EXAT: EXcitonic Analysis Tool

Sandro Jurinovich[*], Lorenzo Cupellini[†], Ciro A. Guido[‡], Benedetta Mennucci[§]

November 4, 2017

Abstract

We introduce EXAT: EXcitonic Analysis Tool, a program able to compute optical spectra of large excitonic systems directly from the output of quantum mechanical calculations performed with the popular Gaussian 16 package. The software is able to combine in an excitonic scheme the single-chromophore properties and exciton couplings to simulate energies, coefficients and excitonic spectra (UV-vis, CD and LD). The effect of the environment can also be included using a Polarizable Continuum Model. EXAT also presents a simple graphical user interface, which shows on-screen both site and exciton properties. To show the potential of the method, we report two applications on a a chiral perturbed BODIPY system and DNA G-quadruplexes, respectively. The program is available online at http://molecolab.dcci.unipi.it/tools/.

Keywords: Circular Dichroism, Excitonic,Coupling, multichromophoric system   ■

---

[*]Dipartimento di Chimica, Università di Pisa, Via G. Moruzzi 13, 56124, Pisa, Italy

[†]Dipartimento di Chimica, Università di Pisa, Via G. Moruzzi 13, 56124, Pisa, Italy, email: lorenzo.cupellini@for.unipi.it

[‡]Laboratoire CEISAM - UMR CNRS 6230, Université de Nantes, 2 Rue de la Houssinire, BP 92208, 44322 Nantes Cedex 3, France

[§]Dipartimento di Chimica, Università di Pisa, Via G. Moruzzi 13, 56124, Pisa, Italy

EXAT
EXcitonic **A**nalysis **T**ool

EXAT: EXcitonic Analysis Tool is a program able to compute optical spectra of large excitonic systems directly from the output of quantum mechanical calculations performed with the popular Gaussian 16 package. The software combines the excitonic properties of single chromophores and their exciton couplings to simulate exciton-coupled optical spectra.

INTRODUCTION

The optical spectra of multichromophoric systems are often determined by the interaction between excitations localized on different chromophores, through the mechanism of exciton coupling. As a result the spectra largely depend on the spatial arrangement of the chromophores. In particular, circular dichroism (CD) spectra are extremely sensitive to three-dimensional disposition of the interacting chromophoric units, both as it regards their distances and the relative orientations. For this reason, CD spectroscopy is a powerful technique to study the structure of supramolecular systems,[1–5] to determine condensed-phase molecular conformation,[6] and to assign absolute configurations.[7–10]

To relate the CD spectrum of chirally arranged chromophores with their structure, and determine their absolute configurations, the so-called exciton chirality method (ECM) has been widely and successfully used.[11–15] The application of the ECM to dimeric systems results in a simple chirality rule, which relates the sign of the CD Cotton effect to the sense of twist between the transition dipole moments of the two chromophores.[2,4,16,17] However, in more complex cases, where many transitions are coupled, the interpretation of CD signatures is less straightforward.[18,19] In these situations, applying the ECM requires a detailed analysis of the properties of the individual subunits, namely the excitation energies and the transition dipoles, and the excited-state interaction between them, namely the exciton coupling.

The first formulation of the exciton method in a general matrix form has been proposed by Schellman et al. and applied to the calculation of the rotatory strength of a dipeptide.[20] Since then the exciton model has been applied to systems of increasing complexity and in particular it has been successfully used to predict the CD spectra of proteins,[21–24] pigment-protein complexes[25–27] and nucleic acids.[28–31] In most applications, quantum mechanical (QM) methods have been employed to determine the fundamental quantities of the exciton model, with various approaches for the calculation of exciton couplings,[21,23,32–40] and possibly extending the model to charge transfer states.[41] The QM version of the exciton model is particularly interesting as, in principle, it is parameter-free; however, the quality of the results strongly depends on the level of QM theory adopted to describe the subunits and their interactions.

A further aspect to consider when simulating CD spectra is the role of the environment surrounding the chromophores. Even if it does not interact excitonically with it, the environment exerts a twofold effect on the excitonic system. First of all, it can polarize the chromophores, thus changing their excitation energies and transition dipole moments. Secondly, it acts as a dielectric in screening the electronic couplings.[42] In this way, all the "ingredients" of the exciton model, that is, site energies, couplings, and transition dipoles, largely depend on the environment. Among the approaches to account for the environment effects in electronic structure calculations, the Polarizable Continuum Model (PCM)[43] is a well established method that treats the environment as a dielectric continuum, characterized by its bulk dielectric constant and refractive index. The QM system of interest is embedded in a cavity that follows its (supra)molecular shape, and the response of the environment is represented by a set of induced charges spreading on the cavity surface. Despite its simplicity, PCM can describe both the polarization effects on the site properties and the screening effects on the electronic couplings,[33,44] thus giving a balanced description of the exciton Hamiltonian.

Here, we present the EXAT (EXcitonic Analysis Tool) program which is based on such a QM exciton model and allows one to compute optical spectra of excitonic systems directly from the output of QM (/PCM) calculations performed with the popular Gaussian (R) 16 package.[45] EXAT eases the analysis of exciton properties also allowing a straightforward interpretation of exciton coupled spectra, especially in the case of large supramolecular systems.

METHODOLOGY

Theory

Here we present the exciton model at the basis of EXAT. Let us consider a system composed by N chromophoric units, each bearing a number n_i of excited states. The units are assumed to be independent, namely there is no electron exchange between different groups, but the groups are electronically coupled in the excited state. Within this approximation, the excited

states $|K\rangle$ of the entire system can be expressed as a linear combination of singly excited states of the chromophores:

$$|K\rangle = \sum_{i}^{N}\sum_{a}^{n_i} C_{ia}^{K} |ia\rangle \tag{1}$$

where the first sum runs over the chromophores while the second sum runs over all the n_i excited states of the chromophore i. For the expansion in eq. 1, the Hamiltonian of the system reads:

$$\hat{H}_{ex} = \sum_{i}^{N}\sum_{a}^{n_i} \mathcal{E}_{ia} |ia\rangle\langle ia| + \sum_{i}^{N}\sum_{j}^{N}\sum_{a}^{n_i}\sum_{b}^{n_j} V_{ij}^{ab} |ia\rangle\langle jb| \tag{2}$$

where $\mathcal{E}_{ia}$ is the a-th excitation energy of the isolated molecule i, namely $\langle ia|\hat{H}|ia\rangle$, and V_{ij}^{ab} is the electronic coupling between two different units i and j in their relative excited states a and b, namely $\langle ia|\hat{H}|jb\rangle$. The excitation energies of the isolated chromophores, $\mathcal{E}_{ia}$, are also known as *site energies*, and are easily accessible through an excited-state electronic structure calculation.

The electronic couplings V_{ij}^{ab} are often divided in Coulomb and short-range interactions, the latter being often negligible because of the exponential decay with distance.[46,47] The Coulomb term is the electrostatic interaction between the transition densities corresponding to the excitations $0 \rightarrow a$ and $0 \rightarrow b$ in chromophores i and j. In the literature, it is often approximated with the interaction between the transition dipoles of i and j, but this approximation is only valid when the distance between i and j is much larger than the chromophores themselves. In our approach, the electronic couplings V_{ij}^{ab} is calculated directly from the transition densities[33,46] through the integral:

$$V_{ij}^{ab,Coul} = \int d\mathbf{r}_1 d\mathbf{r}_2 \, \rho_{i,0a}^{tr*}(\mathbf{r}_1)\frac{1}{r_{12}}\rho_{j,0b}^{tr}(\mathbf{r}_2) \tag{3}$$

where, $\rho_{i,0a}^{tr}$ and $\rho_{j,0b}^{tr}$ are the transition densities of excitations $0 \rightarrow a$ and $0 \rightarrow b$ in chromophores i and j.

In condensed phase, the polarizable medium that surrounds the chromophores results in an additional term, which can be regarded as the interaction between transition densities mediated by the dielectric response of the environment V_{env}. This formalism has been implemented in Gaussian 16,[45] introducing the mediated term within the PCM description of

5

the solvent, namely:[33]

$$V_{ij}^{ab,\text{env}} = \sum_t \left[\int d\mathbf{r} \rho_{i,0a}^{tr*}(\mathbf{r}) \frac{1}{|\mathbf{r} - \mathbf{r}_t|} \right] q_t(\epsilon_\infty; \rho_{j,0b}^{tr}) \tag{4}$$

where index t runs over the set of surface charges, q_t, induced by the density $\rho_{j,0b}^{tr}$ and mediated by the optical dielectric permittivity ϵ_∞.

The eigenvalues of the Hamiltonian of eq. 2 represent the excitation energies of the entire system, and the corresponding eigenvectors represent expansion coefficients C_{ia}^K of the exciton wavefunction on the basis of the localized states $|ia\rangle$. Within the same formalism, the transition probability of the UV-Vis absorption is expressed in terms of the squared transition electric dipole moment of the excitonic state $\boldsymbol{\mu}_{0K}$, which can be expanded over the transition dipole moments of the localized excited states:

$$\boldsymbol{\mu}_{0K} = \langle 0|\hat{\boldsymbol{\mu}}|K\rangle = \sum_i^N \sum_a^{n_i} C_{ia}^K \langle i0|\hat{\boldsymbol{\mu}}|ia\rangle = \sum_i^N \sum_a^{n_i} C_{ia}^K \boldsymbol{\mu}_{i0a} \tag{5}$$

As a result, the absorption spectrum due to M transitions can be obtained as a sum over the exciton states:[16]

$$\epsilon(\tilde{\nu}) = \frac{8\pi^3 N_A}{3000 \ln(10) hc} \tilde{\nu} \sum_k^M |\boldsymbol{\mu}_{0K}|^2 S(\tilde{\nu} - \tilde{\nu}_K) = 1.089 \times 10^{38} \, \tilde{\nu} \sum_k^M |\boldsymbol{\mu}_{0K}|^2 S(\tilde{\nu} - \tilde{\nu}_K) \tag{6}$$

where $\tilde{\nu}$ is the absorption wavenumber, N_A is the Avogadro number, and the transition dipole is expressed in $c.g.s.$ units ($i.e.$ esu·cm). $S(\tilde{\nu})$ is a normalized lineshape function that accounts for homogeneous and inhomogeneous broadening, and can be set $ad\text{-}hoc$ for each transition in order to reproduce experimental bandwidths. A common choice is a Gaussian lineshape:

$$S(\tilde{\nu}) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\tilde{\nu}}{\sigma}\right)^2} \tag{7}$$

In eq. 7 σ is the standard deviation of the Gaussian function. Usually, the broadening of the absorption bands is measured as half-width at half maximum (HWHM $= \sigma\sqrt{2ln2}$) of the peak.

In CD spectroscopy, the measured quantity is the differential molar absorption $\Delta\epsilon = \epsilon^l - \epsilon^r$ between left and right circularly-polarized light. The rotatory strength R represents the sign and the intensity of a CD signal. The rotatory strength for the generic electronic transition

from the ground state 0 to the excited state K, is given by the Rosenfeld equation:

$$R_{0K} = \Im \langle 0 \, | \hat{\boldsymbol{\mu}} | \, K \rangle \cdot \langle K \, | \hat{\mathbf{m}} | \, 0 \rangle \tag{8}$$

where $\Im$ denotes the imaginary part, and the $\hat{\boldsymbol{\mu}}$ and $\hat{\mathbf{m}}$ are the operators of electric and magnetic moment vectors respectively. The magnetic moment can be expanded in the same way as in eq. (5); the magnetic moments $\mathbf{m}_{i0a}$ of the site transitions depend on the position $\mathbf{R}_i$ of chromophore i, as

$$\mathbf{m}_{i0a} = \mathbf{m}_{i0a}^{\text{int}} + \frac{ie\hbar}{2m_e c} \mathbf{R}_i \times \boldsymbol{\nabla}_{i0a} \tag{9}$$

where $\mathbf{m}_{i0a}^{\text{int}}$ is the intrinsic magnetic moment of the transition $0 \to a$ in chromophore i.[48] If the intrinsic magnetic moments of the chromophores are negligible, eq. (8) reduces to:[49]

$$R_{0K} = -\frac{\pi \tilde{\nu}_{0K}}{2c} \sum_{i,j}^{N} \sum_{a,b}^{n_i, n_j} C_{ia}^{K} C_{jb}^{K} \mathbf{R}_{ij} \cdot (\boldsymbol{\mu}_{i0a} \times \boldsymbol{\mu}_{j0b}) \tag{10}$$

where $\tilde{\nu}_{0K}$ is the excitation wavenumber of state K and $\mathbf{R}_{ij} = \mathbf{R}_j - \mathbf{R}_i$ is the distance vector between the chromophore i and j. Notably, this formulation is gauge-invariant in that the result does not change if the system is translated. However, eq. (10) does not account for the presence of the intrinsic transition magnetic moments $\mathbf{m}_{i0a}^{\text{int}}$.

A gauge-invariant expression that accounts for the magnetic transition moments can be obtained by passing to the velocity formulation of the electric dipole:[48]

$$R_{0K} = -\frac{e\hbar}{2\pi m_e \tilde{\nu}_{0K}} \Im \left\{ \sum_{i,j}^{N} \sum_{a,b}^{n_i, n_j} C_{ia}^{K} C_{jb}^{K} \left[\boldsymbol{\nabla}_{i0a} \cdot \mathbf{m}_{j0b} \right] \right\} \tag{11}$$

where $\boldsymbol{\nabla}_{i0a}$ represents the electric dipole moment expressed in the velocity formulation. Equation (11) does not change if the origin on the coordinate system is translated. Moreover, by decomposing the magnetic transition dipole moment as in eq. (9), we can divide R_{0K} into two terms:

$$R_{0K} = -\frac{e\hbar}{2\pi m_e \tilde{\nu}_{0K}} \Im \left\{ \sum_{i,j}^{N} \sum_{a,b}^{n_i, n_j} C_{ia}^{K} C_{jb}^{K} \left[\boldsymbol{\nabla}_{i0a} \cdot \mathbf{m}_{j0b}^{\text{int}} \right] + \sum_{i,j}^{N} \sum_{a,b}^{n_i, n_j} C_{ia}^{K} C_{jb}^{K} \left[\boldsymbol{\nabla}_{i0a} \cdot \frac{ie\hbar}{2m_e c} \mathbf{R}_j \times \boldsymbol{\nabla}_{j0b} \right] \right\}$$

$$= R_{0K}^{\text{int}} + R_{0K}^{\text{ext}}$$

$$\tag{12}$$

Here R_{0K}^{int} is the intrinsic rotatory strength due to the vectorial combination of the intrinsic transition magnetic dipole moments, whereas R_{0K}^{ext} is the rotatory strength coming from displaced electric dipoles, similarly to eq. (10). In literature, R_{0K}^{ext} and R_{0K}^{int} are sometimes also denoted as "$\mu\mu$" and "μm" terms respectively.[20,22,48]

The CD spectrum can be finally obtained by assuming a lineshape function, similarly to what discussed in the case of the absorption:[16]

$$\Delta\epsilon(\tilde{\nu}) = \frac{32\pi^3 N_A}{3000\ln(10)hc}\,\tilde{\nu}\sum_K^M R_{0K}S(\tilde{\nu}-\tilde{\nu_K}) = 4.355\times10^{38}\,\tilde{\nu}\sum_K^M R_{0K}S(\tilde{\nu}-\tilde{\nu_K}) \qquad (13)$$

A widely used quantity in CD spectroscopy is the dissymmetry factor g_{abs}, which describes the ratio between the CD and absorption spectra:[1,16]

$$g_{abs,0K} = \frac{\Delta\epsilon(\tilde{\nu})}{\epsilon(\tilde{\nu})} = 4\frac{R_{0K}}{\mu_{0K}^2} \qquad (14)$$

This quantity is particularly useful for determining the chiral electronic response of a molecule or an aggregate. Experimentally, g_{abs} is determined at the CD peaks, and assumes the same lineshape for absorption and CD.

Finally, oriented aggregates have a differential response to linearly polarized light. This allows to use linear dichroism (LD) spectroscopy as a tool to measure the orientation of chromophores in an aggregate. LD is defined as the differential absorption of linearly polarized light parallel and perpendicular to the orientation axis, i.e. LD $= \epsilon_{\parallel} - \epsilon_{\perp}$.[50] The LD spectrum of a perfectly oriented excitonically coupled aggregate can be easily calculated from the excitonic transition dipole moments obtained as in eq. (5):

$$\text{LD}(\tilde{\nu}) = \frac{8\pi^3 N_A}{3000\ln(10)hc}\,\tilde{\nu}\sum_k^M \frac{3}{2}|\boldsymbol{\mu}_{0K}|^2\left(3\cos(\alpha_{0K})-1\right)S(\tilde{\nu}-\tilde{\nu}_K) \qquad (15)$$

where α_{0K} is the angle that the transition dipole $\boldsymbol{\mu}_{0K}$ forms with the orientation axis. EXAT allows to choose a general orientation axis to compute the LD. Note that this LD formula assumes that the molecular system is perfectly oriented; a parameter should be used to take into account possible fluctuations of the general orientation.[50]

EXAT workflow

The full computational strategy used to obtain the excitonic spectra can be summarized in the workflow illustrated in Fig. 1.[49] The geometry of the entire system may be obtained from different sources, such as crystallographic data, NMR solution structures, or modeled geometries. Experimental structures may need further refinement or optimization in order to be suitable for QM calculations.[49]
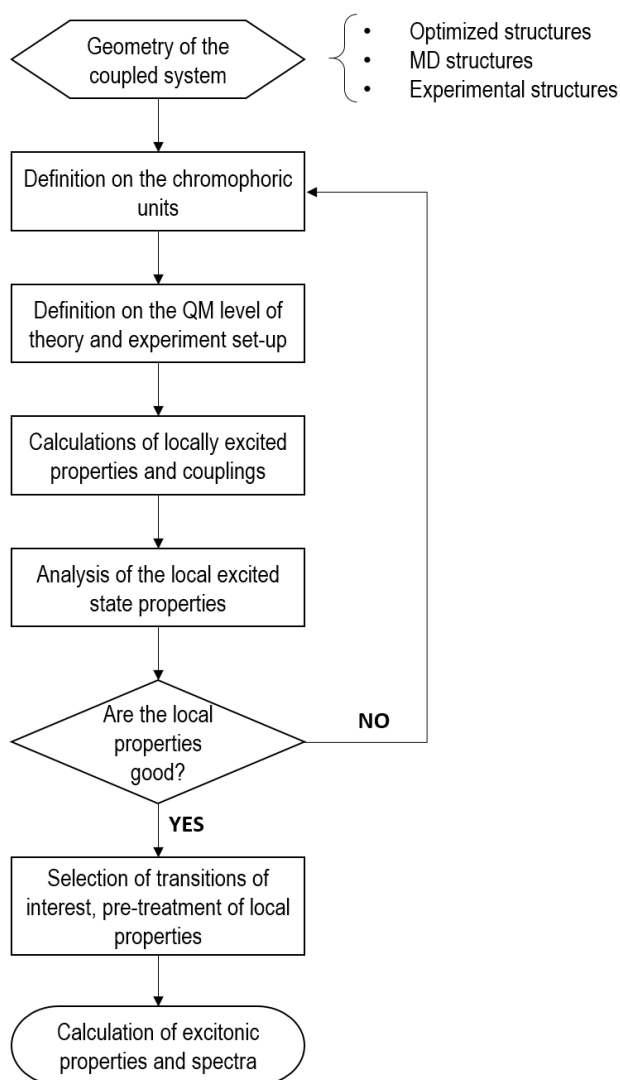

Figure 1: Workflow to simulate the excitonic spectra.

The whole interacting system needs to be split into chromophoric units. When the different chromophores are covalently bonded, one needs to cut covalent bonds and saturate

the free valence, commonly using a hydrogen atom. The definition of chromophoric units and the calculation of local properties and exciton couplings can be done in Gaussian 16[45] by adding the `eet(fragment=N)` keyword, where N is the number of fragments in which the entire system is divided. Once all local properties and couplings are collected, one can decide to select only some specific transitions of interest from each monomeric unit to build up the excitonic Hamiltonian. This selection can be very useful to decompose the excitonic spectra into different contributions coming from different interactions, or to eliminate spurious excitations. The Hamiltonian composed of the selected excitations is built from site energies and exciton couplings, and then diagonalized to yield the exciton energies E_K and the coefficients C_{i0a}^K. The local dipole moments are combined through eq. (5) to determine excitonic electric dipole moments, and through eq. (10) or (11) to determine the rotatory strengths. Finally, the absorption, CD and LD spectra can be computed by means of eqs. (6)-(15), respectively.

The workflow is implemented in a Python 2.6 package that requires the NUMPY and SCIPY libraries, and the MATPLOTLIB library for the visualization of the spectra. The program can be run from the command line, but a graphical user interface is also available (see below). EXAT can read the output of a Gaussian 16[45] EET calculation, and collect site energies, electric and magnetic transition dipoles, and electronic couplings. It is possible to modify these quantities and also to provide user-defined quantities. After the calculations, EXAT produces several output files, among which `results.out` contains the list of excitonic states, their energy, and the corresponding properties. This file can be used as an input to calculate absorption and CD spectra.

The program also outputs a geometry file and an input file for VMD,[51] that allows the user to visually analyze the local electric and magnetic transition dipoles.

EXAT is available online at http://molecolab.dcci.unipi.it/tools/.

PCM and excitonic properties

As reported in the introduction, environment effects can be introduced in EXAT using the PCM solvation model. One of the main ingredient of such a model is the cavity which

embeds the QM system. As here, the system is divided into subunits, there are two possible schemes to define the PCM cavity. In the "fragment cavity" scheme, the excitonic properties of each chromophore (site energy, transition density and transition dipoles), are calculated using the cavity of the chromophore itself. However, the couplings within each pair i, j of chromophores, are calculated with a cavity that follows the shape of i and j together. This cavity can be united or separated depending on the distance between i and j. On the contrary, the "common cavity" scheme always use the cavity defined on the entire system: this means that the calculation of both site properties and couplings is performed with the same "multichromophoric" cavity.

The specific approach used in defining the cavity influences both site energies and couplings in the exciton Hamiltonian. In particular, the fragment cavity scheme treats the other chromophores on the same basis as the rest of the environment, that is, with the same dielectric properties. For quite small molecular systems, where chromophores are close and directly bonded, the common cavity has to be preferred, while for larger systems a fragmented cavity helps one to recover for each chromophore the correct electrostatic effect of the rest of the system.

EXAT-GUI

EXAT includes a graphical user interface (GUI) implemented in python using the GTK-Glade libraries and the MATPLOTLIB library to display spectra and graphs. The GUI allows the user to directly load the output file of a Gaussian 16 `eet` calculation, showing on-screen the site and exciton properties. Different transitions can be selected, and it is possible to switch between the approximate formulation of eq. (10) and the full formulation of eq. (11). Finally, the excitonic absorption and CD spectra can be visualized in a secondary window, along with the stick spectrum that represents the position and intensity of every transition. An example of the GUI is reported in Figure 2.
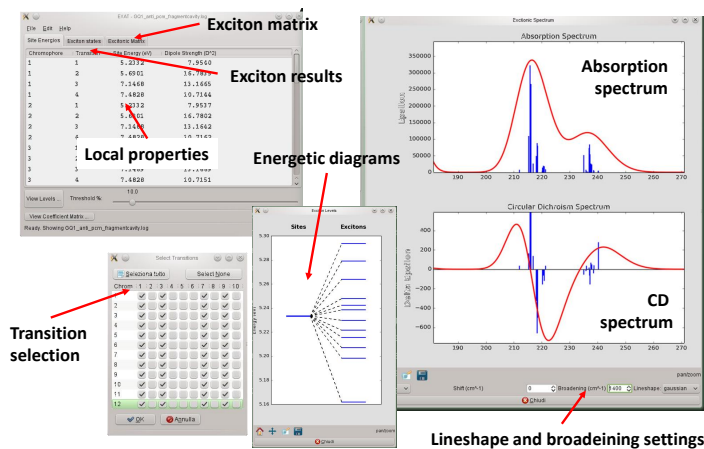
Figure 2: A representation of the EXAT graphical user interface.

RESULTS

In this section we report two applications of EXAT to highlight the potential of the program. In the first example we simulate the absorption and CD spectra of a chiral perturbed BODIPY system using the gauge-independent formulation, which includes the electric-magnetic term in the rotatory strength calculation. In the second example the CD spectra of different conformations of nucleic acid have been simulated showing the applicability of the EXAT in the simulation of large multichromophoric systems.

Chiral perturbation of a BODIPY

Boron dipyrromethene (BODIPY) complexes represent an important family of organic dyes with special absorption and emission properties that can be tuned by functionalization of the BODIPY core.[52] A dissymetrization of the BODIPY core achieved through geometrical strain can turn it into a chiral chromophore; moreover peripheral chromophoric units that present off-resonance transitions can chirally perturb the symmetric BODIPY chromophore. The latter mechanism can be considered as an excitonic perturbation of the BODIPY transition. As a test case we considered the enantiomeric (R)-spiroBODIPY recently synthesized and characterized by Sanchez-Carnerero *et al.* (see Figure 3a).[53] Even though the spiro-

BODIPY does not present a clear excitonic circular dichroism, it is possible to model the chiral perturbation through the EXAT excitonic scheme: the higher-lying excited states of the napthalene rings couple with the BODIPY-centered transition to create a chiral exciton state.

The spiroBODIPY geometry was obtained from a QM optimization at the B3LYP/6-311G(d) level of theory. The system was fragmented as shown in Figure 3a, saturating the free valences with hydrogen atoms. Excitation energies and couplings were computed at the PBE0/6-311+G(d,p) level of theory, including the solvent (Chloroform) through PCM. We computed six excited states for every fragment, in order to include the effect of higher lying states on the CD spectrum. The excitonic analysis was carried out using EXAT-GUI.

The optimized structure of the spiroBODIPY is shown in Figure 3b, where some of the main transition dipoles are depicted. In particular, we highlight the presence of a non-negligible magnetic transition dipole moment of the first BODIPY excited state (blue vector in Figure 3b).

The calculated absorption (OD) and CD spectra are reported in Figure 3c. Apart from a general $\sim$90 nm blue shift, the calculated spectra well reproduce the experiments.[53] In particular we focus on the longest-wavelength negative CD peak at 430 nm, which corresponds to a BODIPY-centered transition.[53] This transition is highly localized on the BOPDIY fragment ($C^2 > 0.998$), and weakly perturbed by all the states of the naphthalene moieties. Its rotatory strength, calculated with eq. (11), is $-63 \cdot 10^{-40}$ esu^{2}cm^2, which corresponds to a dissymmetry factor $g_{abs} = 4 \cdot 10^{-4}$, in qualitative agreement with the experimental value of $7 \cdot 10^{-4}$–$9 \cdot 10^{-4}$ depending on the enantiomer.[53] The approximate treatment of eq. (10), which neglects the magnetic-electric contribution to the CD, yields a much lower rotatory strength of -23$\cdot 10^{-40}$ cgs units ($g_{abs} = 1.5 \cdot 10^{-4}$). The CD signature in the BODIPY transition comes from the coupling of the magnetic transition dipole moment of the BOPDIY moiety to the bright electric transition dipole moments of the naphthalene moieties (See Figure 3b). Using EXAT, we can exclude selected naphthalene transitions to assess their effect on the rotatory strength of the BODIPY-centered transition. The exclusion of states 2, 3, and 5 of the naphthalenes does not change R significantly, while states 4 and 6 account for $\sim$ 41% and 33% of the total rotatory strength, respectively. In fact, states 4 and 6 are the
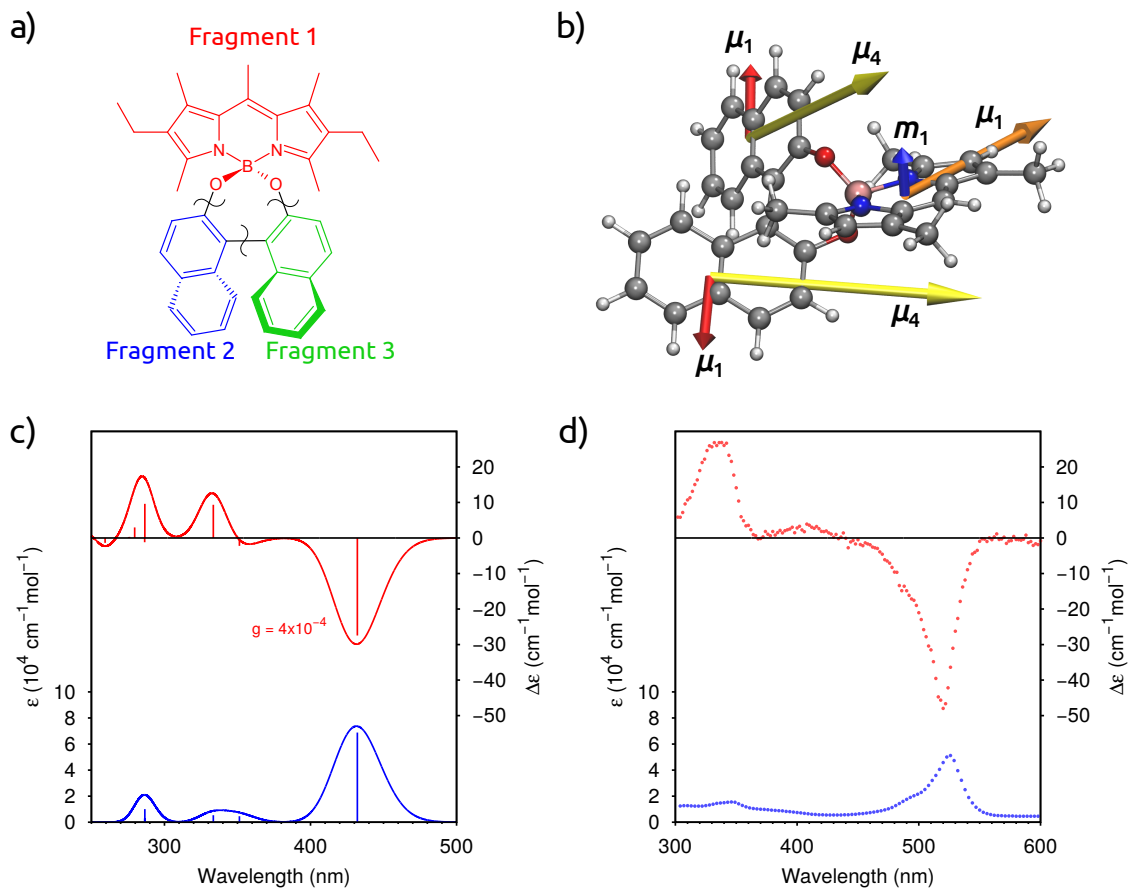
Figure 3: a) Chemical structure and fragmentation scheme of the chiral (R)-spiroBODIPY. b) Optimized structure of chiral (R)spiroBODIPY. The main electric ($\boldsymbol{\mu}$, in yellow, red and orange) and magnetic ($\mathbf{m}$, in blue) transition dipole moments of the fragments are also shown. $\boldsymbol{\mu}_6$ is omitted for clarity, but it lies on the same axis as $\boldsymbol{\mu}_1$. c) Simulated excitonic OD (blue) and CD (red) spectra with a Gaussian lineshape (HWHM $= 1000$ cm^{-1}). Vertical lines indicate the position and relative intensities of the transitions. d) Experimental OD (blue) and CD (red) spectra of (R)-spiroBODIPY taken from Ref. 53. Note the same y-axis scale in c) and d).

14

brightest states of the naphthalene moiety, as their transition dipole magnitudes are 8.0 D and 3.3 D respectively, while the first excitation has a transition dipole of 2.2 D only. The values of these transition dipoles contribute not only to the increase of the coupling with the BODIPY state, but also to the rotatory strength in both terms of eq. (12).

DNA G-quadruplexes

DNA and RNA guanine-rich sequences can fold into tetrahelical structures stabilized by hydrogen bonds between guanine tetrads and electrostatic interactions with monovalent cations. Such structures, denoted as G-quadruplexes, adopt various folding topologies depending on the specific sequence and folding conditions.[54] The topologies are classified in parallel and anti-parallel depending on the relative direction of the four guanine strands. Parallel and anti-parallel topologies are easily distinguished by CD spectroscopy, and show particular fingerprints: parallel G-quadruplexes are characterized by a positive band at 260 nm, whereas antiparallel G-quadruplexes have a positive band at 290 nm and a negative band at 260 nm.[55]

In order to demonstrate the potential of the EXAT workflow for larger systems, we show how the excitonic approach can be used to compute the CD spectra of two G-quadruplexes. We selected one parallel (PDB code: 2MB2)[56] and one anti-parallel (PDB code: 143D)[57] G-quadruplex, determined by solution NMR studies. In both structures the G-quadruplex core is formed by three guanine planes (see Figure 4).

We first refined the NMR structures by projecting the MP2/cc-pVDZ optimized geometry of the guanine base to the NMR structure.[29,31] Site properties and couplings were then computed at the M062X/6-31G(d) level, including the solvent (water) with PCM, in the fragment cavity approach, but scaling the atomic radii by 1.8 to account for the DNA backbone and the less polarizable waters in the first solvation shell.[58] Only the first four bright $\pi - \pi^*$ states of the guanine forming the G-tetrads were included in the exciton Hamiltonian. In order to correct the intrinsic error of the method here employed (*i.e.* the particular combination of TD-DFT, the used functional, and the basis set), we compared the obtained site energy of the first transition (5.23 eV) to the experimental average ($\sim$4.45 eV)[59]
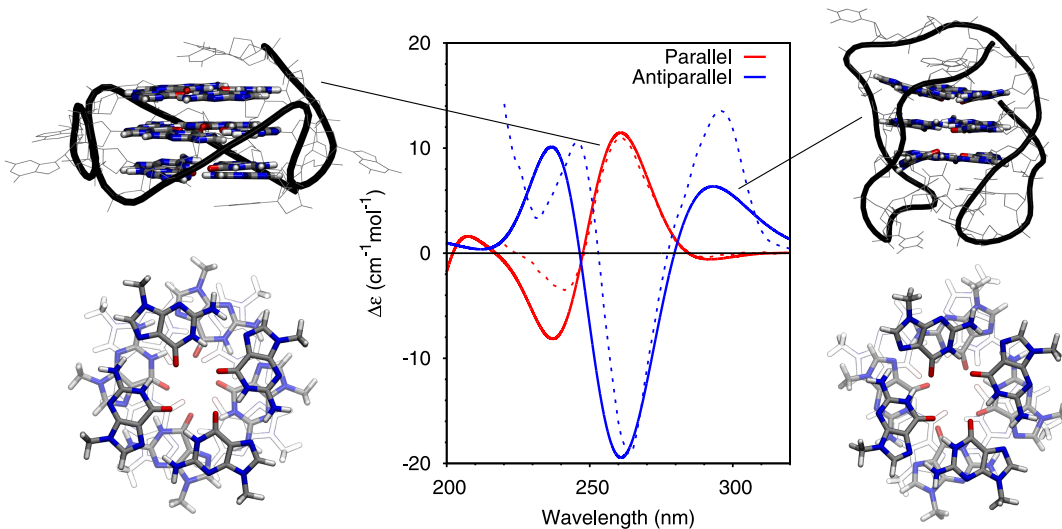
Figure 4: Parallel (left panel) and anti-parallel (right panel) G-quadruplex structures depicted in side and top view. Only guanine core bases are shown in the top view. The central panel shows CD spectra of the two structures. Both spectra were shifted by -0.78 eV and divided by the number of bases, to compare $\Delta\epsilon$ values with the experiment, where the molar concentration is expressed as concentration of nucleobases. Dashed lines represent the corresponding experimental spectra.[29,56]

and we used this difference to shift the resulting spectrum.

The calculated spectra, reported in Figure 4, show the main fingerprints of the parallel and antiparallel structures in the 200–320 nm range. Both spectra are in very good agreement with experiments on large G-quadruplex structures[55] as well as experiments on these two particular structures.[29,56]

It should be noted that fully *ab initio* calculations showed a non-negligible charge-transfer (CT) character for low-lying bright states of nucleic acids,[60] which is a consequence of CT states mixing with exciton states, as CT states themselves are optically forbidden. Such CT mixing is neglected in the excitonic approach here presented. However, the inclusion of CT states should not change qualitatively the result of the exciton model, as the main effect is a shift in the positions of the exciton states.[27,41]

SUMMARY AND FURTHER DEVELOPMENTS

The simulation of optical properties in large, multichromophoric systems can be greatly simplified by using an excitonic model, while retaining the accuracy of fully *ab initio* methods, provided that the system can be divided into independent (but excitonically coupled) units, and the mixing of CT states to exciton states can be neglected. We developed the EXAT program to build up the excitonic Hamiltonian reading all the needed "ingredients" directly from the output of calculations performed with Gaussian 16, and obtain the resulting excitonic spectra. The program also allows users to analyze the exciton properties of the system and understand the physical origin of characteristic signatures in its optical spectra. In addition to absorption spectra, EXAT is particularly useful in simulating and analyzing circular dichroism spectra, also including intrinsic magnetic moments in a gauge-independent formulation.[48] The graphical interface helps the user to interactively manage several aspects of the model and visualize spectra in real time. These features make EXAT suited for computational chemists as well as experimental spectroscopists.

Future developments of the EXAT framework will be directed to include an atomistic description of the environment of the chromophores through a polarizable QM/MM embedding,[35] and to define a robust model for systems, such as proteins, containing hundreds of chromophores. Another important improvement will be the inclusion of CT states in the exciton Hamiltonian,[41] to capture the effect of CT mixing especially in the case of closely packed chromophores. Further developments include models to account for a more realistic optical line shape and the coupling of vibronic transitions.[5,61,62]

ACKNOWLEDGMENTS

References

1. N. Berova, L. Di Bari, and G. Pescitelli, Chem. Soc. Rev. **36**, 914 (2007).

2. G. Pescitelli, L. Di Bari, and N. Berova, Chem. Soc. Rev. **43**, 5211 (2014).

3. G. Charalambidis, E. Georgilis, M. K. Panda, C. E. Anson, A. K. Powell, S. Doyle, D. Moss, T. Jochum, P. N. Horton, S. J. Coles, et al., Nat. Commun. **7**, 12657 (2016).

4. F. D. Lewis, L. Zhang, X. Liu, X. Zuo, D. M. Tiede, H. Long, and G. C. Schatz, J. Am. Chem. Soc. **127**, 14445 (2005).

5. F. C. Spano, J. Am. Chem. Soc. **131**, 4267 (2009).

6. F. Eker, K. Griebenow, and R. Schweitzer-Stenner, J. Am. Chem. Soc. **125**, 8178 (2003).

7. N. Berova, P. L. Polavarapu, K. Nakanishi, and R. W. Woody, *Comprehensive chiroptical spectroscopy, applications in stereochemical analysis of synthetic compounds, natural products, and biomolecules*, vol. 2 (John Wiley & Sons, Hoboken, NJ, USA, 2012).

8. J. Ściebura and J. Gawroski, Chem. Eur. J. **17**, 13138 (2011).

9. D. M. McCann and P. J. Stephens, J. Org. Chem. **71**, 6074 (2006).

10. A. Wakai, H. Fukasawa, C. Yang, T. Mori, and Y. Inoue, J. Am. Chem. Soc. **134**, 4990 (2012).

11. N. Harada, A. Saito, H. Ono, J. Gawronski, K. Gawronska, T. Sugioka, H. Uda, and T. Kuriki, J. Am. Chem. Soc. **113**, 3842 (1991).

12. W. Adam, Z. Lukacs, K. Viebach, H.-U. Humpf, C. R. Saha-Möller, and P. Schreier, J. Org. Chem. **65**, 186 (2000).

13. S. Superchi, D. Casarini, C. Summa, and C. Rosini, J. Org. Chem. **69**, 1685 (2004).

14. S. G. Telfer, T. M. McLean, and M. R. Waterland, Dalton Trans. **40**, 3097 (2011).

15. S. Tartaglia, F. Pace, P. Scafato, and C. Rosini, Org. Lett. **10**, 3421 (2008).

16. N. Berova, P. L. Polavarapu, K. Nakanishi, and R. W. Woody, *Comprehensive Chiroptical Spectroscopy: Instrumentation, Methodologies, and Theoretical Simulations*, vol. 1 (John Wiley & Sons, Hoboken, NJ, USA, 2012).

17. N. Harada, K. Nakanishi, and N. Berova, *Electronic CD Exciton Chirality Method: Principles and Applications* (John Wiley & Sons, Inc., 2012), pp. 115–166.

18. T. Hattori, K. Sakurai, N. Koike, S. Miyano, H. Goto, F. Ishiya, and N. Harada, J. Am. Chem. Soc. **120**, 9086 (1998).

19. S. Superchi, E. Giorgio, and C. Rosini, Chirality **16**, 422 (2004).

20. J. A. Schellman, Acc. Chem. Res. **1**, 144 (1968).

21. N. A. Besley and J. D. Hirst, J. Am. Chem. Soc. **121**, 9636 (1999).

22. N. Sreerama and R. W. Woody, Methods Enzymol. **383**, 318 (2004).

23. M. T. Oakley, B. M. Bulheller, and J. D. Hirst, Chirality **18**, 340 (2006).

24. B. M. Bulheller, A. Rodger, and J. D. Hirst, Phys. Chem. Chem. Phys. **9**, 2020 (2007).

25. M. Schmidt am Busch, F. Müh, M. El-Amine Madjet, and T. Renger, J. Phys. Chem. Lett. **2**, 93 (2011).

26. S. Jurinovich, L. Viani, I. G. Prandi, T. Renger, and B. Mennucci, Phys. Chem. Chem. Phys. **17**, 14405 (2015).

27. L. Cupellini, S. Jurinovich, M. Campetella, S. Caprasecca, C. A. Guido, S. M. Kelly, A. T. Gardiner, R. Cogdell, and B. Mennucci, J. Phys. Chem. B **120**, 11348 (2016).

28. S. Patwardhan, S. Tonzani, F. D. Lewis, L. D. A. Siebbeles, G. C. Schatz, and F. C. Grozema, J. Phys. Chem. B **116**, 11447 (2012).

29. D. Loco, S. Jurinovich, L. Di Bari, and B. Mennucci, Phys. Chem. Chem. Phys. **18**, 866 (2016).

30. H. Gattuso, X. Assfeld, and A. Monari, Theor. Chem. Acc. **134** (2015).

31. D. Padula, S. Jurinovich, L. Di Bari, and B. Mennucci, Chem. Eur. J. **22**, 17011 (2016).

32. B. P. Krueger, G. D. Scholes, and G. R. Fleming, J. Phys. Chem. B **102**, 5378 (1998).

33. M. F. Iozzi, B. Mennucci, J. Tomasi, and R. Cammi, J. Chem. Phys. **120**, 7029 (2004).

34. M. E. Madjet, A. Abdurahman, and T. Renger, J. Phys. Chem. B **110**, 17268 (2006).

35. C. Curutchet, A. Muñoz-Losa, S. Monti, J. Kongsted, G. D. Scholes, and B. Mennucci, J. Chem. Theory Comput. **5**, 1838 (2009).

36. A. Sisto, D. R. Glowacki, and T. J. Martinez, Acc. Chem. Res. **47**, 2857 (2014).

37. A. F. Morrison, Z.-Q. You, and J. M. Herbert, J. Chem. Theory Comput. **10**, 5366 (2014).

38. A. Sisto, C. Stross, M. W. van der Kamp, M. OConnor, S. McIntosh-Smith, G. T. Johnson, E. G. Hohenstein, F. R. Manby, D. R. Glowacki, and T. J. Martínez, Phys. Chem. Chem. Phys. pp. 5–12 (2017).

39. A. A. Kocherzhenko, X. A. Sosa Vazquez, J. M. Milanese, and C. M. Isborn, J. Chem. Theory Comput. **13**, 3787 (2017).

40. C. Steinmann and J. Kongsted, J. Chem. Theory Comput. **11**, 4283 (2015).

41. X. Li, R. M. Parrish, F. Liu, S. I. L. Kokkila Schumacher, and T. J. Martínez, J. Chem. Theory Comput. **13**, 3493 (2017).

42. T. Förster, Discuss. Faraday Soc. **27**, 7 (1959).

43. J. Tomasi, B. Mennucci, and R. Cammi, Chem. Rev. **105**, 2999 (2005).

44. C. Curutchet, B. Mennucci, G. D. Scholes, and D. Beljonne, J. Phys. Chem. B **112**, 3759 (2008).

45. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, et al., *Gaussian 16 Revision A.03* (2016), Gaussian Inc. Wallingford CT.

46. C.-P. Hsu, G. R. Fleming, M. Head-Gordon, and T. Head-Gordon, J. Chem. Phys. **114**, 3065 (2001).

47. Z.-Q. You and C.-P. Hsu, Int. J. Quantum Chem. **114**, 102 (2014).

48. S. Jurinovich, C. A. Guido, T. Bruhn, G. Pescitelli, and B. Mennucci, Chem. Commun. **51**, 10498 (2015).

49. S. Jurinovich, G. Pescitelli, L. Di Bari, and B. Mennucci, Phys. Chem. Chem. Phys. **16**, 16407 (2014).

50. A. Rodger, G. Dorrington, and D. L. Ang, Analyst **141**, 6490 (2016).

51. W. Humphrey, A. Dalke, and K. Schulten, J. Molec. Graphics **14**, 33 (1996).

52. H. Lu, J. Mack, Y. Yang, and Z. Shen, Chem. Soc. Rev. **43**, 4778 (2014).

53. E. M. Sánchez-Carnerero, F. Moreno, B. L. Maroto, A. R. Agarrabeitia, M. J. Ortiz, B. G. Vo, G. Muller, and S. D. L. Moya, J. Am. Chem. Soc. **136**, 3346 (2014).

54. S. Burge, G. N. Parkinson, P. Hazel, A. K. Todd, and S. Neidle, Nucleic Acids Res. **34**, 5402 (2006).

55. J. Kypr, I. Kejnovska, D. Renciuk, and M. Vorlickova, Nucleic Acids Res. **37**, 1713 (2009).

56. A. Sengar and B. Heddi, Biochemistry **53**, 7718 (2014).

57. Y. Wang and D. J. Patel, Structure **1**, 263 (1993).

58. A. Biancardi, T. Biver, F. Secco, and B. Mennucci, Phys. Chem. Chem. Phys. **15**, 4596 (2013).

59. M. P. Fülscher, L. Serrano-Andrés, and B. O. Roos, J. Am. Chem. Soc. **119**, 6168 (1997).

60. J. J. Nogueira, F. Plasser, and L. González, Chem. Sci. **8**, 5682 (2017).

61. K. A. Kistler, C. M. Pochas, H. Yamagata, S. Matsika, and F. C. Spano, J. Phys. Chem. B **116**, 77 (2012).

62. Z. Li and J. D. Hirst, Chem. Sci. **8**, 4318 (2017).