.50	84.96	81.61	81.00	87.92	77.06
	a^*	-37.54	-38.50	-42.80	-39.24	-41.95	28.35
	b^*	-10.92	-24.83	-31.63	-32.55	-20.80	-39.07
	C	39.10	45.81	53.22	50.98	46.83	48.27
	h	196.22	212.82	216.46	219.68	206.37	305.96
	d_{eucl}		14.03	21.92	22.38	10.91	72.27
IN3	L^*	88.00	85.67	82.64	82.14	88.62	77.11
	a^*	-34.93	-38.83	-43.47	-40.29	-41.65	32.57
	b^*	-6.77	-23.64	-29.87	-30.64	-19.59	-38.89
	C	35.58	45.46	52.74	50.62	46.03	50.72
	h	190.97	211.33	214.49	217.25	205.19	309.94
	d_{eucl}		17.47	25.20	25.16	14.49	75.54
IN4	L^*	79.73	84.14	77.23	76.98	87.10	77.76
	a^*	-40.84	-37.69	-37.63	-34.66	-41.90	43.13
	b^*	-22.36	-26.21	-39.05	-39.20	-22.22	-37.42
	C	46.56	45.91	54.23	52.32	47.42	57.10
	h	208.70	214.82	226.06	228.52	207.94	319.06
	d_{eucl}		6.65	17.18	18.15	7.45	85.33

IN5	L^*	86.93	88.87	83.53	83.05	91.74	77.24
	a^*	-36.07	-36.64	-42.36	-37.84	-36.78	36.28
	b^*	-6.56	-18.20	-28.31	-29.03	-14.23	-38.57
	C	36.66	40.92	50.95	47.70	39.44	52.95
	h	190.30	206.42	213.76	217.49	201.15	313.24
	d_{eucl}		11.82	22.90	22.87	9.08	79.71
IN6	L^*	83.56	87.11	84.53	83.70	90.04	77.23
	a^*	-41.57	-38.55	-43.83	-39.17	-40.12	35.97
	b^*	-15.35	-21.19	-26.63	-27.96	-17.15	-38.61
	C	44.32	43.99	51.28	48.12	43.63	52.77
	h	200.27	208.80	211.28	215.52	203.14	312.98
	d_{eucl}		7.47	11.54	12.83	6.87	81.20
IN7	L^*	78.26	86.63	84.11	83.17	89.57	77.23
	a^*	-41.22	-38.78	-43.89	-38.93	-40.75	35.97
	b^*	-21.09	-22.00	-27.34	-28.85	-17.95	-38.61
	C	46.31	44.59	51.71	48.45	44.53	52.77
	h	207.09	209.57	211.92	216.54	203.77	312.98
	d_{eucl}		8.77	8.96	9.46	11.75	79.16
IN8	L^*	80.00	89.75	87.82	86.74	92.56	77.09
	a^*	-40.51	-35.12	-41.55	-36.70	-34.59	31.59
	b^*	-16.81	-16.71	-20.96	-22.76	-12.80	-38.95
	C	43.85	38.90	46.54	43.19	36.88	50.14
	h	202.53	205.45	206.77	211.81	200.30	309.04
	d_{eucl}		11.14	8.92	9.77	14.45	75.47
IN9	L^*	79.45	88.70	86.92	85.88	91.57	77.11
	a^*	-41.98	-36.90	-42.75	-37.87	-37.18	32.57
	b^*	-19.67	-18.50	-22.52	-24.25	-14.51	-38.89
	C	46.36	41.28	48.31	44.97	39.91	50.72
	h	205.11	206.62	207.78	212.63	201.32	309.94
	d_{eucl}		10.61	8.03	8.90	14.02	77.02

7.3 Appendix C

Table SI3.14 AQ Colors from the experimental and theoretical spectra, evaluated using different computational protocol: PBE0^{vib}, ωB97XD^{vib} and ωB97XD/PBE0(vert): fully vibronic coupling; LQ2 and LQ2: approximate vibronic models; PBE0^{G} and wB97XD^{G}: Gaussian convolution; PBE0^{PV}: pseudo-Voigt convolution, **X** no possible to compute vibronic spectra.

	exp	PBE0^{PV}	LQ2	LQ3	ωB97XD/PBE0 (vert)	PBE0^{G}	ωB97XD^{G}	PBE0^{vib}	ωB97XD^{vib}
AQ1									
AQ2									
AQ3									
AQ4								X	X
AQ5								X	X
AQ6								X	X
AQ7								X	X
AQ8								X	X

Table SI3.15 IN Colors from the experimental and theoretical spectra, evaluated using different computational protocol: PBE0^{vib}, ωB97XD^{vib} and ωB97XD/PBE0(vert): fully vibronic coupling; LQ2 and LQ2: approximate vibronic models; PBE0^{G} and wB97XD^{G}: Gaussian convolution; PBE0^{PV}: pseudo-Voigt convolution, **X** no possible to compute vibronic spectra.

	exp	PBE0^{PV}	LQ2	LQ3	ωB97XD/PBE0(vert)	PBE0^{G}	ωB97XD^{G}	PBE0^{vib}	ωB97XD^{vib}
IN1									
IN2					X			X	X
IN3					X			X	X
IN4					X			X	X
IN5					X			X	X
IN6					X			X	X
IN7					X			X	X
IN8					X			X	X
IN9					X			X	X
IN10									

7.4 Appendix D

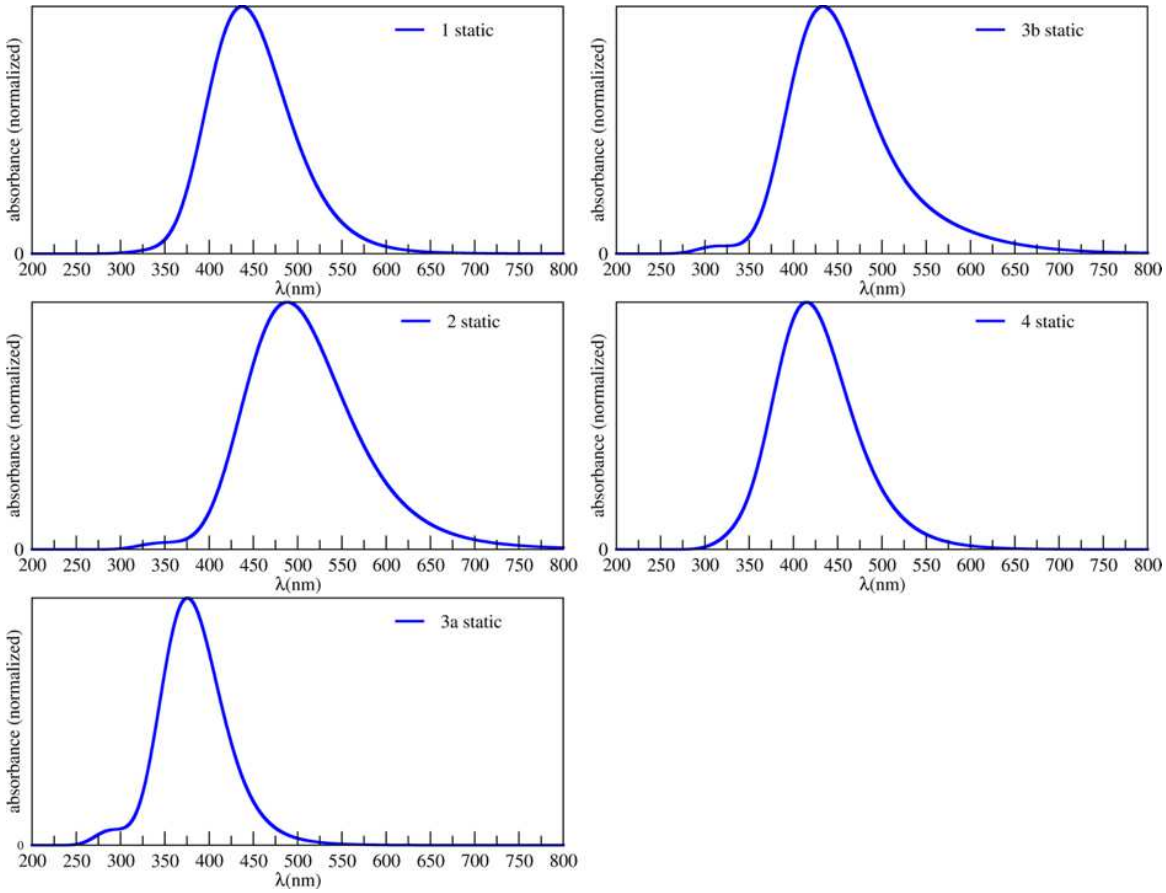


Figure SI4.1 computed spectra for merocyanine dimers

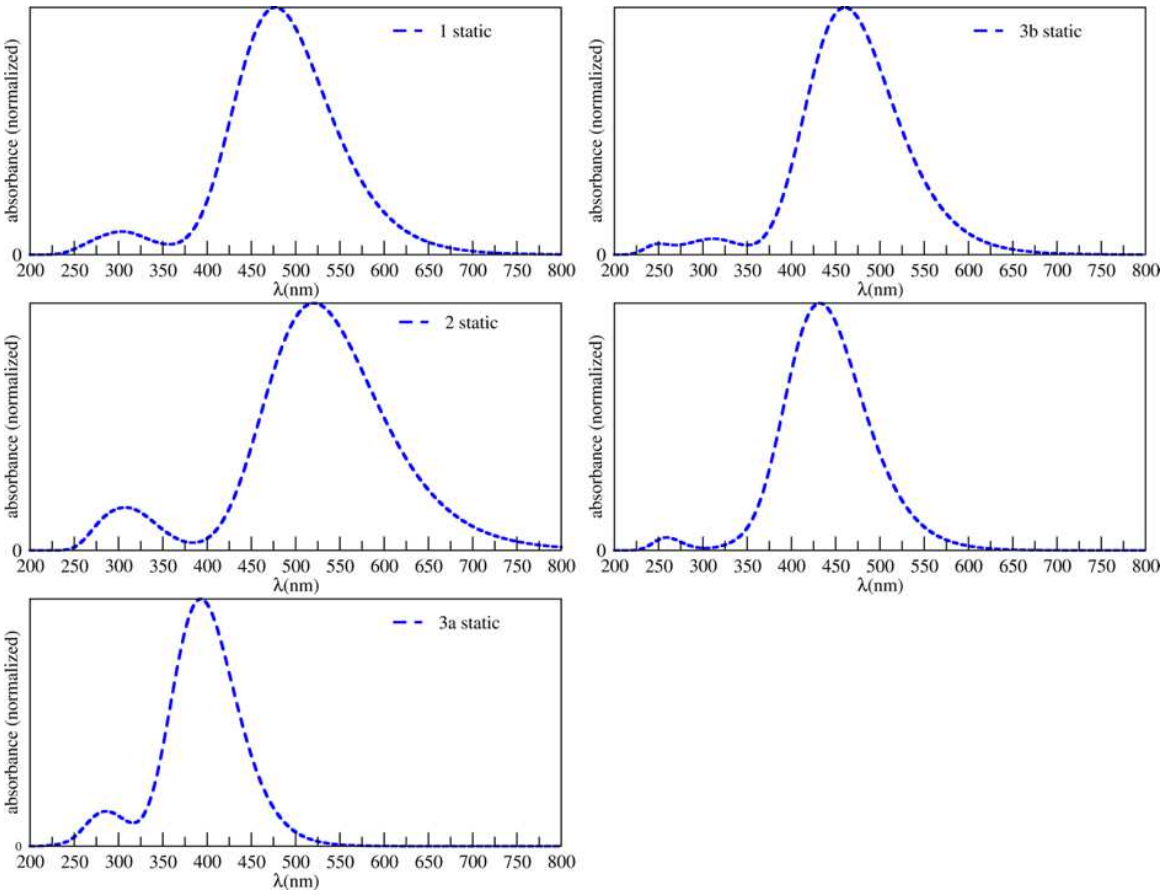


Figure SI4.2 computed spectra for merocyanine monomers

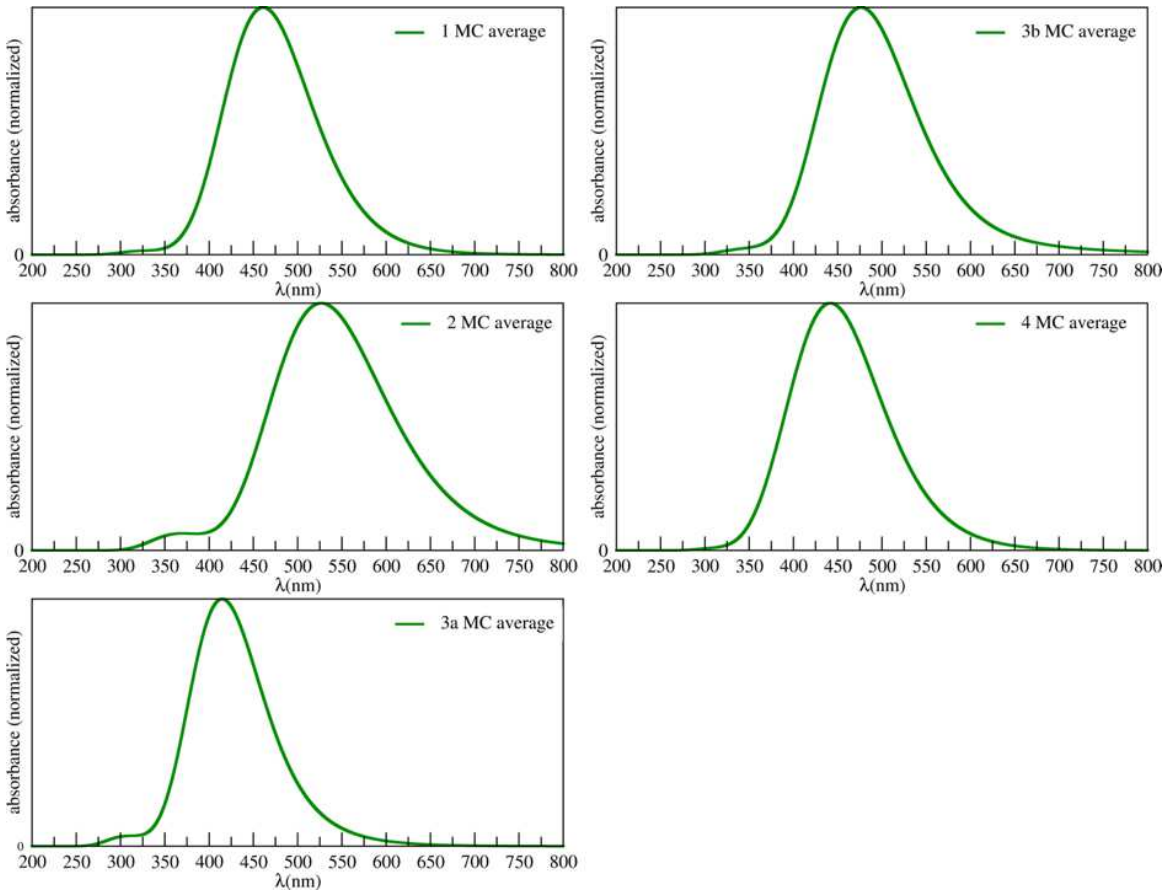


Figure SI4.3 computed spectra for merocyanine dimers

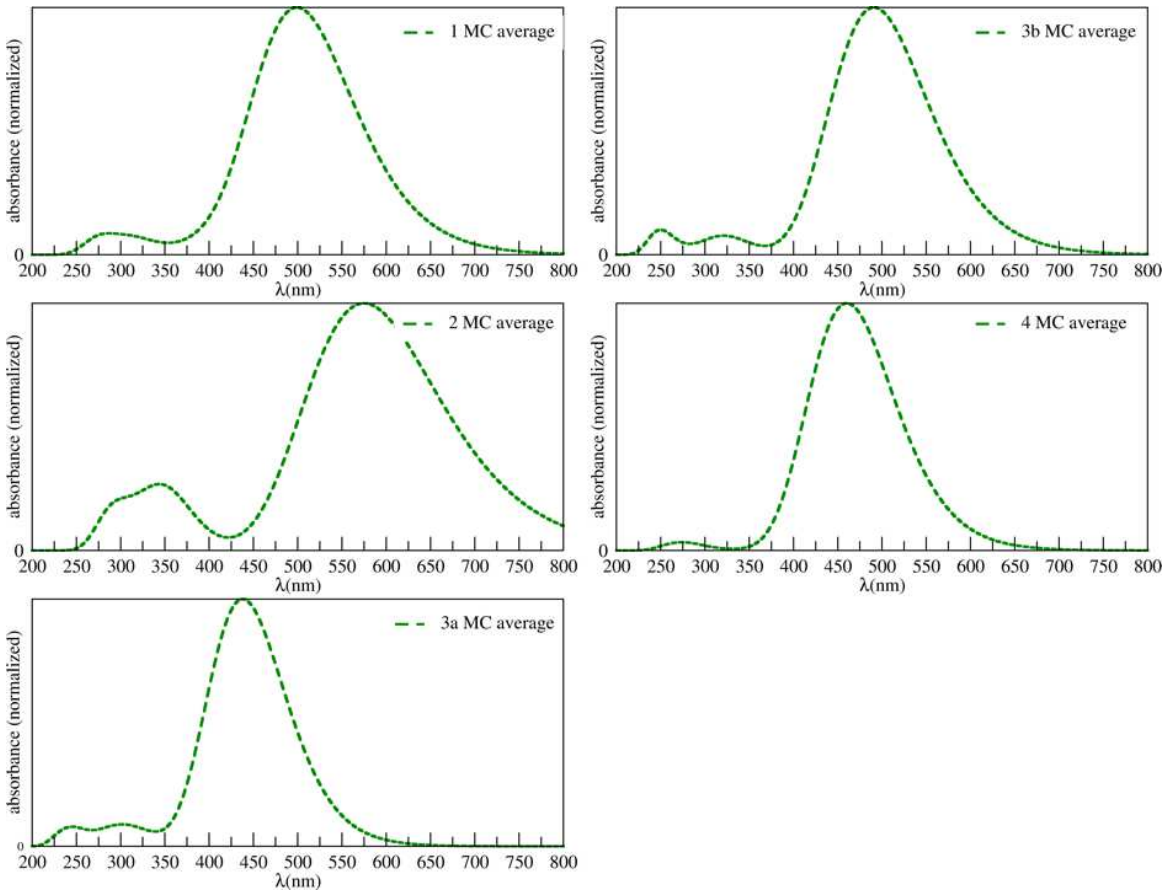


Figure SI4.4 computed spectra for merocyanine monomers

Chapter 8

8. Résumé en français

Les colorants sont attrayants en raison de leurs propriétés électroniques modulables, ce qui leur permet d'avoir de vastes applications allant de la conception de technologies de pointe aux industries textiles. Selon les sources de production, les colorants peuvent être classés comme naturels ou synthétiques [1]. Dérivés de plantes et de micro-organismes, les colorants naturels sont perçus comme inoffensifs pour l'environnement, mais ils sont chers et peuvent être toxiques. D'autre part, plusieurs colorants synthétiques sont dérivés de composés pétrochimiques [2] et peuvent être produits en grande quantité. Cependant, leur impact sur les êtres humains et l'environnement ne peut être négligé. Une question reste ouverte : comment faire le bon choix ?
L'industrie a récemment pris des mesures pour réduire l'impact environnemental de la production de teintures, notamment en s'approvisionnant en matières premières renouvelables. La mise au point de colorants plus écologiques est une entreprise opportune qui pourrait avoir un impact considérable, notamment sur l'industrie des colorants textiles. En passant de la synthèse pétrochimique à la biotechnologie renouvelable, la production pourrait permettre de diviser par 100 la consommation d'énergie et d'éviter 100 kilogrammes d'huiles lourdes et 800 litres d'eau par kilogramme de teinture [3]. Par rapport aux pigments pétrochimiques ordinaires, les colorants biosourcés divisent par neuf l'empreinte carbone. Appliquée à l'industrie des teintures, qui pèse des millions de tonnes, cette transition a le même potentiel de réduction des gaz à effet de serre que la réduction des émissions annuelles d'un million de véhicules. Une stratégie verte possible consiste à utiliser un procédé enzymatique [4]. Aujourd'hui, plusieurs micro-organismes sont connus pour fabriquer des couleurs allant de l'orange (blakeslea trispora) au violet (Chromobaterium violaceum), en passant par le rouge (Serratia marcesccens) [5]. C'est l'objectif de la société PILI. Ici, les composés bio-sourcés sont modifiés en utilisant des processus de chimie verte pour atteindre une diversité dans la palette de couleurs. À cet égard, les schémas de catégorisation sont utiles pour distinguer les couleurs [6]. Bien qu'ils soient utilisés à des fins différentes [7], ces schémas sont fortement liés au spectre d'absorption électronique. En bref, lorsqu'une molécule absorbe un groupe de radiations visibles, elle apparaît colorée et sa couleur

est due à la résultante de toutes les radiations transmises. Le spectre transmis enregistré est lié à celui de l'absorption, et donc les couleurs "visibles" et complémentaires sont également liées entre elles. Ces couleurs doivent être caractérisées, car les êtres humains ont des perceptions différentes des couleurs. Pour ce faire, les distances euclidiennes entre les couleurs, la chroma (c'est-à-dire le degré de saturation) et l'angle de teinte (c'est-à-dire le ton de la couleur) constituent les éléments clés pour distinguer les couleurs et permettent aux industries des colorants de classer les chromophores.

En tenant compte de l'impact sur l'environnement, les industries des colorants doivent élargir leur spectre d'intérêt en rendant plus efficace l'utilisation de leurs recherches, y compris l'optimisation des processus. Pour accélérer la chaîne industrielle, on peut étudier les processus photophysiques (processus physique induit par la lumière) et photochimiques (la lumière induit des transformations chimiques) afin de concevoir de nouveaux chromophores ayant la couleur souhaitée. Lors des processus induits par la lumière, les molécules laissent des signatures spectroscopiques qui sont liées à leur structure électronique moléculaire et à sa perturbation. Dans ce cadre, les méthodes visant à résoudre le problème de la structure électronique sont utiles à la fois pour comprendre et améliorer la synthèse d'un colorant donné.

Au fil des années, les méthodes de premiers principes se sont révélées être des outils valables, et il n'est pas surprenant qu'elles soient couramment utilisées pour le criblage moléculaire et la conception in silico de nouveaux chromophores d'intérêt industriel. Parmi ces méthodes, la théorie de la fonctionnelle de la densité (DFT) et son extension en fonction du temps (TD-DFT) sont actuellement utilisées pour étudier les propriétés des états fondamental et excité de ces systèmes. Au cours des dernières décennies, la TD-DFT a largement démontré sa capacité à fournir d'excellentes solutions concernant les états excités pour plusieurs classes de composés organiques commerciaux d'intérêt industriel, y compris les anthraquinones et les indigos [8,9,10].

Outre le niveau de théorie utilisé pour obtenir les énergies d'absorption, le spectre électronique est affecté par deux autres facteurs : l'environnement et le réarrangement 3D du chromophore. Une modélisation précise de ces facteurs permet de prédire les spectres Uv-vis avec une excellente précision. En ce qui concerne l'environnement chimique, l'inclusion du solvant s'avère décisive pour les énergies de transition électronique, en particulier lorsque le solvant polaire est pris en compte. Le solvant a été modélisé en conjonction avec les simulations DFT au moyen d'un modèle de continuum [11]. Enfin, une transition électronique est modulée par un léger changement de la configuration nucléaire de l'état fondamental à l'état excité. Plus précisément, le couplage entre les degrés de liberté électronique et nucléaire (le couplage vibratoire) influence profondément l'élargissement de la bande. Une approche pour calculer la structure fine vibronique consiste à

utiliser l'approximation de Franck-Condon (FC) qui entraîne un coût de calcul important. Malgré cet inconvénient, les études informatiques ont montré que la structure fine vibronique ne peut pas être négligée dans la prédiction quantitative des couleurs [12]. Cependant, comme les étapes chronophages ne peuvent pas être ignorées, de nouveaux protocoles informatiques doivent être développés.

Alors que pour les chromophores rigides (c'est-à-dire les systèmes caractérisés par une flexibilité moléculaire minuscule), la progression vibratoire peut être prédite avec précision en utilisant l'approximation FC, les chromophores flexibles sont plus complexes à modéliser. Pour ces systèmes, les géométries de l'état fondamental et de l'état excité sont très différentes, et les surfaces d'énergies potentielles peuvent présenter plusieurs minima locaux proches les uns des autres, ce qui empêche l'utilisation de l'approximation FC et, par conséquent, la simulation de leurs spectres UV-vis représente un défi actuel du point de vue informatique. Pour résoudre ce problème complexe, deux approches possibles, qui tournent autour de l'échantillonnage de l'espace de phase, ont été développées. Les simulations de dynamique moléculaire (MD), classiques ou au niveau ab initio [13], sont couramment utilisées pour échantillonner l'espace de phase, ici, sur la trajectoire, le spectre d'absorption est obtenu en extrayant à la volée certaines configurations moléculaires. A la fin, le spectre final est obtenu comme une moyenne de tous les spectres calculés. Comme les simulations MD s'exécutent sur des durées trop courtes (généralement de l'ordre de la nanoseconde), une petite zone de phase est explorée. Par conséquent, les propriétés ne sont pas correctement reproduites le long de la trajectoire dynamique. Malgré cet inconvénient, les simulations MD sont largement appliquées. Une autre stratégie possible pour échantillonner l'espace des phases est basée sur son échantillonnage aléatoire. C'est l'idée derrière la méthode de Monte Carlo (MC), qui permet d'explorer un immense espace de phase [14]. Comme pour les simulations MD, la méthode MC présente également des inconvénients. Par exemple, la génération d'une structure moléculaire peut être traumatisante lorsque de grandes étapes de perturbation sont prises en compte, provoquant des collisions d'atomes ou des structures (aromatiques) trop tendues. Malgré les problèmes liés à la génération de structures potentiellement non physiques, les simulations MC nécessitent moins de temps que les simulations MD. L'objectif est de surmonter les défis combinés de la précision et du temps de calcul. Il est donc essentiel d'élaborer des méthodes simples capables de reproduire facilement le spectre d'absorption pour des molécules caractérisées par un certain degré de flexibilité.

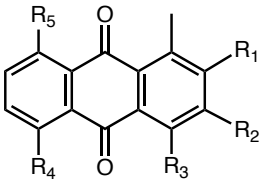
Dans ce contexte général, la société PILI s'intéresse aux approches permettant d'obtenir les signatures spectroscopiques moléculaires qui vont de la prédiction de la couleur des colorants organiques sans métal au développement de nouveaux protocoles de calcul pour modéliser le

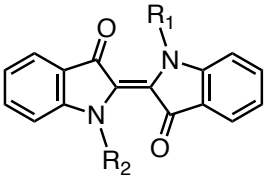
mécanisme de photoréaction. Mon projet s'inscrit dans ce cadre. L'objectif est de fournir et de tester de nouveaux outils de calcul à utiliser dans la simulation des colorants et de rationaliser la relation structure/propriété impliquant les énergies potentielles de l'état fondamental et de l'état excité.

Dans ce but, nous avons décidé d'étudier et de modéliser les phénomènes induits par la lumière dans le cadre de la TD-DFT. Avec cet outil, nous avons mis en place trois procédures de calcul rentables pour trois classes de chromophores. Les deux premières classes moléculaires, les anthraquinones et les indigos, sont actuellement utilisées dans l'industrie textile, car elles présentent une variété de couleurs. La troisième classe, les mérocyanines, est couramment employée dans différents domaines allant des peintures industrielles aux filtres solaires.

8.1 Prédiction des couleurs pour les colorants industriels

La première étude porte sur les deux premières classes moléculaires étudiées dans cette thèse : les anthraquinones et les indigo. Les premiers possèdent deux groupes C=O installés sur l'anneau central du squelette anthracénique et des groupes fonctionnels qui remplacent les atomes d'hydrogène du noyau aromatique (figure 8.1), tandis que les seconds chromophores possèdent deux structures dérivées de l'indole (ou du thioindole) reliées par une double liaison C=C et des groupes fonctionnels qui remplacent l'atome d'hydrogène du groupe amino (Figure 8.1).





AQ1 $R_1 = H; R_2 = NH_2; R_3 = H; R_4 = H; R_5 = OH$

AQ2 $R_1 = H; R_2 = OH; R_3 = NH_2; R_4 = H; R_5 = OH$

AQ3 $R_1 = H; R_2 = OH; R_3 = NH_2; R_4 = NH_2; R_5 = OH$

AQ4 $R_1 = COOH; R_2 = OH; R_3 = NH_2; R_4 = NH_2; R_5 = OH$

AQ5 $R_1 = H; R_2 = NHC_6H_{12}; R_3 = H; R_4 = H; R_5 = NHC_6H_{12}$

IN1 $R_1 = H; R_2 = H$

IN2 $R_1 = CH_3; R_2 = CH_3$

IN3 $R_1 = CH_2CH_2CH_3; R_2 = CH_2CH_2CH_3$

IN4 $R_1 = CO_2tBu; R_2 = CO_2tBu$

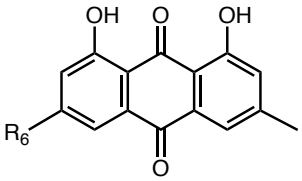
IN5 $R_1 = Bn; R_2 = Bn$

IN6 $R_1 = CH_3; R_2 = p\text{-}CF_3Bn$

IN7 $R_1 = CO_2tBu; R_2 = p\text{-}CNBn$

IN8 $R_1 = CO_2tBu; R_2 = p\text{-}OCH_3Bn$

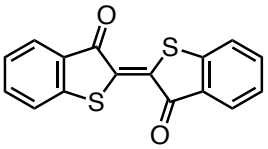
IN9 $R_1 = CO_2tBu; R_2 = p\text{-}OCH_3Bn$



AQ6 $R_6 = OH$

AQ7 $R_6 = H$

AQ8 $R_6 = OCH_3$



IN10

Figure 8.1 Représentation schématique des molécules étudiées dans cette thèse.

Afin de prédire la couleur de ces colorants naturels, nous avons effectué une investigation spectroscopique au niveau DFT/PCM/TD-DFT qui a fourni des spectres Uv-Vis. Nous avons testé deux fonctionnelles XC : la PBE0 et la ωB97XD. Nos calculs ont montré que la fonctionnelle PBE0 reproduit mieux les pics d'absorption expérimentaux que ωB97XD, par conséquent la fonctionnelle PBE0 a été choisie pour obtenir l'énergie d'excitation verticale tout au long de ce travail. Pour rendre compte de la forme de la bande, nous avons utilisé deux méthodes. La première est basée sur une convolution de fonctions gaussiennes ou de fonctions Pseudo-Voigt, tandis que

la seconde méthode tient compte de la structure fine vibratoire. Concernant cette dernière, les progressions vibratoires ont été obtenues en utilisant trois approches: Adiabatic Hessian, la LQ2 e LQ3. Seulement pour la méthode Adiabatic Hessian, nous avons décidé de décaler la bande ωB97XD à la longueur d'onde d'absorption verticale PBE0, puisque la forme de la bande vibronique est mieux décrite par la fonctionnelle ωB97XD que celle de PBE0. Nous avons appelé l'approche résultante ωB97XD /PBE0(vert). Les distances euclidiennes, calculées entre les couleurs expérimentales et calculées dans l'espace de couleurs CIE-Lab, ont détecté des erreurs plus faibles lorsque les approches LQ2 et LQ3 sont utilisées au lieu de ωB97XD /PBE0(vert). Il est intéressant de noter que les résultats ont montré que l'application des approches pseudo-Voigt et LQ2 et LQ3 permet la reproduction précise du spectre d'absorption et donc des couleurs (voir Tableau 8.1 et Tableau 8.2). Globalement, les erreurs sur les distances euclidiennes entre les couleurs calculées et les couleurs expérimentales sont faibles pour les méthodes pseudo-Voigt, LQ2 et LQ3. Notre analyse nous a permis de mettre en place deux procédures peu coûteuses pour obtenir les spectres d'absorption et les caractéristiques spectrales des colorants organiques en solution.

Tableau 8.1 Couleur calculée et expérimentale pour toutes les anthraquinones fonctionnalisées avec tous les protocoles, **X** pas possible de calculer les spectres vibroniques.

	exp	PBE0^{PV}	LQ2	LQ3	ωB97XD/PBE0 (vert)	PBE0^{G}	ωB97XD ^{G}	PBE0^{vib}	ωB97XD^{vib}
AQ1									
AQ2									
AQ3									
AQ4								**X**	**X**
AQ5								**X**	**X**
AQ6								**X**	**X**
AQ7								**X**	**X**
AQ8								**X**	**X**

Tableau 8.2 Couleur calculée et expérimentale pour toutes les indigos fonctionnalisées avec tous les protocoles. **X** pas possible de calculer les spectres vibroniques.

	exp	PBE0^{PV}	LQ2	LQ3	ωB97XD/PBE0(vert)	PBE0^{G}	ωB97XD^{G}	PBE0^{vib}	ωB97XD^{vib}
IN1									
IN2					X			X	X
IN3					X			X	X
IN4					X			X	X
IN5					X			X	X
IN6					X			X	X
IN7					X			X	X
IN8					X			X	X
IN9					X			X	X
IN10									

8.2 Modélisation des spectres UV-vis à l'aide d'une approche multi-échelle

Le deuxième projet est axé sur la prédiction de l'énergie d'absorption maximale d'une classe de composés flexibles, à savoir la mérocyanine. Cette classe de molécules est caractérisée par la présence d'un groupe donneur et d'un groupe accepteur reliés par un pont éthylène ou polyéthylène, Figure 8.2.

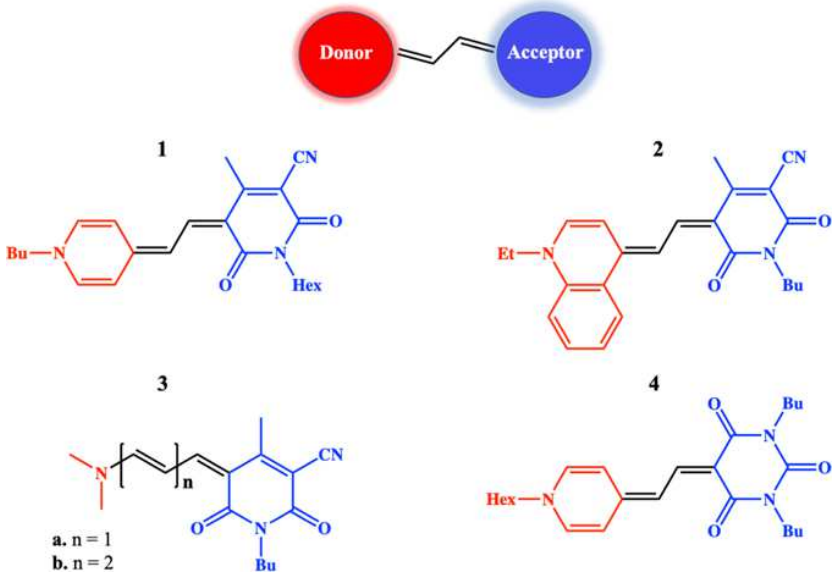


Figure 8.2 Représentation schématique des mérocyanines étudiées dans cette thèse. But, butyle; Hex, hexil; Et, etil.

Afin de considérer l'impact des possibles réarrangements moléculaires sur la propriété d'absorption, deux approches de calcul ont été choisies. La première stratégie fait référence à une approche statique, ainsi, des calculs verticaux TD-DFT/ IEF-PCM ont été effectués sur la structure optimisée de l'état fondamental au niveau DFT. La deuxième stratégie correspond à une combinaison de la mécanique moléculaire et d'une procédure DFT (Figure 8.3).

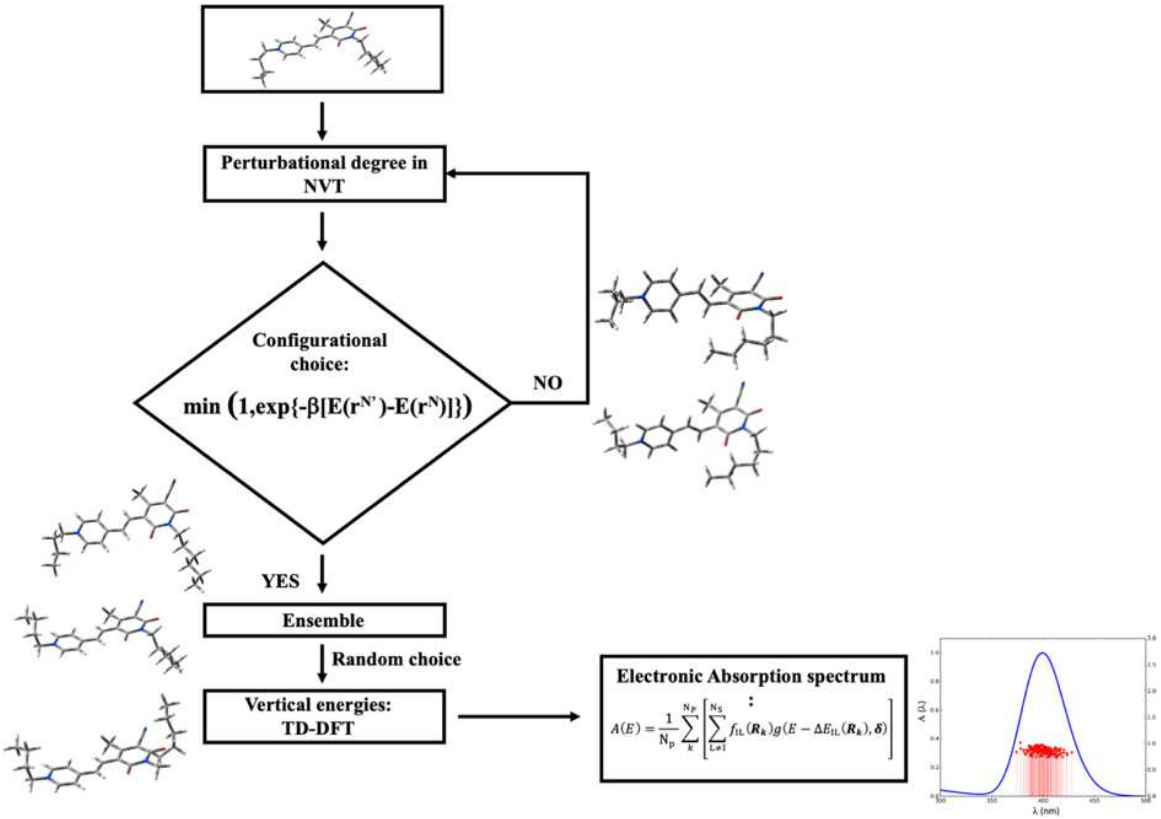


Figure 8.3 Représentation schématique sous forme d'organigramme des étapes du protocole Monte Carlo/DFT.

Précisément, pour ce dernier, l'énergie des configurations de Monte Carlo a été calculée au niveau de la mécanique moléculaire en utilisant le standard Generalized Amber Force Field (GAFF), ensuite, le TD-DFT/ IEF-PCM a été utilisé pour estimer les énergies des excitations verticales pour le calcul des spectres d'absorption. Enfin, l'approche de la convolution gaussienne a été utilisée afin de simuler la forme de la bande.

Les principaux résultats concernent les énergies d'absorption pour les mérocyanines sous leur forme monomère et dimère. Nos calculs ont montré que l'écart absolu moyen (MAD) des PBE0 est plus grand que celui obtenu au niveau MC, même lorsque les énergies d'absorption verticale sont calculées comme la valeur moyenne des deux monomères à leur géométrie statique dans le dimère. La tendance trouvée pour les monomères est conservée pour les dimères (agrégat H).

De plus, afin d'examiner comment les caractéristiques structurelles des mérocyanines et de leurs espèces dimères influencent le λ_{max}, nous avons pris en compte plusieurs paramètres géométriques: l'alternance des longueurs de liaison (BLA), l'orientation entre la partie donneur accepteur liée par le pont π, et la distance entre les fragments monomères qui composent l'agrégat dimère. Ici, nos résultats ont montré que les décalages observés sont dus à de petites déformations de la géométrie

moléculaire. Bien que les simulations MC classiques ne permettent pas de reconstruire la progression vibronique, elles prennent en compte la nature flexible des molécules et son effet sur le λmax qui est globalement décalé conduisant à un meilleur accord avec les spectres expérimentaux. Dans l'ensemble, les résultats ont montré que malgré sa simplicité, une approche combinée MC/TD-DFT est fiable pour décrire des systèmes flexibles qui ne peuvent pas être décrits correctement dans un cadre statique FC.

8.3 Modélisation de l'abstraction d'hydrogène par les premiers principes

Le troisième travail se concentre sur la réactivité dans l'état excité. Plus précisément, la photolyse des anthraquinones fonctionnalisées qui est l'une des principales méthodes pour synthétiser un produit d'hydroxylation [165,166]. Le photoproduit peut être utilisé comme point de départ pour fabriquer de nouveaux chromophores, puisque la partie alcoolique du photoproduit organique peut être facilement transformée [161]. Dans ce cadre, il est crucial de comprendre la réaction photoinduite, qui constitue l'étape initiale de la production de nouveaux colorants d'intérêt pour la société PILI. L'absorption provoque la promotion d'un électron de l'état fondamental à l'état électronique excité, puis l'anthraquinone fonctionnalisée excitée subit un croisement intersystème (ISC) vers l'état triplet, favorisant la formation du diradical 1-4 anthrahydroquinone. Ce diradical est capable de réagir avec l'oxygène moléculaire pour donner l'endoperoxyde. La principale limitation de l'hydroxilation photochimique est due à la conversion de l'intermédiaire diradicalaire en cétone de départ, également connue sous le nom de dismutation interne [163]. Ainsi, l'étape clé de cette réaction à plusieurs étapes est la production d'anthrahydroquinone 1-4 diradicalaire, dont la barrière énergétique régit la formation de l'endoperoxyde. Il est fondamental pour la production de diradicaux d'étudier les voies de désactivation de l'état excité initialement singulet vers l'état triplet. Dans ce contexte, nous avons modélisé la voie de désactivation non radiative pour deux anthraquinones substituées caractérisées par l'acétate (Ac) et le triflate (Tf) comme unités protectrices, Figure 8.4. Ces deux anthraquinones substituées ont été choisies comme composés prototypiques pour l'obtention de produits diradicaux (AcHAQ et TfHAQ).

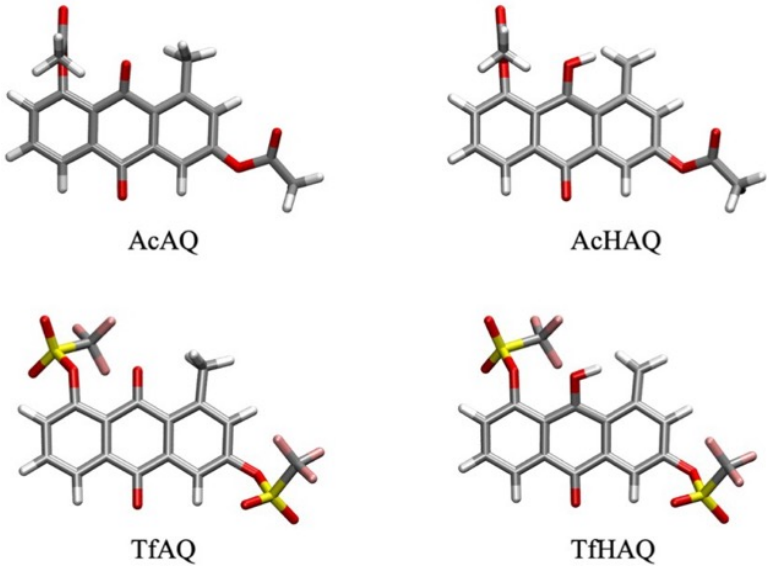


Figure 8.4 Représentation en bâton de deux anthraquinones/anthrahydroquinones disubstituées étudiées dans ce travail.

Afin de dériver un modèle pour la désactivation de l'état excité et d'identifier le canal de réaction, nous avons utilisé une approche basée sur la méthode du transit synchrone linéaire (LST). La méthode LST consiste à construire un chemin d'interpolation linéaire à partir de la structure optimisée du réactif et du produit [167]. En entrant dans les détails, dans la première étape de la réaction, l'absorption occupe un état singulet (c'est-à-dire avec la plus grande force d'oscillation, Sk) tandis que le produit (AcAHQ ou TfHAQ) est dans son état triplet (T). Ensuite, le chemin a été conçu en maintenant les structures AcAQ (TfAQ) et AcAHQ (TfHAQ) dans leur minimum Sk. Étant donné qu'un mécanisme de désintégration implique différents états électroniques avec une multiplicité de spin identique ou différente (par exemple Sn avec Tn ; 1 <n< k), la désactivation de l'état excité a été détectée en faisant varier l'état électronique mais en maintenant fixes les structures LST obtenues à partir de l'état de référence (Sk). Cette approche est cohérente avec la modélisation efficace du mécanisme de désintégration, puisque les changements structurels sont beaucoup plus lents que la désintégration électronique. La Figure 8.5 montre les voies possibles impliquant les états excités de l'AcAQ (TfAQ) pour produire l'anthrahydroquinone protégée correspondante. Selon notre modèle, les abstractions d'hydrogène ne pourraient pas se produire dans la première étape, car les barrières énergétiques calculées sur la surface T_1 pour AcAQ et TfAQ sont $\Delta E_{T_1(I)-T_1(VII)}$= 2.15 eV et $\Delta E_{T_1(I)-T_1(VII)}$= 2.91 eV.

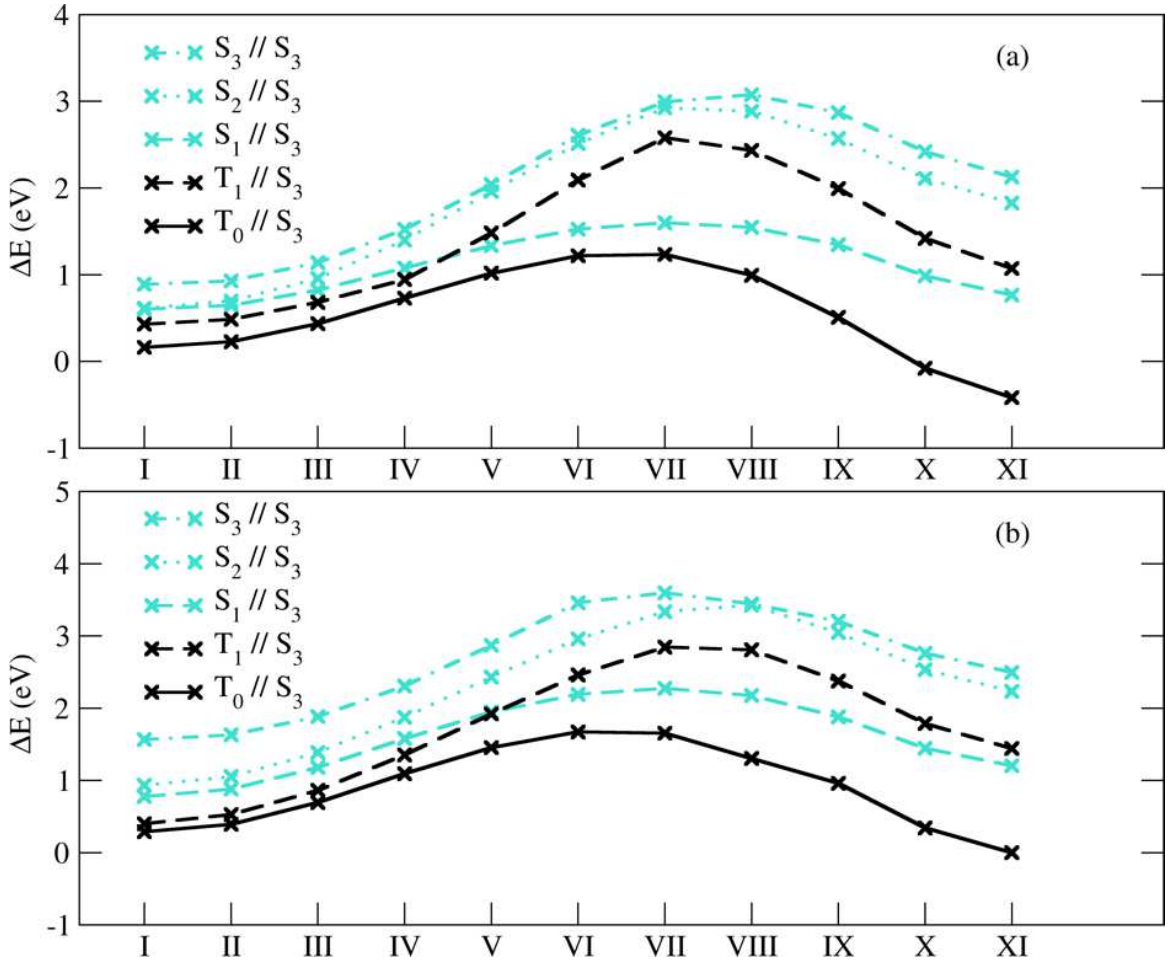


Figure 8.5 Variation de l'énergie électronique en fonction des coordonnées de réaction. Les courbes identifient l'état et la géométrie : Ai//Bj, Ai identifie l'état électroniques; Bj identifie la géométrie optimisée pour un état électronique donné utilisée pour la modélisation de la voie LST. Niveau de théorie : PBE0/631++G(2d,2p)/ IEF-PCM (benzène). Courbes mises à l'échelle en référence à l'énergie minimale $T_0//S_3$. Profils d'énergie d'abstraction d'hydrogène: le panneau supérieur (a) se réfère à une QA substituée par de l'acétate (Ac), le panneau inférieur (b) se réfère à une QA fonctionnalisée par du triflate (Tf).

Afin d'identifier qualitativement la région d'énergie potentielle où un croisement entre les états est suspecté, l'indice Π a été utilisé [178]. Des valeurs élevées de Π définiront les régions correspondant à la plus forte probabilité de désintégration. Bien que l'écart d'énergie entre les deux états sera le terme principal de nombreuses réactions photochimiques impliquant un croisement d'états ($\Delta E \rightarrow 0$), pour une valeur donnée de l'écart d'énergie, une désintégration sera d'autant plus efficace que les densités électroniques des états de départ et final sont similaires, c'est-à-dire lorsque le produit $\boldsymbol{D}\mathrm{CT} \cdot \boldsymbol{q}\mathrm{CT}$ est petit. Autrement dit, plus la réorganisation des électrons entre les deux états est faible, plus la désintégration est probable. Par conséquent, des voies de désintégration

efficaces sont susceptibles d'émerger dans les régions où deux états se rapprochent en énergie ($\Delta E \rightarrow 0$), ou lorsque les densités électroniques des deux états impliqués sont similaires ($DCT \cdot qCT \rightarrow 0$), ou lorsque ces deux critères sont vérifiés simultanément. La figure 8.6 montre les désintégrations les plus probables identifiées par l'indice Π sur toutes les étapes du LST.

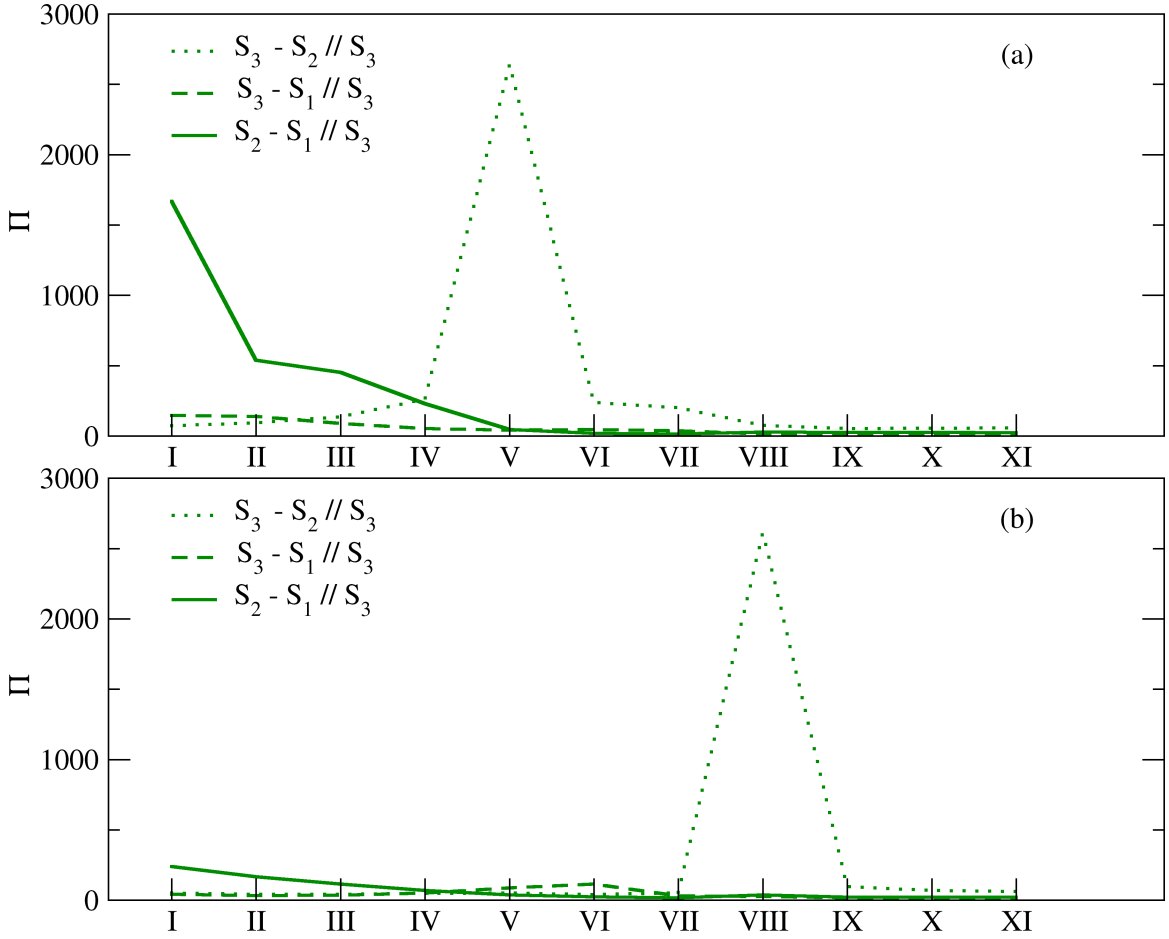


Figure 8.6 Evolution de l'indice Π le long de la coordonnée de réaction. Le panneau supérieur (a) se réfère à l'AQ substitué par de l'acétate (Ac), le panneau inférieur (b) se réfère à l'AQ fonctionnalisé par du triflate (Tf).

En ce qui concerne la QA protégeant l'Ac, la courbe Π S3-S2 atteint son maximum global autour de la structure V. En continuant l'abstraction d'hydrogène, un croisement intersystème est observé autour de la structure VI et la barrière d'énergie calculée sur l'état S2 s'élève à $\Delta E_{S_2(VII)-S_2(VI)}$= 0.41 eV. D'autre part, pour le système TfAQ, un canal de décroissance plus raisonnable a lieu autour de l'étape numéro VIII impliquant les états S3 et S2. Dans l'ensemble, toutes les étapes ont $\Delta E_{S_2-T_1}$ inférieur à un nécessaire pour activer la molécule d'oxygène de son état fondamental triplet à l'état excité singulet 0,97 eV [162, 180]. Ainsi, en accord avec les données expérimentales

[161], les deux molécules photoactives testées réagissent avec le $^3O_2$. Cette observation est cohérente avec les résultats expérimentaux déduits par Elkazaz et ses collaborateurs [161], étant donné qu'un quencher de $^1O_2$ n'avait aucun effet sur le taux de photoclivage conduisant à la formation d'endoperoxyde [161]. Contrairement à l'hypothèse selon laquelle le AQ protégé subit activement la photoréduction par l'abstraction de l'atome d'hydrogène dans l'état T_1, nos calculs ont montré que le croisement direct entre les systèmes est défavorable, en effet, les SOC_{S_3,T_1} calculés sur la première structure sont plus petits que ceux obtenus sur la structure située dans le stade final du LST. Ainsi, l'abstraction de l'hydrogène se déroule sans conservation du spin conduisant à un biradical triplet. Globalement, nos résultats nous permettent à la fois d'identifier un canal de désintégration non-radiative reliant S_3 à T_1 via S_2, et de démontrer l'existence d'un canal étroit qui conduit à un mécanisme apparent sans barrière.

En conclusion, les protocoles de calcul développés dans ce projet se sont avérés efficaces pour modéliser les phénomènes complexes qui impliquent les états excités et sont capables de soutenir efficacement la société PILI dans la conception de nouveaux colorants verts pour réduire l'impact environnemental.

References

[1] Ardila-Leal, L.D., Poutou-Piñales, R.A., Pedroza-Rodríguez, A.M. and Quevedo-Hidalgo, B.E. A Brief History of Colour, the Environmental Impact of Synthetic Dyes and Removal by Using Laccases. *Molecules*, *26*, 3813-3853, **2021**.

[2] Slama, H.B., Chenari Bouket, A., Pourhassan, Z., Alenezi, F.N., Silini, A., Cherif-Silini, H., Oszako, T., Luptakova, L., Golińska, P. and Belbahri, L. Diversity of synthetic dyes from textile industries, discharge impacts and treatment methods. *Applied Sciences*, *11*, 6255-6276, **2021**.

[3] Hassaan, M.A., El Nemr, A. and Hassaan, A. Health and environmental impacts of dyes: mini review. *American Journal of Environmental Science and Engineering*, *1*, 64-67, **2017**.

[4] Fried, R., Oprea, I., Fleck, K. and Rudroff, F. Biogenic colourants in the textile industry–a promising and sustainable alternative to synthetic dyes. *Green Chemistry*, *24*, 13-35, **2022**.

[5] Sen, T., Barrow, C.J. and Deshmukh, S.K.. Microbial pigments in the food industry—challenges and the way forward. *Frontiers in nutrition*, *6*, 1-7., **2019**.

[6] Durmus, D. CIELAB color space boundaries under theoretical spectra and 99 test color samples. *Color Research & Application*, *45*, 796-802, **2020**.

[7] Tkalcic, M., Tasic, J. F. Colour spaces: perceptual, historical and applicational background, *The IEEE Region 8 EUROCON 2003. Computer as a Tool.*, *1*, 304-308, **2003**.

[8] Jacquemin, D., Brémond, E., Planchat, A., Ciofini, I. and Adamo, C. TD-DFT vibronic couplings in anthraquinones: from basis set and functional benchmarks to applications for industrial dyes. *Journal of chemical theory and computation*, *7*, 1882-1892, **2011**.

[9] Barone, V., Biczysko, M., Latouche, C. and Pasti, A., 2015. Virtual eyes for technology and cultural heritage: towards computational strategy for new and old indigo-based dyes. *Theoretical chemistry accounts*, *134*, 1-14, **2015**.

[10] Laurent, A.D., Adamo, C. and Jacquemin, D. Dye chemistry with time-dependent density functional theory. *Physical Chemistry Chemical Physics*, *16*,14334-14356, **2014**.

[11] Tomasi, J., Mennucci, B. and Cammi, R. Quantum mechanical continuum solvation models. *Chemical reviews*, *105*, 2999-3094, **2005**.

[12] Jacquemin, D., Brémond, E., Ciofini, I. and Adamo, C. Impact of vibronic couplings on perceived colors: two anthraquinones as a working example. *The journal of physical chemistry letters*, *3*, 468-471, **2012**.

[13] Deuflhard, P., Hermans, J., Leimkuhler, B., Mark, A.E., Reich, S. and Skeel, R.D. eds.*Computational Molecular Dynamics: Challenges, Methods, Ideas: Proceeding of the 2nd International Symposium on Algorithms for Macromolecular Modelling, Berlin,* Springer Science & Business Media, **1997**.

[14] Barp, A., Briol, F.X., Kennedy, A.D. and Girolami, M. Geometry and dynamics for markov chain monte carlo. *Annual Review of Statistics and Its Application*, *5*, 451-471, **2018**.

[15] Tirri, B., Mazzone, G., Ottochian, A., Gomar, J., Raucci, U., Adamo, C. and Ciofini, I. A combined Monte Carlo/DFT approach to simulate UV-vis spectra of molecules and aggregates: Merocyanine dyes as a case study. *Journal of Computational Chemistry*, *42*, 1054-1063, **2021**.

[16] Cohen, A.J., Mori-Sánchez, P. and Yang, W. Challenges for density functional theory. *Chemical reviews*, *112*, 289-320, **2012**.

[17] Sousa, S.F., Fernandes, P.A. and Ramos, M.J.. General performance of density functionals. *The Journal of Physical Chemistry A*, *111*, 10439-10452, **2007**.

[18] Davidson, E.R. Molecular mechanics and modeling: Overview. *Chemical Reviews*, *93*, 2337-2337, **1993**.

[19] Born, M. and Oppenheimer, R. Zur quantentheorie der molekeln. *Annalen der physik*, *389*(20), pp.457-484, **1927**.

[20] Szabo, A., Ostlun, N. S. *Modern Quantum Chemistry*, Dover Publishing (New York), **1996**.

[21] Slater, J.C.. The theory of complex spectra. *Physical Review*, *34*, 1293-1322, **1929**.

[22] Jensen, F.. Density Functional Theory. *Introduction to Computational Chemistry*, **1999**.

[23] Löwdin, P.O.. Correlation Problem in Many-Electron Quantum Mechanics I. Review of Different Approaches and Discussion of Some Current Ideas. *Advances in chemical physics*, 207-322, **1958**.

[24] Hohenberg, P. and Kohn, W.. Inhomogeneous electron gas. *Physical review*, *136*, B864, **1964**.

[25] Kohn, W. and Sham, L.J.. Self-consistent equations including exchange and correlation effects. *Physical review*, *140*, .A1133, **1965**.

[26] Cramer, C.J. *Essentials of computational chemistry: theories and models*. John Wiley & Sons, **2013**.

[27] Perdew, J.P. and Schmidt, K.. Jacob's ladder of density functional approximations for the exchange-correlation energy. In *AIP Conference Proceedings* , Vol. 577, 1-20, **2001**.

[28] Beck, A.D.. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys*, *98*, 5648-5652, **1993**.

[29] (a) Langreth, D.C. and Mehl, M.J.. Beyond the local-density approximation in calculations of ground-state electronic properties. *Physical Review B*, *28*, 1809, **1983**.
(b) Becke, A.D.. Density-functional exchange-energy approximation with correct asymptotic behavior. *Physical review A*, *38*, 3098, **1988**.

[30] Perdew, J.P., Burke, K. and Ernzerhof, M.. Generalized gradient approximation made simple. *Physical review letters*, *77*(18), p.3865–3868, **1996**.

[31] Adamo, C. and Barone, V.. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *The Journal of chemical physics*, *110*, 6158-6170, **1999**.

[32] Tao, J., Perdew, J.P., Staroverov, V.N. and Scuseria, G.E.. Climbing the density functional ladder: Nonempirical meta–generalized gradient approximation designed for molecules and solids. *Physical Review Letters*, *91*(14), 146401, **2003**.

[33] Lee, C., Yang, W. and Parr, R.G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical review B*, *37*, 785, **1988**.

[34] Zhao, Y. and Truhlar, D.G. A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions. *The Journal of chemical physics*, *125*(19), 194101-19419, **2006**.

[35] Brémond, É., Pérez-Jiménez, Á.J., Sancho-García, J.C. and Adamo, C. Range-separated hybrid density functionals made simple. *The Journal of chemical physics*, *150*, 201102-201108, **2019**.

[36] Chai, J.D. and Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Physical Chemistry Chemical Physics*, *10*(44), pp.6615-6620, **2008**.

[37] Marques, M. A. L. and Gross, E. K. U. Time-Dependent Density Functional Theory, *Annual Review Physical Chemistry*, *55*, 427-55, **2004**.

[38] Casida, M.E. Time-dependent density functional response theory of molecular systems: theory, computational methods, and functionals. *Theoretical and Computational Chemistry*, *4*, 391-439, **1996**.

[39] Laurent, A.D. and Jacquemin, D. TD-DFT benchmarks: a review. *International Journal of Quantum Chemistry*, *113*, 2019-2039, **2013**.

[40] Jacquemin, D., Perpete, E.A., Scuseria, G.E., Ciofini, I. and Adamo, C., 2008. TD-DFT performance for the visible absorption spectra of organic dyes: conventional versus long-range hybrids. *Journal of chemical theory and computation*, *4*, 123-135, **2008**.

[41] Jacquemin, D., Wathelet, V., Preat, J. and Perpète, E.A., 2007. Ab initio tools for the accurate prediction of the visible spectra of anthraquinones. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, *67*, 334-341, **2007**.

[42] Miertuš, S., Scrocco, E. and Tomasi, J. Electrostatic interaction of a solute with a continuum. A direct utilizaion of AB initio molecular potentials for the prevision of solvent effects. *Chemical Physics*, *55*, 117-129, **1981**.

[43] Santoro, F., Improta, R., Lami, A., Bloino, J. and Barone, V. Effective method to compute Franck-Condon integrals for optical spectra of large molecules in solution. *The Journal of chemical physics*, *126*, 084509-084522, **2007**.

[44] Bloino, J., Biczysko, M., Santoro, F. and Barone, V. General approach to compute vibrationally resolved one-photon electronic spectra. *Journal of Chemical Theory and Computation*, *6*(, 1256-1274, **2010**.

[45] Barone, V., Bloino, J., Biczysko, M. and Santoro, F. Fully integrated approach to compute vibrationally resolved optical spectra: from small molecules to macrosystems. *Journal of chemical theory and computation*, *5*, 540-554, **2009**.

[46] Herzberg, G. and Teller, E. Schwingungsstruktur der Elektronenübergänge bei mehratomigen Molekülen. *Zeitschrift für Physikalische Chemie*, *21*, 410-446, **1933**.

[47] Barone, V., Bloino, J. and Biczysko, M. Vibrationally-resolved electronic spectra in GAUSSIAN 09. *Revision a*, *2*, 1-20, **2009**.

[48] Santoro, F. and Jacquemin, D. Going beyond the vertical approximation with time-dependent density functional theory. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, *6*, 460-486, **2016**.

[49] Di Tommaso, S., Bousquet, D., Moulin, D., Baltenneck, F., Riva, P., David, H., Fadli, A., Gomar, J., Ciofini, I. and Adamo, C. Theoretical approaches for predicting the color of rigid dyes in solution. *Journal of computational chemistry*, *38*, 998-1004, **2017.**

[50] Ferrer, F.J.A. and Santoro, F. Comparison of vertical and adiabatic harmonic approaches for the calculation of the vibrational structure of electronic spectra. *Physical Chemistry Chemical Physics*, *14*, 13549-13563, **2012**.

[51] Muniz-Miranda, F., Pedone, A., Battistelli, G., Montalti, M., Bloino, J. and Barone, V. Benchmarking TD-DFT against vibrationally resolved absorption spectra at room temperature: 7-aminocoumarins as test cases. *Journal of chemical theory and computation*, *11*, 5371-5384, **2015**.

[52] Hodecker, M., Biczysko, M., Dreuw, A. and Barone, V. Simulation of vacuum UV absorption and electronic circular dichroism spectra of methyl oxirane: the role of vibrational effects. *Journal of chemical theory and computation*, *12*, 2820-2833, **2016**.

[53] Baiardi, A., Bloino, J. and Barone, V. General time dependent approach to vibronic spectroscopy including Franck–Condon, Herzberg–Teller, and Duschinsky effects. *Journal of chemical theory and computation*, 9, 4097–4115, **2013**.

[54] Li, S.L. and Truhlar, D.G. Franck–Condon models for simulating the band shape of electronic absorption spectra. *Journal of chemical theory and computation*, 13, 2823–2830, **2017**.

[55] Heller, E.J., 1975. Time-dependent approach to semiclassical dynamics. *The Journal of Chemical Physics*, *62*, 1544, **1975**.

[56] Tannor, D.J. and Heller, E.J. Polyatomic Raman scattering for general harmonic potentials. *The Journal of Chemical Physics*, 77, 202, **1982**.

[57] Antonov, L. and Stoyanov, S. Analysis of the overlapping bands in UV-vis absorption spectroscopy. *Applied spectroscopy*, *47*,1030-1035, **1993**.

[58] Kauppinen, J.K., Moffatt, D.J., Mantsch, H.H. and Cameron, D.G. Fourier self-deconvolution: a method for resolving intrinsically overlapped bands. *Applied Spectroscopy*, *35*, 271-276, **1981**.

[59] Nassau, K. *The physics and chemistry of color: the fifteen causes of color* , Jhon Wiley & Son, **2001.**

[60] Wright, W.D. A re-determination of the mixture curves of the spectrum. *Transactions of the Optical Society*, 31, 201-218, **1930**

[61] Luo, M.R. A review of chromatic adaptation transforms. *Review of Progress in Coloration and Related Topics*, *30*, 77-92, **2000**.

[62] Tkalcic, M. and Tasic, J.F. *Colour spaces: perceptual, historical and applicational background* , 1, 304-308, **2003.**

[63] M. R. McLellan, L. R. Lind, R. W. Kime, Hue angle determinations and statistical analysis for multiquadrant Hunter L, a, b data*, Journal of food quality*, *18*, 235-240, **1995**.

[64] Ponder, J.W. and Case, D.A. Force fields for protein simulations. *Advances in protein chemistry*, *66*, 27–85, **2003**.

[65] Tuckerman, M. *Statistical mechanics: theory and molecular simulation*. Oxford university press, **2010**.

[66] Metropolis, N. The beginning of the Monte Carlo method. *Los Alamos Science*, *15*, 125-130, **1987**.

[67] Frenkel, D. and Smit, B. Understanding molecular simulation: From algorithms to applications, **2001**.

[68] Easton, J.R. The dye maker's view. *Colour in dyehouse effluent*, The Alden Press: Oxford, *11,* 9−21, **1995**.

[69] Hunger, K. Industrial Dyes, Chemistry, Properties, Applications Wiley, **2003**.

[70] Gregory, P. Anthraquinone Chromophore, in Industrial Dyes: Chemistry, Properties, Applications*, ed. K. Hunger, Wiley- VCH, Weinheim*, **2003**.

[71] Franchi, D.; Calamante, M.; Coppola, C.; Mordini, A.; Reginato, G.; Sinicropi, A.; Zani, L. Synthesis and Characterization of New Organic Dyes Containing the Indigo Core. *Molecules*, *25*, 3377, **2020**.

[72] Venil, C.K., Zakaria, Z.A. and Ahmad, W.A. Bacterial pigments and their applications. *Process Biochemistry*, *48*, 1065-1079, **2013**.

[73] (a)Recombinant host cells to produce anthraquinone derivatives, L. Crepin, C. Schiavon, A. Baylach, É. Cordierg, G. Boissonnat, WO2020161354A1.

[74] Loos, P.F., Galland, N. and Jacquemin, D.Theoretical 0–0 energies with chemical accuracy. *The journal of physical chemistry letters*, *9*(16), pp.4646-4651, 2108.

[75] Li, H., Tirri, B., Brémond, E., Sancho-García, J.C. and Adamo, C. Beyond Chemical Accuracy for Alkane Thermochemistry: The DH thermo Approach. *The Journal of Organic Chemistry*, *86*, 5538-5545, **2021**.

[76] Nattino, F., Migliorini, D., Kroes, G.J., Dombrowski, E., High, E.A., Killelea, D.R. and Utz, A.L. Chemically accurate simulation of a polyatomic molecule-metal surface reaction. *The journal of physical chemistry letters*, *7*, 2402-2406, **2016.**

[77] Lier, B., Poliak, P., Marquetand, P., Westermayr, J. and Oostenbrink, C. BuRNN: A Buffer Region Neural Network Approach for QM/MM Simulations of Complex Systems. *arXiv preprint arXiv:2112.11395*, *121*, 9873-9926, **2021.**

[78] Ramakrishnan, R., Hartmann, M., Tapavicza, E. and Von Lilienfeld, O.A. Electronic spectra from TDDFT and machine learning in chemical space. *The Journal of chemical physics*, *143*, 084111, **2015**.

[79] Beck, M.E. Estimation of physiologically perceived color from TDDFT-derived excitation spectra. *International journal of quantum chemistry*, *101*, 683-689,**2005.**

[80] Cysewski, P., Jeliński, T., Przybyłek, M. and Shyichuk, A. Color prediction from first principle quantum chemistry computations: a case of alizarin dissolved in methanol. *New Journal of Chemistry*, *36*(9), pp.1836-1843, **2012**.

[81] Matsuura, A., Sato, H., Sotoyama, W., Takahashi, A. and Sakurai, M. AM1, PM3, and PM5 calculations of the absorption maxima of basic organic dyes. *Journal of Molecular Structure: THEOCHEM*, *860*, 119-127, **2008**.

[82] Jacquemin, D., Perpete, E.A., Ciofini, I. and Adamo, C. Accurate simulation of optical properties in dyes. *Accounts of chemical Research*, *42*, 326-334, **2009**.

[83] Marković, Z., Manojlović, N. and Zlatanović, S. Electronic absorption spectra of substituted anthraquinones and their simulation using ZINDO/S method. *Int. J. Serb. Soc. Comput. Mechan*, *2*, 73-79, 2008.

[84] Lechner, M.H., Neese, F. and Izsák, R. An excited state coupled-cluster study on indigo dyes. *Molecular Physics*, p.e1965235, **2021**.

[85] Schreiber, M., Buß, V. and Fülscher, M.P. The electronic spectra of symmetric cyanine dyes: A CASPT2 study. *Physical Chemistry Chemical Physics*, *3*, 3906-3912, **2001**.

[86] Moser, M., Thorley, K.J., Moruzzi, F., Ponder, J.F., Maria, I.P., Giovannitti, A., Inal, S. and McCulloch, I. Highly selective chromoionophores for ratiometric Na+ sensing based on an oligoethyleneglycol bridged bithiophene detection unit. *Journal of Materials Chemistry C*, *7*, 5359-5365, **2019**.

[87] Guillaumont, D. and Nakamura, S. Calculation of the absorption wavelength of dyes using time-dependent density-functional theory (TD-DFT). *Dyes and Pigments*, *46*, 85-92, **2000**.

[88] Kawauchi, S., Antonov, L. and Okuno, Y. Prediction of the color of dyes by using time-dependent density functional theory (TD-DFT). *Bulg. Chem. Commun*, *46*, 228-237, **2014**.

[89] Dreuw, A. and Head-Gordon, M. Single-reference ab initio methods for the calculation of excited states of large molecules. *Chemical reviews*, *105*, 4009-4037, **2005**.

[90] Jacquemin, D., Wathelet, V., Perpete, E.A. and Adamo, C. Extensive TD-DFT benchmark: singlet-excited states of organic molecules. *Journal of Chemical Theory and Computation*, *5*(9), pp.2420-2435, **2009**.

[91] Silva-Junior, M.R., Schreiber, M., Sauer, S.P. and Thiel, W. Benchmarks for electronically excited states: Time-dependent density functional theory and density functional theory based multireference configuration interaction. *The Journal of chemical physics*, *129*, 104103, **2008**.

[92] Perpète, E.A. and Jacquemin, D. TD-DFT benchmark for indigoïd dyes. *Journal of Molecular Structure: THEOCHEM*, *914*, 100-105, **2009**.

[93] (a) Franck, J. and Dymond, E.G. Elementary processes of photochemical reactions. *Transactions of the Faraday Society*, *21*, 536-542, **1926**. (b) Condon, E.U. Nuclear motions associated with electron transitions in diatomic molecules. *Physical Review*, *32*, 858, **1928**.

[94] Dierksen, M. and Grimme, S. An efficient approach for the calculation of Franck–Condon integrals of large molecules. *The Journal of chemical physics*, *122*, 244101, **2005**.

[95] Petrenko, T. and Neese, F. Analysis and prediction of absorption band shapes, fluorescence band shapes, resonance Raman intensities, and excitation profiles using the time-dependent theory of electronic spectroscopy. *The Journal of chemical physics*, *127*, 164319, **2007**.

[96] De Mitri, N., Monti, S., Prampolini, G. and Barone, V. Absorption and emission spectra of a flexible dye in solution: A computational time-dependent approach. *Journal of chemical theory and computation*, *9*, 4507-4516, **2013**.

[97] Raucci, U., Perrella, F., Donati, G., Zoppi, M., Petrone, A. and Rega, N. Ab-initio molecular dynamics and hybrid explicit-implicit solvation model for aqueous and nonaqueous solvents: GFP chromophore in water and methanol solution as case study. *Journal of Computational Chemistry*, *41*, 2228-2239, **2020**.

[98] Kuhn, W. and Braun, E. Messung und deutung der rotationsdispersion einfacher stoffe; Zf Phys. *Chem.(B)*, pp.281-313, **1930**.

[99] Electronic Spectra and Quantum Chemistry C. Sandorfy, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, **1964**.

[100] Brémond, É.A., Kieffer, J. and Adamo, C. A reliable method for fitting TD-DFT transitions to experimental UV–visible spectra. *Journal of Molecular Structure: THEOCHEM*, *954*, 52-56, **2010**.

[101] Jain, V., Biesinger, M.C. and Linford, M.R. The Gaussian-Lorentzian Sum, Product, and Convolution (Voigt) functions in the context of peak fitting X-ray photoelectron spectroscopy (XPS) narrow scans. *Applied Surface Science*, *447*, 548-553, **2018**.

[102] Adamo, C. and Jacquemin, D. The calculations of excited-state properties with Time-Dependent Density Functional Theory. *Chemical Society Reviews*, *42*(3), pp.845-856, **2013**.

[103] Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith,

T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian, Inc., Wallingford CT, **2016**.

[104] Huet, K. and Boissonnat, G. Anthraquinonic derivatives and their use as colouring agents, WO 2020/016245.

[105] Miliani, C., Romani, A. and Favaro, G.. A spectrophotometric and fluorimetric study of some anthraquinoid and indigoid colorants used in artistic paintings. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, *54*(4), pp.581-588, **1998**.

[106] Pummerer, R. and Marondel, G. Indigofarbstoffe, VII. Über den optischen Nachweis der reversiblen Bildung von cis-Formen symmetrischer und unsymmetrischer Indigoide. *Chemische Berichte*, *93*, 2834-2839, **1960**.

[107] Pinheiro, D., Pineiro, M., Galvão, A.M. and de Melo, J.S.S. Deep in blue with green chemistry: influence of solvent and chain length on the behaviour of N-and N, N'-alkyl indigo derivatives. *Chemical Science*, *12*, 303-313, **2021**.

[108] Huang, C.Y., Bonasera, A., Hristov, L., Garmshausen, Y., Schmidt, B.M., Jacquemin, D. and Hecht, S. N, N'-disubstituted indigos as readily available red-light photoswitches with tunable thermal half-Lives. *Journal of the American Chemical Society*, *139*, 15205-15211, **2017**.

[109] Serr, A. and O'Boyle, N., 2009. Convoluting UV-Vis spectra using oscillator strengths.

[110] Stancik, A.L. and Brauns, E.B. A simple asymmetric lineshape for fitting infrared absorption spectra. *Vibrational Spectroscopy*, *47*, 66-69, **2008**.

[111] Proctor, A. and Hercules, D.M. Inelastic background and peak area determination in X-ray photoelectron spectroscopy (ESCA). *Applied spectroscopy*, *38*, 505-518, **1984**.

[112] Ansell, R.O., Dickinson, T., Povey, A.F. and Sherwood, P.M.A. X-ray photoelectron spectroscopic studies of electrode surfaces using a new controlled transfer technique: Part II. Results for a molybdenum electrode and the curve fitting procedure. *Journal of Electroanalytical Chemistry and Interfacial Electrochemistry*, *98*, 79-89, **1979**.

[113] Schmid, M., Steinrück, H.P. and Gottfried, J.M. A new asymmetric Pseudo-Voigt function for more efficient fitting of XPS lines. *Surface and Interface Analysis*, *46*, 505-511, **2014**.

[114] Korepanov, V.I. and Sedlovets, D.M. An asymmetric fitting function for condensed-phase Raman spectroscopy. *Analyst*, *143*, 2674-2679, **2018**.

[115] S. L. Li and D. G. Truhlar, FCBand, http://comp.chem.umn.edu/fcband/, **2017.**

[116] Mishra, A., Behera, R.K., Behera, P.K., Mishra, B.K. and Behera, G.B. Cyanines during the 1990s: a review. *Chemical reviews*, *100*, 1973-2012,**2000**.

[117] Brückner, C. and Engels, B. Benchmarking ground-state geometries and vertical excitation energies of a selection of P-type semiconducting molecules with different polarity. *The Journal of Physical Chemistry A*, *119*, 12876-12891, **2015**.

[118] Ciofini, I., Adamo, C. and Chermette, H., 2005. Effect of self-interaction error in the evaluation of the bond length alternation in trans-polyacetylene using density-functional theory. *The Journal of chemical physics*, *123*, 121102, **2005**.

[119] Jacquemin, D., Perpète, E.A., Ciofini, I. and Adamo, C. Assessment of recently developed density functional approaches for the evaluation of the bond length alternation in polyacetylene. *Chemical physics letters*, *405*, 376-381, **2005**.

[120] Barone, V. The virtual multifrequency spectrometer: a new paradigm for spectroscopy. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, *6*, 86-110, **2016**.

[121] Bloino, J., Baiardi, A. and Biczysko, M. Aiming at an accurate prediction of vibrational and electronic spectra for medium-to-large molecules: an overview. *International Journal of Quantum Chemistry*, *116*, 1543-1574, **2016**.

[122] Autschbach, J. Charge-Transfer Excitations and Time-Dependent Density Functional Theory: Problems and Some Proposed Solutions. *ChemPhysChem*, *10*, 1757-1760, **2009**.

[123] Moore, B. and Autschbach, J. Longest-wavelength electronic excitations of linear cyanines: the role of electron delocalization and of approximations in time-dependent density functional theory. *Journal of chemical theory and computation*, *9*, 4991-5003, **2013**.

[124] Le Guennic, B. and Jacquemin, D. Taking up the cyanine challenge with quantum tools. *Accounts of chemical research*, *48*, 530-537, **2015**.

[125] Mustroph, H., Reiner, K., Senns, B., Mistol, J., Ernst, S., Keil, D. and Hennig, L. The Effects of Substituents and Solvents on the Ground-State π-Electronic Structure and Electronic Absorption Spectra of a Series of Model Merocyanine Dyes and Their Theoretical Interpretation. *Chemistry–A European Journal*, *18*, 8140-8149, **2012**.

[126] Mustroph, H. and Towns, A. Fine Structure in Electronic Spectra of Cyanine Dyes: Are Sub-Bands Largely Determined by a Dominant Vibration or a Collection of Singly Excited Vibrations?. *ChemPhysChem*, *19*, 1016, **2018**.

[127] Ishchenko, A.A., Kulinich, A.V., Bondarev, S.L. and Raichenok, T.F. UV–Vis absorption spectra and electronic structure of merocyanines in the gas phase. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, *190*, 332-335, **2018**.

[128] Capobianco, A., Borrelli, R., Landi, A., Velardo, A. and Peluso, A. Absorption band shapes of a push–pull dye approaching the cyanine limit: A challenging case for first principle calculations. *The Journal of Physical Chemistry A*, *120*, 5581-5589, **2016**.

[129] Champagne, B., Guillaume, M. and Zutterman, F. TDDFT investigation of the optical properties of cyanine dyes. *Chemical physics letters*, *425*, 105-109, **2006**.

[130] Masunov, A.E., Anderson, D., Freidzon, A.Y. and Bagaturyants, A.A. Symmetry-breaking in cationic polymethine dyes: part 2. Shape of electronic absorption bands explained by the thermal fluctuations of the solvent reaction field. *The Journal of Physical Chemistry A*, *119*, 6807-6815, **2015**.

[131] Kulinich, A.V., Kurdyukov, V.V. and Ishchenko, A.A. Effect of bulky substituents in the donor and acceptor terminal groups on solvatochromism of Brooker's merocyanine. *New Journal of Chemistry*, *43*, 7379-7385, **2019**.

[132] Wada, T., Nakano, H. and Sato, H. Solvatochromic Shift of Brooker's Merocyanine: Hartree–Fock Exchange in Time Dependent Density Functional Calculation and Hydrogen Bonding Effect. *Journal of chemical theory and computation*, *10*, 4535-4547, **2014**.

[133] Tanaka, Y., Kawashima, Y., Yoshida, N. and Nakano, H. Solvatochromism and preferential solvation of Brooker's merocyanine in water–methanol mixtures. *Journal of computational chemistry*, *38*, 2411-2419, **2017**.

[134] Rösch, U., Yao, S., Wortmann, R., Würthner, F. Aggregates of Merocyanine Dyes. *Angewandte Chemie International Edition*, *45*, 7026, **2006**.

[135] Würthner, F. Dipole–dipole interaction driven self-assembly of merocyanine dyes: from dimers to nanoscale objects and supramolecular materials, *Accounts of chemical research*, 49, 868-876, **2016.**

[136] Santoro, F., Lami, A., Improta, R., Bloino, J. and Barone, V. Effective method for the computation of optical spectra of large molecules at finite temperature including the Duschinsky and Herzberg–Teller effect: The Q x band of porphyrin as a case study. *The Journal of chemical physics*, *128*, 224311, **2008**.

[137] Peltier, C., Lainé, P.P., Scalmani, G., Frisch, M.J., Adamo, C. and Ciofini, I. Environmental effects on electronic absorption spectra using DFT: An organic and positively charged fused polycyclic chromophore as a case study. *Journal of Molecular Structure: THEOCHEM*, *914*, 94-99, **2009**.

[138] Cerezo, J., Avila Ferrer, F.J., Prampolini, G. and Santoro, F. Modeling solvent broadening on the vibronic spectra of a series of coumarin dyes. From implicit to explicit solvent models. *Journal of chemical theory and computation*, *11*, 5810-5825, **2015**.

[139] De Mitri, N., Monti, S., Prampolini, G. and Barone, V. Absorption and emission spectra of a flexible dye in solution: A computational time-dependent approach. *Journal of chemical theory and computation*, *9*, 4507-4516, **2013**.

[140] Cascella, M., Cuendet, M.A., Tavernelli, I. and Rothlisberger, U. Optical Spectra of Cu (II)− Azurin by Hybrid TDDFT-Molecular Dynamics Simulations. *The Journal of Physical Chemistry B*, *111*, 10248-10252, **2007**.

[141] Chiariello, M.G. and Rega, N. Exploring Nuclear Photorelaxation of Pyranine in Aqueous Solution: an Integrated Ab-Initio Molecular Dynamics and Time Resolved Vibrational Analysis Approach. *The Journal of Physical Chemistry A*, *122*, 2884-2893, **2018**.

[142] Chiariello, M.G., Raucci, U., Coppola, F. and Rega, N. Unveiling anharmonic coupling by means of excited state ab initio dynamics: application to diarylethene photoreactivity. *Physical Chemistry Chemical Physics*, *21*, 3606-3614, **2019**.

[143] Haverkort, F., Stradomska, A., de Vries, A.H. and Knoester, J. Investigating the structure of aggregates of an amphiphilic cyanine dye with molecular dynamics simulations. *The Journal of Physical Chemistry B*, *117*, 5857-5867, **2013**.

[144] Branduardi, D., Bussi, G. and Parrinello, M. Metadynamics with adaptive Gaussians. *Journal of chemical theory and computation*, *8*, 2247-2254, **2012**.

[145] De Meyer, T., Ensing, B., Rogge, S.M., De Clerck, K., Meijer, E.J. and Van Speybroeck, V. Acidity Constant (pK a) Calculation of Large Solvated Dye Molecules: Evaluation of Two Advanced Molecular Dynamics Methods. *ChemPhysChem*, *17*, 3447, **2016**.

[146] Allen, M. P., and Tildesley, D. J. Computer Simulation of Liquids, Clarendon, Oxford **1987**.

[147] Georg, H.C., Coutinho, K. and Canuto, S. Solvent effects on the UV-visible absorption spectrum of benzophenone in water: A combined Monte Carlo quantum mechanics study including solute polarization. *The Journal of chemical physics*, *126*, p.034507, **2007**.

[148] Coutinho, K. and Canuto, S. Solvent effects in emission spectroscopy: A Monte Carlo quantum mechanics study of the n← π* shift of formaldehyde in water. *The Journal of Chemical Physics*, *113*, 9132-9139, **2000**.

[149] Oliveira, L.B., Prado, R.C., Júnior, L.A. and Colherinhas, G. The influence of flexibility on the spectroscopic properties for organic molecules in solution: A theoretical study applied to A3R polypeptide. *Journal of Molecular Liquids*, *263*, 334-341, **2018**.

[150] Ditchfield, R.H.W.J., Hehre, W.J. and Pople, J.A., 1971. Self-consistent molecular-orbital methods. IX. An extended Gaussian-type basis for molecular-orbital studies of organic molecules. *The Journal of Chemical Physics*, *54*, 724-728, **1971**.

[151] Grimme, S., Antony, J., Ehrlich, S. and Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *The Journal of chemical physics*, *132*, 154104, **2010**.

[152] Scalmani, G. and Frisch, M.J. Continuous surface charge polarizable continuum models of solvation. I. General formalism. *The Journal of chemical physics*, *132*, 114110, **2010**.

[153] Würthner, F., Yao, S., Debaerdemaeker, T. and Wortmann, R. Dimerization of merocyanine dyes. Structural and energetic characterization of dipolar dye aggregates and implications for nonlinear optical materials. *Journal of the American Chemical Society*, *124*, 9431-9447, **2002**.

[154] Wang, J., Wolf, R.M., Caldwell, J.W., Kollman, P.A. and Case, D.A. Development and testing of a general amber force field. *Journal of computational chemistry*, *25*, 1157-1174, **2004**.

[155] Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics, J Comp Phys, *117*, 1, **1995**.

[156] Jacquemin, D. and Adamo, C. Bond length alternation of conjugated oligomers: wave function and DFT benchmarks. *Journal of chemical theory and computation*, *7*, 369-376, **2011**.

[157] Wykes, M., Su, N.Q., Xu, X., Adamo, C. and Sancho-Garcia, J.C. Double hybrid functionals and the Π-system bond length alternation challenge: rivaling accuracy of post-HF methods. *Journal of chemical theory and computation*, *11*, 832-838, **2015**.

[158] Jacquemin, D., Zhao, Y., Valero, R., Adamo, C., Ciofini, I. and Truhlar, D.G. Verdict: time-dependent density functional theory "not guilty" of large errors for cyanines. *Journal of chemical theory and computation*, *8*, 1255-1259, **2012**.

[159] Ma, J., Su, T., Li, M.D., Du, W., Huang, J., Guan, X. and Phillips, D.L. How and when does an unusual and efficient photoredox reaction of 2-(1-hydroxyethyl) 9, 10-anthraquinone occur? A combined time-resolved spectroscopic and DFT study. *Journal of the American Chemical Society*, *134*, 14858-14868, **2012**.

[160] Neufeind, S., Hülsken, N., Neudörfl, J.M., Schlörer, N. and Schmalz, H.G. Total Synthesis of cyclo-Mumbaistatin Analogues through Anionic Homo-Fries Rearrangement. *Chemistry–A European Journal*, *17*, 2633-2641, **2011**.

[161] Elkazaz, S. and Jones, P.B., 2010. Photochemical Hydroxylation of 1-Methyl-9, 10-anthraquinones: Synthesis of 9'-Hydroxyaloesaponarin II. *The Journal of organic chemistry*, *75*, 412-416, **2010**.

[162] Sarma, S.J. and Jones, P.B. Photochemistry of 1, n-Dibenzyloxy-9, 10-anthraquinones. *The Journal of organic chemistry*, *75*, 3806-3813, **2010**.

[163] Wagner, P.J. Type II photoelimination and photocyclization of ketones. *Accounts of Chemical Research*, *4*, 168-177, **1971**.

[164] Liu, N. and Sun, G.Production of reactive oxygen species by photoactive anthraquinone compounds and their applications in wastewater treatment. *Industrial & Engineering Chemistry Research*, *50*, 5326-5333, **2011**.

[165] Font-Sanchis, E., Miranda, M.A., Pérez-Prieto, J. and Scaiano, J.C. Laser Flash Photolysis of [3, n] Paracyclophan-2-ones. Direct Observation and Chemical Behavior of 4, 4 '-(1, n-Alkanediyl) bisbenzyl Biradicals. *The Journal of organic chemistry*, *67*, 6131-6135, **2002**.

[166] Moorthy, J.N., Samanta, S., Koner, A.L., Saha, S. and Nau, W.M. Intramolecular O− H··· O Hydrogen-Bond-Mediated Reversal in the Partitioning of Conformationally Restricted Triplet 1, 4-Biradicals and Amplification of Diastereodifferentiation in Their Lifetimes. *Journal of the American Chemical Society*, *130*, 13608-13617, **2008**.

[167] Halgren, T.A. and Lipscomb, W.N. The synchronous-transit method for determining reaction pathways and locating molecular transition states. *Chemical Physics Letters*, *49*, 225-232, **1977**.

[168] Le Bahers, T., Adamo, C. and Ciofini, I. A qualitative index of spatial extent in charge-transfer excitations. *Journal of chemical theory and computation*, *7*, 2498-2506, **2011**.

[169] Marian, C.M. Spin–orbit coupling and intersystem crossing in molecules. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, *2*, 187-203, **2012**.

[170] Shi, H., Song, L., Ma, H., Sun, C., Huang, K., Lv, A., Ye, W., Wang, H., Cai, S., Yao, W. and Zhang, Y. Highly efficient ultralong organic phosphorescence through intramolecular-space heavy-atom effect. *The journal of physical chemistry letters*, *10*, 595-600, **2019**.

[171] Heß, B.A., Marian, C.M., Wahlgren, U. and Gropen, O. A mean-field spin-orbit method applicable to correlated wavefunctions. *Chemical Physics Letters*, *251*, 365-371, **1996**.

[172] Alberto, M.E., De Simone, B.C., Mazzone, G., Quartarolo, A.D. and Russo, N. Theoretical determination of electronic spectra and intersystem spin–orbit coupling: the case of isoindole-BODIPY dyes. *Journal of chemical theory and computation*, *10*, 4006-4013, b.

[173] DALTON: A Molecular Electronic Structure Program, http://daltonprogram.org, **2011**.

[174] Simone, B.C.D., Mazzone, G., Russo, N., Sicilia, E. and Toscano, M. Metal atom effect on the photophysical properties of Mg (II), Zn (II), Cd (II), and Pd (II) tetraphenylporphyrin complexes proposed as possible drugs in photodynamic therapy. *Molecules*, *22*, 1093, **2017**.

[175] Alberto, M.E., Pirillo, J., Russo, N. and Adamo, C. Theoretical exploration of type I/Type II dual photoreactivity of promising Ru (II) dyads for PDT approach. *Inorganic chemistry*, *55*, 11185-11192, **2016**.

[176] Bedini, A., De Laurentiis, E., Sur, B., Maurino, V., Minero, C., Brigante, M., Mailhot, G. and Vione, D. Phototransformation of anthraquinone-2-sulphonate in aqueous solution. *Photochemical & Photobiological Sciences*, *11*, 1445-1453, **2012**.

[177] Maurino, V., Bedini, A., Borghesi, D., Vione, D. and Minero, C. Phenol transformation photosensitised by quinoid compounds. *Physical Chemistry Chemical Physics*, *13*, 11213-11221, **2011**.

[178] Maschietto, F., Sanz García, J., Campetella, M. and Ciofini, I. Using density based indexes to characterize excited states evolution. *Journal of computational chemistry*, *40*, 650-656, **2019**.

[179] Perfetto, A., Maschietto, F. and Ciofini, I., 2019. Following excited states in molecular systems using density-based indexes: A dual emissive system as a test case. *Journal of Photochemistry and Photobiology A: Chemistry*, *383*, 111978, **2019**.

[180] Min, D.B. and Boff, J.M. Chemistry and reaction of singlet oxygen in foods. *Comprehensive reviews in food science and food safety*, *1*, 58-72, **2002**.

[181] Gritsan, N.P., Khmelinski, I.V. and Usov, O.M. Experimental and theoretical study of photoenolization mechanism for 1-methylanthraquinone. *Journal of the American Chemical Society*, *113*, 9615-9620, **1991**.

[182] Wagner, P.J., Kelso, P.A., Kemppainen, A.E., McGrath, J.M., Schott, H.N. and Zepp, R.G. Type II photoprocesses of phenyl ketones. Glimpse at the behavior of 1, 4 biradicals. *Journal of the American Chemical Society*, *94*(21), pp.7506-7512, **1972**.

RÉSUMÉ

Le présent travail se concentre sur le développement et l'application de protocoles de calcul à faible coût pour prédire et étudier les caractéristiques spectrales et les transformations chimiques photoinduites des colorants en solution. Le but de notre étude est de développer des outils informatiques pour comprendre et modéliser des processus complexes qui aideront la société PILI dans son objectif de concevoir de nouveaux colorants écologiques.
Brièvement, le Chapitre 1 et le Chapitre 2 se concentrent sur le sujet de la thèse et sur les méthodes. Ensuite, nous proposons trois procédures conçues pour les chromophores organiques rigides et flexibles, à savoir les anthraquinones, les indigos et les mérocyanines. Pour les deux premières classes moléculaires, nous présentons un schéma de calcul pour obtenir les couleurs de ces colorants en modélisant leurs spectres d'absorption au moyen de méthodes empiriques et non empiriques. Ce sera le sujet du Chapitre 3.
Pour les colorants mérocyanines, nous essayons traiter la complexité entre la structure et les propriétés de cette classe de systèmes flexibles en utilisant une approche multi-échelle basée sur l'échantillonnage de Monte Carlo et des méthodes de structure électronique. Cette approche sera discutée au Chapitre 4.
Dans le Chapitre 5, nous modélisons une réaction photochimique par interpolation linéaire de transit synchrone et nous suivons l'évolution de l'état excité en utilisant à la fois des indices basés sur la densité et des données de spin-orbite afin de décrire quantitativement le mécanisme de réaction dans le but d'améliorer la synthèse des colorants anthraquinones.
Enfin, dans le Chapitre 6, nous tirons des conclusions et des perspectives.
Dans l'ensemble, avec le protocole proposé, nous sommes en mesure de surmonter en partie certains défis informatiques clés pour prédire les signatures spectroscopiques des colorants industriels. Les schémas développés offrent une stratégie robuste, fiable et rentable pour la modélisation computationnelle des processus qui impliquent les états excités.

MOTS CLÉS

TD-DFT, Monte Carlo, méthode multi-échelle, prédiction de couleur, réaction photochimique, molécules rigides et flexibles.

ABSTRACT

The present work focuses on the development and application of low-cost computational protocols for predicting and studying spectral features and photoinduced chemical transformations of dyes in solution. The purpose of our study is to develop computational tools for understanding and modelling complex processes that will support the PILI company in its goal of designing new eco-friendly dyes.
Briefly, Chapter 1 and Chapter 2 focus on the thesis topic and on the methods. After that, we propose three procedures designed for both rigid and flexible metal-free chromophores, namely anthraquinones, indigos and merocyanines. For the first two molecular classes we present computational scheme for obtaining the colors of these dyes by modelling their absorption spectra by means of empirical and non-empirical methods. This will be the topic of Chapter 3.
For the merocyanine dyes we try to disentangle the complex node between structure and properties of this class of flexible systems employing a multiscale approach based on Monte Carlo sampling and electronic structure methods. This will be discussed in Chapter 4.
In Chapter 5, we model a photochemical reaction via linear synchronous transit interpolation and we follow the excited state evolution by employing both the density based indexes and spin-orbit data in order to quantitatively describe the reaction mechanism with the goal of improving the synthesis of anthraquinone dyes.
Finally, in Chapter 6, we draw conclusions and perspectives.
Overall, with the protocol proposed we are able to partially overcome some relevant computational challenges for predicting the spectroscopic signatures of industrial dyes. The schemes developed offer a robust, reliable and cost-effective strategy for the computational modelling of processes which involve the excited states.

KEYWORDS

TD-DFT, Monte Carlo, multiscale method, color prediction, photochemical reaction, rigid and flexible dyes