




Cite this: *Photochem. Photobiol. Sci.*, 2018, **17**, 552

The modeling of the absorption lineshape for embedded molecules through a polarizable QM/MM approach†

Daniele Loco,* Sandro Jurinovich, Lorenzo Cupellini,  Maximilian F. S. J. Menger‡ and Benedetta Mennucci *

We present a computational strategy to simulate the absorption lineshape of a molecule embedded in a complex environment by using a polarizable QM/MM approach. This strategy is presented in two alternative formulations, one based on a molecular dynamics simulation of the structural fluctuations of the system and the other using normal modes and harmonic frequencies calculated on optimized geometries. The comparison for the case of a chromophore within a strongly inhomogeneous and structured environment, namely the intercalation pocket of DNA, shows that the MD-based approach is able to reproduce the experimental spectral bandshape. In contrast, the static approach overestimates the vibronic coupling, resulting in a much broader band.

Received 25th January 2018

Accepted 2nd March 2018

DOI: 10.1039/c8pp00033f

rsc.li/pps

1 Introduction

Solvatochromism is a well-known phenomenon that describes the change in the position and intensity of an electronic absorption (or emission) band due to a change in the polarity of the medium. As a matter of fact, the lineshape is not only sensitive to polarity effects but also extremely sensitive to the environment. Due to the complexity of the interactions and the multiple dynamics involved, the accurate computational simulation of the lineshape of an embedded system remains a challenge.¹ First, a proper description of the vibronic structure is needed, even if it often hides beneath the inhomogeneous broadening. In particular, the latter is generally dominant in polar solvents, where the static disorder is too large to cover all the details of the underlying vibronic structure. Instead the situation is different for other environments such as biomatrices, which in many cases act as binding pockets which keep the solvent molecules away from the chromophoric unit making the vibronic structure visible again. In all cases, an accurate description of static disorder has to be coupled to that of the vibronic structure. We note, however, that when the

system is embedded in a flexible and responsive environment, a net separation into homogeneous and inhomogeneous broadening is not always possible as, also depending on the temperature, the two effects can merge.

In the last few years, the quantum chemical simulation of the optical spectra of solvated molecules has largely progressed both in terms of accuracy and completeness. A large part of this improvement is due to the diffusion of hybrid QM/classical models, either in their atomistic or continuum formulations.^{2–6} These methods have been shown to be able to describe the effects of the solvent not only in the band position but also in its shape.^{7–12} In the case of heterogeneous environments such as a protein or DNA, the modelling has been mostly based on the atomistic version of the classical models introducing a molecular mechanics (MM) description of the environment.^{13,14} In most cases, a sequential approach is used where a MM force field (FF) molecular dynamics (MD) is run to generate different configurations of the coupled chromophore–environment system to be used in the following QM/MM calculations of the properties (transition energies and dipoles) to be used to reconstruct the absorption spectrum. The main concern regarding this approach is related to the accuracy that a MM-MD can achieve. Available FFs, in fact, are generally well suited to sample the structural fluctuations of the biomatrix but they are not accurate enough to properly describe those internal to the chromophore. If the static disorder is dominant and the vibronic structures have a minor effect on the final bandshape, a possible strategy to avoid possible artifacts due to this limitation is to run the MD simulation with the geometry of the chromophore kept frozen within a dynamic environment. This strategy is rather effective

Department of Chemistry, University of Pisa, via G. Moruzzi 13, 56124 Pisa, Italy.

E-mail: daniele.loco01@gmail.com, benedetta.mennucci@unipi.it

† Electronic supplementary information (ESI) available: Details of the classical molecular dynamics simulation and analysis of the structural parameters; experimental lineshapes of free and intercalated TO; homogeneous lineshape obtained from the low-frequency spectral densities; comparison of homogeneous and total lineshapes for intercalated TO; comparison of the optimized geometry of TO with the MD average geometry. See DOI: 10.1039/C8PP00033F

‡ Present address: Institute of Theoretical Chemistry, University of Vienna, Währinger Str. 17, A-1090 Vienna.

for quite rigid chromophores but clearly fails for floppy molecules whose dynamics strongly couple with that of the environment.

In all cases, however, when a classical MD is used, the quantum nature of the vibrational effects on the resulting simulated spectra is lost.

A possible effective strategy to recover this aspect, still keeping the computational cost feasible, involves the frequency dependent spectral density (SD) function.¹⁵ This is an exact approach if we assume that the excitations are linearly coupled to a continuous bath of harmonic oscillators. When instead, a more general definition of the SD as the Fourier transform of the time autocorrelation function of the excitation energy fluctuations is used, only an approximated description can be achieved but, at the same time, anharmonic effects can be recovered.¹⁶ Moreover, if the excitation energy fluctuations are obtained through QM/MM calculations along a MD trajectory of the whole system, environment effects are automatically included.^{17–19} Unfortunately, due to the large computational cost, customarily classical FF-based MD trajectories are used. In such a case, however, what was said before about the limits of standard FFs becomes even stronger. To obtain correct fluctuations of the excitation energies, the FF-based MD should in fact give a ground state potential energy surface as close as possible (both in the shape and position) to the QM/MM one. If a mismatch between the corresponding minima is present, or if the shape of the PESs is different, artificially large fluctuations in the excitation energies can be found which inevitably lead to unphysical SDs.^{18–23}

To avoid this risk, an alternative normal-mode analysis has been proposed. Within this framework, a harmonic approximation is assumed and the SD is reconstructed from the Huang–Rhys factors obtained from the calculations of the excited-state gradients in the ground state equilibrium geometry.^{22,24} This approach, from now indicated as *vertical gradient*, or VG, has been successfully applied to the determination of spectral densities in pigment–protein complexes.^{22,25} The VG approach is surely an effective one but the MD-based formulation is, in principle, more accurate as it can go beyond the harmonic approximation, and naturally include the effect of the temperature. These important advantages, however, require the removal of the ambiguity intrinsic to FF based methods and bring the MD simulation to higher accuracy.

Here we present a possible method of solving this issue by using a QM/MM Born–Oppenheimer MD which combines Density Functional Theory (DFT) with the AMOEBA polarizable FF.²⁶ The resulting trajectory is then used to calculate the SD and obtain an accurate simulation of the vibronic coupling. The same polarizable QM/MM approach is also used within the VG formulation, and the two models are compared in the simulation of a chromophore intercalated in DNA. The effects of static disorder are included in the two alternative models by combining a fully classical MD with a quantum chemical refinement of the internal structure of the chromophore either by QM/MM dynamics or QM/MM geometry optimization.

2 Theory and methods

2.1 Spectral densities and absorption lineshape

From a classical trajectory, the spectral density can be obtained by applying a temperature quantum correction to the classical autocorrelation function (ACF) of the excitation energy $C_{\text{cl}}(t)$:¹⁷

$$\tilde{C}''(\omega) = \frac{\beta\omega}{\pi} \int_{-\infty}^{\infty} e^{i\omega t} C_{\text{cl}}(t) dt \quad (1)$$

where $\beta = 1/k_{\text{B}}T$ refers to the temperature of the classical trajectory, and the classical ACF is calculated from the excitation energies $E(t_i)$ at given times t_i along the MD trajectory:¹⁷

$$C_{\text{cl}}(t_j) = \frac{1}{N-j} \sum_{i=1}^{N-j} \Delta E(t_i + t_j) \Delta E(t_i) \quad (2)$$

From the spectral density we can also measure the total coupling to vibrations in terms of the reorganization energy λ , defined as:

$$\lambda = \int_0^{\infty} \frac{\tilde{C}''(\omega)}{\pi\omega} d\omega \quad (3)$$

The reorganization energy measures the difference between the vertical and adiabatic excitation energies, namely the displacement between the minima of the ground and excited-state potential energy surfaces (PES), assuming that the two PES are described by the same harmonic oscillator, but with displaced equilibrium.

Alternatively, the spectral density can be obtained with a time-independent Vertical Gradient (VG) approach.²² Within this framework, the geometry of the chromophore in its ground state is optimized within a frozen environment and the corresponding vibrational frequencies and normal modes are obtained using the standard harmonic approximation. The Huang–Rhys factor S_k of each mode k was computed by projecting the gradient of the excited state \tilde{f} on the ground-state normal mode k , namely $S_k = \frac{\tilde{f}_k^2}{2\hbar\omega_k^3}$. Applying a Gaussian broadening, the SD becomes:

$$\tilde{C}''_{\text{VG}}(\omega) = \pi\omega \sum_{k=0}^M \frac{S_k\omega_k}{\sigma\sqrt{2\pi}} e^{-\frac{(\omega-\omega_k)^2}{2\sigma^2}} \quad (4)$$

where ω_k and S_k are the frequency and the Huang–Rhys factor of mode k . Within this framework, the reorganization energy becomes

$$\lambda^{\text{VG}} = \sum_{k=0}^M S_k\omega_k \quad (5)$$

The computed spectral density can be finally used to define a lineshape function $g(t)$:¹⁵

$$g(t) = - \int_0^{\infty} d\omega \frac{\tilde{C}''(\omega)}{\omega^2} \left[\coth\left(\frac{\beta\hbar\omega}{2}\right) (\cos(\omega t) - 1) - i(\sin(\omega t) - \omega t) \right] \quad (6)$$

Using eqn (6), the homogeneous absorption lineshape can be obtained as:

$$S(\omega - \omega_{eg}) = \Re \int_0^{\infty} dt e^{i(\omega - \omega_{eg})t - g(t)}, \quad (7)$$

where $|e\rangle$ and $|g\rangle$ denote the electronic excited and ground state respectively. Notably, in eqn (6) β is related to the temperature at which the spectrum is to be computed.

In the condensed phase, in addition to the homogeneous broadening, we have to also consider that there is a large number of chromophores, each with its own environment, slightly different from the others. Every chromophore, therefore, has a probability $P(\omega_{eg})$ to absorb light at a frequency ω_{eg} . The resulting absorption spectrum is inhomogeneously broadened by the convolution of the homogeneous spectrum with $P(\omega_{eg})$. The inhomogeneous distribution of excitation energies and transition dipole moments can be obtained from a set of N_{runs} separate dynamics runs. Namely, from each dynamics run an average excitation energy ($\omega_{eg,i}$) and an average transition dipole ($\mu_{eg,i}$) are obtained and finally combined with the lineshape in eqn (7) to simulate the absorption spectrum:

$$A(\omega) = \frac{\omega}{N_{\text{runs}}} \sum_{i=1}^{N_{\text{runs}}} |\mu_{eg,i}|^2 S(\omega - \omega_{eg,i}). \quad (8)$$

2.2 The polarizable QM/MM model

All the QM/MM calculations make use of a polarizable embedding based on an induced point-dipole model, within the smeared Thole damping interaction scheme.²⁷ Within this framework, at each classical atom is attached a charge (or a set of fixed multipoles), and an isotropic scalar atomic polarizability.

The global QM/MM energy functional for the polarizable embedding can be written as

$$\mathcal{E}(P, \mu) = \mathcal{E}^{\text{QM}}(P) + \mathcal{E}^{\text{MM}} + \mathcal{E}^{\text{ele}}(P) + \mathcal{E}^{\text{Pol}}(P, \mu), \quad (9)$$

where we have explicitly written, for each term, the functional dependency on the induced dipoles μ and the electronic density, here represented by the one-body density matrix P . The first term is the purely QM energy functional (equivalent to that of the isolated system), while the second represents purely MM interactions (bonding, dispersion/repulsion, electrostatic and polarization) between classical sites, according to their definition within the selected FF. Such a term does not depend on either the electronic density or the induced dipoles, being a constant for any given nuclear arrangement. The coupling term \mathcal{E}^{ele} is given by the electrostatic interaction between the MM charges (or static multipoles) and the QM density, whereas \mathcal{E}^{Pol} accounts for the mutual polarization contribution to the QM/MM energy.

Analytical derivatives of eqn (9) have been implemented and coupled with an integrator of the classical Newton's equations of motion for the nuclei.²⁸ To improve the performances of the SCF procedure, which needs to be solved on-the-fly for each nuclear configuration along the trajectory, we also implemented

an extended Lagrangian approach to exploit information on the electronic density from previous MD steps, starting from an improved density for a faster SCF convergence.²⁹

Energies and gradients of the polarizable QM/MM have been finally extended to excited states³⁰ by using the extended Lagrangian scheme within a time-dependent DFT formulation.

2.3 Computational details

Classical MD. A 100 ns classical MD run was performed with AMBER³¹ on the TO-DNA intercalated system, starting from the ¹H-NMR in the solution structure for the DNA oligonucleotide d(5'-CGCTAGCG-3') complexed with the bis-intercalating TO dimer (TOTO) (pdb entry: 108D).³² All the details about the classical MD simulation are reported in section S1 of the ESI.†

Polarizable QM/MM MD. The polarizable QM/MM Born-Oppenheimer (BO) MD simulations^{28,33} were performed starting from geometries extracted from the classical MD simulation. Initial velocities were assigned according to the Boltzmann distribution at 300 K. QM/MM MD were run in an NVE ensemble with the B3LYP/6-31G level of theory for the TO and the polarizable AMOEBA FF²⁶ for the rest of the environment. An integration time step of 0.3 fs was employed, for a total simulation time of 1 ps.

In our simulation, only the TO chromophore is allowed to move, whereas the rest of the environment is kept frozen in the initial configuration as extracted from the classical MD. We included in the MM region all the solvent molecules and ions within 30 Å from the TO chromophore. All the MM sites included in an inner shell of 15 Å around TO were treated as polarizable sites, while the external MM sites are only described with fixed multipoles. The DNA helix and the counter-ions are always treated as polarizable sites. All the QM/MM BOMD simulations were run by interfacing the MM program TINKER³⁴ with a locally modified version of the Gaussian09 package.³⁵

Spectral density. In order to compute the spectral density, one of the QM/MM MD simulations was extended up to 8 ps. In parallel, a BOMD dynamics of the free TO, *i.e.* removing the classical environment, was performed starting from the same initial point of the corresponding QM/MM MD run. Excited state calculations used for computing the ACF in eqn (2) were performed by extracting structures every 1.2 fs for a total of 5000 single-point calculations.

Geometry optimizations. We performed polarizable QM/MM geometry optimizations for the ground state (GS) of TO embedded in a frozen environment, followed by frequency and vertical gradient calculations. The polarizable embedding scheme, as implemented in our polarizable version of the ONIOM approach,³⁶ was employed for both GS optimizations and vertical gradients.^{30,37} GS optimizations were performed at the B3LYP/6-31G level in analogy with the QM/MM MD, whereas the 6-31+G(d) basis set was used for vertical gradients. Geometries were optimized starting from the same structures that were used for the QM/MM MD, preserving the same definition of the QM and MM regions.

Excited-state calculations. All excited state calculations were performed at the TD-B3LYP/6-31+G(d)/AMOEBA, by using a locally modified version of the Gaussian09 package.³⁵

3 A test case: Thiazole Orange intercalated in the DNA

Thiazole Orange (TO) is a very well-known fluorescent probe used to detect DNA;^{38–40} in fact, in solution its quantum yield is almost null, while, when intercalated in DNA, it becomes strongly fluorescent. The transition responsible of this behavior is a $\pi\pi^*$ transition, with a partial charge-transfer character, polarized along the long axis of the molecule, which is almost planar in its ground state (see Fig. 1).

Both experiments and theoretical simulations seem to indicate that the fluorescence enhancing effects that occur when TO passes from solution to single and double stranded DNA are related to the twisting of the two rings around the central methine bridge: this motion is possible for free TO, but is strongly constrained by the DNA pocket.^{41–43}

The effects due to the DNA binding on the structural degrees of freedom of TO are also responsible of the change observed in the absorption bandshape when going from the free (solvated) TO to the chromophore intercalated in DNA (see Fig. S5†). The intercalated TO shows a structured band, with a side band shifted towards the high energy with respect to the absolute maximum of the spectrum. In the case of the “free” TO in water solution, a broader and less structured spectrum is instead observed. This change in the bandshape shows that upon intercalation in DNA, the vibronic structure and the inhomogeneous broadening significantly change.

To investigate the influence of the DNA environment on these two aspects, we apply different techniques, mixing quantum and classical methods at different levels. Furthermore, we make use of both an MD-based method, *i.e.* atomistic simulation of the dynamics, and approaches suited for an implicit wavepacket dynamics, to retrieve information about the vibronic relaxation of the excited state. In particular, the polarizable QM/MM approach outlined in section 2.2 is used together with its extension to a Born–Oppenheimer (BO) MD^{28,33} of the TO in its ground state.

3.1 Spectral density and reorganization energies

As a first step of the analysis we have calculated the spectral densities (SD) of free (@VAC) and intercalated (@DNA) TO according to eqn. (1) and (2). To do that we have used two QM (/AMOEBA) BOMD simulations, one for the isolated TO and the other for the TO intercalated in DNA as described in the Computational details. In the case of the intercalated TO, we have assumed that there is not any mixing of vibrations between TO and the environment; this is a reasonable approximation as TO is embedded in a rather rigid pocket within the DNA scaffold from which solvent molecules are kept away.

The two resulting SDs, shown in Fig. 2(a), present peaks in very similar positions, but with often completely different

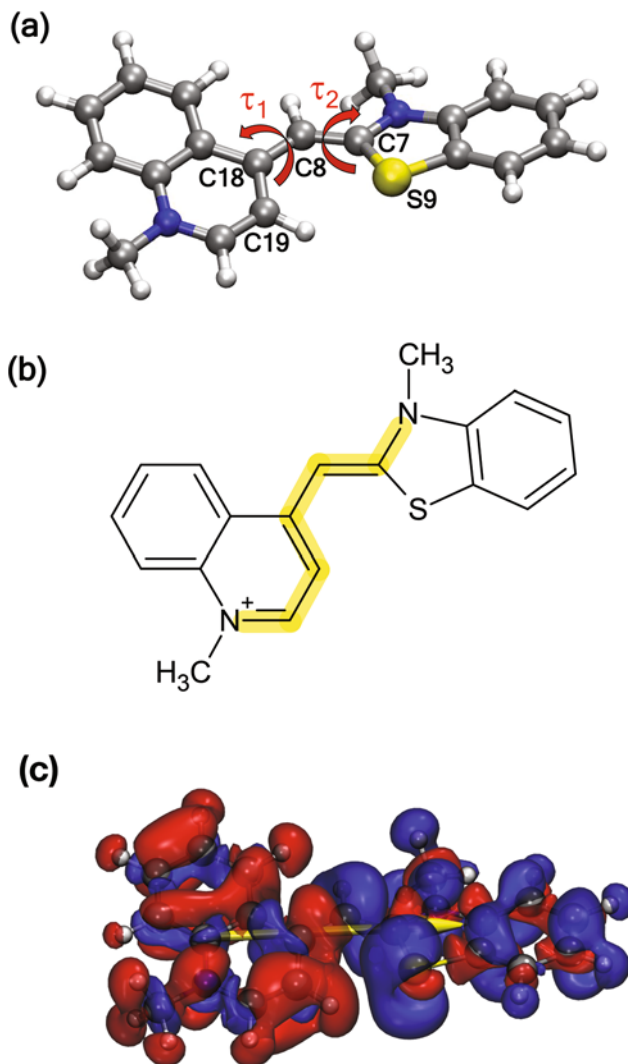


Fig. 1 (a) Ground state geometry of Thiazole Orange (TO). Red arrows denote the two dihedral angles, τ_1 (C19–C18–C8–C7) and τ_2 (C18–C8–C7–S9), which define the twisting between quinolinium and the benzothiazole moieties. (b) The pattern highlighted in yellow has been used to define the Bond Length Alternation (BLA), computed as the difference between the average length of single and double bonds. (c) Transition density and transition dipole corresponding to the lowest $\pi\pi^*$ excitation.

intensities. The peaks around 1600 and 1400 cm^{-1} , the two features between 1300 and 1100 cm^{-1} , and the two peaks around 800 and 650 cm^{-1} show the most significant differences. In the low-frequency part, below 150 cm^{-1} , the SD of free TO is systematically larger than that of intercalated TO.

The intensity differences between the two SDs arise from environment effects, which cause the excitation to couple more strongly to some normal modes than to others. While the coupling around 1400 cm^{-1} is enhanced in DNA, the coupling to other modes is largely suppressed by the intercalation. The overall effect of intercalation is a substantial suppression of the vibronic coupling, as evidenced in the reorganization energy, which is almost 40% smaller in DNA. However, the differences between the two SDs are restricted to particular fre-

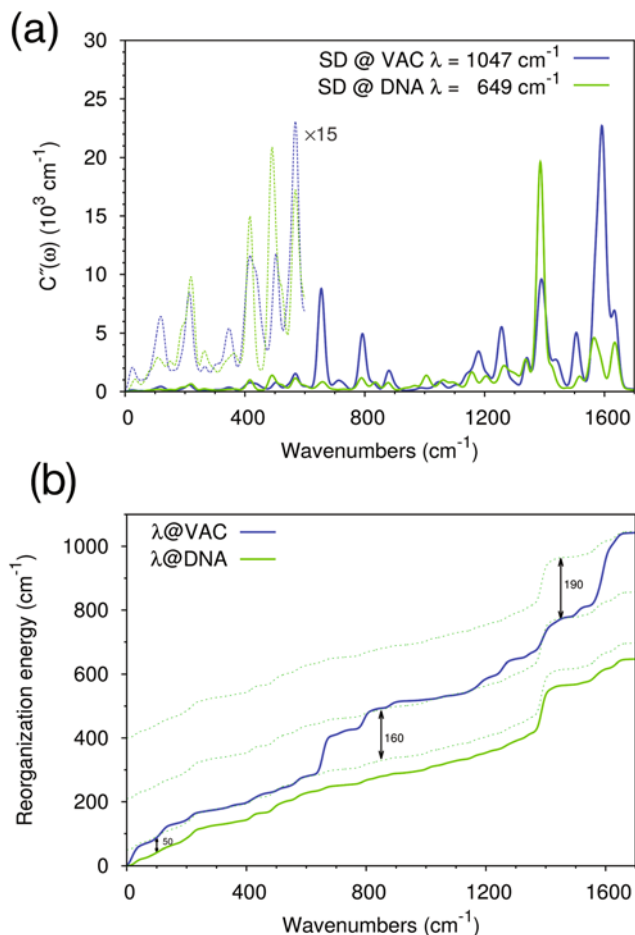


Fig. 2 (a) SDs computed from the MD of free (@VAC) and intercalated (@DNA) TO with the corresponding reorganization energies (λ). Dashed lines show the low-frequency part magnified by a factor of 15. (b) Cumulative reorganization energy for the spectral densities of free and intercalated TO. The three dashed lines show the plot of @DNA offset by 50, 210, and 400 cm^{-1} , in order to highlight the similarities and differences between the SDs.

frequency regions, namely, around 600–800 cm^{-1} and around 1600 cm^{-1} . By analyzing the cumulative reorganization energy distribution, shown in Fig. 2(b), we found that the difference in the reorganization energy distribution is related to the normal modes in three frequency regions: below 100 cm^{-1} , 600–800 cm^{-1} , and 1500–1700 cm^{-1} .

We used a frequency calculation of the TO *in vacuo* to assign the peaks in the SD to normal modes. Since the QM/MM MD trajectory motion is anharmonic, the modes that contribute to the SDs are not independent of each other, and we can expect that the harmonic modes of similar frequencies mix together due to anharmonic perturbations. However, we can identify groups of normal modes within a certain frequency window, which are responsible for the peaks in the SD. Indeed, in the low-frequency region between 300 and 800 cm^{-1} , most normal modes have contributions of out-of-plane C–C and C–H bending (Fig. 3(a and b)), whereas in the 800–1600 cm^{-1} region most normal modes involve mixed

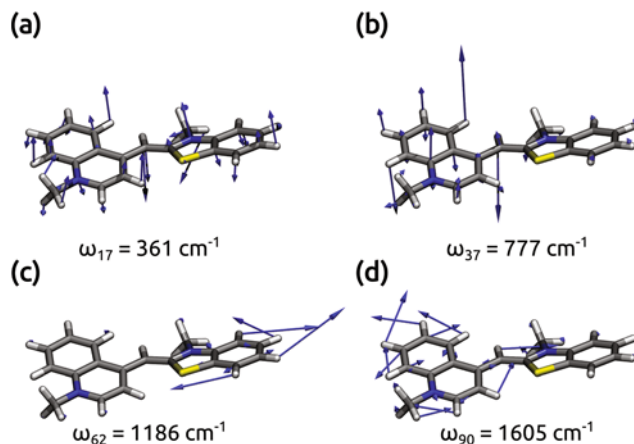


Fig. 3 Examples of the representative normal modes from the *in vacuo* frequency analysis, for the 300–800 cm^{-1} region (a, b) and the 800–1600 cm^{-1} region (c, d).

C=C stretching and C–H in-plane bending (Fig. 3(c and d)). The out-of-plane modes in the 600–800 cm^{-1} region are the ones most strongly suppressed by the intercalation in DNA. In contrast, among in-plane modes, some are enhanced while others are suppressed.

From the normal-mode analysis we were able to identify collective relative motions of the two rings around the central methine bridge in the low-frequency region (below 200 cm^{-1}): these motions involve twisting deformations that were suggested to be responsible for the fluorescence quenching in free TO.^{41,42} The cumulative reorganization energy for these modes is 86 cm^{-1} in DNA, compared with 143 cm^{-1} for free TO, indicating that intercalation suppresses the coupling between the excitation and the torsional modes of TO, which are hindered by the steric constraints of the DNA pocket. The coupling to low-frequency modes is reflected in the homogeneous line broadening, as shown in Fig. S6 of the ESI.†

As described in the Methods section, the SD can be alternatively obtained with a time-independent approach using the vertical gradient formulation.²² The VG results are reported in Fig. 4 together with the MD-based SD.

The two SDs strongly differ, in particular in the 400–800 cm^{-1} region and around 1200 and 1600 cm^{-1} . The VG reorganization energy is twice as large as the MD-based one. The reason for this discrepancy may lie either in the anharmonicity of the TO modes, which is captured only by the MD approach, or in the vibronic coupling, which for some modes is not a linear function of the nuclear displacements. As the MD explores a region of space around the equilibrium structure, the energy fluctuations are not necessarily given by the gradient of the excited state at the equilibrium geometry, especially for anharmonic modes.

3.2 Static disorder and the final spectrum

To obtain the final spectrum to be compared with experiments, the bandshape coming from the calculated spectral

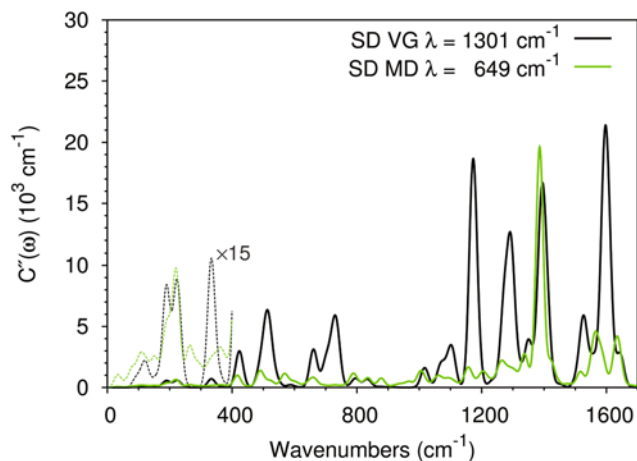


Fig. 4 Comparison of the SDs computed from the MD (green) and VG (black) approaches for intercalated TO. Dashed lines show the low-frequency part magnified by a factor of 15.

densities has to be combined with the static disorder. To do that, two different strategies have been compared.

In the first strategy (from now indicated as MD-based and depicted in Fig. 5), a fully classical MD trajectory was generated to obtain a sampling of the environment. From that trajectory, 32 uncorrelated configurations were extracted and used as a starting point for short (1 ps) QM/MM MD runs of a flexible TO in a frozen environment. From each of these short runs, 20 structures were extracted every 50 fs and used to compute QM/MM excitation energies and transition dipoles. The resulting average values were used to build homogeneously broadened spectra, with different positions and intensities, which were finally convoluted in the absorption spec-

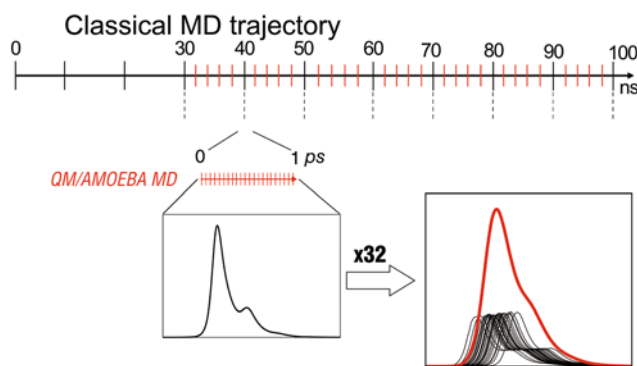


Fig. 5 Schematic representation of the approach used to compute the absorption bandshape. The horizontal time line represents the classical MD trajectory. Vertical red ticks indicate the 32 structures sampled from classical MD that were used as startingpoint geometries for the short 1 ps QM/MM MD runs. The average excitation energy and electronic transition dipole were computed sampling 20 structures from each of those trajectories every 50 fs starting from 100 fs and a homogeneously broadened spectrum has been created from the calculated SD (a small Gaussian broadening has been added with the only purpose of smoothing the lines). All the 32 spectra were finally convoluted to obtain the final bandshape (in red).

trum. The total lineshape retains the vibronic features of the homogeneous one, with additional broadening that arises from static disorder (Fig. S7 in the ESI†).

In the alternative static strategy, the same 32 configurations extracted from the classical MD have been used as starting points for the QM/MM geometry optimizations of the TO within a different frozen environment. On each of the resulting relaxed structures, QM/MM excitation energies and transition dipoles have been computed and used to build homogeneously broadened spectra, with different positions and intensities, which were finally convoluted in the final spectrum.

We note that the average excitation energy coming from the optimized structures is $\sim 500 \text{ cm}^{-1}$ higher than the average of the QM/MM MD runs (see Table 1), indicating that the average geometry in the MD picture is slightly offset from the minimum due to the anharmonicity of the PES. Moreover, the distribution of vertical excitation energies is $\sim 30\%$ narrower in the static approach, meaning that the MD average geometries show more variability than the optimized geometries.

If we now compare the spectra obtained by the two methods (Fig. 6), we clearly see that the static approach overestimates the vibronic coupling in both low and high-frequency regions. In fact, the intensity of the vibronic sideband is significantly greater in the static approach, due to the larger vibronic coupling in the high-frequency part of the SD. The larger coupling in the low-frequency part is reflected in the broadening of the total lineshape, which is much larger in the static approach despite the narrower inhomogeneous disorder.

Table 1 Average excitation energies and structural parameters computed over different optimized structures (OPT) and along the different QM/AMOEBA MD runs. The standard deviations are also shown

	MD	OPT
Exc (cm^{-1})	$21\,706 \pm 336$	$22\,150 \pm 264$
BLA (pm)	1.1 ± 0.8	0.8 ± 0.4
τ_1 (degrees)	-0.4 ± 4.3	-0.3 ± 5.0
τ_2 (degrees)	3.7 ± 6.6	2.0 ± 5.7

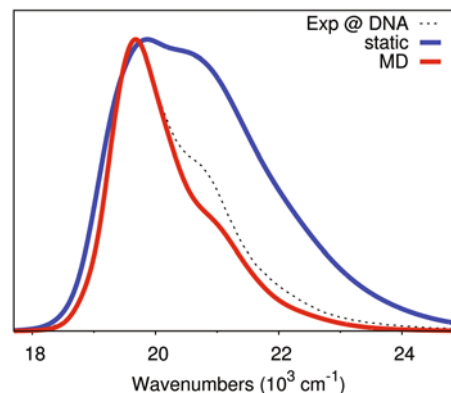


Fig. 6 Comparison between the MD-based approach, the time-independent approach (static) and the experimental absorption spectra. All spectra have been normalized to the maximum. Both calculated spectra are shifted by -1450 cm^{-1} .

The line broadening obtained from the static approach is clearly overestimated if compared with the experiment, as also is the vibronic sideband. The larger reorganization energy from the VG spectral density, however, compensates for the difference in vertical excitation energies (Table 1), so that the position of the 0–0 band is almost the same for both approaches.

To better understand the origin of these spectral differences, we selected some structural parameters, namely, the two dihedral angles involved in the methine bridge, and Bond Length Alternation (BLA), defined in Fig. 1. For each short QM/MM MD, we extracted the average of such structural parameters, to compare with those obtained for the corresponding optimized geometries. Their average and standard deviations are reported in Table 1. The obtained structural parameters are very similar, suggesting that, despite the distortions highlighted above, the QM/MM MD sample structures around the optimized geometry.

However, even though the BLAs are very small in the two approaches (as expected for cyanine-like systems), their effect on the excitation energy is different. In fact, for the optimized geometries, the excitation energies strongly correlate with the BLA value, meaning that the BLA is the main parameter that accounts for excitation energy variability (see Fig. 7). In contrast, for the QM/MM MD geometries, only 30% of the excitation energy variance can be explained by the BLA, showing that other degrees of freedom contribute to the excitation energy fluctuations.

A possible explanation of this behavior can be related to distortions on the two aromatic groups which are present in the MD-based model but are lost in the static description using optimized geometries. This suggestion is confirmed by the average structure from the MD trajectory which is more distorted compared to the corresponding optimized geometry (see Fig. S8 in the ESI†). The presence of distortions in the rings could in fact make the BLA index less exhaustive in describing fluctuations in the conjugation.

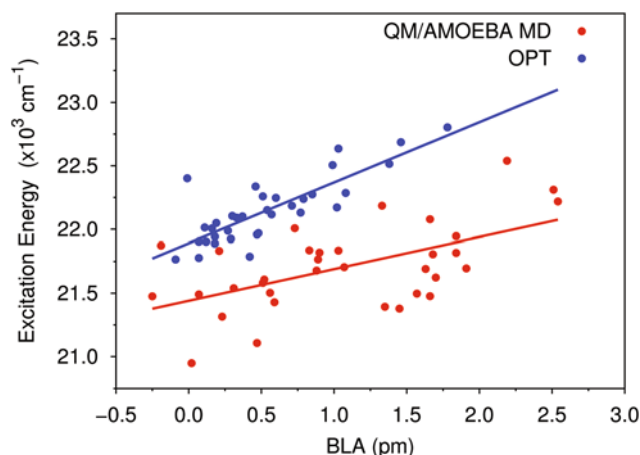


Fig. 7 Comparison of the correlation between the excitation energies and the BLA obtained either using optimized structures (OPT) or QM/MM MD trajectories.

4. Conclusions

The absorption lineshape of embedded molecules is the result of various coupled effects due to different types of interactions between the molecular system and the environment. Not only can the transition energies and dipoles determining the position and the intensity of the band be largely affected by the environment but also the coupling of the excitation process with nuclear motions can be changed. While the former are mostly due to electrostatic and polarization effects, in the latter, steric effects such as those acting in binding pockets within biomatrices can become relevant and the vibronic structure can be largely modified. Finally, the environment introduces an additional effect leading to the inhomogeneous broadening not present in isolated molecules.

The strategy here presented combines classical and polarizable QM/MM MD to obtain a reliable estimation of the homogeneous and inhomogeneous contributions to the lineshape of a chromophore embedded in a complex environment. The case of TO intercalated in DNA shows that this strategy can accurately reproduce the experimental spectrum. In contrast, the time-independent approach based on the vertical gradient calculation of the spectral density severely overestimates the vibronic coupling, resulting in a much broader band.

Unfortunately, the large computational cost associated with the calculation of the spectral density through a polarizable QM/MM MD still prevents an extensive use of this strategy. However, this technical limit will certainly be reduced, if not completely solved, with future optimizations of the algorithms. Still within the present formulation, though, the method can represent an important benchmark to validate both static approaches and the MD-based approaches that rely on classical FFs.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

MM gratefully acknowledges financial support from the EU Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 642294 (ITN-EJD: Theoretical Chemistry and Computational Modelling).

References

- 1 F. Santoro and D. Jacquemin, *WIREs Comput. Mol. Sci.*, 2016, **6**, 460–486.
- 2 C. J. Cramer and D. G. Truhlar, *Chem. Rev.*, 1999, **99**, 2161–2200.
- 3 J. Tomasi, B. Mennucci and R. Cammi, *Chem. Rev.*, 2005, **105**, 2999–3094.

- 4 J. M. H. Olsen and J. Kongsted, in *Advances in Quantum Chemistry*, ed. J. R. Sabin and E. Brandas, Academic Press, 2011, vol. 61, pp. 107–143.
- 5 M. S. Gordon, D. G. Fedorov, S. R. Pruitt and L. V. Slipchenko, *Chem. Rev.*, 2012, **112**, 632–672.
- 6 B. Mennucci, *Phys. Chem. Chem. Phys.*, 2013, **15**, 6583–6512.
- 7 F. J. A. Ferrer, R. Improta, F. Santoro and V. Barone, *Phys. Chem. Chem. Phys.*, 2011, **13**, 17007–17012.
- 8 N. De Mitri, S. Monti, G. Prampolini and V. Barone, *J. Chem. Theory Comput.*, 2013, **9**, 4507–4516.
- 9 A. Charaf-Eddin, A. Planchat, B. Mennucci, C. Adamo and D. Jacquemin, *J. Chem. Theory Comput.*, 2013, **9**, 2749–2760.
- 10 A. Petrone, J. Cerezo, F. J. A. Ferrer, G. Donati, R. Improta, N. Rega and F. Santoro, *J. Phys. Chem. A*, 2014, **119**, 5426–5438.
- 11 B. Mennucci, *Int. J. Quantum Chem.*, 2015, **115**, 1202–1208.
- 12 T. J. Zuehlsdorff and C. M. Isborn, *J. Chem. Phys.*, 2018, **148**, 024110.
- 13 H. M. Senn and W. Thiel, *Angew. Chem.*, 2009, **48**, 1198–1229.
- 14 E. Brunk and U. Rothlisberger, *Chem. Rev.*, 2015, **115**, 6217–6263.
- 15 S. Mukamel, *Principles of Nonlinear Optical Spectroscopy*, Oxford University Press, New York, 1995.
- 16 Y. Georgievskii, C.-P. Hsu and R. A. Marcus, *J. Chem. Phys.*, 1999, **110**, 5307–5317.
- 17 S. Valleau, A. Eisfeld and A. Aspuru-Guzik, *J. Chem. Phys.*, 2012, **137**, 224103.
- 18 M. C. Zwier, J. M. Shorb and B. P. Krueger, *J. Comput. Chem.*, 2007, **28**, 1572–1581.
- 19 S. Chandrasekaran, M. Aghtar, S. Valleau, A. Aspuru-Guzik and U. Kleinekathöfer, *J. Phys. Chem. B*, 2015, **119**, 9995–10004.
- 20 C. W. Kim, J. W. Park and Y. M. Rhee, *J. Phys. Chem. Lett.*, 2015, **6**, 2875–2880.
- 21 A. M. Rosnik and C. Curutchet, *J. Chem. Theory Comput.*, 2015, **11**, 5826–5837.
- 22 M. K. Lee, P. Huo and D. F. Coker, *Annu. Rev. Phys. Chem.*, 2016, **67**, 639–668.
- 23 O. Andreussi, I. G. Prandi, M. Campetella, G. Prampolini and B. Mennucci, *J. Chem. Theory Comput.*, 2017, **13**, 4636–4648.
- 24 M. K. Lee and D. F. Coker, *J. Phys. Chem. Lett.*, 2016, **7**, 3171–3178.
- 25 F. Segatta, L. Cupellini, S. Jurinovich, S. Mukamel, M. Dapor, S. Taioli, M. Garavelli and B. Mennucci, *J. Am. Chem. Soc.*, 2017, **139**, 7558–7567.
- 26 J. W. Ponder, C. Wu, P. Ren, V. S. Pande, J. D. Chodera, M. J. Schnieders, I. Haque, D. L. Mobley, D. S. Lambrecht, R. A. DiStasio Jr., M. Head-Gordon, G. N. I. Clark, M. E. Johnson and T. Head-Gordon, *J. Phys. Chem. B*, 2010, **114**, 2549–2564.
- 27 B. T. Thole, *Chem. Phys.*, 1981, **59**, 341–350.
- 28 D. Loco, L. Lagardère, S. Caprasecca, F. Lipparini, B. Mennucci and J.-P. Piquemal, *J. Chem. Theory Comput.*, 2017, **13**, 4025–4033.
- 29 A. M. N. Niklasson, P. Steneteg, A. Odell, N. Bock, M. Challacombe, C. J. Tymczak, E. Holmström, G. Zheng and V. Weber, *J. Chem. Phys.*, 2009, **130**, 214109.
- 30 M. F. S. J. Menger, S. Caprasecca and B. Mennucci, *J. Chem. Theory Comput.*, 2017, **13**, 3778–3786.
- 31 D. Case, J. Berryman, R. Betz, D. Cerutti, T. Cheatham III, T. Darden, R. Duke, T. Giese, H. Gohlke, A. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T. Lee, S. LeGrand, P. Li, T. Luchko, R. Luo, B. Madej, K. Merz, G. Monard, P. Needham, H. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, D. Roe, A. Roitberg, R. Salomon-Ferrer, C. Simmerling, W. Smith, J. Swails, R. Walker, J. Wang, R. Wolf, X. Wu, D. York and P. Kollman, *AMBER 2016*, University of California, San Francisco, 2016.
- 32 H. P. Spielmann, D. E. Wemmer and J. P. Jacobsen, *Biochemistry*, 1995, **34**, 8542–8553.
- 33 D. Loco, É. Polack, S. Caprasecca, L. Lagardère, F. Lipparini, J.-P. Piquemal and B. Mennucci, *J. Chem. Theory Comput.*, 2016, **12**, 3654–3661.
- 34 J. W. Ponder, *TINKER, Software Tools for Molecular Design*, <http://dasher.wustl.edu/tinker>.
- 35 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, B. G. Janesko, F. Lipparini, G. Zheng, J. L. Sonnenberg, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, T. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, P. V. Parandekar, N. J. Mayhall, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian Development Version, Revision H.36*, Gaussian Inc., Wallingford, CT, 2010.
- 36 L. W. Chung, W. M. C. Sameera, R. Ramozzi, A. J. Page, M. Hatanaka, G. P. Petrova, T. V. Harris, X. Li, Z. Ke, F. Liu, H.-B. Li, L. Ding and K. Morokuma, *Chem. Rev.*, 2015, **115**, 5678–5796.
- 37 S. Caprasecca, S. Jurinovich, L. Viani, C. Curutchet and B. Mennucci, *J. Chem. Theory Comput.*, 2014, **10**, 1588–1598.
- 38 A. N. Glazer and H. S. Rye, *Nature*, 1992, **359**, 859–861.
- 39 H. S. Rye, S. Yue, D. E. Wemmer, M. A. Quesada, R. P. Haugland, R. A. Mathies and A. N. Glazer, *Nucleic Acids Res.*, 1992, **20**, 2803–2812.

- 40 E. Privat, T. Melvin, F. Mérola, G. Schweizer, S. Prodhomme, U. Asseline and P. Vigny, *Photochem. Photobiol.*, 2002, **75**, 201–210.
- 41 J. Nygren, N. Svanvik and M. Kubista, *Biopolymers*, 1998, **46**, 39–51.
- 42 A. Biancardi, T. Biver, A. Marini, B. Mennucci and F. Secco, *Phys. Chem. Chem. Phys.*, 2011, **13**, 12595–12602.
- 43 E. E. Rastede, M. Tanha, D. Yaron, S. C. Watkins, A. S. Waggoner and B. A. Armitage, *Photochem. Photobiol. Sci.*, 2015, **14**, 1703–1712.