Open access • Journal Article • DOI:10.1021/ACS.JCTC.8B00930

UV Absorption and Magnetic Circular Dichroism Spectra of Purine, Adenine, and Guanine: A Coupled Cluster Study in Vacuo and in Aqueous Solution

Sarah Karbalaei Khani, Sarah Karbalaei Khani, Rasmus Faber, Fabrizio Santoro ...+2 more authors

Institutions: Technical University of Denmark, Ruhr University Bochum

Published on: 12 Feb 2019 - Journal of Chemical Theory and Computation (American Chemical Society)

Topics: Solvent effects, Solvation, Magnetic circular dichroism, Coupled cluster and Water cluster

UV Absorption and Magnetic Circular Dichroism Spectra of Purine, Adenine, and Guanine: a Coupled Cluster Study in Vacuo and in Aqueous Solution

Karbalaei Khani, Sarah; Faber, Rasmus; Santoro, Fabrizio; Hattig, Christof; Coriani, Sonia

Published in:
Journal of Chemical Theory and Computation

Link to article, DOI:
10.1021/acs.jctc.8b00930

Publication date:
2019

Document Version
Peer reviewed version

UV Absorption and Magnetic Circular Dichroism Spectra of Purine, Adenine, and Guanine: a Coupled Cluster Study in Vacuo and in Aqueous Solution

Sarah Karbalaei Khani,[*,†,‡] Rasmus Faber,[‡] Fabrizio Santoro,[¶] Christof Hättig,[*,†] and Sonia Coriani[*,‡]

†*Arbeitsgruppe Quantenchemie, Ruhr-Universität, Bochum D-44780, Germany*

‡*DTU Chemistry, Technical University of Denmark, Kemitorvet Build. 207, DK-2800 Kongens Lyngby, Denmark*

¶*Istituto di Chimica dei Composti Organo-Metallici, Consiglio Nazionale delle Ricerche (ICCOM-CNR), Area della Ricerca, via G. Moruzzi 1, I-56124 Pisa, Italy*

E-mail: Sarah.Karbalaeikhani@rub.de; Christof.Haettig@rub.de; soco@kemi.dtu.dk

Abstract

The absorption and magnetic circular dichroism (MCD) spectra of purine and of the purine nucleobases adenine and guanine have been calculated in the gas phase at the Coupled Cluster Singles and Doubles (CCSD) and Resolution-of-Identity Singles and Approximate Doubles (RI-CC2) levels of theory. Exploiting a new development in the TURBOMOLE program package for computing vertical excitation energies and Faraday $\mathcal{B}$ terms in an implicit solvent approximated by the conductor-like screening model (COSMO) at the CC2 level, we

have investigated the solvent effects on the relative positions of the $\pi\pi^*$ and $n\pi^*$ electronic transitions in these three molecules and compared them to the corresponding vacuum results. In the case of adenine, we also included specific solvent effects with a small water cluster. The spectra obtained with the implicit model COSMO are in qualitative agreement with those obtained with explicit water molecules both with and without the inclusion of the bulk solvent effects via the continuum solvent model. This suggests that the inclusion of the electrostatic contributions of the solvent can provide a sufficiently accurate description of the absorption spectra for adenine. The results for purine, adenine, and guanine show that, after the inclusion of bulk solvation, the $\pi\pi^*$ states shift to lower energies while at the same time $n\pi^*$ states show a reversed behavior. The computed MCD spectra show the characteristic bi-signate profile found experimentally in all cases, despite, for adenine, remarkable differences in the origin of the individual peaks for different computational methods. Therefore, the ability (or inability) of MCD for determining the relative stability of the L_a and L_b states is critically re-assessed. According to our best estimate for adenine in aqueous solution, the L_a state is more stable than L_b.

1 Introduction

The excited states of DNA bases have in the past years been subject of a plethora of theoretical and experimental studies with the objective of gaining new insights into the photochemistry and dynamics of these systems.[1–4] Purinic (adenine and guanine) nucleobases are one of the key groups of light-absorbing chromophores in the 200-300 nm region which contain the information in DNA and RNA. Due to the short excited-state lifetimes of purine bases and the ultrafast non-radiative deactivation of excited states, photochemical reactions, and thus the formation of primary photoproducts, are disfavored. This is important for the stability of the DNA against photodamage that can e.g. cause skin cancer.

Because of its biological significance, the photo-stability of nucleobases against UV-

initiated damage and photoactivated processes in purine bases have been and still are intensively studied. In general, providing accurate energetic and spectroscopic data for these nucleobases is essential for understanding the processes in which they are involved. In recent years, the relative stability of the two lowest bright states with $\pi\pi^*$ character of 9H-adenine has been extensively investigated with different methods.[5–10] In the case of purines, these two $\pi\pi^*$ states – which are labeled as L_a and L_b – are characterized as a HOMO $\rightarrow$ LUMO excitations for L_a and a combination of HOMO$-2 \rightarrow$ LUMO and HOMO $\rightarrow$ LUMO$+1$ excitations for L_b. In addition to the absorption, also the MCD spectra have been calculated both in vacuo and in aqueous solution, [6,10] and compared with existing experimental data[5,11,12] in order to help to discriminate between these two states. A number of TD-DFT (time-dependent density functional theory) investigations have suggested that L_a is energetically more stable than L_b.[6,13] For instance, spectroscopic studies of adenine at the TD-DFT level[6] have predicted a lower energy for the L_a compared to the L_b state by $\approx$ 0.1–0.3 eV in vacuum and in aqueous solution (describing solvent effects by cluster models and by the polarizable continuum model (PCM)), where the excitation to L_a carries most of the oscillator strength. Contrary to TD-DFT, most wave-function based methods as EOM-CCSD, EOM-CCSD(T) in vacuo[14–16] and CASSCF/CASPT2 in both vacuum[17,18] and a DNA environment[19] have predicted the L_b state at lower energy than L_a by $\approx$ 0.2 eV. An exception is CC2, which gives in vacuum for the vertical excitation energies to the L_a and L_b states virtually degenerate results.[20]

Guanine is a purine nucleobase that exists in several tautomeric forms of which the most stable ones are in the gas phase very close in energy. Quantum mechanical studies at the EOMEE-CCSD(T) level on the biologically most relevant tautomer of guanine, keto-N9H, have predicted that the lowest excitation is a $\pi\pi^*$ transition, while at the EOMEE-CCSD level the first and second transitions were found to be virtually degenerate and to have πRyd (where Ryd stands for Rydberg) and $\pi\pi^*$ characters.[14] An assignment based on CC2 (using a triple-ζ basis with diffuse functions) predicted a weak πRyd transition along with a

substantial $\pi\pi^*$ character for the lowest excited state and a close-lying πRyd transition for the second excited state while the third and the fourth excited states with, respectively, $n\pi^*$ and πRyd character appear at significantly higher energies.[20] At the TD-DFT level, the two low-lying states have been assigned as $\pi\pi^*$ transitions with the L_b state lower than L_a at 4.64 eV,[21] whereas at the CASPT2 level L_a has been found lower than L_b by ≈ 0.74 eV.[22]

However, the energy differences between these states are of about the same order as the remaining errors of the currently available methods and basis sets that can be applied to molecules of the size of purinic nucleobases. Thus, despite numerous previous theoretical studies, there are still unsolved important questions as *e.g.* the exact order of their lowest excited states in solution and biological environments. In a biological setting, molecules are embedded in a complex environment with multiple interactions including electrostatic ones. The energy of electronically excited states, particularly those with large transition dipole moments or significant changes in the charge distribution upon excitation, is sensitive to polar environments and solvents. At the same time solvents also induce remarkable inhomogeneous broadenings that smear out vibronic progressions and make the identification of different close lying electronic states more difficult. In this regard, it has been suggested to use Magnetic Circular Dichroism (MCD) spectroscopy on the purine bases in the UV-Vis region to assist in the interpretation of their electronic spectra,[6] since, due to their signed nature, MCD signals improve the capability of locating and separating transitions which are overlapping or have weak intensity in the one-photon absorption (OPA) spectra.[23,24] It was assumed that the signed nature of MCD spectra could elucidate the energy ordering of excited states especially when, as *e.g.* in the adenine case, the debate evolves around the energy ordering of bright excited states.[6] We will in the following critically reassess this assumption.

To facilitate the study of solvent effects on the relative positions, intensities and even the nature of different electronic states in MCD and UV/Vis absorption spectra, we report in the current contribution for the first time an implementation of the conductor-like screen-

ing model (COSMO)[25] to include the solvation effects for oscillator strengths and for the Faraday $\mathcal{B}$ term of MCD at the level of the second-order approximate coupled-cluster singles and doubles model CC2[26,27] within the post–SCF reaction field scheme.[28–33] We apply the COSMO-RI-CC2 approach to investigate the UV/Vis absorption and MCD spectra of purine and the nucleobases adenine and guanine in aqueous solution. To account for higher-order correlation effects beyond CC2, we also performed calculations of excitation energies, oscillator strengths and Faraday $\mathcal{B}$ terms in vacuum at the CCSD level of theory for both purine and adenine.

2 Theory

Within coupled cluster (CC) response theory, excitation energies ω_j and left and right excitation vectors, $\bar{E}^j(-\omega_j)$ and $E^j(\omega_j)$, for an excited state j are obtained as solutions of the non-symmetric eigenvalue equations

$$\mathbf{A}E^j(\omega_j) = \omega_j E^j(\omega_j) \quad \text{and} \quad \bar{E}^j(-\omega_j)\mathbf{A} = \omega_j \bar{E}^j(-\omega_j) \tag{1}$$

for the CC Jacobian matrix, $\mathbf{A}$, which is the derivative of the cluster equations with respects to the cluster amplitudes t_ν

$$A_{\mu\nu} = \langle\mu| \exp(-\hat{T})[\hat{H}, \tau_\nu] \exp(\hat{T})|\mathrm{HF}\rangle \ . \tag{2}$$

In the last equation, $\hat{T} = \sum_\mu t_\mu \tau_\mu$ is the cluster operator. The transition strength matrix elements (for electric dipole moment components d_α and d_β) for an excitation out of the ground state 0 are determined from the residues of the linear response function and take the form

$$S^{\alpha\beta}_{0\to j} = \frac{1}{2}\left\{ M^{d_\alpha}_{0\leftarrow j} M^{d_\beta}_{j\leftarrow 0} + (M^{d_\beta}_{0\leftarrow j} M^{d_\alpha}_{j\leftarrow 0})^* \right\} \ , \tag{3}$$

where the left and right transition moments are given by

$$M^{d_\alpha}_{0\leftarrow j} = \eta^{d_\alpha} E^j(\omega_j) + \bar{M}^j(\omega_j)\xi^{d_\alpha} \quad \text{and} \quad M^{d_\alpha}_{j\leftarrow 0} = \bar{E}^j(-\omega_f)\xi^{d_\alpha} \ . \tag{4}$$

In the last equation, ξ^{d_α} and η^{d_α} are, respectively, the second derivatives of the Lagrange function with respect to the field strength and Lagrange multipliers and the field strength and cluster amplitudes. $\bar{M}^j(\omega_j)$ is a vector containing auxiliary (zero-order) Lagrange multipliers obtained as solution of the linear equation

$$\bar{M}^j \left(\mathbf{A} + \omega_j \mathbf{1}\right) = -\mathbf{F} E^j(\omega_j) \ \ , \tag{5}$$

where $\mathbf{F}$ is the second derivative of the CC Lagrangian with respect to the cluster amplitudes. The oscillator strengths $f(0 \rightarrow j)$ that determine the intensity of the UV/Vis absorption bands, are obtained directly from the diagonal elements of the electric dipole transition strength matrix

$$f(0 \rightarrow j) = \tfrac{2}{3}\omega_j (S^{xx}_{0\rightarrow j} + S^{yy}_{0\rightarrow j} + S^{zz}_{0\rightarrow j}) \ . \tag{6}$$

A derivation of the above equations can be found for instance in Ref. 34. For details on the implementation of the above equations at the RI-CC2 and CCSD levels in gas phase we refer to Refs. 35–37, and to Refs. 38,39 for the details on their implementation at the CC2 level in vacuo. In the following, we briefly discuss the modifications needed to include solvent effects in the CC2 or RI-CC2 calculations with the conductor-like screening model (COSMO)[25] as continuum solvation model.[40]

The Lagrange function for the approximate coupled cluster singles and doubles model CC2 including the environment contribution within the framework of COSMO can be written

as

$$L = \mathcal{G}_{\mathrm{HF}} + \langle \mathrm{HF}|\hat{\tilde{F}}_N^{pol} + \hat{\tilde{W}}_N^{\mathrm{QM}} + [\hat{\tilde{W}}_N^{\mathrm{QM}}, \hat{T}_2]|\mathrm{HF}\rangle + \langle \bar{t}_1|\hat{\tilde{F}}_N^{pol} + \hat{\tilde{W}}_N^{\mathrm{QM}} + [\hat{F}_N^{pol} + \hat{\tilde{W}}_N^{\mathrm{QM}}, \hat{T}_2]|\mathrm{HF}\rangle$$
$$+ \langle \bar{t}_2|[\hat{F}_N^{pol}, \hat{T}_2] + \hat{\tilde{W}}_N^{\mathrm{QM}}|\mathrm{HF}\rangle + \frac{1}{2}\sum_{\kappa}\langle \mathrm{HF} + \bar{t}_1|\hat{\tilde{Q}}_{N,\kappa}|\mathrm{HF}\rangle\langle \mathrm{HF} + \bar{t}_1|\hat{\tilde{V}}_{N,\kappa}|\mathrm{HF}\rangle \ . \tag{7}$$

In the above equation, $\bar{t}_i$ are the Lagrange multipliers and the tilde accent on the operators indicates a similarity transformation according to $\hat{\tilde{O}} = \exp(-\hat{T}_1)\hat{O}\exp(\hat{T}_1)$ and a lower index N normal ordering. $\mathcal{G}_{\mathrm{HF}}$ is the COSMO-Hartree-Fock free energy. The Fock operator for the polarized system $\hat{\tilde{F}}^{pol}$ and the fluctuation potential $\hat{\tilde{W}}^{QM}$, along with the density-dependent surface charge and potential operators, $\hat{\tilde{Q}}_{N,\kappa}$ and $\hat{\tilde{V}}_{N,\kappa}$, respectively, for the cavity segments κ are explained in detail in Refs. 40,41. Inclusion of solvent contributions in the calculation of excitation energies and oscillator strengths at the CC2 level within the framework of COSMO and the post–SCF reaction field scheme introduced in Ref. 42 keeps the expressions for η^{d_α} and ξ^{d_α} in Eq. (4) the same as in the vacuum case, while it adds some explicit terms to the Jacobian $\mathbf{A}$

$$A^{pol}_{\mu_1\nu_1} = \langle \mu_1|[\hat{G}^{pol,\mathrm{HF}} + \hat{\tilde{G}}^{\Delta}(\mathbf{D}^{\Delta}), \tau_{\nu_1}] + \hat{\tilde{G}}^{\nu_1}(\mathbf{D}^{\Delta})|\mathrm{HF}\rangle \ , \tag{8}$$

$$A^{pol}_{\mu_2\nu_2} = \langle \mu_2|[\hat{G}^{pol,\mathrm{HF}}, \tau_{\nu_2}]|\mathrm{HF}\rangle \ , \tag{9}$$

where $\hat{G}^{pol,\mathrm{HF}}$ is the Hartree-Fock reaction field potential. The operators $\hat{\tilde{G}}^{\Delta}(\mathbf{D}^{\Delta})$ and $\hat{\tilde{G}}^{\nu_1}(\mathbf{D}^{\Delta})$ in the singles block $A_{\mu_1\nu_1}$ can be rewritten as

$$\hat{\tilde{G}}^{\Delta}(\mathbf{D}^{\Delta}) = \sum_{\kappa}\hat{\tilde{Q}}_{\kappa}(\mathbf{D}^{\Delta})\langle \mathrm{HF} + \bar{t}_1|\hat{\tilde{V}}_{\kappa,N}|\mathrm{HF}\rangle \ , \tag{10}$$

and

$$\hat{\tilde{G}}^{\nu_1}(\mathbf{D}^{\Delta}) = \sum_{\kappa}\hat{\tilde{Q}}_{\kappa}(\mathbf{D}^{\Delta})\langle \mathrm{HF} + \bar{t}_1|[\hat{V}_{\kappa}, \tau_{\nu_1}]|\mathrm{HF}\rangle \ , \tag{11}$$

with the indices κ runs over all cavity segments. The density $\mathbf{D}^{\Delta}$ is defined as

$$\mathbf{D}^{\Delta}_{pq} = \langle \mathrm{HF} + \bar{t}_1 | \exp(-\hat{T}_1) \hat{E}_{pq} \exp(\hat{T}_1) | \mathrm{HF} \rangle - \langle \mathrm{HF} | \hat{E}_{pq} | \mathrm{HF} \rangle \ , \tag{12}$$

and $\hat{E}_{pq}$ are spin-free one particle operators.[43] The matrix $\mathbf{F}$ in Eq. 5 contains the following solvent contributions in the singles block $\mathbf{F}_{\mu_1 \nu_1}$

$$F^{pol}_{\mu_1 \nu_1} = \langle \bar{t}_1 | [[\hat{\tilde{G}}^{\Delta}(\mathbf{D}^{\Delta}), \tau_{\mu_1}], \tau_{\nu_1}] | \mathrm{HF} \rangle + \sum_{\kappa} \langle \mathrm{HF} + \bar{t}_1 | [\hat{\tilde{Q}}_{\kappa}, \tau_{\mu_1}] | \mathrm{HF} \rangle \times \langle \mathrm{HF} + \bar{t}_1 | [\hat{\tilde{V}}_{\kappa}, \tau_{\nu_1}] | \mathrm{HF} \rangle \ , \tag{13}$$

while the other blocks have the same form as in the vacuum case.

The $\mathcal{B}(0 \rightarrow j)$ terms of MCD can be computed from the derivatives of the one-photon transition strength matrix with respect to the strength of a magnetic field[44,45]

$$\mathcal{B}(0 \rightarrow j) = \frac{1}{2} \sum_{\alpha\beta\gamma} \epsilon_{\alpha\beta\gamma} \Im \left\{ \frac{dS^{\alpha\beta}_{0 \rightarrow j}}{dB_{\gamma}} \right\} \ , \tag{14}$$

where α, β and γ distinguish the three Cartesian components x, y and z, $\epsilon_{\alpha\beta\gamma}$ is the Levi-Civita tensor and B_{γ} is the Cartesian component of the magnetic field strength. For a magnetic-field independent basis set, the derivatives of the left and right transition moments are, in terms of the CC building blocks, formulated as[44]

$$\begin{aligned} \frac{dM^{d_\alpha}_{0 \leftarrow j}}{dB_{\gamma}} = & \left[\mathbf{G} t^{m_\gamma}(0) t^{d_\alpha}(-\omega_j) + \mathbf{F}^{m_\gamma} t^{d_\alpha}(-\omega_j) + \mathbf{F}^{d_\alpha} t^{m_\gamma}(0) \right] \times E^{j}(\omega_j) \\ & + \bar{M}^{j}(\omega_j) [\mathbf{A}^{d_\alpha} t^{m_\gamma}(0) + \mathbf{A}^{m_\gamma} t^{d_\alpha}(-\omega_j) + \mathbf{B} t^{m_\gamma}(0) t^{d_\alpha}(-\omega_j)] \\ & + \bar{\xi}^{d_\alpha}(-\omega_j) {}^{\perp}E^{j m_\gamma}(\omega_j, 0) + \bar{\xi}^{m_\gamma}(0) E^{j d_\alpha}(\omega_j, -\omega_j) \end{aligned} \tag{15}$$

and

$$\frac{dM^{d_\alpha}_{j \leftarrow 0}}{dB_{\gamma}} = {}^{\perp}\bar{E}^{j m_\gamma}(-\omega_j, 0) \xi^{d_\alpha} + \bar{E}^{j}(-\omega_j) \mathbf{A}^{d_\alpha} t^{m_\gamma}(0) \tag{16}$$

with $t^{m_\gamma}(0)$ as the amplitude's first-order response to the static magnetic field, $t^{d_\alpha}(\omega)$ as the amplitude's first-order response to the electric field with frequency ω, and similarly for the excited state vectors $E^{jm_\gamma}(\omega_j, 0)$ and $E^{jd_\alpha}(\omega_j, -\omega_j)$ for state j with excitation energy ω_j. $\mathbf{G}$ is the third derivative of the CC Lagrangian with respect to the cluster amplitudes and $\mathbf{B}$ a mixed third derivative where L has been differentiated once with respect to the Lagrange multipliers and twice with respect to the cluster amplitudes.[34] The left vectors $\bar{\xi}^Y(\omega_Y)$ (for a generic operator $\hat{Y}$) are defined as $\eta^Y + \mathbf{F}t^Y(\omega_Y)$ and the amplitude responses $t^Y(\omega_Y)$ are determined by the equation

$$[\mathbf{A} - \omega_Y \mathbf{1}]\, t^Y(\omega_Y) = -\xi^Y \ , \tag{17}$$

where ω_Y is the frequency associated with the operator Y, which can be a component of the magnetic or the electric field. The first-order responses $E^{jY}(\omega_j, \omega_Y)$ and $\bar{E}^{jY}(-\omega_j, \omega_Y)$ of the eigenvectors $E^j(\omega_j)$ and $\bar{E}^j(-\omega_j)$ with respect to the operator $\hat{Y}$ are obtained solving the linear equations

$$[\mathbf{A} - (\omega_j + \omega_Y)\mathbf{1}]\, E^{jY}(\omega_j, \omega_Y) = -\left[\mathbf{B}t^Y(\omega_Y) + \mathbf{A}^Y\right] E^j(\omega_j) \ , \tag{18}$$

$$\bar{E}^{jY}(-\omega_j, \omega_Y)\, [\mathbf{A} + (\omega_Y - \omega_j)\mathbf{1}] = -\bar{E}^j(-\omega_j)\left[\mathbf{B}t^Y(\omega_Y) + \mathbf{A}^Y\right] \ . \tag{19}$$

The superscript $\perp$ on the derivatives of the eigenvectors indicates a projection onto the orthogonal complement of the undifferentiated eigenvectors to avoid unphysical divergences[44,46,47] for the static magnetic field. See Ref. 44 for further details and the working equations for CCSD. The implementation of perturbation-induced i.e. first derivatives of one-photon transition moments for CC2 in the TURBOMOLE package has been presented in Ref. 48 as part of the implementation of spin-orbit induced oscillator strengths, which for the current work has been extended to account for solvent effects within the COSMO model.

We will here only discuss the additional reaction field contributions for the interaction with the solvent. They are for the magnetic-field induced transition moments structurally

similar to those for two-photon transition moments, which have been discussed in Ref. 49 for the atomistic QM/MM-like polarizable embedded PERI-CC2 approach.

In addition to the solvent contributions to the Jacobian $\mathbf{A}$ and the matrix $\mathbf{F}$ that have been discussed above, explicit solvent contributions arise for CC2 in the singles/singles/singles blocks of $\mathbf{B}$ and $\mathbf{G}$. For the $\mathbf{B}$ matrix transformations in Eqs. 18 and 19, the solvent contributions to the singles/singles/singles block are

$$\begin{aligned} B^{pol}_{\kappa_1\mu_1\nu_1} t^Y_{\mu_1} E^j_{\nu_1} &= \langle \kappa_1 | [[\hat{\tilde{G}}^{\Delta}(\mathbf{D}^{\Delta}), \hat{T}^Y_1], \hat{E}^j_1] | \mathrm{HF} \rangle + \langle \kappa_1 | \hat{\tilde{G}}^{\Delta}\big(\mathbf{D}^F(t^Y, E^j)\big) | \mathrm{HF} \rangle \\ &+ \langle \kappa_1 | [\hat{\tilde{G}}^{\Delta}\big(\mathbf{D}^{\eta}(t^Y)\big), \hat{E}^j_1] | \mathrm{HF} \rangle + \langle \kappa_1 | [\hat{\tilde{G}}^{\Delta}\big(\mathbf{D}^{\eta}(E^j)\big), \hat{T}^Y_1] | \mathrm{HF} \rangle \end{aligned} \tag{20}$$

and

$$\begin{aligned} \bar{E}^j_{\kappa_1} B^{pol}_{\kappa_1\mu_1\nu_1} t^Y_{\nu_1} &= \langle \bar{E}^j | [[\hat{\tilde{G}}^{\Delta}(\mathbf{D}^{\Delta}), \hat{T}^Y_1], \tau_{\nu_1}] | \mathrm{HF} \rangle + \langle \mathrm{HF} + \bar{t}_1 | [[\hat{\tilde{G}}^{\Delta}\big(\mathbf{D}^{\xi}(\bar{E}^j)\big), \hat{T}^Y_1], \tau_{\mu_1}] | \mathrm{HF} \rangle \\ &+ \langle \bar{E}^j | [\hat{\tilde{G}}^{\Delta}\big(\mathbf{D}^{\eta}(t^Y)\big), \tau_{\mu_1}] | \mathrm{HF} \rangle + \langle \mathrm{HF} + \bar{t}_1 | [\hat{\tilde{G}}^{\Delta}\big(\mathbf{D}^{A}(\bar{E}^j, t^Y)\big), \tau_{\mu_1}] | \mathrm{HF} \rangle \ . \end{aligned} \tag{21}$$

The contributions arising from the interaction with the solvent in the $\mathbf{G}$ matrix in Eq. 15 are obtained as:

$$\begin{aligned} G^{pol}_{\mu_1\nu_1\kappa_1} t^X_{\mu_1} t^Y_{\nu_1} E^j_{\kappa_1} &= \langle \bar{t}_1 | [[\hat{\tilde{G}}^{\Delta}(\mathbf{D}^{\eta}(E^j)), \hat{T}^X_1], \hat{T}^Y_1] | \mathrm{HF} \rangle + \langle \bar{t}_1 | [[\hat{\tilde{G}}^{\Delta}(\mathbf{D}^{\eta}(t^Y)), \hat{T}^X_1], \hat{E}^j_1] | \mathrm{HF} \rangle \\ &+ \langle \bar{t}_1 | [[\hat{\tilde{G}}^{\Delta}(\mathbf{D}^{\eta}(t^X)), \hat{T}^Y_1], \hat{E}^j_1] | \mathrm{HF} \rangle \ . \end{aligned} \tag{22}$$

Detailed expressions for the auxiliary reaction field operators and densities can be found in Refs. 49 and 40.

3 Computational details

For all investigated compounds, the ground state structures have been optimized at the MP2[50] level of theory in vacuum and in solution with the cc-pVTZ[51] basis set combined

with the corresponding optimized auxiliary basis set.[52] In addition, the B3LYP[53]/cc-pVTZ optimized geometries in vacuum of Ref. 6,10 were considered. The geometry optimizations in solution for both calculations with and without explicit water molecules were carried out at PTE-COSMO-MP2 level employing the `ricc2` module of the TURBOMOLE program package version 7.2.[54,55] The frozen core approximation was applied with the $1s^2$ electrons kept frozen for all C, O and N atoms. The excitation energies, oscillator strengths, and Faraday $\mathcal{B}$ terms were computed at the CC2[27] level of theory using a local development version of the TURBOMOLE program package. Two basis sets were used for CC2 calculations of the spectra: aug-cc-pVDZ and aug-cc-pVTZ[56] combined with their auxiliary basis sets.[52] The post–SCF scheme was applied for the CC2 calculations in solution at the COSMO level.[40,42] For the cavity construction,[25] for a set of atom positions with corresponding atom radii R_α, we used the union of spheres of R_α and the solvent-accessible surface R_{SOLV}. In all COSMO calculations, water has been chosen as the solvent with the dielectric constant of $\varepsilon = 80.4$ and optical dielectric constant of $n^2 = 1.78$ (where n is the index of refraction). These values are taken from Ref. 57.

In addition, frozen-core CCSD calculations using the aug-cc-pVDZ[56] basis set in vacuo were carried out with the Dalton code[58] for the same vacuum structures mentioned above.

Since all three compounds studied here (see Fig. 1) are closed-shell molecules without degenerate excited states, of the three Faraday parameters $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ used to rationalize the MCD spectral shapes only the $\mathcal{B}$ term is of relevance.[23,45] For the calculation of the $\mathcal{B}$ term, the origin of the coordinate system was placed in the center of mass of the investigated system.

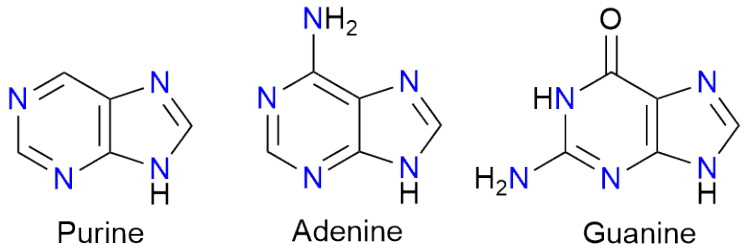

Figure 1: Schematic structures of purine and the nucleobases adenine and guanine.

The absorption spectra reported below display the decadic molar extinction coefficient[59]

$$\epsilon(\omega) = \frac{2e^2\pi^2 N_A}{1000 \times \ln(10)(4\pi\epsilon_0) n m_e c_0} \sum_j a_j(\omega) \frac{3}{2} \frac{f(0 \to j)}{\omega_j} = 7.03301 \times 10^2 \omega \sum_j a_j(\omega) \frac{3}{2} \frac{f(0 \to j)}{\omega_j} \tag{23}$$

in units of $\mathrm{M^{-1}cm^{-1}}$. Similarly, the MCD spectra display the anisotropy of the extinction coefficient (per unit of external magnetic field)

$$\frac{\Delta\epsilon(\omega)}{B_{ext}} = \frac{-8\pi^2 N_A \omega}{3 \times 1000 \times \ln(10)(4\pi\epsilon_0)\hbar c_0} \sum_j a_j(\omega) \mathcal{B}(0 \to j) = -5.98442 \times 10^{-3} \omega \sum_j a_j(\omega) \mathcal{B}(0 \to j) \tag{24}$$

in units of $\mathrm{M^{-1}cm^{-1}T^{-1}}$. In the above formulas, ω is the circular frequency of the radiation, N_A is the Avogardo number, c_0 is the speed of light in vacuo, ϵ_0 is the dielectric constant, n is the refractive index, $\hbar$ is Planck's constant, m_e is the electron mass, $\hbar\omega_j = (E_j - E_0)$ is the excitation energy and $a_j(\omega)$ is a Lorentzian function used to simulate the spectral broadening

$$a_j(\omega) = \frac{1}{\pi} \frac{\gamma}{(\omega_j - \omega)^2 + \gamma^2} \ . \tag{25}$$

The value $\gamma = 1000\ \mathrm{cm^{-1}}$ was adopted to facilitate the comparison with previous investigations.[6,10]

4 Purine

In the present study we include 9H-purine, which is the simplest member of the purine family. For purine, we have investigated independently the effects of using DFT with the B3LYP functional for ground state geometry optimizations for the following CC2 and CCSD calculations and also the basis set size for vertical excitation calculations. Fig. 2 presents the calculated excitation energies and oscillator strengths, along with natural transition orbitals (NTOs) and the corresponding assignment of the transitions. Comparison of adjacent columns in Fig. 2 shows that the excitation energies of the first and second nπ^* transitions at the CC2/aug-cc-pVDZ level at 4.43 eV and 5.48 eV are lower by 0.30 eV and 0.22 eV than the corresponding CCSD results for the same ground state geometry (optimized at the B3LYP level). This is not unexpected since CC2 excitation energies are usually lower than the corresponding CCSD values. The second $\pi\pi^*$ (L_b) transition from CC2 (at 5.69 eV for the B3LYP geometry and at 5.65 eV for the MP2 geometry) appears lower than the CCSD one by 0.20 eV, while for the first $\pi\pi^*$ (L_a) state at $\approx$ 5.25 eV the agreement between the two methods is good. The main differences between the two CCSD columns, which differ in the level for the ground state geometry optimization, is that the two $\pi\pi^*$ (L_a and L_b) states are lower by 0.05 eV and 0.04 eV for MP2-optimized ground state geometry, whereas the second nπ^* state is blued shifted by 0.03 eV to 5.73 eV.

For the same ground state geometry (optimized with MP2) the largest divergence between CCSD and CC2 can be observed for the first and second nπ^* and also the second $\pi\pi^*$ (L_b) state for which CC2 gives energies which are, respectively, 0.30 eV, 0.23 eV and 0.2 eV lower. For the larger triple-ζ basis set, the excitation energies for the lowest five states decrease by 0.02-0.06 eV and, thereupon the highest two states (L_b and nπ^*) become degenerate at 5.63 eV.

Inclusion of solvent effects in the framework of COSMO both for the geometry optimizations and in the excitation energy calculations lowers the excitation energies for the $\pi\pi^*$ (L_a and L_b) transitions by 0.13 eV and 0.18 eV, while those of the nπ^* transitions rise from 4.39,

5.44 and 5.63 eV to 4.58, 5.61 and 5.79 eV. As a consequence, the order of the second nπ^* and the second $\pi\pi^*$ states becomes reversed.

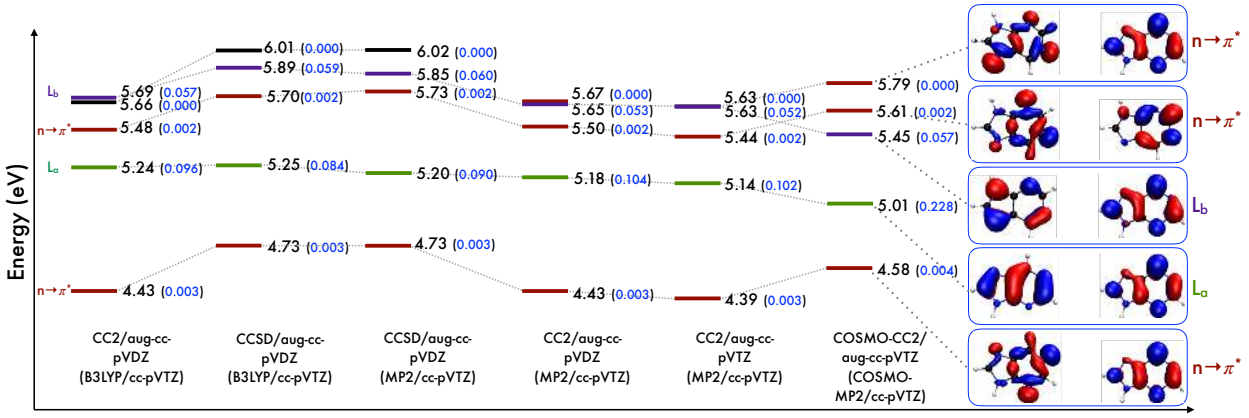

Figure 2: Purine: CCSD/aug-cc-pVDZ vertical singlet excitation energies in vacuum and CC2 vertical singlet excitation energies with the aug-cc-pVDZ and aug-cc-pVTZ basis sets in vacuum and in solution using the COSMO solvent model. Oscillator strengths of the transitions are given in blue, the levels of theory used for the (ground state) geometry optimizations are indicated in parentheses.

According to the results shown in Fig. 2, as both B3LYP and MP2 predict similar planar geometries for the ground state, changing the level of geometry optimization from DFT to MP2 changes the excitation energies at the CCSD level only very slightly and the order of states remains unchanged. Including the electrostatic contribution of the environment via COSMO stabilizes the $\pi\pi^*$ states from 5.14 to 5.01 eV for L_a and from 5.63 to 5.45 eV for L_b due to the larger transition dipole moments and the increase of the static dipole moments upon excitation by 2.13 debye for L_a and 4.88 debye for L_b. The interaction with the solvent also increases the intensity of L_a by a factor of ≈ 2. At variance with the $\pi\pi^*$ excitations, the nπ^* states are, as usual, destabilized (blue shifted) in aqueous solution because of the stabilization of the lone pairs. Given the experimental analysis of the absorption and MCD spectra of purine, the absorption band maximum in aqueous solution is located at 4.71 eV (263 nm) along with a shoulder at 5.17 eV (240 nm).[5] Our calculations are in agreement with this and associate the maximum band to the L_a state and the weak shoulder to L_b. The COSMO-CC2 results for the vertical excitation energies agree with the experimental band

maxima within 0.3 eV, which is roughly the intrinsic accuracy of the CC2 method. The deviation of the COSMO-CC2 results from the experimental values has also the expected sign: usually CC2 gives vertical excitation energies which are slightly blue-shifted and for purine in aqueous solution the COSMO model probably underestimates the solvent effects on the $\pi\pi^*$ states L_a and L_b by about 0.1 eV compared to calculations that include explicit water molecules.

The MCD spectra obtained for the gas phase at the CCSD and CC2 levels, and in aqueous solution at the COSMO-RI-CC2 level, from the Faraday $\mathcal{B}$ terms along with the OPA spectra of purine are depicted in Fig. 3. The $\mathcal{B}$ values used to simulate the spectra are collected in Table 1. The computed MCD spectra predict opposite signs for L_a and L_b and yield bi-signate spectra with a positive band on the red and a negative band on the blue side in agreement with the experimental spectrum, and also obtained with the B3LYP and CAM–B3LYP functionals in Ref. 10.

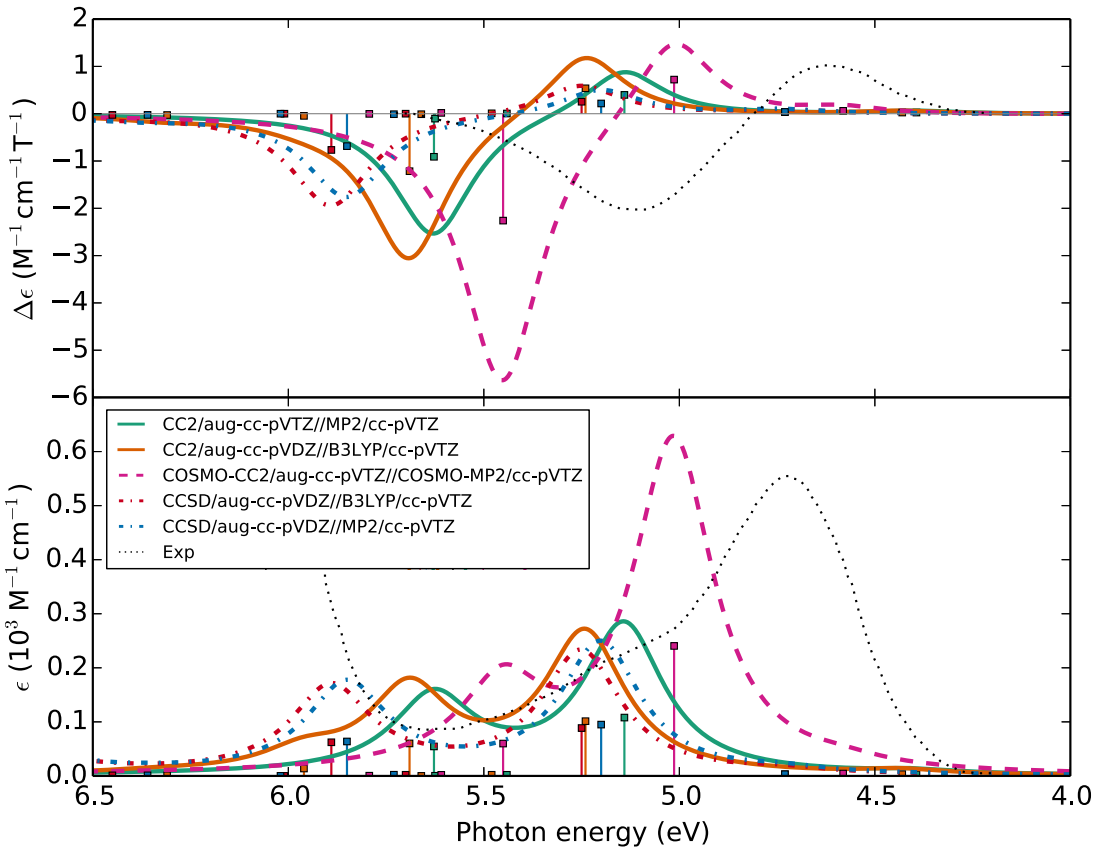

Figure 3: Purine. OPA (lower panel) and MCD (upper panel) spectra computed in vacuum at the CCSD and CC2 levels, and in water solution as approximated at the COSMO-RI-CC2/aug-cc-pVTZ level for the corresponding optimized structures (at B3LYP, MP2 and COSMO-MP2). All geometry optimizations were done in the cc-pVTZ basis set. Experimental spectra were measured in water at pH=7 and digitized from Ref. 5 .

Table 1: Purine. Computed values of the excitation energies ω_j (eV) and Faraday $\mathcal{B}(0 \to j)$ terms (a.u.) for the different methods discussed in the text. Note that the sign of the $\mathcal{B}$ term is opposite to the sign of the corresponding MCD peak (cmp. Eq. (24)).[24]

CC2/ aug-cc-pVDZ @B3LYP geom		CCSD/ aug-cc-pVDZ @B3LYP geom		CCSD/ aug-cc-pVDZ @MP2 geom		CC2/ aug-cc-pVTZ @MP2 geom		COSMO-RI-CC2/ aug-cc-pVTZ @COSMO-MP2 geom	
ω_j(type)	$\mathcal{B}(0 \to j)$	ω_j	$\mathcal{B}(0 \to j)$	ω_j	$\mathcal{B}(0 \to j)$	ω_j	$\mathcal{B}(0 \to j)$	ω_j	$\mathcal{B}(0 \to j)$
4.43 ($n\pi^*$)	−0.933	4.73	−1.19	4.73	−1.31	4.39	−0.93	4.58	−2.23
5.24 ($\pi\pi^*(L_a)$)	−17.1	5.25	−8.05	5.29	−6.96	5.14	−12.9	5.01	−24.0
5.48 ($n\pi^*$)	−0.25	5.70	−0.006	5.73	0.29	5.44	−0.13	5.61	−0.52
5.69 ($\pi\pi^*(L_b)$)	35.7	5.89	21.7	5.85	19.6	5.63	27.0	5.45	69.3
5.66 ($n\pi^*$)	0.31	6.01	0.003	6.02	−0.005	5.63	3.09	5.79	0.11

5 Adenine

On the basis of the CCSD/aug-cc-pVDZ results for Adenine at two different geometries (a planar one optimized at the B3LYP level and a non-planar one optimized at the RI-MP2 level), the three lowest excited states, two of $\pi\pi^*$ and one of $n\pi^*$ character, are rather close in energy, see Table 2.

Table 2: Adenine: Vertical excitation energies ω_j (eV), oscillator strengths $f(0 \to j)$ and Faraday $\mathcal{B}(0 \to j)$ terms (a.u.) of Adenine in vacuum calculated by different methods. Note that the sign of the $\mathcal{B}$ term is opposite to the sign of the corresponding MCD peak (cmp. Eq. (24)).[24] The geometry optimizations in all cases were done in the cc-pVTZ basis.

CCSD aug-cc-pVDZ @B3LYP geom			CCSD aug-cc-pVDZ @MP2 geom			CC2 aug-cc-pVDZ @MP2 geom	CC2 aug-cc-pVTZ @MP2 geom		
ω_j(type)	f	$\mathcal{B}(0 \to j)$	ω_j(type)	f	$\mathcal{B}(0 \to j)$	ω_j	ω_j(type)	f	$\mathcal{B}(0 \to j)$
5.36(L_b)	0.008	40.72	5.34(L_b)	0.030	24.48	5.14	5.10($n\pi^*$)	0.009	0.78
5.49(L_a)	0.280	−11.95	5.46($n\pi^*$)	0.064	−3.12	5.29	5.24(L_a)	0.155	313.14
5.50($n\pi^*$)	0.002	−11.68	5.55(L_a)	0.174	−13.35	5.29	5.27(L_b)	0.126	−310.06
5.54	0.009	−16.31	5.61	0.023	−5.52	5.39	5.56	0.011	−0.75
5.89	0.001	−0.41	5.95	0.005	−0.35	5.71	5.73	0.003	−0.71

To identify the character of these three excitations in CCSD, we have computed the NTOs for each geometry at this level, which are reported in Tables S3 and S4 of the Supporting Information. The most significant geometrical difference between adenine optimized at B3LYP and MP2 arises from the amino group; B3LYP predicts a planar geometry while MP2 gives some degrees of pyramidalization on NH_2. The level of theory employed for the geometry optimization has a remarkably large effect on the relative ordering of the second $\pi\pi^*$ and $n\pi^*$ states found by CCSD: at the planar B3LYP geometry, the CCSD calculations predict the two close-lying $\pi\pi^*$ and $n\pi^*$ transitions at 5.49 eV and 5.50 eV, while at the non-planar MP2 geometry the latter states appear in reversed ordering at 5.46 eV and 5.55 eV, with an increase in their energy gap. At both geometries the lowest excitation has $\pi\pi^*$ character and, based on the NTOs, is assigned to L_b, while the second $\pi\pi^*$ state, which carries most of the oscillator strength, is attributed to L_a. Remarkably, at the MP2 geometry, the CCSD excitation characterized as $n\pi^*$ has a relatively large oscillator strength — larger than for the L_b state. This indicates a mixing of the $n\pi^*$ with the L_a state induced by the non-planarity

of the geometry.[60]

Of particular significance is to see how well the CCSD and CC2 results agree for the valence $\pi\pi^*$ and $n\pi^*$ transitions. From the theoretical point of view, CC2 is an approximation to CCSD where the coupling between doubly excited determinants through the fluctuation potential is neglected. It gives usually excitation energies that are lower than for CCSD. Since higher-order correlation effects beyond CCSD often red-shift excitation energies, CC2 profits sometimes from a fortunate partial compensation of these two errors. Comparing the second and third columns of Table 2, the CC2 excitation energy for the $n\pi^*$ state at 5.14 eV is considerably lower than the CCSD result obtained with the same basis set and ground state geometry. The pair of close-lying $\pi\pi^*$ transitions in adenine's spectrum (first and third excitations with CCSD at the MP2-optimized geometry) causes CC2 to fail in resolving them in the aug-cc-pVDZ basis: it gives a pair of complex eigenvalues, as also observed in an earlier study.[20]

In the larger aug-cc-pVTZ basis, CC2 predicts, in agreement to what is obtained in the double-ζ basis, the lowest excitation at 5.10 eV to have $n\pi^*$ character while the two $\pi\pi^*$ transitions are resolved but close in energy at 5.24 eV and 5.27 eV and large values for the Faraday $\mathcal{B}$ terms. Note that the order of these $\pi\pi^*$ transitions is reversed with respect to the CCSD results, but at the same time also the sign of the respective $\mathcal{B}$ terms is inversed. For CC2/aug-cc-pVTZ the in the absorption spectrum brighter transition (L_a) is lower in energy. A comparison of the results obtained for CC2 in the triple-ζ basis with those at the EOM-CCSD(T) level[14] in the same basis shows that CC2 underestimates the excitation energy for the first $n\pi^*$ transition compared to EOM-CCSD(T) by ≈ 0.18 eV and shifts it down below the first two $\pi\pi^*$ states, while at the EOM-CCSD(T) level the ordering is flipped. Experimental results for the vertical excitation energies are not available. Molecular beam experiments showed that the band origin (0-0 transition) for the $n\pi^*$ state is slightly below that of a bright $\pi\pi^*$ state.[61]

To investigate the magnitude of solvent effects for aqueous solution – and in particular,

how important they are to get the correct ordering of the lowest states – we applied again the COSMO continuum model in both the geometry optimization and the excitation energy and transition strength calculations. In the case of adenine, since we have observed large solvation effects, we have independently examined the effects of COSMO contributions on the ground state geometry and also on vertical transition energies. The results are summarized, along with the vacuum values, in Fig. 4. Once again, the assignment of the excited states is based on NTOs.

In the COSMO-RI-CC2/aug-cc-pVDZ calculations the lowest two $\pi\pi^*$ excitations are, in contrast to the vacuum case, resolved with one state at 5.04 eV and the second at 5.27 eV for the vacuum geometry. Note that the inclusion of solvent effects causes red shifts of both $\pi\pi^*$ states while, in contrast, the nπ^* excitation is blue shifted from 5.14 eV to 5.40 eV and became the third excited state. These solvent shifts are typical for nπ^* and bright $\pi\pi^*$ states (cmp. e.g. Ref. 41) and have the same direction as those observed in Ref. 6 at the TD-DFT level.

Furthermore, when the structure is optimized with COSMO for bulk water (the fifth column from left), the energies do not change significantly (in the order of 0.01-0.03 eV) with respect to the same calculations at the vacuum geometry (the forth column from left). The ordering of the states remains unchanged. As expected, the excitation energies are slightly lower with the aug-cc-pVTZ basis (the last column from left) except for the πRyd state, which is blue shifted by 0.14 eV to 5.68 eV. Besides the three valence states, there is a πRyd state around 5.56 eV in vacuum at the CC2/aug-cc-pVTZ level, and around 5.68 eV in aqueous solution at the same level of theory. For the ordering of the first five transitions found at the COSMO-RI-CC2 level with both double-ζ and triple-ζ basis sets, using optimized structures in vacuum and solution, there is a definite analogy: $\pi\pi^*$ (L_a) $<$ $\pi\pi^*$ (L_b) $<$ nπ^* $<$ πRyd $<$ nπ^*. At the higher energy end of the studied region, CC2/aug-cc-pVTZ in vacuum predicts another nπ^* state at 5.73 eV, which is blue-shifted in solution to 5.89 eV.

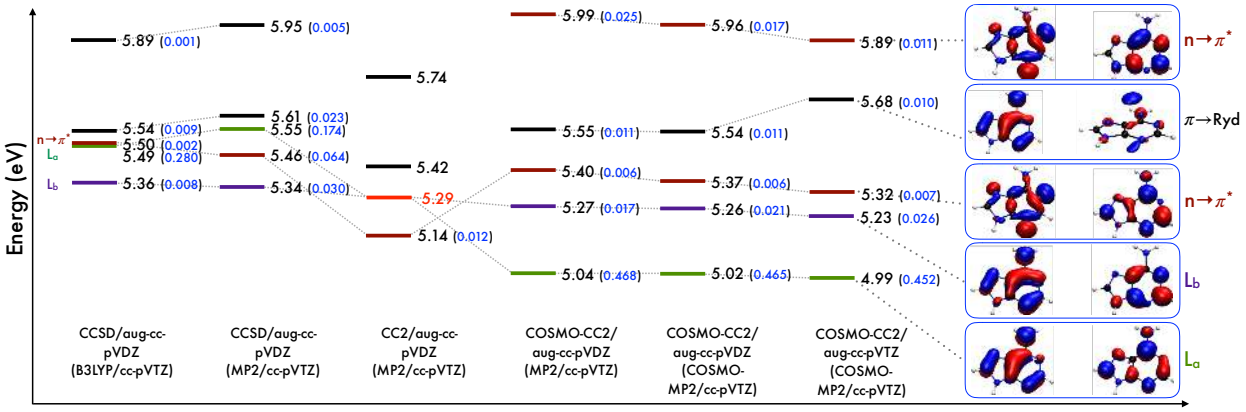

Figure 4: Adenine: CCSD/aug-cc-pVDZ singlet vertical excitation energies in vacuum and CC2 singlet vertical excitation energies with the double-ζ and triple-ζ basis sets in vacuum and in solution using the COSMO solvent model at the corresponding ground state optimized geometries. Oscillator strengths of the transitions are given in blue. For CC2/aug-cc-pVDZ the two lowest $\pi\pi^*$ states are obtained as a conjugated pair of degenerate roots with complex eigenvalues.

For adenine, we decided to study in more detail the effect of Hydrogen bonds on the lowest excited states. To this end, we performed additional calculations on a cluster with five explicit water molecules to model part of adenine's first solvation shell, both with and without the COSMO continuum solvation model to include bulk solvent effects. Such a cluster was already adopted in Ref. 6, and the number of water molecules had been chosen so to saturate the hydrogen bond ability of adenine in the molecular plane, where solute-solvent interactions are expected to be stronger. It is known that, in DNA nucleobases, the nπ^* states are destabilized in water by an extent not fully captured by continuum models.[4] Note that a single cluster cannot be considered fully representative of all the possible solute/solvent arrangements that should be studied with a proper Molecular Dynamics simulation, but such treatment goes beyond the scopes of the present article.

Due to its computational cost, the study on the cluster model was performed only for adenine. This nucleobase was selected because, in gas phase, some of the methods we applied predict a remarkable role of nπ^* states in its MCD response, see Table 3. Moreover, the assignment of the relative position of the L_a and L_b states in adenine is controversial and

important in the context of studying DNA photodamage.

As far as the excitation energies are concerned, just the inclusion of the explicit water molecules (without COSMO) results in an excitation energy for the first $\pi\pi^*$ (L_a) state of 5.02 eV, which is lower than the value in vacuum (by ≈ 0.22 eV) and only slightly higher (by ≈ 0.03 eV, cmp. Table 3) than the result obtained with COSMO without explicit water. For the second $\pi\pi^*$ (L_b) transition, the explicit water molecules lower the excitation energy to 5.16 eV, that is ≈ 0.11 eV and 0.07 eV lower than, respectively, in vacuum and with the continuum solvation model. The combination of explicit and implicit water models lowers the excitation energies for the $\pi\pi^*$ states further and at the same time increases the energy gap between them by a factor of 2 compared to the case where only the explicit water molecules are included.

Table 3: Adenine. Vertical excitation energies ω_j (eV), oscillator strengths $f(0 \to j)$ and Faraday $\mathcal{B}(0 \to j)$ terms (a.u.) calculated in vacuo and in solution at the CC2/aug-cc-pVTZ level of theory (geometries are optimized at MP2/cc-pVTZ level in the specified environment vacuo, COSMO, vacuo+$5H_2O$, COSMO+$5H_2O$).

Vacuum			COSMO			+$5H_2O$			COSMO+$5H_2O$		
ω_j(type)	f	$\mathcal{B}(0 \to j)$	ω_j	f	$\mathcal{B}(0 \to j)$	ω_j	f	$\mathcal{B}(0 \to j)$	ω_j	f	$\mathcal{B}(0 \to j)$
5.24 ($\pi\pi^*$ (L_a))	0.155	313.14	4.99	0.452	77.13	5.02	0.303	117.42	4.87	0.504	101.77
5.27 ($\pi\pi^*$ (L_b))	0.126	−310.06	5.23	0.026	−75.76	5.16	0.063	−121.56	5.14	0.041	−112.00
5.10 ($n\pi^*$)	0.009	0.78	5.32	0.007	−4.55	5.59	0.000	−0.02	5.61	0.002	−1.69
5.56 (πRyd)	0.011	−0.75	5.68	0.010	0.71	5.57	0.005	−1.12	5.64	0.005	0.72
5.73 ($n\pi^{*,a}$)	0.003	−0.71	5.89	0.011	−0.11	6.10	0.006	0.95	6.06	0.645	103.39

[a] The fifth state has an $n\pi^*$ character for the vacuum and COSMO calculations while it is a πRyd and $\pi\pi^*$ state for adenine in water cluster without and with COSMO contribution, respectively.

The $n\pi^*$ states, however, are significantly blue shifted — by ≈ 0.49 eV to 5.59 eV for the first $n\pi^*$ state — when explicit water molecules are added to the vacuum calculation. With the continuum solvation included, the addition of explicit water molecules still leads for the 1 $n\pi^*$ transition to an additional blue shift of 0.29 eV to 5.61 eV, confirming that a relevant part of the $n\pi^*$ destabilization in water can only be captured with an explicit description of the solute-solvent H-bonds.

The fifth excited state in vacuum and in the COSMO calculation is the 2 $n\pi^*$ state, but when explicit water molecules are added this state is shifted upwards and a πRyd transition at 6.10 eV becomes the fifth excited states without COSMO and a $\pi\pi^*$ transition at 6.06 eV

when also the COSMO model is included.

For the vertical excitation energies in the gas phase, EOM-CCSD(T)/aug-cc-pVTZ results computed at a CCSD/cc-pVDZ geometry are available from Ref. 14. They are 5.04 eV for the L_b, 5.23 eV for the L_a, and 5.28 eV for the 1 $n\pi^*$ state and 0.26, 0.24, and 0.26 eV lower than the respective CCSD results for the same basis set and structure. CCSD/cc-pVDZ structures are for organic closed-shell molecules inferior to MP2/cc-pVTZ structures. Therefore, we prefer to add the CCSD $\rightarrow$ EOM-CCSD(T) shifts from Ref. 14 and the aug-cc-pVDZ $\rightarrow$ aug-cc-pVTZ shifts from the CC2 calculations (0.05, 0.02, and 0.04 eV for respectively L_a, L_b and 1 $n\pi^*$) to the current CCSD/aug-cc-pVDZ results for the MP2/cc-pVTZ geometry to obtain best estimates for gas phase vertical excitation energies. This gives 5.06 eV for the L_b, 5.26 eV for the L_a, and 5.16 eV for the 1 $n\pi^*$ state. Compared to these best estimates, the gas-phase CC2/aug-cc-pVDZ results overestimate the excitation energies for the L_b and L_a states by, respectively, 0.23 eV and 0.03 eV and underestimate that for the 1 $n\pi^*$ state by 0.02 eV. If we add these shifts to the COSMO-CC2 results for adenine with five explicit water molecules, we arrive at best estimates for the excitation energies in aqueous solution of 4.84 eV for L_a, 4.91 eV for L_b and 5.63 eV for 1 $n\pi^*$.

From the above discussion and the results displayed in Fig. 4, one can draw the following conclusions: due to remaining correlation errors in the order of 0.2 eV, which have different trends for different types of states, CC2 can not reproduce for the gas phase the ordering of states obtained in EOM-CCSD(T). The position of the 1 $n\pi^*$ is, however, also sensitive to the structure: at the CCSD/cc-pVDZ structure, this excitation appears, for EOM-CCSD and EOM-CCSD(T) at, respectively, 0.07 eV and 0.05 eV above the L_a state; for the B3LYP/cc-pVTZ structure it is, at the CCSD level, almost degenerate with L_a and at the MP2/cc-pVTZ structure it is found at 0.09 eV and 0.15 eV below L_a, for CCSD and CC2 respectively. Experimentally the band origin of the $n\pi^*$ state is slightly below that of the lowest $\pi\pi^*$ excitation.[61] However, these uncertainties are small compared to the solvent shifts for aqueous solution, which are -0.37 eV for the L_a, -0.13 eV for the L_b and $+0.51$ eV

for the 1 nπ^* state. As also found in a previous TD-DFT study,[6] continuum solvation models underestimate in this case the solvent effects because of hydrogen bond interactions, although they already shift the 1 nπ^* above the L_a and L_b states and bring the brighter L_a state below L_b. The inclusion of explicit water molecules is necessary to obtain quantitative results and further boosts the energy gap between the L_b and the 1 nπ^* state by 0.38 eV. In synthesis both the CC2 results and our "best estimates" predict that for adenine in aqueous solution L_a is more stable than L_b, confirming previous TD-DFT results.[6] This result is partially at variance with a CASPT2/MM study[62] which, adopting a non-polarizable molecular mechanics (MM) force field, predicted L_b slightly more stable than L_a by 0.02 eV.

Experimentally, the absorption band maximum of adenine in aqueous solution is located at 4.77 eV (260 nm) compared with 4.92 eV (252 nm) in the gas phase, which corresponds to a redshift of 0.15 eV (8 nm).[63,64] In the same way as the experimental results, the L_a state, which carries most of the oscillator strength, is, in our COSMO-RI-CC2 calculations including the water cluster, red-shifted by $\approx$ 0.37 eV (18 nm) with respect to the results in the gas phase. The best estimate of 4.84 eV obtained by including higher-order correlation effects agrees very well with the experimental band maximum in aqueous solution.

The computed absorption and MCD spectra of adenine in vacuum and aqueous solution including the water cluster at the CCSD and CC2 levels are shown in Fig. 5. They were obtained from the values of the excitation energies, oscillator strengths and Faraday $\mathcal{B}$ terms provided by our calculations in Table 2 and Table 3. In the vacuum CC2/aug-cc-pVTZ case, even though the two $\pi\pi^*$ states have large oppositely signed $\mathcal{B}$ terms, the energy gap between them is so small that the broadening will result in an inconspicous signal. In all CC2 calculations, L_a is at lower energy than L_b and has a positive $\mathcal{B}$ term (i.e., a negative MCD peak), whereas L_b is at higher energy with a negative $\mathcal{B}$ term (positive MCD peak). The CCSD calculations, on the other hand, predict L_b in the gas phase energetically lower than L_a, and it is in this case L_b that carries the positive $\mathcal{B}$ term (negative MCD peak) and L_a the negative one. Moreover, the intensity of the positive MCD peak originates, in

the gas phase CCSD calculation, not only from the $\pi\pi^*$ state, but also from the $n\pi^*$ states (cmp. Table 2). As a consequence, the spectral profile of the computed MCD spectra is, for CCSD in the gas phase and CC2 in the gas phase and in solution, compatible with that from experiment[5,11,12] and the TD-DFT calculations,[6] showing the characteristic pair of negative and positive bands in the region of 4.87–5.27 eV, despite the different relative positions of the L_a and L_b states, and the different composition of the second band yielded by CCSD.

The above observation has led us to reconsider whether MCD can indeed give direct evidence on the relative position of the L_a and L_b states, as proposed in Ref. 6. In this regard, it is useful to inspect the sum-over-state (SOS) expression of the $\mathcal{B}(0 \rightarrow j)$ term[65] and look in particular at the following term

$$\sum_{k \neq j} \frac{\langle j|\mathbf{m}|k\rangle \cdot \langle 0|\mathbf{d}|j\rangle \times \langle k|\mathbf{d}|0\rangle}{\omega_k - \omega_j} \equiv \epsilon_{\alpha\beta\gamma} \sum_{k \neq j} \frac{m_\alpha^{jk} d_\beta^{0j} d_\gamma^{k0}}{\omega_k - \omega_j} \tag{26}$$

where 0 is the ground state, j is the final excited state, and k runs on all states except j; $d_\alpha^{jk} = \langle j|d_\alpha|k\rangle$, and similarly for the other terms.

If, for two close lying excited states, the $\mathcal{B}$ terms are dominated by this contribution, they fulfill the relation:

$$\mathcal{B}(0 \rightarrow 1) \approx \epsilon_{\alpha\beta\gamma} \frac{m_\alpha^{12} d_\beta^{01} d_\gamma^{20}}{\omega_2 - \omega_1} \tag{27}$$

$$\mathcal{B}(0 \rightarrow 2) \approx \epsilon_{\alpha\beta\gamma} \frac{m_\alpha^{21} d_\beta^{02} d_\gamma^{10}}{\omega_1 - \omega_2} = \epsilon_{\alpha\beta\gamma} \frac{-m_\alpha^{12} d_\beta^{20} d_\gamma^{01}}{\omega_1 - \omega_2} = \epsilon_{\alpha\gamma\beta} \frac{-m_\alpha^{12} d_\gamma^{20} d_\beta^{01}}{\omega_1 - \omega_2} = +\epsilon_{\alpha\beta\gamma} \frac{m_\alpha^{12} d_\beta^{01} d_\gamma^{20}}{\omega_1 - \omega_2} \tag{28}$$

Thus, the two $\mathcal{B}$ terms only differ in sign because of their energy difference in the denominator and a bi-signate spectral profile with a positive and a negative band of equal magnitude will be obtained. The characters of the two states determine through the transition moments the magnitude and the sign of the numerator. Since the MCD spectra computed at the CC2 level exhibit such an antisymmetric relation between the $\mathcal{B}$ terms for the L_a and L_b states, it is reasonable to assume that this contribution is the dominant one.

To test this hypothesis, we evaluated the SOS expression for a three state (S_0, L_a, and L_b) model with transition moments evaluated at the COSMO-CC2/aug-cc-pVTZ level for the COSMO structure. Including all SOS terms gave for this model 74.1 and -73.4 au compared to 77.13 and -75.76 au from the CC2 response calculation. The contribution from the magnetic dipole moment between the L_a and L_b states to this SOS model is 72.9 and -72.9 au. These results support the above assumption that the Faraday $\mathcal{B}$ term of the L_a and L_b states in adenine are in fact dominated by the contribution given in Eq. (26). The SOS procedure for non-variational CC methods is outlined in the Supplementary Information.

As a consequence, and contrary to what was assumed in Ref. 6, the positive or negative sign of the two red peaks in the MCD spectrum of adenine is not related to whether the underlying electronic states are L_a or L_b, and MCD cannot be used to ascertain the relative position of these two states. In other words, even if different methods predict opposite relative positions of L_a and L_b, the MCD profiles are still correctly reproduced for all of them.

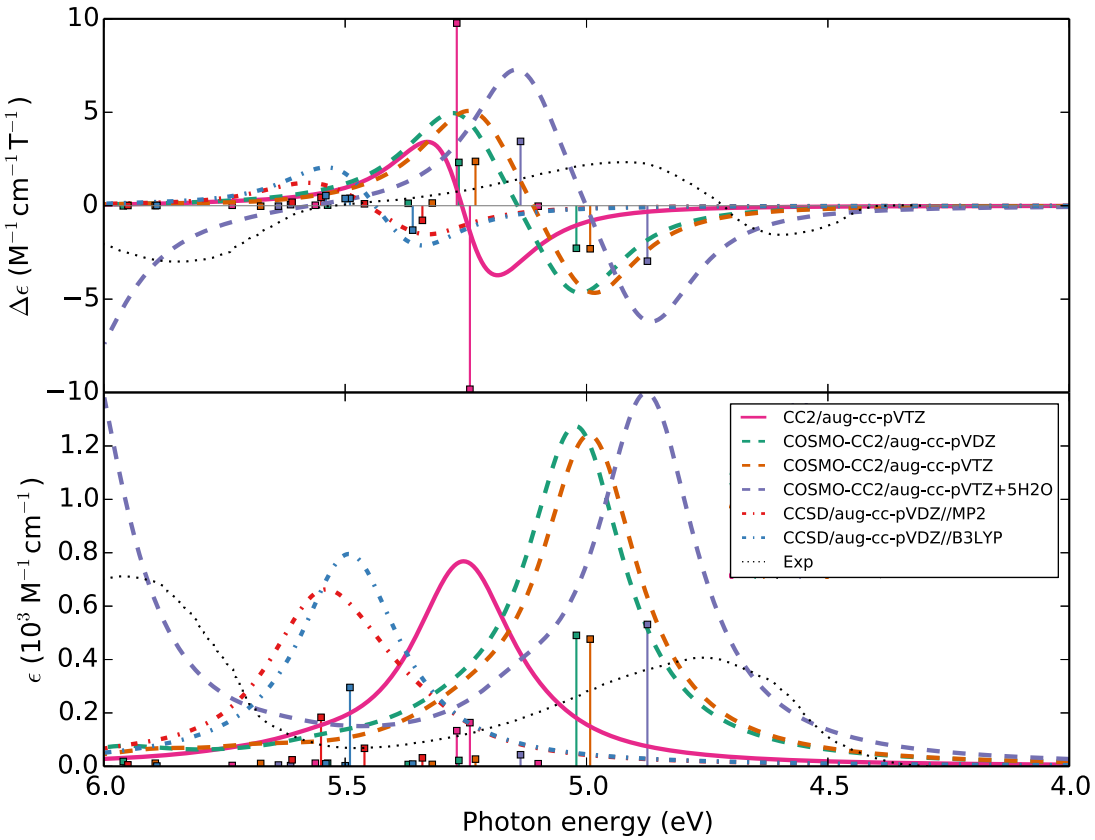

Figure 5: Adenine. MCD and OPA spectra computed for a single optimized structure (at MP2/cc-pVTZ level) in vacuum and in water solution, as approximated at COSMO, at CCSD and CC2 levels using double-ζ and triple-ζ basis sets. Experimental spectra were measured in water at pH=7 digitized from Ref. 11.

The wall-time for the calculations of vertical excitation energies plus transition moments and magnetic Faraday $\mathcal{B}$ terms in solution at COSMO-RI-CC2 is very close to that of corresponding calculations for vacuum case. In the case of adenine, the calculations of these quantities for five states in solution using the aug-cc-pVDZ basis set with no explicit water cluster took 5 hours and 17 minutes with OpenMP parallelization on 8 CPUs of an Intel Xeon X5670 CPU with 8 GiB core memory.

6 Guanine

The structure of 9H-Guanine has been optimized in vacuo and in aqueous solution via COSMO at the MP2/cc-pVTZ level. The geometries are basically planar, except for a slight pyramidalization at the amino group (coordinates are available in the Supporting Information). Excitation energies of the first five transitions have been obtained in vacuum and also in aqueous solution in the framework of the COSMO solvent model at the CC2 level using the aug-cc-pVTZ basis set. Fig. 6 shows the results of these calculations along with the corresponding NTOs and the assignments.

As shown in Fig. 6, the lowest-energy transition in vacuum is a $\pi\pi^*$ (L_a) transition at 5.02 eV, in agreement with previous calculations on Guanine with EOM-CCSD(T),[14] CC2[20] and CASPT2,[22] while the assignment based on EOM-CCSD[14] and TD-DFT are given as πRyd and $\pi\pi^*$ (L_b), respectively.

The CC2 results in Fig. 6 predict πRyd character for the two transitions above the L_a state located at 5.11 eV and 5.44 eV, with non-zero intensities. It is interesting to note that CCSD(T) and the current CC2 calculations in vacuum agree well on the order of the three lowest excited states, although CC2 predicts higher energies for those states. The nπ^* state, which involves a lone pair of the oxygen atom, lies at 5.46 eV, just below the second $\pi\pi^*$ (L_b) state at 5.49 eV. The energy of this nπ^* transition is obtained slightly higher than with EOM-CCSD(T) (by $\approx$ 0.03 eV) and, due to near-degeneracy in vacuum at the CC2 level (cmp. Fig 6), the oscillator strength of the $\pi\pi^*$ (L_b) transition is partially distributed to the nπ^* and πRyd transitions. Although the nπ^* and $\pi\pi^*$ (L_b) transitions are very close in energy (their gap is $\approx$ 0.03 eV), underestimation of the nπ^* excitation energy at the CC2 level yields the nπ^* transition below L_b, contrary to CCSD and CCSD(T) where the reversed ordering is predicted.[14]

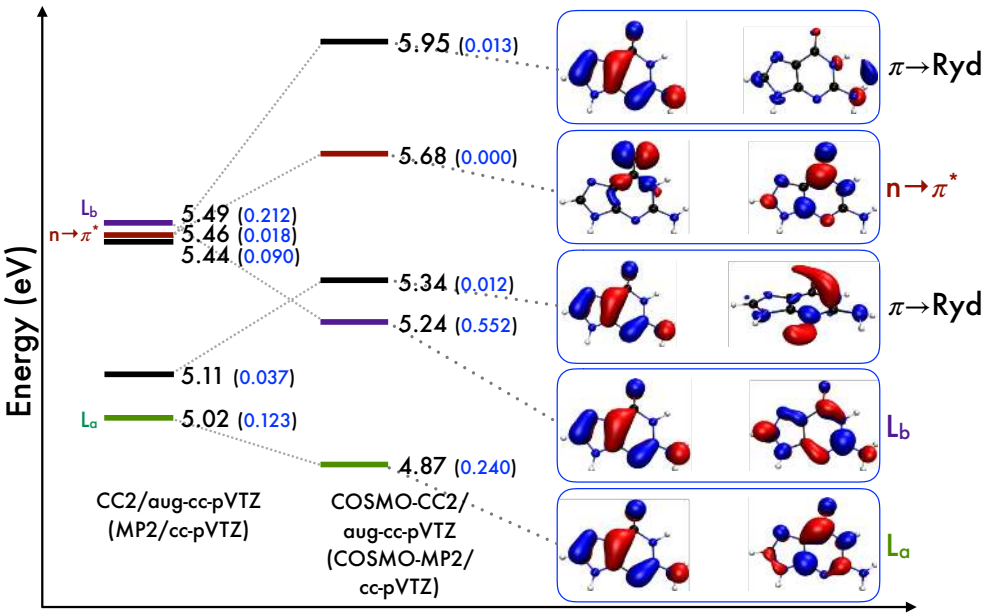

Figure 6: Guanine: CC2/aug-cc-pVTZ vertical singlet excitation energies in vacuum and solution using COSMO solvent model at the corresponding ground state optimized geometries. Oscillator strengths of the transitions are given in blue.

In aqueous solution the lowest transitions are energetically more separated than in vacuum (cmp. Fig. 6). Inclusion of the electrostatic effects of an isotropic environment lowers the energies of the most intense $\pi\pi^*$ type transitions to 4.87 eV and 5.24 eV, which can be attributed to the large transition dipole moments. The intensities are also doubled with respect to vacuum. Similar as for purine and adenine, the different behavior of the $\pi\pi^*$ and $n\pi^*$ transitions in aqueous solution affects the relative ordering of the states by blue shifting the $n\pi^*$ transition to 5.68 eV, i.e. above the $\pi\pi^*$ states, while its intensity decreases to zero. Finally, the πRyd transitions show similar trends as $n\pi^*$: in aqueous solution they appear at higher energies (5.34 eV and 5.95 eV) and have lower intensities. For guanine, given that the assignments of the transitions are clear, no further calculations have been done.

Fig. 7 shows a simulation of the absorption and MCD spectra obtained from the computed excitation energies, oscillator strengths and MCD $\mathcal{B}$ terms summarized in Table 4. The CC2/aug-cc-pVTZ results in vacuum and in solution for 9H-Guanine lead again to a bisignate spectrum, analogous to the one obtained at the TD-DFT level (B3LYP and CAM–

Table 4: Guanine. Vertical excitation energies ω_j (eV), oscillator strengths $f(0 \rightarrow j)$ and Faraday $\mathcal{B}(0 \rightarrow j)$ terms (a.u.) calculated in vacuo and in aqueous solution (COSMO) at two different geometries.

CC2/aug-cc-pVTZ			COSMO-CC2/aug-cc-pVTZ		
ω_j (type)	f	$\mathcal{B}(0 \rightarrow j)$	ω_j	f	$\mathcal{B}(0 \rightarrow j)$
5.02 ($\pi\pi^*$ (L_a))	0.123	39.0	4.87	0.240	118.
5.11 (πRyd)	0.037	14.6	5.34	0.012	−2.15
5.44 (πRyd)	0.090	−20.5	5.95	0.013	1.41
5.46 (nπ^*)	0.018	−7.39	5.68	0.000	−0.09
5.49 ($\pi\pi^*$)	0.212	−32.2	5.24	0.552	−135.

B3LYP functionals) in Ref. 6, and in agreement with both the experimental spectrum of the guanine derivative guanosine,[11] and the experimental peak maxima of guanine at pH = 2 reported in Ref. 12.

In the COSMO-CC2 calculation the signal originates exclusively from the oppositely signed $\mathcal{B}$ terms of the L_a and L_b states, whereas in the vacuum case the MCD intensity is distributed on both $\pi\pi^*$, πRyd and nπ^* states. In aqueous solution both $\pi\pi^*$ states, L_a and L_b, are red shifted where L_a is ≈ 0.37 eV more stable than L_b.

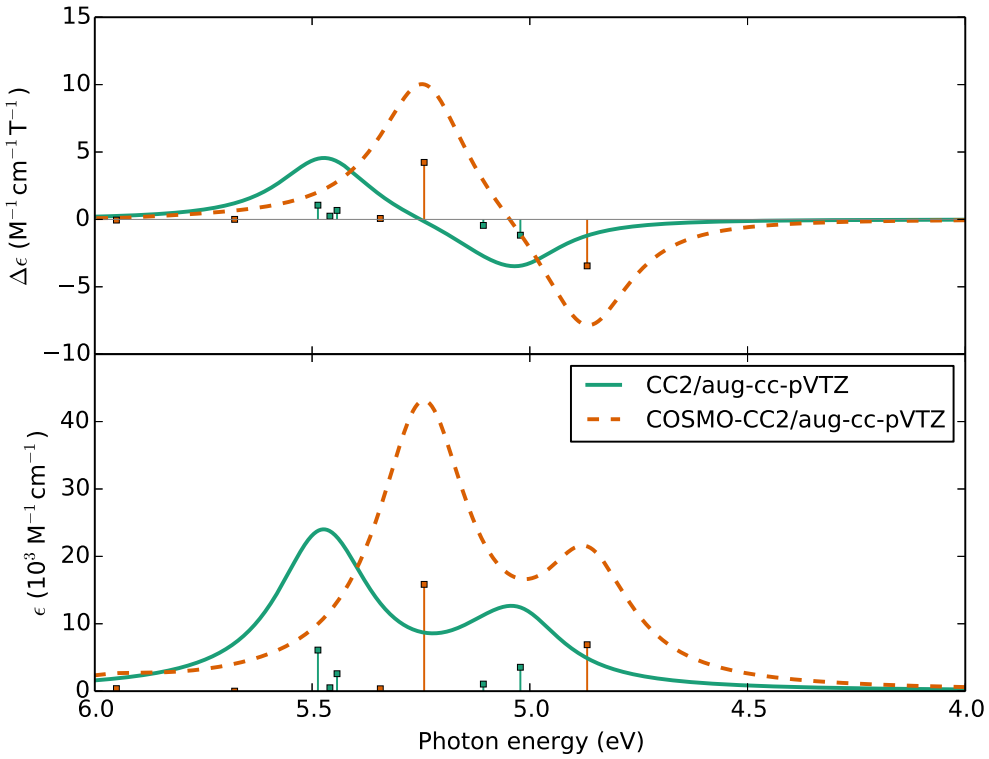

Figure 7: Guanine. Absorption and MCD spectra computed for a single structure (optimized at MP2/cc-pVTZ level) in vacuo and in water solution as approximated by COSMO at CC2/aug-cc-pVTZ level.

7 Conclusions

The vertical transition energies, oscillator strengths and Faraday $\mathcal{B}$ terms of purine and of the two primary nucleobases adenine and guanine have been investigated by the ab initio methods CCSD and CC2 in gas phase. The electrostatic effects of an isotropic environment on the absorption and MCD spectra were taken into account using the implicit continuum solvation model COSMO at the CC2 level of theory. The main aim of this study was to evaluate the performance of CCSD and CC2 in vacuum along with the solvent effects on the ordering of the transitions and on the MCD spectra. In the case of purine, in the gas phase, the latter methods agree well on the order of the two lowest $\pi\pi^*$ transitions, where L_a is

lower than L_b. Although in some calculations for purine at the CC2 level the $n\pi^*$ excitation energies are underestimated compared to CCSD (sometimes by as much as 0.35 eV), there is a reasonable agreement between these two methods even for the $n\pi^*$ transitions and connected triple excitations are expected to red shift the CCSD results by ≈ 0.2 eV. Our COSMO-RI-CC2 calculations for purine in aqueous solution show that the energy ordering between the $\pi\pi^*$ and $n\pi^*$ transitions is affected by the solvent as the $\pi\pi^*$ states are red-shifted and the $n\pi^*$ states are blue-shifted.

Our results for adenine reproduce the discrepancy in the energy order of the $\pi\pi^*$ transitions in gas phase between CCSD and CC2: the former method predicts the L_b state below L_a, while this order is flipped at the CC2 level. Inclusion of bulk solvent effects as described by the COSMO model has significant effects. It enlarges the energy gap between the two $\pi\pi^*$ transitions and destabilizes the $n\pi^*$ transitions. Concerning specific solvent effects from hydrogen bonding to the solute, we observed a qualitative agreement for the transition energies and the energy order of the L_a and L_b states in adenine between the COSMO calculations with and without explicit water molecules and calculations with explicit water molecules without COSMO. We can thus conclude that for the nucleobase adenine including the bulk effect of the environment can be sufficient for a qualitative description of the vertical excitation energies, but for quantitative results explicit water molecules are necessary. For Guanine, the absolute values of the CC2 excitation energies in vacuum agree well with those from CCSD(T),[14] however, there is a discrepancy in the ordering of the transitions: CC2 gives an $n\pi^*$ below L_b,[20] while CCSD(T) predicts a reversed order for the latter transitions. On the other hand, the inclusion of solvent effects for aqueous solution in the CC2 calculations for guanine changes the arrangement of the transitions by lowering the $\pi\pi^*$ and up-shifting the $n\pi^*$ transitions.

The spectral profiles of the computed MCD spectra with all methods used (vacuum CCSD, vacuum CC2 and COSMO-RI-CC2) are found to match the (bi-signate) experimental profiles, apart from some shifts on the energy scale and variations in intensity.

For adenine in water, CC2 with any of the solvent models predicts the lower energy $\pi\pi^*$ to be of L_a type and to bear a positive $\mathcal{B}$ term (negative MCD peak) in agreement with what previously found by two of us at the TD-DFT level.[6] In vacuo, at variance with CC2, CCSD predicts the MCD negative band to be due to the L_b state and the positive one to a sum of contributions from both L_a, $n\pi^*$ and πRyd states. Stimulated by this conflicting predictions, we performed an analysis of the sum-over-states expression of the $\mathcal{B}$ term, finding strong arguments that indicate that the sign of the MCD bands cannot be used to ascertain the relative stability of the L_a and L_b states of adenine.[66]

Acknowledgement

S.C. thanks J. Michl for discussions. R.F. and S.C. acknowledge financial support from DTU Chemistry (start-up post-doctoral grant) and from the Independent Research Fund Denmark through the DFF-Forskningsprojekt2 grant no. 7014-00258B. S.K. and C.H. acknowledge financial support from the Cluster of Excellence RESOLV (EXC 1069) and project Ha 2588/8-1 funded by the Deutsche Forschungsgemeinschaft.

Supporting Information: Table S1 and S2: Values for Faraday $\mathcal{B}$ terms at RI-CC2 and CCSD in vacuum and solution. Table S3 and S4: NTOs for adenine at CCSD. Table S5: NTOs for adenine with water cluster including COSMO at RI-CC2. Cartesian coordinates of the nucleobases.

References

(1) Kohler, B. Nonradiative Decay Mechanisms in DNA Model Systems. *J. Phys. Chem. Lett.* **2010**, *1*, 2047–2053.

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

(3) 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*.

(4) Improta, R.; Santoro, F.; Blancafort, L. Quantum Mechanical Studies on the Photophysics and the Photochemistry of Nucleic Acids and Nucleobases. *Chem. Rev.* **2016**, *116*, 3540–3593.

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

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

(7) Fülscher, M. P.; Serrano-Andrés, L.; Roos, B. O. A Theoretical Study of the Electronic Spectra of Adenine and Guanine. *J. Am. Chem. Soc.* **1997**, *119*, 6168–6176.

(8) Lan, Z.; Lu, Y.; Fabian, E.; Thiel, W. QM/MM nonadiabatic decay dynamics of 9H-adenine in aqueous solution. *ChemPhysChem* **2011**, *12*, 1989–1998.

(9) Mondal, S.; Puranik, M. Ultrafast structural dynamics of photoexcited adenine. *Phys. Chem. Chem. Phys.* **2017**, *19*, 20224–20240.

(10) Fahleson, T.; Kauczor, J.; Norman, P.; Santoro, F.; Improta, R.; Coriani, S. TD-DFT Investigation of the Magnetic Circular Dichroism Spectra of Some Purine and Pyrimidine Bases of Nucleic Acids. *J. Phys. Chem. A* **2015**, *119*, 5476–5489.

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

(12) 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.

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

(14) Szalay, P. G.; Watson, T.; Perera, A.; Lotrich, V. F.; Bartlett, R. J. 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.

(15) Szalay, P. G.; Watson, T.; Perera, A.; Lotrich, V. F.; Fogarasi, G.; Bartlett, R. J. Benchmark Studies on the Building Blocks of DNA. 2.Effect of Biological Environment on the Electronic Excitation Spectrum of Nucleobases. *J. Phys. Chem. A* **2012**, *116*, 8851–8860.

(16) Szalay, P. G.; Watson, T.; Perera, A.; Lotrich, V. F.; Bartlett, R. J. Benchmark Studies on the Building Blocks of DNA. 3. *J. Phys. Chem. A* **2013**, *117*, 3149–3157.

(17) Silva-Junior, M. R.; Schreiber, M.; Sauer, S. P. A.; Thiel, W. Benchmarks for Electronically Excited States: Time-Dependent Density Functional Theory and Density Functional Theory Based Multireference Configuration Interaction. *J. Chem. Phys.* **2008**, *129*, 104103/1–104103/14.

(18) Serrano-Andrés, L.; Merchán, M.; Borin, A. Adenine and 2-Aminopurine: Paradigms of Modern Theoretical Photochemistry. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 8691–8696.

(19) Conti, I.; Altoé, P.; Stenta, M.; Garavelli, M.; Orlandi, G. Adenine Deactivation in DNA Resolved at the CASPT2//CASSCF/ AMBER Level. *Phys. Chem. Chem. Phys.* **2010**, *12*, 5016–5023.

(20) Fleig, T.; Knecht, S.; Hättig, C. Quantum-Chemical Investigation of the Structures and Electronic Spectra of the Nucleic Acid Bases at the Coupled Cluster CC2 Level. *J. Phys. Chem. A* **2007**, *111*, 5482–5491.

(21) Shukla, M. K.; Leszczynski, J. TDDFT Investigation on Nucleic Acid Bases: Comparison with Experiments and Standard Approach. *J. Comput. Chem.* **2004**, *25*, 768–778.

(22) Yamazaki, S.; Domcke, W.; Sobolewski, A. L. Nonradiative Decay Mechanisms of the Biologically Relevant Tautomer of Guanine. *J. Phys. Chem. A* **2008**, *112*, 11965–11968.

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

(24) Michl, J. Magnetic Circular Dichroism of Aromatic Molecules. *Tetrahedron* **1984**, *40*, 3845–3934.

(25) Klamt, A.; Schüürmann, G. COSMO:A new approach to dielectric screening in solvents with explicit expressions. *J. Chem. Soc. Perkin Trans. 2* **1993**, 799–805.

(26) Christiansen, O.; Koch, H.; Jørgensen, P. The second-order approximate coupled cluster singles and doubles model CC2. *Chem. Phys. Lett.* **1995**, *243*, 409–418.

(27) Hättig, C.; Weigend, F. CC2 excitation energy calculations on large molecules using the resolution of the identity approximation. *J. Chem. Phys.* **2000**, *113*, 5154–5161.

(28) Cammi, R. Quantum cluster theory for the polarizable continuum model.I. The CCSD level with analytical first and second derivatives. *J. Chem. Phys.* **2009**, *131*, 164104.

(29) Cammi, R. Coupled-Cluster theories for the polarizable continuum model.II. Analytical gradients for excited states of molecular solutes by the equation of motion coupled-cluster method. *Int. J. Quantum Chem.* **2010**, *110*, 3040–3052.

(30) Cammi, R. Coupled-Cluster theory for the polarizable continuum model. III. A response theory for molecules in solution. *Int. J. Quantum Chem.* **2012**, *112*, 2547–2560.

(31) Caricato, M. Exploring potential energy surfaces of electronic excited states in solution with the EOM-CCSD-PCM method. *J. Chem. Theory Comput.* **2012**, *8*, 5081–5091.

(32) Caricato, M. A corrected-linear response formalism for the calculation of electronic excitation energies for solvated molecules with the CCSD-PCM method. *J. Comput. Theor. Chem.* **2014**, *1040*, 99–105.

(33) Caricato, M. Linear response coupled cluster theory with the polarizable continuum model within the singles approximation for the solvent response. *J. Chem. Phys.* **2018**, *148*, 134113.

(34) Christiansen, O.; Jørgensen, P.; Hättig, C. Response functions from Fourier component variational perturbation theory applied to a time-averaged quasienergy. *Int. J. Quantum Chem.* **1998**, *98*, 1.

(35) Christiansen, O.; Koch, H.; Jørgensen, P. The second-order approximate coupled cluster singles and doubles model CC2. *Chem. Phys. Lett.* **1995**, *243*, 409–418.

(36) Christiansen, O.; Halkier, A.; Koch, H.; Jørgensen, P.; Helgaker, T. Integral-direct coupled cluster calculations of frequency-dependent polarizabilities, transition probabilities and excited-state properties. *J. Chem. Phys.* **1998**, *108*, 2801–2816.

(37) Hättig, C.; Christiansen, O.; Jørgensen, P. Multiphoton transition moments and absorption cross sections in coupled cluster response theory employing variational transition moment functionals. *J. Chem. Phys.* **1998**, *108*, 8331–8354.

(38) Hättig, C.; Weigend, F. CC2 excitation energy calculations on large molecules using the resolution of the identity approximation. *J. Chem. Phys.* **2000**, *113*, 5154–5161.

(39) Friese, D. H.; Hättig, C.; Ruud, K. Calculation of two-photon absorption strengths with the approximate coupled cluster singles and doubles model CC2 using the resolution-of-identity approximation. *Phys. Chem. Chem. Phys.* **2012**, *14*, 1175–1184.

(40) Khani, S. K.; Khah, A. M.; Hättig, C. COSMO-RI-ADC(2) Excitation Energies And Excited State Gradients. *Phys. Chem. Chem. Phys.* **2018**, *20*, 16354–16363.

(41) Lunkenheimer, B.; Köhn, A. Solvent Effects on Electronically Excited States Using the Conductor-Like Screening Model and the Second-Order Correlated Method ADC(2). *J. Chem. Theory Comput.* **2013**, *9*, 977–994.

(42) Schwabe, T.; Sneskov, K.; Olsen, J. M. H.; Kongsted, J.; Christiansen, O.; Hättig, C. PERI-CC2: A Polarizable Embedded RI-CC2 Method. *J. Chem. Theory Comput.* **2012**, *8*, 3274–3283.

(43) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*; John Wiley & Sons: Chichester, 2000.

(44) 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.

(45) Kjærgaard, T.; Coriani, S.; Ruud, K. Ab initio calculation of magnetic circular dichroism. *WIREs Comput Mol Sci* **2012**, *2*, 443–455.

(46) Hättig, C.; Jørgensen, P. Derivation of coupled cluster excited states response functions and multiphoton transition moments between two excited states as derivatives of variational functionals. *J. Chem. Phys.* **1998**, *109*, 9219.

(47) Hättig, C.; Christiansen, O.; Coriani, S.; Jørgensen, P. Static and frequency-dependent polarizabilities of excited singlet states using coupled cluster response theory. *J. Chem. Phys.* **1998**, *109*, 9237.

(48) Helmich-Paris, B.; Hättig, C.; Van Wüllen, C. A CC2 Implementation of Induced Transitions between Singlet Ground and Triplet Excited States. *J. Chem. Theory Comput.* **2016**, *12*, 1892–1904.

(49) Hrsak, D.; Khah, A. M.; Christiansen, O.; Hättig, C. Polarizable Embedded RI-CC2 Method for Two-Photon Absorption Calculations. *J. Chem. Theory Comput.* **2015**, *11*, 3669–3678.

(50) Head-Gordon, M.; Pople, J. A.; Frisch, M. J. MP2 energy evaluation by direct methods. *Chem. Phys. Lett.* **1988**, *153*, 503–506.

(51) Feller, D. Application of systematic sequences of wave functions to the water dimer. **1992**, *96*, 6104–6114.

(52) Weigend, F.; Köhn, A.; Hättig, C. Efficient use of the correlation consistent basis sets in resolution of the identity MP2 calculations. *J. Chem. Phys.* **2002**, *116*, 3175–3183.

(53) Becke, A. D. Density-Functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988**, *38*, 3098–3100.

(54) TURBOMOLE V7.2 2017, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from `http://www.turbomole.com`.

(55) Furche, F.; Ahlrichs, R.; Hättig, C.; Klopper, W.; Sierka, M.; Weigend, F. *WIREs Computational Molecular Science* **2014**, *4*, 91–100.

(56) Dunning Jr., T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

(57) Mewes, J.-M.; You, Z.-Q.; Wormit, M.; Kriesche, T.; Herbert, J. M.; Dreuw, A. Experimental benchmark data and systematic evaluation of two *a posteriori*, polarizable-continuum corrections for vertical excitation energies in solution. *J. Phys. Chem. A* **2015**, *119*, 5446–5464.

(58) Aidas, K.; Angeli, C.; Bak, K. L.; Bakken, V.; Bast, R.; Boman, L.; Christiansen, O.; Cimiraglia, R.; Coriani, S.; Dahle, P.; Dalskov, E. K.; Ekström, U.; Enevoldsen, T.;

Eriksen, J. J.; Ettenhuber, P.; Fernández, B.; Ferrighi, L.; Fliegl, H.; Frediani, L.; Hald, K.; Halkier, A.; Hättig, C.; Heiberg, H.; Helgaker, T.; Hennum, A. C.; Hettema, H.; Hjertenæs, E.; Høst, S.; Høyvik, I.-M.; Iozzi, M. F.; Jansik, B.; Jensen, H. J. A.; Jonsson, D.; Jørgensen, P.; Kauczor, J.; Kirpekar, S.; Kjærgaard, T.; Klopper, W.; Knecht, S.; Kobayashi, R.; Koch, H.; Kongsted, J.; Krapp, A.; Kristensen, K.; Ligabue, A.; Lutnæs, O. B.; Melo, J. I.; Mikkelsen, K. V.; Myhre, R. H.; Neiss, C.; Nielsen, C. B.; Norman, P.; Olsen, J.; Olsen, J. M. H.; Osted, A.; Packer, M. J.; Pawlowski, F.; Pedersen, T. B.; Provasi, P. F.; Reine, S.; Rinkevicius, Z.; Ruden, T. A.; Ruud, K.; Rybkin, V. V.; Salek, P.; Samson, C. C. M.; de Merás, A. S.; Saue, T.; Sauer, S. P. A.; Schimmelpfennig, B.; Sneskov, K.; Steindal, A. H.; Sylvester-Hvid, K. O.; Taylor, P. R.; Teale, A. M.; Tellgren, E. I.; Tew, D. P.; Thorvaldsen, A. J.; Thøgersen, L.; Vahtras, O.; Watson, M. A.; Wilson, D. J. D.; Ziolkowski, M.; Ågren, H. The Dalton Quantum Chemistry Program System. *WIREs Comput. Mol. Sci.* **2014**, *4*, 269.

(59) 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.

(60) Picconi, D.; Avila, 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.

(61) Kim, N. J.; Kang, H.; Park, Y. D.; Kim, S. K. Dispersed fluorescence spectroscopy of jet-cooled adenine. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2802–2805.

(62) Ludwig, V.; da Costa, Z. M.; do Amaral, M. S.; Borin, A. C.; Canuto, S.; Serrano-Andrés, L. Photophysics and photostability of adenine in aqueous solution: A theoretical study. *Chemical Physics Letters* **2010**, *492*, 164 – 169.

(63) Clark, L. B.; Peschel, G. G.; Tinoco Jr, I. Vapor Spectra and Heats of Vaporization of Some Purine and Pyrimidine Bases. *J. Phys. Chem.* **1965**, *69*, 3615–3618.

(64) Voet, D.; Gratzer, W. B.; Cox, R. A.; Doty, P. Absorption spectra of nucleotides, polynucleotides, and nucleic acids in the far ultraviolet. *Biopolymers* **1963**, *1*, 193–208.

(65) Buckingham, A.; Stephens, P. Magnetic Optical Activity. *Ann. Rev. Phys. Chem.* **1966**, *17*, 399.

(66) Michl, J. Private Communication, 2018.

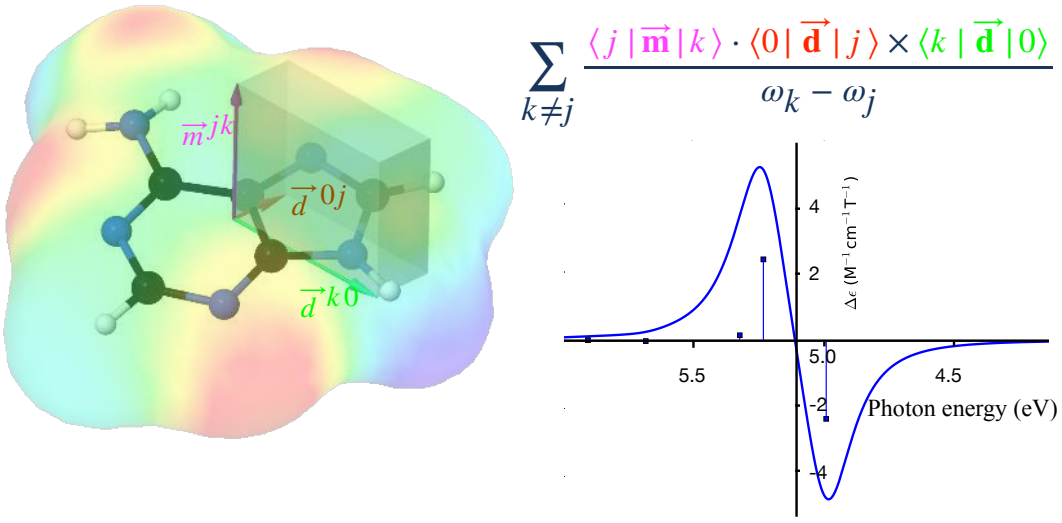

Figure 8: Table of contents