TD-DFT Investigation of the Magnetic Circular Dichroism Spectra of Some Purine and Pyrimidine Bases of Nucleic Acids

Tobias Fahleson,[†] Joanna Kauczor,[†] Patrick Norman,[†] Fabrizio Santoro,[‡] Roberto Improta,[¶] and Sonia Coriani[*,§,‖]

Department of Physics, Chemistry and Biology, Linköping University, Sweden, Istituto di Chimica dei Composti Organometallici (ICCOM–CNR), Area della Ricerca del CNR, via Moruzzi 1, I-56124, Pisa, Italy, Istituto di Biostrutture e Bioimmagini-CNR, Via Mezzocannone 6, I-80134 Napoli, Italy, Dipartimento di Scienze Chimiche e Farmaceutiche, Università degli Studi di Trieste, via L. Giorgieri 1, I-34127, Trieste, Italy, and Aarhus Institute of Advanced Studies, Aarhus University, DK-8000 Aarhus C, Denmark

E-mail: coriani@units.it

Abstract

We present a computational study of the magnetic circular dichroism (MCD) spectra in the 200-300 nm wavelength region of purine and its derivative hypoxanthine, as well as of the

[*]To whom correspondence should be addressed

[†]Department of Physics, Chemistry and Biology, Linköping University, Sweden

[‡]Istituto di Chimica dei Composti Organometallici (ICCOM–CNR), Area della Ricerca del CNR, via Moruzzi 1, I-56124, Pisa, Italy

[¶]Istituto di Biostrutture e Bioimmagini-CNR, Via Mezzocannone 6, I-80134 Napoli, Italy

[§]Dipartimento di Scienze Chimiche e Farmaceutiche, Università degli Studi di Trieste, via L. Giorgieri 1, I-34127, Trieste, Italy

[‖]Aarhus Institute of Advanced Studies, Aarhus University, DK-8000 Aarhus C, Denmark

pyrimidine bases of nuclear acids uracil, thymine and cytosine, using the B3LYP and CAM-B3LYP functionals. Solvent effects are investigated within the polarizable continuum model and by inclusion of explicit water molecules.

In general, the computed spectra are found to be in good agreement with the experimental ones, apart from some overall blue shifts. Both the pseudo-*A* term shape of the MCD spectra of the purines and the *B*-type shape of the ones of pyrimidine bases are reproduced. Our calculations also correctly reproduce the reversed phase of the MCD bands in purine compared to that of its derivatives present in nucleic acids. Solvent effects are sizeable and system specific, but they do not in general alter the qualitative shape of the spectra. The bands are dominated by the bright $\pi \rightarrow \pi^*$ transitions, and our calculations in solution nicely reproduce their energy differences, improving the estimates obtained in gas phase. Shoulders are predicted for purine and uracil due to n$\rightarrow \pi^*$ excitations, but they are too weak to be observed in the experiment.

Introduction

Magnetic circular dichroism (MCD) is a powerful spectroscopic technique for probing the electronic states of a molecule[1] where the absorption spectrum is often characterized by broad, unresolved bands hiding several electronic excitations. The MCD spectrum—also thanks to the fact that its signals are signed (i.e. can be both positive and negative)—can reveal such hidden states. MCD can thus be very useful in detecting transitions unresolved in the unpolarized absorption spectrum, extending to achiral systems the potentialities of electronic circular dichroism for chiral molecules. MCD has also been used for analytic applications, for instance to detect small quantities of specific compounds in a sample, or to discriminate between compounds that have very similar absorption spectra.[2,3]

The last two decades have witnessed renewed interest in MCD spectroscopy, in particular from the theoretical and computational side. Efficient computational protocols for the determination of MCD spectra and MCD parameters (the so-called *A*, *B* and *C* terms) have been proposed within Response Theory starting from a study on ethylene and p-benzoquinone in 1999 at the Hartree-

Fock and Multiconfigurational Self-Consistent Field (MCSCF) levels of approximation.[4] After a few relatively scarce studies at Hartree-Fock, MCSCF[4] and Coupled Cluster[5,6] levels of theory, a boost in the applications has come from the extension of the computational methodology within the realm of Time-Dependent Density Functional Theory (TD-DFT), resulting in the ability to address larger systems.[7–19] Beside the residue-based approach of Refs. 4,6,11,16, so-called complex polarization propagator (also known as damped response theory) methods[12,13,20]—as well as real-time methods[21]—have emerged in the last decade along with modern versions of sum-over-states approaches;[22,23] relativistic schemes have also been proposed.[24] We refer to Refs. 10 and 17 for reviews up to 2011. The complex polarization propagator (CPP) approach,[12,13,17,25,26] in particular, has been proposed as a cost-effective way to obtain the MCD spectrum without the need to decompose the calculation into individual A and B terms.[12] Moreover, it is expected to be numerically more stable that the residue-based approach when addressing broader spectral regions and dealing with a high density of excited states.

Solvent effects can alter the relative position and intensity of different electronic states, and even change their nature. Since many measurements are carried out in solution, the inclusion of such effects is extremely important for a direct comparison between experiments and theoretical predictions. The first extension of the computational methodology for MCD to include solvent effect dates back to 2008 with the Polarizable Continuum Model (PCM)[27] generalization of the residue-based methodology presented in Ref. 11. In addition, vibronic effects have also been considered in Ref. 28 and more recently in Ref. 29.

The electronic structure of purine and pyrimidine derivatives has been the subject of both experimental and theoretical extensive investigations in the past. Among them, the nucleobases occurring in nucleic acids either DNA or RNA are constantly in focus, as one aims at clarifying the nucleic acids photo-reactivity.[30–32] The excited state dynamics of all the nucleobases have been thoroughly studied by steady-state and time-resolved spectroscopy techniques,[30–32] revealing a complex behavior, with multiple states involved in the deactivation path, both in the gas phase and in solution. Conversely, a huge number of theoretical studies, too many for being exhaustively

reviewed here, has been applied to compute the potential energy surfaces of their lowest energy excited states, in order to interpret and assign the experimental spectra, leading often to contradictory results. This is not surprising since, even in the simplest bases, several excited states, both $\pi\pi^*$ and $n\pi^*$, lie close in energy in the Franck-Condon region, their energy being extremely sensitive to the details of the adopted computational method and to the nature of the embedding medium (gas phase, apolar or hydrogen bonding solvent). In this scenario, nucleobases have become one of the most important *playground* for the test and the comparison of different computational methods, and their study has acquired a significant methodological interest. Unfortunately, standard optical techniques can hardly shed light on many of the most debated issues, like the relative energy of two close-lying transitions with very different oscillator strength. In fact, the nucleobases absorption spectra in solution are too broad and unresolved, thus providing little information on the exact relative position of this kind of states. In this respect, the application of MCD spectroscopy can be very useful, as shown by a very recent investigation of the UV-vis and MCD spectra of adenine, guanine and their nucleosides; the ability of MCD to resolve multiple transitions in a single absorption band provides direct evidence on the much debated ordering of the L_a and L_b states of adenine.[33] Here we extend our investigation to encompass purine and its naturally occurring derivative hypoxanthine, present in tRNA, as well as the pyrimidine bases of RNA, uracil, and of DNA, thymine and cytosine. We will focus on the relative energy of the lowest energy $\pi\pi^*$ states, for which different methods provide different estimates (see, for example, Ref. 34). Moreover, we will show that $n\pi^*$ states do not exhibit significant MCD intensities, so that they cannot be easily located from experiments even if, in principle, the vibronic interaction between $n\pi^*$ and $\pi\pi^*$ can affect the MCD spectrum.

The main scope of the present contribution is the detailed analysis of the effect of the aqueous solvent on the MCD spectra of these compounds and the evaluation of the relative performance of implicit (PCM), explicit (cluster), and hybrid models (PCM+cluster). Besides that, this study will allow to further benchmark the accuracy of the B3LYP and CAM-B3LYP functionals in the determination of MCD spectra of these compounds, encompassing a rather large energy range. We

conclude this section emphasizing that, beyond their interest for the photophysics of the isolated bases,[35–38] our results could also be useful for the study of the behavior of photoexcited oligonucleotides, a topic very lively debated in recent literature. The nature and the relative energy of the excited states of single and double strand nucleic acids will indeed strongly depend on the nature and on the interaction existing between the excited states of their components.[39–44]

Methodology

Experimentally, MCD spectra are obtained measuring the differential absorption of right and left circularly polarized light in presence of an external magnetic field aligned to the direction of propagation of the probing light. The spectra are conventionally rationalized in terms of three "strength" parameters, called A, B and C terms, which are associated to three specific spectral features. Thus, for a sample of randomly moving closed-shell molecules in presence of an external magnetic field in the Z direction, the anisotropy of the molar decadic coefficient in the MCD experiment (where the strength of the external magnetic field is B_{ext}) is given by[45]

$$\Delta\varepsilon(\omega) = -\frac{8\pi^2 N_A \omega B_{\text{ext}}}{3\times 1000\times \ln 10(4\pi\varepsilon_0)\hbar c_0}\sum_j \left\{ \frac{1}{\hbar}\frac{\partial a_j(\omega)}{\partial \omega} A(0\rightarrow j) + a_j(\omega)\left[B(0\rightarrow j) + \frac{C(0\rightarrow j)}{kT}\right]\right\} \tag{1}$$

Above ω is the frequency of the probing light, T is the temperature, N_A is the Avogadro number, $(0\rightarrow j)$ indicates a transition from the ground state 0 to the excited state j, $\omega_j = E_j - E_0$ is the corresponding excitation energy, and the lineshape function $a_j(\omega)$ is taken as a Lorentzian

$$a_j(\omega) = \frac{1}{\pi}\frac{\gamma}{(\omega_j-\omega)^2+\gamma^2} \tag{2}$$

will full width at half maximum, FWHM, equal to 2γ. The other quantities are well-know fundamental constants. B terms are present for all systems irrespective of their spacial symmetry and spin degeneracy, whereas the occurrence of A and C terms depends on the existence of degenerate excited and ground states, respectively.

The microscopic theory of MCD, which forms the basis for the computational methodologies in the literature, was first derived by Buckingham and Stephens,[46] and later fully developed by Stephens.[47] As thoroughly reviewed in Ref. 18, the two dominating approaches for calculating the strength parameters of MCD spectra by ab initio methods are explicit excited-state methods [including limited sum-over- states procedures] and response-theory based methods. Although the former has been used with success in particular in the study of transition-metal complexes, wherein often only a few transition-metal *d* orbitals dominate the MCD signal,[48,49] the latter are more common for compounds involving lighter elements, as in the cases studied here.

In the framework of response theory, MCD spectra can be computed in two ways. In the first one, the sum-over-state expressions of the *A*, *B* and *C* terms derived by Buckingham and Stephens[46,47] are related to and computed from the poles and residues of appropriate response functions. Since we are here only concerned with molecules without degenerate states, we limit ourselves to considering the *B* term, which is obtainable from the residue of the quadratic response function (imaginary part thereof)[4]

$$B(0 \to j) = \varepsilon_{\alpha\beta\delta} \lim_{\omega \to \omega_j} (\omega - \omega_j) \Im \langle\langle \mu_\delta; m_\beta, \mu_\alpha \rangle\rangle_{0,\omega} \tag{3}$$

or, equivalently, as magnetic-field derivative of the left and right dipole transition moments between two electronic states evaluated at zero perturbation strength[5]

$$B(0 \to j) = \frac{1}{2} \varepsilon_{\alpha\beta\delta} \frac{d \Im \{ \langle 0 | \mu_\alpha | j \rangle \langle j | \mu_\delta | 0 \rangle \}}{dB_\beta} \bigg|_{B=0} \tag{4}$$

Above, α, β and δ distinguish the three Cartesian components, $\varepsilon_{\alpha\beta\delta}$ is the Levi-Civita tensor, and μ_α and m_α are the Cartesian components of the electric and magnetic dipole moment operator, respectively. Einstein's implicit summation over repeated indices is assumed here and throughout. Once the strength parameters have been computed, the spectrum can be simulated by attaching suitable lineshape functions to each individual term.

According to the second approach, alternative to the computation of the individual strength

parameters, the MCD spectrum is obtained directly from the real (i.e. absorptive) part of the complex (also known as "damped") quadratic response function,

$$\Delta\varepsilon(\omega) = -\frac{8\pi N_A B_{\text{ext}}}{3\times 1000\times \ln 10(4\pi\varepsilon_0)\hbar c_0}\omega\varepsilon_{\alpha\beta\delta}\Re\langle\langle\mu_\alpha;\mu_\beta,m_\delta\rangle\rangle^{\gamma}_{\omega,0} \tag{5}$$

where a phenomenological finite lifetime $\frac{1}{2}\gamma^{-1}$ of all the excited states is introduced in the response function to get a correct physical behavior in the absorptive regions of the sample. In other words, the damping terms make the quadratic response function convergent even in the absorptive regions of the sample (where the standard response function otherwise would diverge), so that the damped response function goes also under the name of "resonance-convergent" response function. Introducing the damping term is equivalent to considering complex excitation energies according to $\omega_n \to \omega_n - i\gamma$. Thus, for species where only B terms are allowed, the response function in Eq. (5) and the B term are related by

$$\pi^{-1}\varepsilon_{\alpha\beta\delta}\Re\{\langle\langle\mu_\alpha;\mu_\beta,m_\delta\rangle\rangle^{\gamma}_{\omega,0}\} \equiv 2\sum_j B(0\to j)a_j(\omega) \tag{6}$$

The computation of the MCD spectrum by the CPP approach, see Eq. (5), requires the ability to solve complex linear response equations, which, for Hartree-Fock and TD-DFT, take the form

$$\left(\mathbf{E}^{[2]} - \hbar(\omega+i\gamma)\mathbf{S}^{[2]}\right)(\mathbf{N}^C_R + i\mathbf{N}^C_I) = \mathbf{C}^{[1]}_R + i\mathbf{C}^{[1]}_I \tag{7}$$

where $\mathbf{E}^{[2]}$ is the (generalized) Hessian matrix, $\mathbf{S}^{[2]}$ is the metric matrix, $\mathbf{N}^C$ is the solution vector, and $\mathbf{C}^{[1]}$ is the property gradient vector, for a given operator (component) C. Subscripts R and I indicate real and imaginary components of the given vectors, respectively. The algorithm used in the present study for solving the above equation exploits a symmetrized trial vector basis and an efficient preconditioner explicitly considering coupling between different components, and it is extensively described in Refs. 26,50.

The final expression for the resonance-convergent quadratic response function becomes[25]

$$\begin{aligned}\langle\langle \mu_A;\mu_B,m_C\rangle\rangle^{\gamma}_{\omega_1,\omega_2} &= \sum \mathscr{P}_{1,2}\Big[N^A_n(\omega_\sigma)B^{[2]}_{nm}N^C_m(\omega_2) \\ &\quad +N^B_n(\omega_1)A^{[2]}_{nm}N^C_m(\omega_2) \\ &\quad +N^B_n(\omega_1)C^{[2]}_{nm}N^A_m(\omega_\sigma)\Big] \\ &\quad +N^A_n(\omega_\sigma)\left[E^{[3]}_{nml}+E^{[3]}_{nlm}\right]N^B_m(\omega_1)N^C_l(\omega_2),\end{aligned} \tag{8}$$

where the three response vectors equal

$$\begin{aligned}N^A_n(\omega_\sigma) &= A^{[1]}_l\left[E^{[2]}-\hbar(\omega_\sigma+i\gamma)S^{[2]}\right]^{-1}_{ln}, \\ N^B_n(\omega_1) &= \left[E^{[2]}-\hbar(\omega_1+i\gamma)S^{[2]}\right]^{-1}_{nl}B^{[1]}_l, \\ N^C_n(\omega_2) &= \left[E^{[2]}-\hbar(\omega_2+i\gamma)S^{[2]}\right]^{-1}_{nl}C^{[1]}_l.\end{aligned} \tag{9}$$

with $\omega_1=\omega$, $\omega_2=0$ and $\omega_\sigma=\omega_1+\omega_2$. The explicit expressions for the elements of the matrices $\mathbf{E}^{[3]}$, $\mathbf{A}^{[2]}$, $\mathbf{B}^{[2]}$ and $\mathbf{C}^{[2]}$ may be found for example in Ref. 25.

The generalization of the above CPP equations to the IEF-PCM framework follows straightforwardly from the generalization of the standard response equations for the linear and quadratic response functions and residues.[11,51,52] Thus, for the description of the environment, solvent effects are included by adding the relevant solute-solvent one- and two-electron interaction terms into the gas phase Hamiltonian, which yields to the replacement of the $\mathbf{E}^{[2]}$ and $\mathbf{E}^{[3]}$ terms with the so-called $\mathbf{G}^{[2]}$ and $\mathbf{G}^{[3]}$ terms, whose explicit expressions are given in Refs. 51 and 52, respectively.

More details on the unification of the PCM and CPP formalisms can be found in another contribution of this Festschrift.[53]

Computational Details

For all the considered compounds and clusters, the ground-state structures have been optimized at B3LYP/cc-pVTZ level of theory. In order to compute gas-phase spectra, we used gas-phase optimized geometries. For the solvated species (herewith also including clusters of each given species coordinated to a number of explicit water molecules) we optimized the structures in a solvated environment described by PCM. The number of water molecules in each cluster was determined following a pragmatic approach where we saturated the H bond in the molecular plane. Specifically, for all O and N atoms in the molecules purine, hypoxanthine and cytosine:

1. For each oxygen, we included 2 water molecules donating a H each to one of the two oxygen's lone pairs;

2. For each N atom with lone pairs (the N atom involved in 2 σ bonds) we included 1 water molecule that donates a H to the N lone pair;

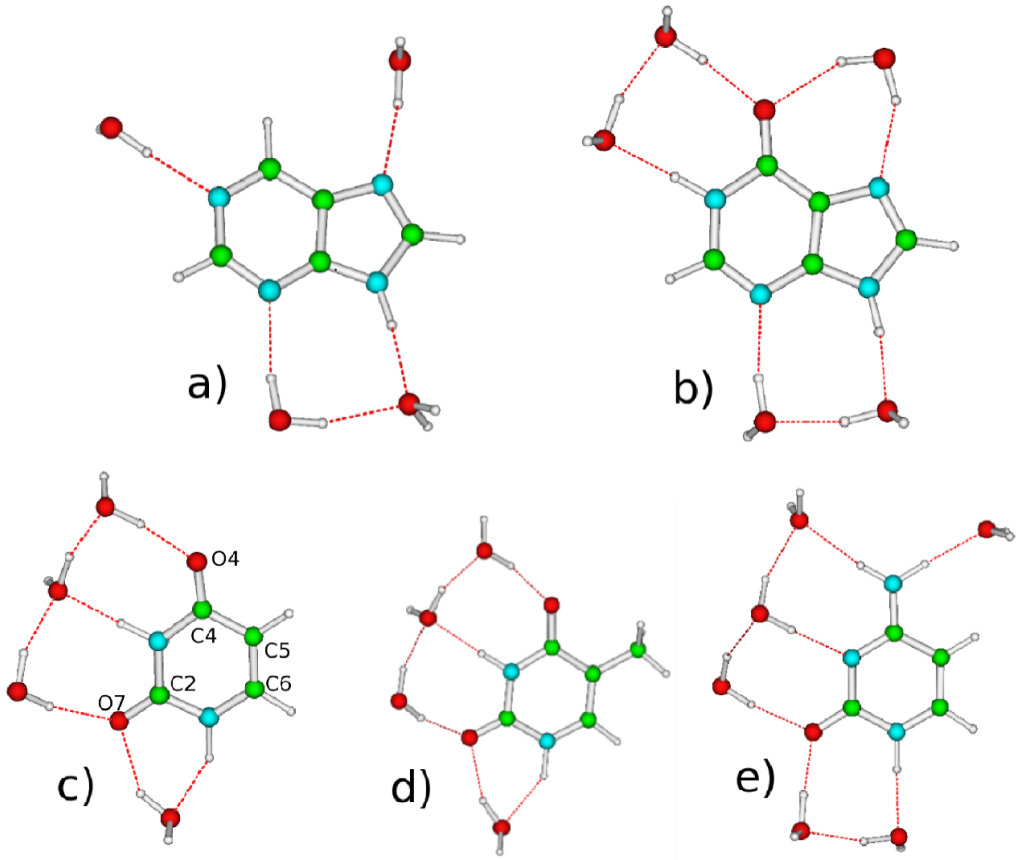
3. For each N with 3 σ bonds, we included one water molecule that accepts the donation of the H atom bound to the N atom;

4. We checked the molecules' steric encumbrance, and whether it was realistic that a given molecule played more than one role, or whether the various H_2O molecules formed H bonds among them. If that was the case, we oriented them as to allow it.

For uracil and thymine, we adopted the cluster model of Ref. 54, which, as discussed in Ref. 36, is consistent with the available experimental and computational results, as starting geometry and re-optimized it at the B3LYP/cc-pVTZ level of theory. A schematic representation of the different species including their coordinating water molecules is given in Figure 1.

The absorption and MCD spectra were computed with both B3LYP and CAM-B3LYP functionals and the aug-cc-pVDZ basis set. All geometry optimizations were performed with Gaussian09[55] adopting default settings for PCM. Absorption and MCD spectra calculations were carried out with Dalton.[56,57] They were done in gas phase on the gas-phase optimized species and on

Figure 1: Schematic representation of the investigated molecules with the coordinating water molecules included in some of the calculations: (a) purine + $4H_2O$; (b) hypoxantine + $5H_2O$; (c) uracil + $4H_2O$; (d) thymine + $4H_2O$; (e) cytosine + $6H_2O$.



the PCM optimized clusters, and in the solvated environment described by PCM in the case of the PCM optimized isolated basis and of the PCM optimized clusters with explicit water molecules.

In the PCM calculation with Dalton, the code's recommended spherical cavities of 1.2, 1.7, 1.6, and 1.5 Å and default cavity scaling factor of 1.2 were used for H, C, N and O, respectively. In the case of C-H bonding, the hydrogen was encapsulated within an enlarged sphere centered at the carbon atom, assuming values of 1.9, 2.0, and 2.1 Å for 1, 2, and 3 hydrogen atoms, respectively. A similar case holds for N-H bonds, where an enlarged sphere exhibits a radius of 1.8 Å for 1 additional hydrogen; in the case of 2 hydrogen atoms, an exception was made and the three atoms were treated separately. The coordinating water molecules were enclosed in spheres of 1.8 Å. For the cluster of purine with 4 water molecules, due to convergence problems using the above radii, a different set of values was adopted, namely 1.925 Å for C, 2.125 Å for CH, 1.83 Å for N, 1.93 Å for NH and 1.95 Å for the coordinating water molecules. They were taken from the default PCM settings of the Gaussian03 code.

A non-equilibrium PCM solvent model was adopted, with optical dielectric constant value of 1.7760 and static dielectric constant of 78.39.

Following previous practice,[19,33] a common lifetime parameter γ of 1000 cm^{-1} ($\approx$ 0.004556 a.u.) and a frequency step of 0.0025 a.u. were used in the CPP calculations. The experimental spectra were digitized from the original references using WebPlotDigitizer,[58] and all plots have been generated using Matplotlib.[59]

To conclude this section, a few words are in place on units. Experimental MCD spectra are usually reported either as anisotropy of the molar decadic coefficient $\Delta\varepsilon$, or as molar ellipticity $[\Theta]_M$, either at given values of the external magnetic field, or normalized to 1 unit of magnetic field (Gauss or Tesla). The standard unit of the molar ellipticity (for unit of magnetic field) is deg dm^3 mol^{-1} cm^{-1} T^{-1}. To convert from $[\Theta]_M$ given in deg M^{-1} cm^{-1} T^{-1} to $\Delta\varepsilon$ in M^{-1} cm^{-1} T^{-1} we use the relation[45]

$$[\Theta]_M(\omega) = \frac{18000\ln(10)}{4\pi}\Delta\varepsilon(\omega) \approx 3298\Delta\varepsilon(\omega) \tag{10}$$

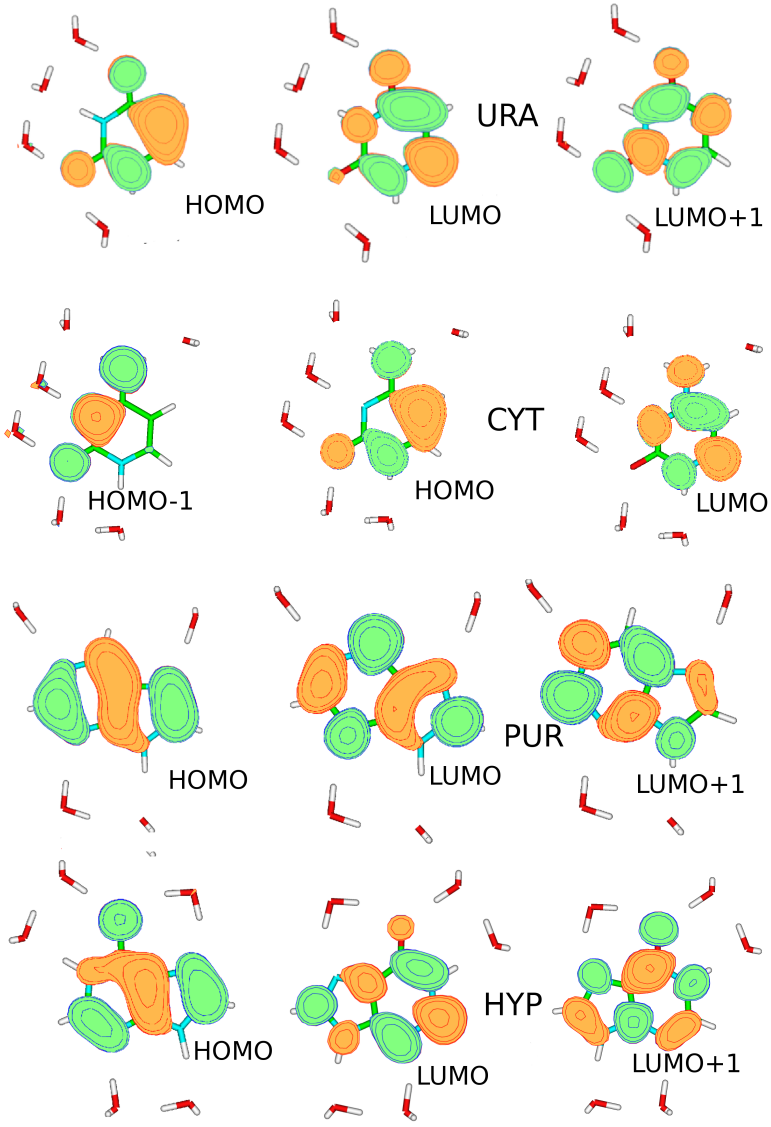
Discussion of Results

Before discussing in detail the MCD spectra of the compounds we examined, it is worthwhile to briefly sketching some general features of their lowest-energy excited states (see Figure 2 for a schematic description of the frontier orbitals).

The spectra of the purines are dominated by two close lying $\pi\pi^*$ transitions. The lowest-energy one, usually labeled as L_a, can be described as a HOMO→LUMO excitation. Another close-lying transition is found, usually labeled as L_b, with a predominant contribution of a HOMO→LUMO+1 excitation (see Figure 2).

As for the purines case, the MCD spectra of pyrimidines are ruled by the lowest energy $\pi\pi^*$ excited states, two for all the compounds examined. The lowest energy bright excited state can always be described as a HOMO→LUMO excitation (the frontier orbitals of uracil and cytosine are shown in Figure 2. The thymine orbitals are very similar to those of uracil). The HOMO has a similar shape in all the compounds examined (though with significant differences, discussed below), with a strong bonding contribution with respect to the C5-C6 bonds, see Figure 1 for the atoms' labeling, and a significant participation from the *p* orbital of the oxygen bound to C2 and of the heteroatom bound to C4 (labeled as R), either the oxygen of a carbonyl group or a nitrogen of an amino group, depending on the compounds examined. The LUMO is instead a π^* orbital, always antibonding between C5-C6, and antibonding/non bonding for C2-O and C4-R bonds. For uracil and thymine a higher lying $\pi\pi^*$ excited state plays a role in the MCD spectra and can be described as a HOMO→LUMO+1 excitation, antibonding with respect to C2-O and C4-R bonds. This state is also present for cytosine (Cyt), with an energy gap with respect to the lowest energy similar to the one of Ura and Thy. For Cyt, however, another bright $\pi\pi^*$ lies between those just described and it can described as an HOMO-1→LUMO transition (see Figure 2).

Figure 2: Schematic drawing of the Molecular Orbitals involved in the most relevant low-energy bright excited states of Uracil (those of Thymine are very similar), Cytosine (CYT), Purine (PUR) and Hypoxanthine (HYP)



Purine and hypoxanthine

As also commented upon by Sutherland and Holmquist,[60] purines are prototypical examples to showcase the ability of MCD spectroscopy to resolve multiple transitions contained in a single absorption envelope, due to the biphasic nature of their MCD peaks.

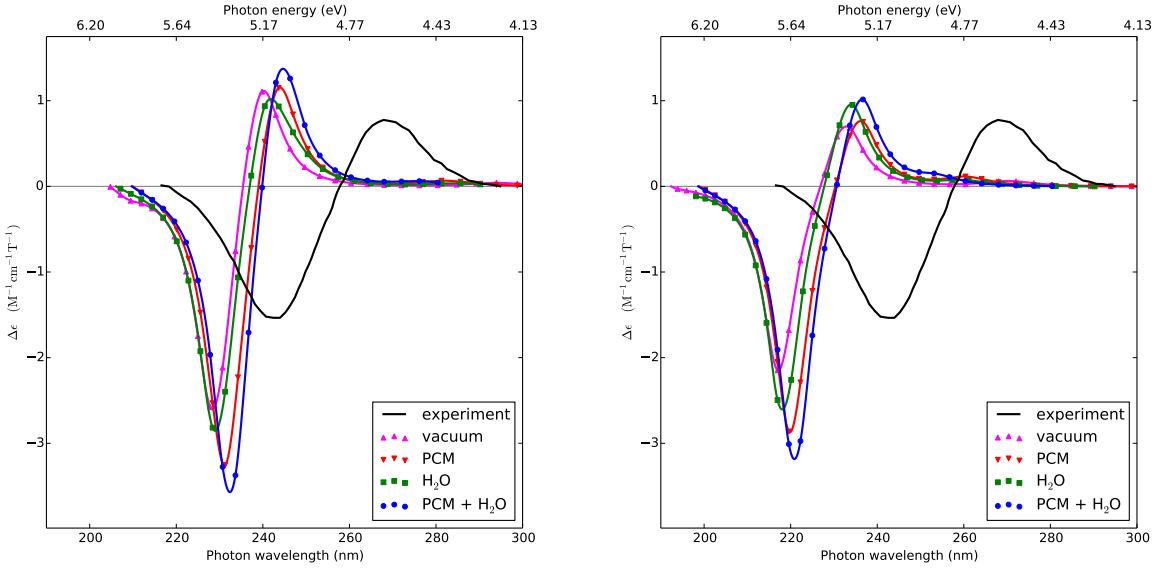
Purine. The absorption and MCD spectra of purine in aqueous solution at pH 7 were measured between 200 and 300 nm by Djerassi and coworkers, and reported in their review in Ref. 3. In the given frequency range, the absorption spectrum is characterized by one band with maximum at 263 nm (classified as B_{2u} in the benzene nomenclature), with a shoulder at around 240 nm (B_{1u}). A pair of E_{1u} bands are expected below 200 nm. The n$\rightarrow \pi^*$ transitions are generally hidden under the more intense $\pi \rightarrow \pi^*$ transitions, but fluorescence measurements place one of these at around 280 nm. The intensity difference between the two MCD peaks is considered an indication of the presence of a n$\rightarrow \pi^*$ transition.[3] Our calculations of the excitation energies and strengths, see Table 1, confirm the presence of n$\rightarrow \pi^*$ transitions both below and above (in energy) the intense $\pi \rightarrow \pi^*$ absorption band.

The experimental MCD spectrum is compared in Figure 3 with our calculated spectra obtained with the two DFT functionals in gas phase, in (bulk) aqueous solution as described by PCM, for the cluster of purine and four water molecules in vacuo, and for the cluster of purine and four water in bulk aqueous solution.

In all cases, the pseudo-*A* term shape of the MCD spectral profile is qualitatively reproduced, including the fact that the first peak is less intense than the second one. Both B3LYP and CAM-B3LYP peaks are however shifted towards higher energies, with the largest shifts observed for CAMB3LYP. The separation between the two peaks is (for the selected broadening) $\approx$0.27–0.29 eV for B3LYP and $\approx$0.37–0.39 eV for CAM-B3LYP, see Table 2, this latter estimate being closer to the experimental value of 0.45 eV.

Selected *B*-term calculations in gas phase (see Table 1) confirm that the observed bands are predominantly due to the strong $\pi \rightarrow \pi^*$ transitions. Both functionals show shoulders in the MCD spectra due to the n$\rightarrow \pi^*$ transition in the low energy region preceding the first intense peak, as

Figure 3: MCD spectra of purine in gas phase and in water solution. Left panel: B3LYP; Right panel: CAM-B3LYP. The black line in both panels is the experimental spectrum in water by Djerassi and coworkers.[3]



anticipated by fluorescence measurements, but their intensity is predicted to be too small to be observed, in agreement with experiment.

Inclusion of environmental effects improves the agreement with experiment, shifting the intense bands toward lower energies; the redshift with respect to gas-phase is $\sim$ 0.1 eV for the cluster with explicit water molecule embedded in PCM. On the contrary, as expected, the weak nπ^* shoulders are moved towards higher energies, since the transition weakens the interaction of the lone pairs with the solvent. The peak intensities tend to increase, with the highest intensities observed for the cluster of purine and four water molecules in bulk water solution.

Hypoxanthine. The absorption and MCD spectra of hypoxanthine in aqueous solution (at pH 6) up to 180 nm were reported (together with those of adenine and guanosine 5'-diphosphate at pH 7) by Sutherland and Griffin.[61] Absorption derivative spectra were also included. The authors observe that in the 300-180 nm region both hypoxanthine and adenine manifest two absorption bands, referred to as the 260-nm and the 200-nm bands. As for purine, the corresponding MCD spectra are biphasic (due to the *B*-type bands of opposite sign of two, unresolved, electronic

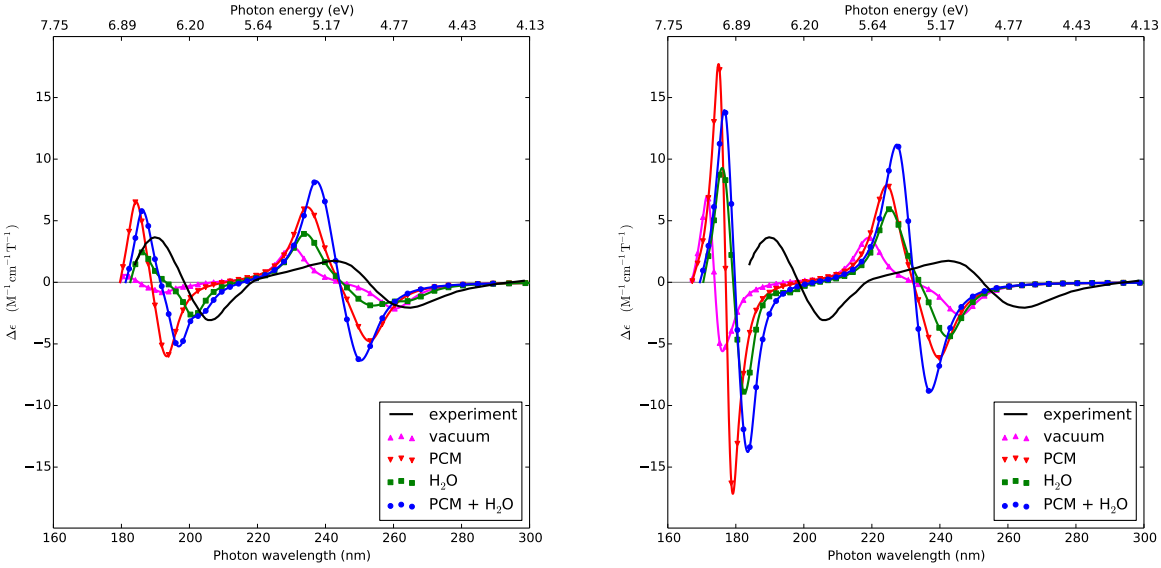
transitions in the absorption envelope with non-parallel transition dipoles), but of opposite sign compared to the purine precursor, i.e. negative on the long-wavelength side, and positive at short-wavelength side, with zero-crossing near the λ of the maximum absorption. According to Sutherland and Griffin, the MCD spectra of both hypoxanthine and adenine demonstrate the presence of at least two unresolved components also in the 200-nm region. From an analysis of the derivative absorption spectra the authors also conclude that the less intense component of the 200-nm band of adenine is on the short wavelength side, and on the long wavelength side for hypoxanthine.

Voelter *et al.*[2] tabulated values of the peak maxima at pH 7 measured between 300 and 200 nm. A maximum absorption peak is given at 248 nm and $\varepsilon = 9.8 \times 10^3$ M^{-1} cm^{-1}. The maxima of the MCD spectrum are found at 263 nm and $[\Theta]_M = -27.5 \times 10^3$ deg mol^{-1} cm^2 (related to a 49.5 kG magnetic field), classified as B_{2u} according to the benzene-like nomenclature; at 243 nm and $[\Theta]_M = +24.2 \times 10^3$ deg mol^{-1} cm^2, classified B_{1u} and a negative shoulder at lower wavelengths (in the 210 region, most probably). Converting in molar ellipticity, $\Delta\varepsilon$, per one Tesla of magnetic field, the maxima at 263 nm corresponds to -1.68 M^{-1} cm^{-1} T^{-1}, and the one at 243 nm corresponds to $+1.45$ M^{-1} cm^{-1} T^{-1}, in reasonable agreement with the measurements of Sutherland and Griffin at pH 6.[61]

Our calculated MCD spectra are compared to the experimental one in Figure 4. Also in this case the computed spectra qualitatively reproduce the biphasic nature of the signals, and are blue-shifted compared to the experimental ones. Both pairs of pseudo-*A* terms, one at low and one at high energy, are found. Yet, with the B3LYP functional, the high-energy pseudo-*A* band tends to be less intense than the low-energy one and, in the vacuum case, extremely weak. For CAM-B3LYP, the low-energy signal is broader and less intense than the high-energy one and hence, in this respect, qualitatively more in line with the experimental spectrum.

Of the two functionals, CAM-B3LYP is the one showing the larger blue shifts and spectral intensities. For both functionals, and opposite to what observed for purine, the inclusion of solvent effects tends to further blue-shift the negative peak at lower energy from to experiment, whereas the opposite occurs for the positive peak. These shifts lead to a net decrease of the separation of the

Figure 4: MCD spectra of hypoxanthine in gas phase and in water solution. Left panel: B3LYP; right panel: CAM-B3LYP. The black line in both panels is the experimental spectrum in water by Sutherland and Griffin.[61]



two peaks that is 0.62 eV in gas phase and 0.27 eV (B3LYP) or 0.23 eV (CAM-B3LYP) according to $6H_2O$+PCM predictions, improving the agreement with experiment (0.39 eV).

For the second pseudo-*A* term at $\sim$ 200 nm, solvent effects always decrease the blue-shift of both peaks compared to the gas phase result. As far as the splitting of the components of this second doublet is concerned, bulk solvent (PCM) and explicit solute-solvent interactions give rise to opposite effects which end up, for the $6H_2O$+PCM model, in a small decrease of the separation predicted in vacuo for B3LYP and in the opposite result according to CAM-B3LYP predictions. In absolute terms the experimental splitting is slightly better reproduced by B3LYP calculations.

Table 1: Purine. Vertical excitation energies (eV), oscillator strengths f (dimensionless) and selected B terms (a.u.)

State	Sym	Energy	f	B	Sym	Energy	f	B
	B3LYP/Gas phase				CAM-B3LYP/Gas phase			
1	A″	4.219	0.002	−0.612	A″	4.587	0.002	−0.8615
2	A′	5.176	0.143	−20.80	A′	5.335	0.157	−11.55
3	A″	5.216	0.002	−0.971	A″	5.540	0.002	−0.0655
4	A″	5.307	<0.0001	0.0721	A′	5.705	0.024	25.94
5	A′	5.419	0.008	36.30	A″	5.773	0.0001	−0.017
6	A″	5.948	0.001		A″	6.466	0.001	
7	A′	5.957	0.009		A″	6.625	0.007	
8	A″	6.265	0.003		A′	6.702	0.327	
9	A′	6.393	0.306		A′	6.851	0.064	
10	A′	6.522	0.007		A′	7.180	0.070	
	B3LYP/aq(PCM)				CAM-B3LYP/aq(PCM)			
1	A″	4.392	0.002		A″	4.759	0.003	
2	A′	5.099	0.173		A′	5.263	0.203	
3	A′	5.358	0.015		A′	5.636	0.015	
4	A″	5.364	0.001		A″	5.689	0.002	
5	A″	5.455	0.0001		A″	5.929	0.0002	
6	A′	6.288	0.411		A′	6.595	0.486	
7	A″	6.353	0.002		A″	6.719	0.001	
8	A″	6.418	0.005		A″	6.790	0.010	
9	A′	6.475	0.043		A′	7.055	0.150	
10	A′	6.815	0.172		A′	7.211	0.304	
	B3LYP/aq(+4H_2O+PCM)				CAM-B3LYP/aq(+4H_2O+PCM)			
1	A	4.491	0.001		A	4.889	0.002	
2	A	5.083	0.155		A	5.255	0.191	
3	A	5.331	0.023		A	5.610	0.015	
4	A	5.437	0.002		A	5.823	0.001	
5	A	5.465	0.001		A	6.011	0.0003	
6	A	5.480	0.0002		A	6.605	0.461	
7	A	5.493	0.000		A	6.837	0.005	
8	A	5.524	0.000		A	6.916	0.0002	
9	A	6.297	0.383		A	7.038	0.003	
10	A	6.325	0.005		A	7.110	0.004	
11	A	6.336	0.004		A	7.124	0.0002	
12	A	6.443	0.002		A	7.141	0.002	
13	A	6.478	0.0001		A	7.168	0.084	
14	A	6.526	0.002		A	7.184	0.374	
15	A	6.550	0.033		A	7.244	0.107	

Table 2: Energy of the peak maxima and their separations (in eV) in purine and hypoxanthine. Computed values are obtained with a broadening parameter of 1000 cm^{-1} and in the aug-cc-pVDZ basis set.

Environment	Peak maxima	Peak separation	Peak maxima	Peak separation
	B3LYP		CAM-B3LYP	
	Purine			
Gas phase	5.162	0.268	5.324	0.386
	5.430		5.710	
aq (PCM)	5.083	0.281	5.250	0.390
	5.364		5.640	
aq (+4H_2O)	5.126	0.288	5.300	0.387
	5.415		5.686	
aq (+4H_2O+PCM)	5.068	0.268	5.242	0.372
	5.336		5.614	
Experiment[3]	4.628	0.454		
	5.082			
	Hypoxanthine			
Gas phase	4.762	0.622	5.039	0.620
	5.384	1.058	5.659	1.382
	6.442	0.402	7.041	0.185
	6.844		7.225	
aq (PCM)	4.902	0.382	5.172	0.353
	5.284	1.129	5.525	1.397
	6.413	0.314	6.921	0.167
	6.727		7.088	
aq (+5H_2O)	4.881	0.417	5.116	0.388
	5.298	0.866	5.504	1.286
	6.164	0.496	6.790	0.255
	6.660		7.045	
aq (+5H_2O+PCM)	4.952	0.274	5.229	0.225
	5.226	1.072	5.454	1.301
	6.298	0.359	6.756	0.264
	6.657		7.020	
Experiment[61]	4.679	0.433		
	5.112	0.893		
	6.005	0.527		
	6.533			

Table 3: Hypoxanthine. Vertical excitation energies (in eV), oscillator strengths f (dimensionless) and selected B terms (a.u.).

State	Sym	Energy	f	B	Sym	Energy	f	B
	B3LYP/Gas phase				CAM-B3LYP/Gas phase			
1	A′	4.762	0.132	31.06	A′	5.041	0.139	36.77
2	A″	5.221	0.0021	0.437	A″	5.512	<0.0001	0.367
3	A″	5.263	0.0002	0.840	A′	5.656	0.210	−42.26
4	A″	5.277	0.0000	−0.004	A″	5.725	0.002	−2.631
5	A″	5.353	0.001	−0.310	A″	5.779	0.001	−0.307
6	A′	5.385	0.150	−37.78	A″	6.028	0.0005	0.0995
7	A″	5.437	0.0002	0.780	A″	6.274	0.001	−0.622
8	A″	6.195	0.001	−4.175	A″	6.346	0.008	0.012
9	A′	6.201	0.002	7.336	A′	6.683	0.016	
10	A′	6.230	0.018		A′	6.964	0.110	
11	A′	6.430	0.051		A′	7.025	0.013	
	B3LYP/aq(PCM)				CAM-B3LYP/aq(PCM)			
1	A′	4.907	0.153		A′	5.179	0.141	
2	A′	5.277	0.220		A′	5.522	0.309	
3	A″	5.492	<0.0001		A″	5.761	<0.0001	
4	A″	5.576	0.001		A″	5.978	0.002	
5	A″	5.662	0.003		A″	6.046	0.002	
6	A″	5.779	0.0003		A″	6.43 9	0.001	
7	A″	5.930	0.0003		A″	6.647	0.001	
8	A′	6.333	0.006		A′	6.801	0.029	
9	A′	6.426	0.088		A′	6.964	0.231	
10	A′	6.725	0.444		A′	7.045	0.421	
	B3LYP/aq(+5H_2O+PCM)				CAM-B3LYP/aq(+5H_2O+PCM)			
1	A	4.960	0.141		A	5.244	0.110	
2	A	5.213	0.231		A	5.442	0.336	
3	A	5.561	<0.0001		A	5.911	0.0001	
4	A	5.654	0.002		A	6.185	0.001	
5	A	5.716	0.0005		A	6.202	0.001	
6	A	5.933	0.0001		A	6.473	0.001	
7	A	5.945	0.0002		A	6.763	0.231	
8	A	6.014	0.001		A	6.800	0.003	
9	A	6.083	0.032		A	6.840	0.006	
10	A	6.140	0.0003		A	6.870	0.060	

Uracil, thymine and cytosine

As discussed by Sutherland and Griffin,[61] and by Sutherland and Holmquist,[60] the experimental MCD spectra of the biologically relevant pyrimidine bases (uracil, thymine, cytosine) are substantially different from that of the purines, as they are more of *B* type instead of pseudo-*A* type. Moreover, the intensity is 3-4 times lower than that of the purines.

Uracil. The experimental absorption and MCD spectra of uracil in water at pH 7 were reported by Voelter *et al.*[2] The absorption spectrum presented two broad bands, one centered at 258 nm and one at 203 nm of similar intensity. The MCD spectrum was characterized by a very broad negative band stretching in between 220 and 290 nm and peaked at 255 nm (with tabulated maximum peak intensities, for a field strength of 49500 Gauss, of $[\Theta]_M = -8.5 \times 10^3$ hence $\Delta\varepsilon$ = -0.52 M^{-1} cm^{-1} T^{-1}), and a narrower positive band peaking at 212 nm (with max intensity of $[\Theta]_M = +6.7 \times 10^3$, i.e. $\Delta\varepsilon$ = 0.41 M^{-1} cm^{-1} T^{-1}).

The experimental MCD spectrum is compared with our computed spectra in Figure 5. Table 4 reports the computed excitation energies and oscillator strengths, and Table 5 the energy separation between the peaks.

In accordance with the experiment, the computed spectra show a negative band at lower energy and a positive one at higher energy. The bands are primarily due to the two intense $\pi\pi^*$ excitations. The separation between the two peak maxima is in all cases larger than the one obtained from the experimental spectrum. As for the other molecules analyzed, all computed spectra are blue shifted, the CAM-B3LYP ones slightly more than the B3LYP ones.

From the qualitative point of view, the spectrum of uracil is more sensitive to bulk-solvent effect than to explicit solute-solvent interactions. This is not surprising for a $\pi\pi^*$ transition with small contribution from the carbonyl groups. In any case, a weak solvent red-shift is predicted, in agreement with previous computational results and experimental indications (see Ref. 62 for a more detailed discussion).

B3LYP also predicts the presence of one strong negative feature (due to the HOMO→LUMO excitation) and a positive one (due to the HOMO→LUMO+1 excitation). On the other hand,

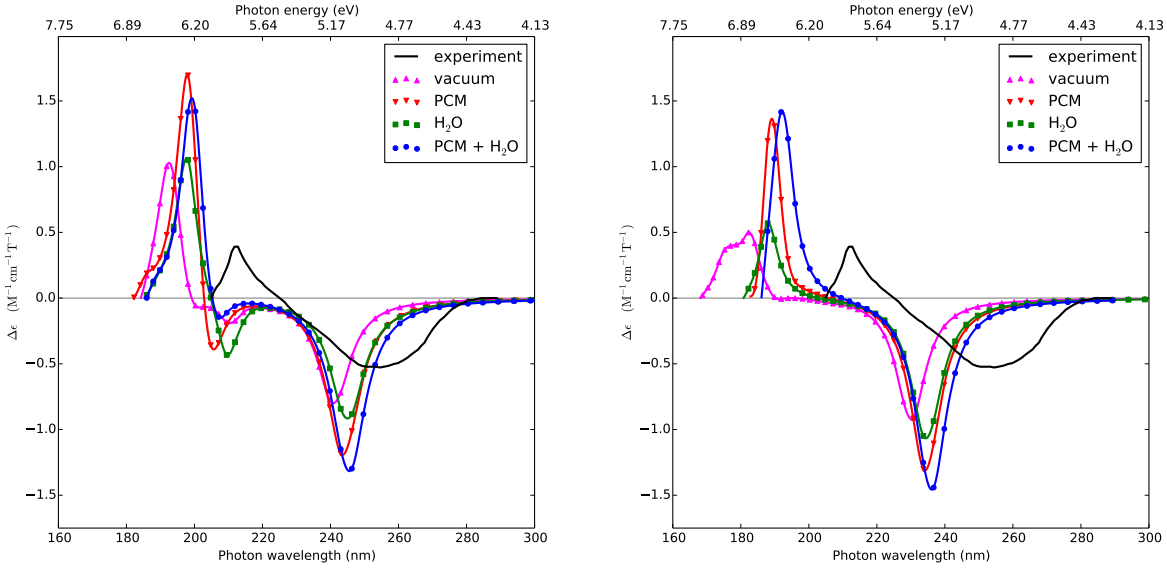
Table 4: Uracil. Vertical excitation energies (in eV), oscillator strengths f (dimensionless), and selected B terms (a.u.).

State	Sym	Energy	f	B	Sym	Energy	f	B
	B3LYP/Gas phase				CAM-B3LYP/Gas phase			
1	A″	4.668	<0.0001	−0.001	A″	5.064	<0.0001	0.002
2	A′	5.144	0.122	10.45	A′	5.389	0.172	11.51
3	A″	5.598	0.002	0.134	A′	6.012	0.003	0.106
4	A″	5.769	<0.0001	−0.251	A′	6.314	<0.0001	−0.041
5	A′	5.890	0.032	2.226	A′	6.535	0.037	0.913
6	A″	6.149	0.0004	0.078	A′	6.778	0.161	−5.064
7	A′	6.221	0.025	2.547	A″	6.940	0.012	−0.621
8	A′	6.425	0.121	−11.57	A″	7.012	0.004	−0.849
9	A″	6.503	0.007		A″	7.103	0.002	−1.753
10	A″				A′	7.161	0.017	0.196
	B3LYP/aq(PCM)				CAM-B3LYP/aq(PCM)			
1	A"	4.9818	<0.00001		A"	5.761	<0.00001	
2	A′	5.085	0.175		A’	5.290	0.230	
3	A″	5.989	0.003		A"	5.978	0.010	
4	A″	5.998	0.004		A"	6.046	<0.00001	
5	A′	6.087	0.087		A"	6.439	0.0004	
6	A′	6.252	0.143		A’	6.561	0.211	
7	A″	6.335	0.0005		A"	6.647	0.004	
8	A″	6.590	<0.0001		A’	6.710	0.092	
9	A′	6.791	0.016		A’	7.374	0.021	
	B3LYP/aq(+4H_2O+PCM)				CAM-B3LYP/aq(+4H_2O+PCM)			
1	A	5.051	0.183		A	5.253	0.237	
2	A	5.124	0.0002		A	5.514	0.0002	
3	A	5.952	0.002		A	6.438	0.223	
4	A	6.005	0.004		A	6.533	0.012	
5	A	6.011	0.001		A	6.607	0.001	
6	A	6.075	0.172		A	6.754	0.104	
7	A	6.214	0.073		A	7.177	0.004	
8	A	6.346	0.001		A	7.210	0.003	
9	A	6.368	0.001		A	7.276	0.001	
10	A	6.577	0.007		A	7.385	0.005	

Table 5: Energy of the peak maxima and their separations (in eV) in uracil, thymine and cytosine.

	Peak maxima	Peak separation	Peak maxima	Peak separation
Environment	B3LYP		CAM-B3LYP	
	Uracil			
Vacuum	5.148	1.291	5.387	1.412
	6.439		6.799	
aq(PCM)	5.090	1.175	5.295	1.260
	6.266		6.555	
aq(+4H_2O)	5.060	1.211	5.289	1.304
	6.271		6.593	
aq(+4H_2O+PCM)	5.050	1.173	5.251	1.203
	6.222		6.453	
Experiment[2]	4.871	0.955		
	5.827			
	Thymine			
Vacuum	4.959	1.345	5.209	1.406
	6.303		6.615	
aq(PCM)	4.907	1.294	5.122	1.305
	6.201		6.427	
aq(+4H_2O)	4.888	1.248	5.106	1.334
	6.137		6.440	
aq(+4H_2O+PCM)	4.881	1.074	5.090	1.215
	5.954		6.305	
Experiment[2]	4.714	1.052		
	5.767			
	Cytosine			
Vacuum	4.633	0.806	4.939	0.952
	5.439		5.891	
aq(PCM)	4.882	0.601	5.114	0.797
	5.483		5.911	
aq(+6H_2O)	4.808	0.682	5.101	0.726
	5.490		5.827	
aq(+6H_2O+PCM)	4.965	0.524	5.186	0.726
	5.490		5.912	
Experiment[2]	4.674	0.571		
	5.245			

Figure 5: MCD spectra of uracil in gas phase and in water solution. Left panel: B3LYP; right panel: CAM-B3LYP. The black line in both panels is the experimental spectrum in water by Voeltner *et al.*[2]



a small negative feature (two in the vacuum case) falls between these two peaks in the B3LYP spectrum, which is absent in the experimental one. B3LYP overestimates indeed the stability of another $\pi\pi^*$ transitions, that can be described as an HOMO-3→LUMO transition, providing that it falls $\sim$ 0.2 eV on the red with respect to the HOMO→LUMO+1 transition, whereas CAM-B3LYP provides the opposite picture.

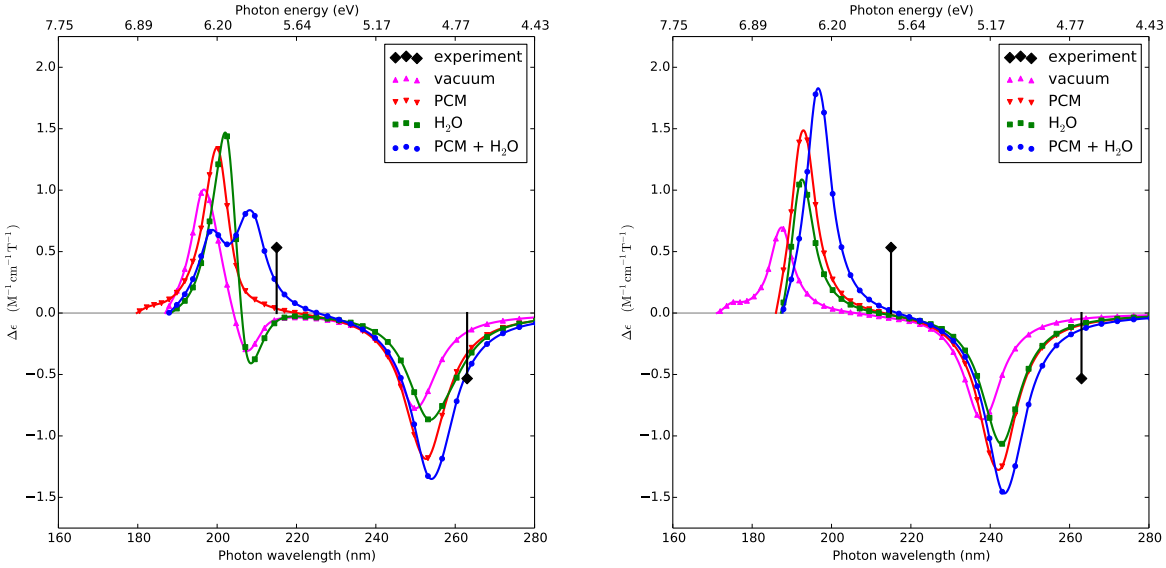
Thymine. The picture obtained for thymine is similar to that of uracil, see Figure 6. Even though the full MCD experimental spectrum in water has not been published, the experimental peak maxima in the UV-vis and MCD were reported by Voelter et al.[2] The latter are indicated as black diamonds in Figure 6. Both functionals provide MCD spectra in qualitative agreement with experiments. B3LYP, indeed, when compared to uracil, provides spectra more consistent with the experimental findings, especially when both bulk effects and explicit solute-solvent interactions are considered.

Cytosine. Experimental absorption and MCD spectra of cytosine in water have been reported both by Voelter *et al.*,[2] and by Kaito *et al.*[63] The MCD ones are shown in Figure 7 together with our

Table 6: Thymine. Vertical excitation energies (in eV), oscillator strengths f (dimensionless), and selected B terms (a.u.).

State	Sym	Energy	f	B	Sym	Energy	f	B
	B3LYP/Gas Phase				CAM-B3LYP/Gas Phase			
1	A″	4.741	0.00001	−0.012	A″	5.103	0.00001	0.035
2	A′	4.954	0.129	10.50	A′	5.206	0.176	11.28
3	A″	5.360	0.0002	0.093	A″	5.772	0.0005	0.089
4	A″	5.830	0.0002	−0.416	A″	6.401	0.00001	−0.109
5	A′	5.978	0.060	5.199	A″	6.539	0.0005	−0.357
6	A″	6.037	0.0000	0.031	A′	6.593	0.042	2.057
7	A″	6.188	0.00004	−0.993	A′	6.617	0.200	−9.335
8	A′	6.198	0.090	1.136	A″	6.696	0.0001	0.5220
9	A″	6.213	0.0003	−2.981	A″	7.105	0.001	−0.7051
10	A′	6.310	0.061	−9.187	A′	7.114	0.008	−0.6231
11	A′	6.731	0.001		A′	7.159	0.013	
	B3LYP/aq(PCM)				CAM-B3LYP/aq(PCM)			
1	A′	4.908	0.177		A′	5.124	0.228	
2	A″	5.007	0.00001		A″	5.356	0.00002	
3	A″	5.724	0.0003		A″	6.126	0.001	
4	A′	6.047	0.233		A′	6.428	0.258	
5	A″	6.062	0.0002		A″	6.589	<0.00001	
6	A″	6.167	0.003		A″	6.699	0.002	
7	A′	6.198	0.053		A′	6.739	0.116	
8	A″	6.289	<0.00001		A″	6.858	0.0003	
9	A′	6.733	0.018		A′	7.165	0.005	
10	A′	6.763	0.003		A′	7.331	0.023	
	B3LYP/aq(+4H_2O+PCM)				CAM-B3LYP/aq(+4H_2O+PCM)			
1	A	4.876	0.184		A	5.085	0.233	
2	A	5.145	0.0001		A	5.505	0.0001	
3	A	5.727	0.0002		A	6.254	0.004	
4	A	5.943	0.235		A	6.300	0.265	
5	A	6.030	0.009		A	6.706	0.0004	
6	A	6.076	0.003		A	6.779	0.135	
7	A	6.078	0.001		A	6.831	0.004	
8	A	6.198	0.0003		A	6.942	0.0001	
9	A	6.242	0.061		A	7.234	0.007	
10	A	6.406	0.0004		A	7.293	0.001	

Figure 6: MCD spectra of thymine computed in gas phase and in water solution. Left panel: B3LYP; right panel: CAM-B3LYP. The black diamonds are the experimental peak maxima in water reported by Voelter *et al.*[2]



corresponding computed spectra. Table 7 reports our computed excitation energies and oscillator strengths, and Table 5 the energy separation between the peaks. Voelter et al.[2] also tabulated one MCD maximum $(\Theta]_M$) at pH 7 of -4.5×10^3 at 270 nm, for a field strength of 49500 Gauss, which corresponds to $\Delta\varepsilon_{270} = -0.275$ $\mathrm{M^{-1}cm^{-1}T^{-1}}$.

Inclusion of solvent effect leads to a decrease of the separation between the two negative bands of the MCD spectrum, that is 0.81 eV in vacuo and 0.52 eV with the $6H_2O$+PCM model according to B3LYP; these values are respectively 0.95 and 0.73 eV considering CAM-B3LYP results. Therefore, according to both functionals, solvent effects improve the agreement with the experimental value (0.65 eV). The MCD spectrum of cytosine, especially the HOMO→LUMO transition, is more sensitive to explicit solute-solvent interaction than that of uracil and thymine and, contrarily to what predicted for Ura and Thy, even if the frontier orbitals are similar, our calculations predict an hypsochromic solvent shift for the lowest energy band, i.e. the negative peak is blue-shifted when going from the gas phase to water. This prediction agrees with what was previously obtained using multiconfigurational self-consistent field (MCSCF) based approaches.[62,64,65] It is

difficult to check whether this prediction is confirmed in experiment. In fact, solvent affects the conformational equilibrium among the different cytosine tautomers, which have different absorption spectra.[66,67]

Besides to the possible small contribution of other excitations, this result could be due to the fact that for Cyt the LUMO does not receive any contribution from O7 *p* orbital and, especially, that hydrogen bonds are expected to have an opposite effect on the contributions of the *p* orbital of Nitrogen of the amino group in Cyt and of the *p* orbital of O4 in Ura and Thy. For Ura/Thy O4 is an H-bond acceptor, whereas for Cyt the amino group is an H-bond donor.

Figure 7: MCD spectra of cytosine computed in gas phase and in water solution. Left panel: B3LYP; right panel: CAM-B3LYP. The full-drawn black line in both panels is the experimental spectrum in water by Voeltner *et al.*[2] The dashed line is the MCD spectrum in water by Kaito *et al.*[63]

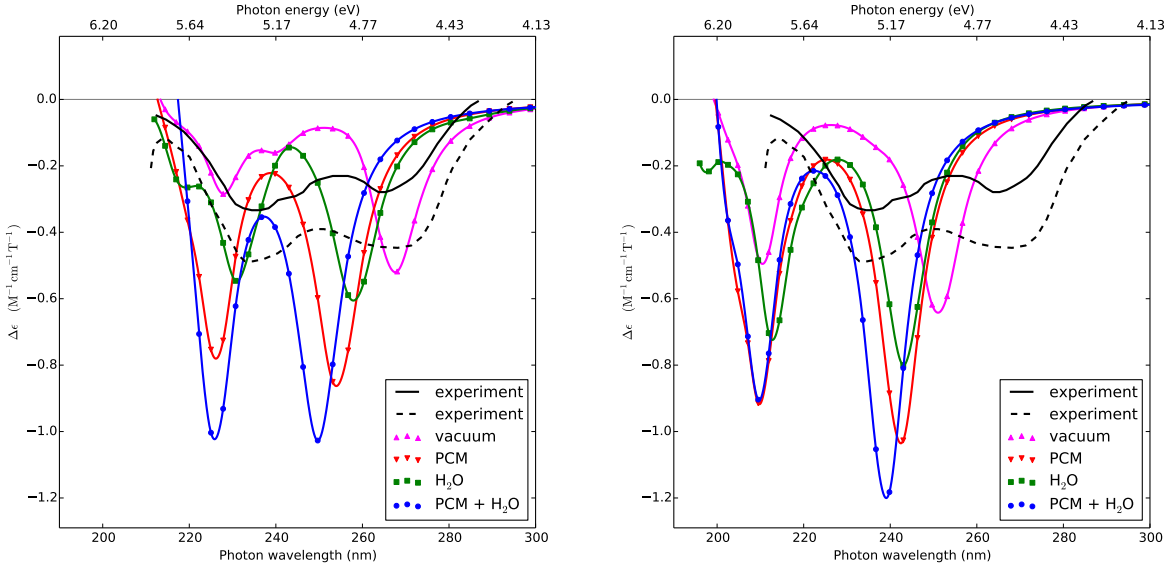


Table 7: Cytosine. Vertical excitation energies (in eV), oscillator strengths f (dimensionless), and selected B terms.

State	Sym	Energy	f	B	Sym	Energy	f	B
	B3LYP/Gas phase				CAM-B3LYP/Gas phase			
1	A′	4.650	0.043	7.840	A′	4.948	0.067	8.495
2	A″	4.783	0.001	−1.344	A″	5.266	0.002	0.448
3	A″	5.126	0.003	−0.9403	A″	5.626	0.004	−0.065
4	A″	5.151	0.003	2.119	A′	5.885	0.125	8.164
5	A′	5.433	0.080	3.391	A″	5.897	0.0001	−2.767
6	A″	5.509	0.0004	−0.5115	A″	6.091	0.0001	0.201
7	A″	5.687	0.0004	−0.286	A″	6.135	0.002	0.426
8	A″	5.694	0.004	0.4669	A″	6.240	0.098	−1.675
	B3LYP/aq(PCM)				CAM-B3LYP/aq(PCM)			
1	A	4.876	0.087		A	5.115	0.128	
2	A	5.323	0.002		A	5.717	0.003	
3	A	5.478	0.130		A	5.901	0.190	
4	A	5.638	0.0002		A	6.080	0.006	
5	A	5.663	0.004		A	6.357	0.0002	
6	A	6.051	0.159		A	6.370	0.330	
7	A	6.083	<0.0001		A	6.529	0.005	
8	A	6.105	0.004		A	6.620	0.0003	
9	A	6.205	0.002		A	6.814	0.418	
10	A	6.373	0.007		A	6.829	0.0001	
	B3LYP/aq(6H_2O+PCM)				CAM-B3LYP/aq(6H_2O+PCM)			
1	A	4.962	0.125		A	5.187	0.178	
2	A	5.473	0.083		A	5.900	0.115	
3	A	5.549	0.033		A	5.991	0.057	
4	A	5.555	0.003		A	6.185	0.006	
5	A	5.840	0.002		A	6.245	0.252	
6	A	5.878	0.135		A	6.593	0.005	
7	A	5.907	0.003		A	6.629	0.005	
8	A	5.945	0.003		A	6.744	0.472	
9	A	6.048	0.003		A	6.834	0.010	
10	A	6.083	0.004		A	6.864	0.031	

Concluding remarks

We have presented a computational investigation at the time-dependent density functional level of theory (B3LYP and CAM-B3LYP functionals) of the MCD spectra of selected purine and pyrimidine bases of nucleic acids in aqueous solution. The importance of solvent interactions was assessed using the polarizable continuum model, model clusters of one solute molecule with water molecules and clusters with water molecules embedded in a polarizable continuum environment. All the spectra computed at the CAM-B3LYP level match very well the experimental band-shape confirming the general reliability of this functional. The computed peaks are systematically blue-shifted with respect to the experimental ones. On the other hand part of this discrepancy would be likely recovered by inclusion of vibrational effects, leading to an even better agreement with experiments. Indeed, the vertical transition does not corresponds to the spectral maximum, and it has been shown[68] that in many cases introduction of vibronic effects leads to a sizeable red-shift of the maximum peak that can frequently be in the range 0.1-0.3 eV (however such shift is of course dependent on the system, the electronic state, the temperature and the envirnonment). Interestingly, the energy gap between the different peaks is also well reproduced, the difference with the experimental estimates (obtained as the energy difference between the band maxima) being 0.1$\sim$0.2 eV; only for hypoxanthine slightly larger values are found. This result shows that the quality of CAM-B3LYP results does not deteriorate for higher energy transitions. B3LYP also provides quite accurate indications, showing qualitative discrepancy with respect to experiments only for Ura. The predicted maxima are also closer to the experimental peak, and the energy gap between the different peaks reproduced with an overall degree of accuracy similar to that obtained at the CAM-B3LYP level.

The most significant discrepancy with the experimental spectra concerns the overestimation of the computed intensities, although this latter is overemphasized by our choice for the γ parameter that, as it can be appreciated in Figure 3–Figure 7, corresponds to a linewidth narrower than what observed in the experiment. It has been shown before that the LR-PCM/TD-DFT method leads in some cases to to an overestimation of the computed intensities for absorption.[69,70] It has also been

observed earlier, namely in the PCM study of Ref. 11 that the introduction of PCM increased (in absolute value) the strength parameter (B term) of MCD compared to the gas phase. Here, as in our previous study on Adenine,[33] the MCD (and absorption) intensities computed in solution are indeed systematically significantly larger than those obtained in the gas phase, the effect of PCM being usually larger than that of explicit solvent molecules. In any case, for what concerns the inclusion of solvent effects, the mixed discrete/continuum procedure we adopted is confirmed to represent a good compromise between accuracy and computational cost.[71]

The peaks computed in solution are much closer to their experimental counterpart than those obtained in the gas phase. Interestingly, inclusion of solvent effects positively affects the accuracy of our prediction of the relative energy of the different peaks, suggesting that the computational protocol here adopted is able to reproduce solvent effects also on the energetic of $\pi\pi^*$ transitions, which are, in principle, less sensitive to the polarity of the embedding medium.

The present results together with those recently communicated in Ref. 33 on the two purine bases of DNA, adenine and guanine, and their nucleosides, show that the adopted computational protocol is able to reproduce the similarities and differences of the MCD signals of these structurally related compounds indicating that it provides a reliable description of the relative energies and of the properties of the electronic states of DNA and RNA nucleobases.

Acknowledgment

It is a pleasure to dedicate this article to Professor Jacopo Tomasi in occasion of his eightieth birthday. Computer time from CINECA-UniTS and National Supercomputer Centre (NSC) in Linköping is acknowledged. This work has been supported by the PRIN 2010-2011 (Project no. 2010ERFKXL_008) funding scheme. SC acknowledges support from the University of Trieste (grant CHIM02-Ricerca). The COST–CMTS Action CM1002 “COnvergent Distributed Environment for Computational Spectroscopy (CODECS)” is also acknowledged.

References

(1) Mason, W. R. *A Practical Guide to Magnetic Circular Dichroism Spectroscopy*; Wiley: New York, 2007.

(2) Voelter, W.; Records, R.; Bunnenberg, E.; Djerassi, C. Magnetic Circular Dichroism Studies. VI. Investigation of Some Purines, Pyrimidines, and Nucleosides. *J. Am. Chem. Soc.* **1968**, *90*, 6163.

(3) Djerassi, C.; Bunnenberg, E.; Elder, D. L. Organic Chemical Applications of Magnetic Circular Dichroism. *Pure Appl. Chem.* **1971**, *25*, 57–90.

(4) Coriani, S.; Jørgensen, P.; Ruud, K.; Rizzo, A.; Olsen, J. Ab Initio Determinations of Magnetic Circular Dichroism. *Chem. Phys. Lett.* **1999**, *300*, 61–68.

(5) Coriani, S.; Hättig, C.; Jørgensen, P.; Helgaker, T. Gauge-origin Independent Magneto-Optical Activity within Coupled Cluster Response Theory. *J. Chem. Phys.* **2000**, *113*, 3561–3572.

(6) Kjærgaard, T.; Jansík, B.; Jørgensen, P.; Coriani, S.; Michl, J. Gauge-origin-independent Coupled Cluster Singles and Doubles Calculation of Magnetic Circular Dichroism of Azabenzenes and Phosphabenzene using London Orbitals. *J. Phys. Chem. A* **2007**, *111*, 11278–11286.

(7) Seth, M.; Ziegler, T.; Banerjee, A.; Autschbach, J.; van Gisbergen, S. J. A.; Baerends, E. J. Calculation of the $\mathscr{A}$ Term of Magnetic Circular Dichroism based on Time Dependent-Density Functional Theory I. Formulation and Implementation. *J. Chem. Phys.* **2004**, *120*, 10942–10954.

(8) Seth, M.; Ziegler, T.; Autschbach, J. Ab Initio Calculation of the C/D Ratio of Magnetic Circular Dichroism. *J. Chem. Phys.* **2005**, *122*, 094112.

(9) Seth, M.; Ziegler, T. Calculation of the $\mathscr{B}$ term of Magnetic Circular Dichroism. A Time-Dependent Density Functional Approach. *J. Chem. Theory. Comput.* **2007**, *3*, 434–447.

(10) Seth, M.; Ziegler, T. Calculation of Magnetic Circular Dichroism Spectra with Time-Dependent Density Functional Theory. *Adv. Inorg. Chem.: Theory Comp. Inorg. Chem.* **2010**, *62*, 41–109.

(11) Solheim, H.; Frediani, L.; Ruud, K.; Coriani, S. An IEF-PCM Study of Solvent Effects on the Faraday B Term of MCD. *Theor. Chem. Accounts* **2008**, *119*, 231–244.

(12) Solheim, H.; Ruud, K.; Coriani, S.; Norman, P. Complex Polarization Propagator Calculations of Magnetic Circular Dichroism Spectra. *J. Chem. Phys.* **2008**, *128*, 094103.

(13) Solheim, H.; Ruud, K.; Coriani, S.; Norman, P. The A and B Terms of Magnetic Circular Dichroism Revisited. *J. Phys. Chem. A* **2008**, *112*, 9615–9618.

(14) Seth, M.; Krykunov, M.; Ziegler, T.; Autschbach, J. Application of Magnetically Perturbed Time-Dependent Density Functional Theory to Magnetic Circular Dichroism. II. Calculation of $\mathscr{A}$ Terms. *J. Chem. Phys.* **2008**, *128*, 234102.

(15) Seth, M.; Krykunov, M.; Ziegler, T.; Autschbach, J.; Banerjee, A. Application of Magnetically Perturbed Time-Dependent Density Functional Theory to Magnetic Circular Dichroism: Calculation of $\mathscr{B}$ Terms. *J. Chem. Phys.* **2008**, *128*, 144105.

(16) Kjærgaard, T.; Jørgensen, P.; Thorvaldsen, A.; Sałek, P.; Coriani, S. Gauge-origin Independent Formulation and Implementation of Magneto-Optical Activity within Atomic-Orbital-Density Based Hartree-Fock and Kohn-Sham Response Theories. *J. Chem. Theory Comp.* **2009**, *5*, 1997–2020.

(17) Kjærgaard, T.; Kristensen, K.; Kauczor, J.; Jørgensen, P.; Coriani, S.; Thorvaldsen, A. Comparison of Standard and Damped Response Formulations of Magnetic Circular Dichroism. *J. Chem. Phys.* **2011**, *135*.

(18) Kjærgaard, T.; Coriani, S.; Ruud, K. Ab Initio Calculation of Magnetic Circular Dichroism. *WIREs Comput Mol Sci* **2012**, *2*, 443–455.

(19) Fahleson, T.; Kauczor, J.; Norman, P.; Coriani, S. The Magnetic Circular Dichroism Spectrum of the C_{60} Fullerene. *Mol. Phys.* **2013**, 1401–1404.

(20) Krykunov, M.; Seth, M.; Ziegler, T.; Autschbach, J. Calculation of the Magnetic Circular Dichroism B Term from the Imaginary Part of the Verdet Constant using Damped Time-Dependent Density Functional Theory. *J. Chem. Phys.* **2007**, *127*, 244102.

(21) Lee, K.-M.; Yabana, K.; Bertsch, G. F. Magnetic Circular Dichroism in Real-Time Time-Dependent Density Functional Theory. *J. Chem. Phys.* **2011**, *134*, 144106.

(22) Stepanek, P.; Bour, P. Computation of Magnetic Circular Dichroism by Sum-Over-States Summations. *J. Comput. Chem.* **2013**, *34*, 1531–1539.

(23) Stepanek, P.; Straka, M.; Andrushchenko, V.; Bour, P. Communication: Fullerene Resolution by the Magnetic Circular Dichroism. *J. Chem. Phys.* **2013**, *138*, 151103.

(24) Honda, Y.; Hada, M.; Ehara, M.; Nakatsuji, H.; Michl, J. Theoretical Studies on Magnetic Circular Dichroism by the Finite Perturbation Method with Relativistic Corrections. *J. Chem. Phys.* **2005**, *123*, 164113.

(25) Norman, P.; Bishop, D. M.; Jensen, H. J. A.; Oddershede, J. Nonlinear Response Theory with Relaxation: The First-order Hyperpolarizability. *J. Chem. Phys.* **2005**, *123*, 194103.

(26) Kauczor, J.; Jørgensen, P.; Norman, P. On the Efficiency of Algorithms for Solving Hartree–Fock and Kohn–Sham Response Equations. *J. Chem. Theory Comput.* **2011**, *7*, 1610.

(27) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum Mechanical Continuum Solvation Models. *Chem. Rev.* **2005**, *105*, 2999.

(28) Zgierski, M. Z. Vibronic Structure of MCD Spectra. I. Non-Condon Effects in Molecules with Nondegenerate Electronic States. *J. Chem. Phys.* **1985**, *83*, 2170–2185.

(29) Lin, N.; Solheim, H.; Zhao, X.; Santoro, F.; Ruud, K. First Principles Studies of the Vibrationally Resolved Magnetic Circular Dichroism Spectra of Biphenylene. *J. Chem. Theory. Comput.* **2013**, *9*, 1557–1567.

(30) Crespo-Hernandez, C. E.; Cohen, B.; Hare, P. M.; Kohler, B. Ultrafast Excited-State Dynamics in Nucleic Acids. *Chem. Rev.* **2004**, *104*, 1977–2020.

(31) Middleton, C. T.; de La Harpe, K.; Su, C.; Law, Y. K.; Crespo-Hernandez, C. E.; Kohler, B. DNA Excited-State Dynamics: From Single Bases to the Double Helix. *Annu. Rev. Phys. Chem.* **2009**, *60*.

(32) Gustavsson, T.; Improta, R.; Markovitsi, D. DNA/RNA: Building Blocks of Life under UV Irradiation. *J. Phys. Chem. Lett.* **2010**, *1*, 2025–2030.

(33) Santoro, F.; Improta, R.; Fahleson, T.; Kauczor, J.; Norman, P.; Coriani, S. Relative Stability of the L_a and L_b States in Adenine and Guanine: Direct Evidence from TD-DFT Salculations of MCD Spectra. *J. Phys. Chem. Lett.* **2014**, *5*, 1806–1811.

(34) Szalay, P.; Watson, T.; Perera, A.; Lotrich, V.; Bartlett, R. Benchmark Studies on the Building Blocks of DNA. 1. Superiority of Coupled Cluster Methods in Describing the Excited States of Nucleobases in the Franck-Condon Region. *J. Phys. Chem. A* **2012**, *116*, 6702–6710.

(35) Hare, P. M.; Crespo-Hernández, C. E.; Kohler, B. Internal Conversion to the Electronic Ground State Occurs via Two Distinct Pathways for Pyrimidine Bases in Aqueous Solution. *PNAS* **2007**, *104*, 435–440.

(36) Santoro, F.; Barone, V.; Gustavsson, T.; Improta, R. Solvent Effect on the Singlet Excited-State Lifetimes of Nucleic Acid Bases: A Computational Study of 5-Fluorouracil and Uracil in Acetonitrile and Water. *J. Am. Chem. Soc.* **2006**, *128*, 16312–16322.

(37) Improta, R.; Barone, V.; Lami, A.; Santoro, F. Quantum Dynamics of the Ultrafast $\pi\pi^*/n\pi^*$

Population Transfer in Uracil and 5-Fluoro-Uracil in Water and Acetonitrile. *J. Phys. Chem. B* **2009**, *113*, 14491–14503.

(38) Picconi, D.; Avila Ferrer, F. J.; Improta, R.; Lami, A.; Santoro, F. Quantum-Classical Effective-Modes Dynamics of the $\pi\pi^* \rightarrow n\pi^*$ Decay in 9H-Adenine. A Quadratic Vibronic Coupling Model. *Faraday Discuss.* **2013**, *163*, 223–242.

(39) Biemann, L.; Kovalenko, S. A.; Kleinermanns, K.; Mahrwald, R.; Markert, M.; Improta, R. Excited State Proton Transfer Is Not Involved in the Ultrafast Deactivation of Guanine-Cytosine Pair in Solution. *J. Am. Chem. Soc.* **2011**, *133*, 19664–19667.

(40) Improta, R.; Barone, V. Interplay between "Neutral" and "Charge-Transfer" Excimers Rules the Excited State Decay in Adenine-Rich Polynucleotides. *Angew. Chemie Int. Ed.* **2011**, *50*, 12016–12019.

(41) Santoro, F.; Barone, V.; Improta, R. The Excited States Decay of the A-T DNA: a PCM/TD-DFT Study in Aqueous Solution of the (9-Methyl-Adenine) 2 (1-Methyl-Thymine) 2 Stacked Tetramer. *J. Am. Chem. Soc.* **2009**, *131*, 15232–15245.

(42) Spata, V.; Matsika, S. Role of Excitonic Coupling and Charge-Transfer States in the Absorption and CD Spectra of Adenine-Based Oligonucleotides Investigated through QM/MM Simulations. *J. Phys. Chem. A* **2014**, *118*, 12021–12030.

(43) Plasser, F.; Aquino, A. J. A.; Hase, W. L.; Lischka, H. UV Absorption Spectrum of Alternating DNA Duplexes. Analysis of Excitonic and Charge Transfer Interactions. *J. Phys. Chem. A* **2012**, *116*, 11151–11160.

(44) Plasser, F.; Lischka, H. Electronic Excitation and Structural Relaxation of the Adenine Dinucleotide in Gas Phase and Solution. *Photochem. Photobiol. Sci.* **2013**, *12*, 1440–1452.

(45) Rizzo, A.; Coriani, S.; Ruud, K. In *Computational Strategies for Spectroscopy. From Small*

Molecules to Nano Systems; Barone, V., Ed.; John Wiley and Sons, 2012; Chapter 2, pp 77–135.

(46) Buckingham, A. D.; Stephens, P. J. Magnetic Optical Activity. *Adv. Res. Chem. Phys.* **1966**, *17*, 399.

(47) Stephens, P. J. Theory of Magnetic Circular Dichroism. *J. Chem. Phys.* **1970**, *52*, 3489.

(48) Ganyushin, D.; Neese, F. First-principles Calculations of Magnetic Circular Dichroism Spectra. *J. Chem. Phys.* **2008**, *128*, 114117.

(49) Bolvin, H. Theoretical Determination of the Excited States and of g-Factors of the Creutz-Taube Ion, $[(NH_3)_5\text{-Ru-pyrazine-Ru-}(NH_3)_5]_5^+$. *Inorg. Chem.* **2007**, *46*, 417.

(50) Kauczor, J.; Norman, P. Efficient Calculations of Molecular Linear Response Properties for Spectral Regions. *J. Chem. Theory Comput.* **2014**, *10*, 2449–2455.

(51) Cammi, R.; Mennucci, B.; Tomasi, J.; Ruud, K. Multiconfigurational Self-Consistent Field Linear Response for the Polarizable Continuum Model: Theory and Application to Ground and Excited-State Polarizabilities of Paranitroaniline in Solution. *J. Chem. Phys.* **2003**, *119*, 5818–5827.

(52) Frediani, L.; Ågren, H.; Ferrighi, L.; Ruud, K. Second-Harmonic Generation of Solvated Molecules Using Multiconfigurational Self-Consistent-Field Quadratic Response Theory and the Polarizable Continuum Model. *J. Chem. Phys.* **2005**, *123*, 144117.

(53) Milne, B. F.; Norman, P. Resonant-Convergent PCM Response Theory Calculations of Second Harmonic Generation in Makaluvamines A-V; Pyrroloiminoquinone Marine Natural Products from Poriferans of Genus Zyzzya. *J. Phys. Chem. A* **2015**, d 10.1021/jp5102362.

(54) Improta, R.; Barone, V. Absorption and Fluorescence Spectra of Uracil in the Gas Phase and in Aqueous Solution: A TD-DFT Quantum Mechanical Study. *J. Am. Chem. Soc.* **2004**, *126*, 14320–1432.

(55) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A. et al. Gaussian 09 Revision C.3. Gaussian Inc. Wallingford CT 2009.

(56) DALTON, a molecular electronic structure program, Release Dalton2013 and LSDalton2013. 2013; See http://daltonprogram.org/.

(57) Aidas, K.; Angeli, C.; Bak, K. L.; Bakken, V.; Bast, R.; Boman, L.; Christiansen, O.; Cimiraglia, R.; Coriani, S.; Dahle, P. et al. The Dalton quantum chemistry program system. *WIREs Comput. Mol. Sci.* **2014**, *4*, 269–284, doi: 10.1002/wcms.1172.

(58) Rohatgi, A. Version 3.3 of WebPlotDigitizer. ZENODO. 10.5281/zenodo.10532. 2014; `http://arohatgi.info/WebPlotDigitizer`.

(59) Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Computing In Science & Engineering* **2007**, *9*, 90.

(60) Sutherland, J. C.; Holmquist, B. Magnetic Circular Dichroism of Biological Molecules. *Ann. Rev. Biophys. Bioeng.* **1980**, *9*, 293–326.

(61) Sutherland, J. C.; Griffin, K. Magnetic Circular Dichroism of Adenine, Hypoxanthine, and Guanosine 5'-Diphosphate to 180 nm. *Biopolymers* **1984**, *23*, 2715–2724.

(62) Improta, R.; Barone, V. Excited States Behavior of Nucleobases in Solution: Insights from Computational Studies. *Top. Curr. Chem.* **2015**, *355*, 329–358.

(63) Kaito, A.; Hatano, M.; Ueda, T.; Shibuya, S. CNDO treatment for Faraday B terms of some azaheterocycles. *Bull. Chem. Soc. Jpn.* **1980**, *53*, 3073–3078.

(64) Kistler, K.; Matsika, S. Solvatochromic Shifts of Uracil and Cytosine Using a Combined Multireference Configuration Interaction/Molecular Dynamics Approach and the Fragment Molecular Orbital Method. *J. Phys. Chem. A* **2009**, *113*, 12396–12403.

(65) Blancafort, L.; Migani, A. Water Effect on the Excited-state Decay Paths of Singlet Excited Cytosine. *J. Photochem. Photobiol.* **2007**, *190*, 283–290.

(66) Bazso, G.; Tarczay, G.; Fogarasi, G.; Szalay, P. G. Tautomers of Cytosine and Their Excited Electronic States: a Matrix Isolation Spectroscopic and Quantum Chemical Study. *Phys. Chem. Chem. Phys.* **2011**, *13*, 6799–6807.

(67) Ferrer, F. J. A.; Santoro, F.; Improta, R. The Excited State Behavior of Cytosine in the Gas Phase: A TD-DFT Study. *Comput. Theor. Chem.* **2014**, *1040-1041*, 186–194.

(68) Avila Ferrer, F. J.; Cerezo, J.; Stendardo, E.; Improta, R.; Santoro, F. Insights for an Accurate Comparison of Computational Data to Experimental Absorption and Emission Spectra : Beyond the Vertical Transition Approximation. *J. Chem. Theory Comp.* **2013**, *9*, 2072–2082.

(69) Corni, S.; Cammi, R.; Mennucci, B.; Tomasi, J. Electronic Excitation Energies of Molecules in Solution within Continuum Solvation Models: Investigating the Discrepancy Between State-Specific and Linear-Response Methods. *J. Chem. Phys.* **2005**, *123*, 134512.

(70) Improta, R.; Barone, V.; Scalmani, G.; Frisch, M. A State-Specific Polarizable Continuum Model Time Dependent Density Functional Theory Method for Excited State Calculations in Solution. *J. Chem. Phys.* **2006**, *125*, 054103.

(71) Improta, R. In *Computational Strategies for Spectroscopy. From Small Molecules to Nano Systems*; Barone, V., Ed.; John Wiley and Sons, 2012; Chapter 1, pp 39–76.

Figure 8: TOC figure