RUHR-UNIVERSITÄT BOCHUM
FAKULTÄT FÜR CHEMIE UND BIOCHEMIE
LEHRSTUHL FÜR THEORETISCHE CHEMIE

RUB

A computational study of the first electronic excited state of Berenil in solutions

Master Thesis

Alireza Marefat Khah

Contents

1 Introduction **3**

2 Theoretical Background **5**
2.1 Response theory and excited state properties 5
2.2 Second order methods 6
2.3 Continuum Solvation models 7
2.3.1 General approximations in continuum solvation models 7
2.3.2 Polarizable continuum models 9
2.3.3 Post Hartree-Fock continuum solvation models 12

3 Stability of Ground State Conformers **14**

4 Excitation properties **18**
4.1 Simulation of UV/Vis Spectrum 18
4.2 Natural transition orbital analysis 20

5 Berenil in the excited state **23**
5.1 The relaxed structure of the excited state 23
5.2 Energy levels at S_1 geometries 25
5.3 Intrinsic reaction coordinate analysis 26
5.4 Electronic charge density analysis 28

6 Geometrical flexibility of the excited state **29**
6.1 Flexibility of Triazenyl group 29
6.2 Flexibility of non-bonded distance of Amidine groups 33

7 Summary and conclusion **36**

8 Supporting Informations **38**
8.1 Excited state properties for 10 excitation levels 38
8.2 S_0-S_1 potential energies 42
8.3 Geometrical analysis 44
8.4 Force field fitted curves 45

1 Introduction

Berenil (Diminazene aceturate) is an aromatic diamidin (4,4'-(diazoamino) benzamidine) which solves in water. There are several studies about the selective binding of Berenil to different secondary structures of DNA and RNA which are shown that it can be considered as a DNA minor groove binder [1],[2].

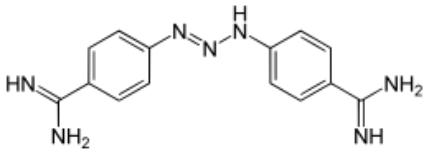


Figure 1: Chemical structure of Diminazene

Because of the high binding affinity to particular parts of DNA and pharmacological properties– Berenil is a bio-compatible molecule– and at the same time its physical properties (UV/Vis absorption and fluorescence emission), it becomes an important molecule for targeting RNA [3].

This study attempts to help physical chemists to analysing the UV/Vis, fluorescence and time-resolved fluorescence spectra of the molecule in solutions by means of *ab initio* quantum chemistry.

Particularly, we are (in both theoretical and physical chemistry groups) interested in studying the role of solvent in the chemistry of the lowest excited state of Berenil and the related spectroscopic effects. In the current project, as a first step, we have studied the lowest electronic excited state of Berenil in vacuum and continuum dielectric environment (as a model for solvent) by *ab initio* quantum chemistry methods.

To this end, various conformers of the electronic ground state of a particular double-protonated state of Berenil are studied with quantum chemistry methods. This study helps us to find the most stable conformers of the molecule.

In the next step, the UV/Vis spectra of berenil is simulated in solutions. The latter spectroscopy technique is used in various experimental studies of Berenil. Natural transition orbitals analysis is used to characterize the nature of electronic transitions.

By finding the local minima on the potential energy surface of the lowest electronically excited state and comparing their energies with the energies of electronic ground state, it is found that there is a possibility for the molecule to de-excited to the ground state without radiation. These de-excitation paths are analyzed by the intrinsic reaction coordinate analysis to show how different local minima are connected to each other. The geometrical and electronic structure of the molecule are studied in details for the involved structures.

In the last chapter, the flexibility of the minima of the excited state are investigated to consider the possibility of having a similar minima if the molecule has to be disordered in some special cases, for instance having explicit solvent molecules or chelate binding to DNA.

Supporting informations (SI) are provided in a separate chapter to present results in more details and to keep the main text concise.

2 Theoretical Background

Several theoretical approaches were introduced to simulate solution behaviours classically or/and quantum mechanically. As far as classical MD simulations focus on the dynamic nature of the solution, quantum chemistry methods like super molecular approaches beside continuum model of solution and QM/MM methods were designed to study quantum nature of the microscopic systems, energy levels and demanding properties, *e.g.* energies, geometries and electronic excitations and spectra of them which can not easily be done with classical molecular dynamic simulations, *etc.*

Therefore to explore the Berenil problem, we need a quantum based methods to study the electronic excited state. Response theory is a framework that let us to study the electronic ground state and excited state at the same level.

Beside that, we need a model for solvent which has been combined with proper quantum methods to study the effect of solvation. This method should be accurate and efficient enough to be applicable for our problem. The second part of the current chapter is alloted to give a briefing about the continuum solvation models.

2.1 Response theory and excited state properties

In the framework of response theory, the propagation of the electronic wavefunction $|0\rangle$ in time finds out by the time-dependent Schrödinger equation,

$$H|0\rangle = i\partial_t|0\rangle \tag{1}$$

where the Hermitian Hamiltonian H involves a time-independent part H_0 and a time-dependent perturbation $V(t)$, equation 2 [4].

$$H = H_0 + V(t) \tag{2}$$

Considering the wavefunction in its phase isolated form,

$$|0\rangle = e^{-iF(t)}|\tilde{0}\rangle \quad \text{with} \quad \langle\tilde{0}|\tilde{0}\rangle = 1, \tag{3}$$

the following equation can be obtained.

$$Q(t) = \dot{F}(t) = \langle\tilde{0}|(H - i\partial_t)|\tilde{0}\rangle \tag{4}$$

On the presence of oscillating perturbations,

$$V(t) = V(t+T) = \sum_{k=-N}^{N} e^{(-i\omega_k t)} V^{\omega_k}, \tag{5}$$

one can derive the variational principle for the time averaged $\{Q\}_T$ that must be stationary with the changes of wavefunction parameters 6.

$$\delta\{Q(t)\}_T = 0 \tag{6}$$

The time-averaged Helmann-Feynman theorem allows one to compute properties, *e.g.*, excitation energies, transition strengths and frequency-dependent properties from the derivatives of $\{Q\}_T$. This time-averaged quantity is called Quasi-energy.
From the variational principle, there is a certain possibility to construct a Lagrangian for a given approximate wave-function model [5],

$$L(\bar{\lambda}_k, \lambda_{k'}, t) = \left\langle H(t) - i\partial_t \right\rangle + \sum_k \bar{\lambda}_k f_k \big[H(t) - i\partial_t\big]\ (\lambda_{k'}, t), \tag{7}$$

that contains the expectation values (the first term) and some constrains f_k for the wave function parameters $\lambda_{k'}$ where the $\bar{\lambda}_k$ parameters are Lagrange multipliers. The response functions can be obtained by taking the derivatives of $L(\bar{\lambda}_k, \lambda_{k'}, t)$. Therefore by considering the fact that the poles of the response functions occur at the eigenvalues of the stability matrix of the Lagrangian, the excitation energies of non-variational methods– like coupled-cluster models– can be calculated as the eigenvalues of the electronic Jacobian matrix **A**. Equation 8 shows elements of the Jacobian matrix.

$$A_{ij} = \big(\frac{d^2 L}{d\bar{\lambda}_i d\lambda_j}\big) \tag{8}$$

2.2 Second order methods

The CC2 model, staying in a hierarchy between coupled-cluster singles (CCS) and coupled-cluster singles and doubles (CCSD) models, is defined by the following cluster equations, 9 and 10. In this model the Hamiltonian is partitioned to two parts, $\widehat{H} = \widehat{F} + \widehat{U}$, where $\widehat{F}$ is the Fock matrix, $\widehat{U}$ is the fluctuation potential and $T_i = \sum_{\mu_i} t_{\mu_i} \tau_{\mu_i}$ with $i = 1, 2$ [6].

$$\Omega_{\mu_1} = \langle \mu_1 | \widehat{H} + [\widehat{H}, T_2] | \mathrm{HF} \rangle = 0 \tag{9}$$

$$\Omega_{\mu_1} = \langle \mu_1 | \widehat{H} + [\widehat{F}, T_2] | \mathrm{HF} \rangle = 0 \tag{10}$$

For medium-sized organic molecules and for singlet excited states it was shown that the CC2 results are close to the CC3 results [7] which has a reliable accuracy. By using the resolution-of-the-identity (RI) approximation (RI-CC2 method) [8] and optimized auxiliary basis sets [9, 10] one can reproduce the CC2 results for a given basis set by a significantly lower computational cost.

$$A_{\mu_i \nu_j} = \begin{pmatrix} \langle \mu_1 | [\widehat{H}, \tau_{\nu_1}] + [[\widehat{H}, \tau_{\nu_1}], T_2] | \mathrm{HF} \rangle & \langle \mu_1 | [\widehat{H}, \tau_{\nu_2}] | \mathrm{HF} \rangle \\ \langle \mu_2 | [\widehat{H}, \tau_{\nu_1}] | \mathrm{HF} \rangle & \delta_{\mu\nu} \omega_{\mu_2} \end{pmatrix} \tag{11}$$

An element of Jacobian matrix of the CC2 model is presented in the former equation 11. The algebraic-diagrammatic construction through second-order ADC(2) method can be considered as an approximation to the CC2 method, albeit naturally derived from the polarization propagator approach [11]. To obtain ADC(2) excitation energies one has to replace the ground-state cluster amplitudes by MP2 parameters– the ground state energy of ADC(2) is identical with MP2– and keeps only the Hermitian part of the CC2 Jacobian matrix [12] which gives accurate results up to the second order.
The ADC(2) method in combination with the resolution-of-the-identity approximation is our method of choice to study the excited state properties of Berenil in vacuum as well as aqueous solutions via a continuum solvation model, COSMO.

2.3 Continuum Solvation models

The idea of using a dielectric continuum model to calculate the free-energy of solvation of ions goes back to Max Born was later extended by Kirkwood and Onsager in 1930s and is based on classical electrodynamics [13]. In the Kirkwood's formula,

$$\Delta G_{\mathrm{sol}}^{\mathrm{mol}} = -\frac{1}{2}\sum_{l=1}^{\infty} f_l(\varepsilon_{\mathrm{sol}})\frac{M_{l,\mathrm{mol}}^2}{R_{\mathrm{mol}}^{2l-1}} \tag{12}$$

with

$$f_l(\varepsilon) = \frac{\varepsilon - 1}{\varepsilon + x_l} \tag{13}$$

and

$$x_l = \frac{l-1}{l}, \tag{14}$$

M_l represents the charge, dipole, and higher-order multipole moments of rank l. R_{mol} is the effective radius of the spherical cavity where $f(\varepsilon_{\mathrm{sol}})$ and x_l are the scaling function of the dielectric constant, and the reaction field (RF) for the general dielectric polarization response, respectively. The factor $\frac{1}{2}$ arises since half of the interaction energy is required to polarize the continuum. The R_{mol} is a pure empirical parameter. The assumption of a spherical cavity is unrealistic for most molecules and all advanced continuum models offer a molecule-like cavity based on a segmentation scheme as a treatment [14]. Although details of segmentation method is important, they are not discussed here in details.

2.3.1 General approximations in continuum solvation models

Another problem of continuum models is the need to define the cavity by some distance criteria, e.g. R_{ion} for atomic ions. Different models use different criteria from 1.0 to 1.4 van der Waals radii for this propose [13]. The electronic wave-functions of molecules have a tail beyond the chosen region which does not contribute to polarize the dielectric continuum. Thus, the continuum sees the molecule as a partially positively charged species. The artificial polarization due to this effect is referred to as "outlying charge error". We

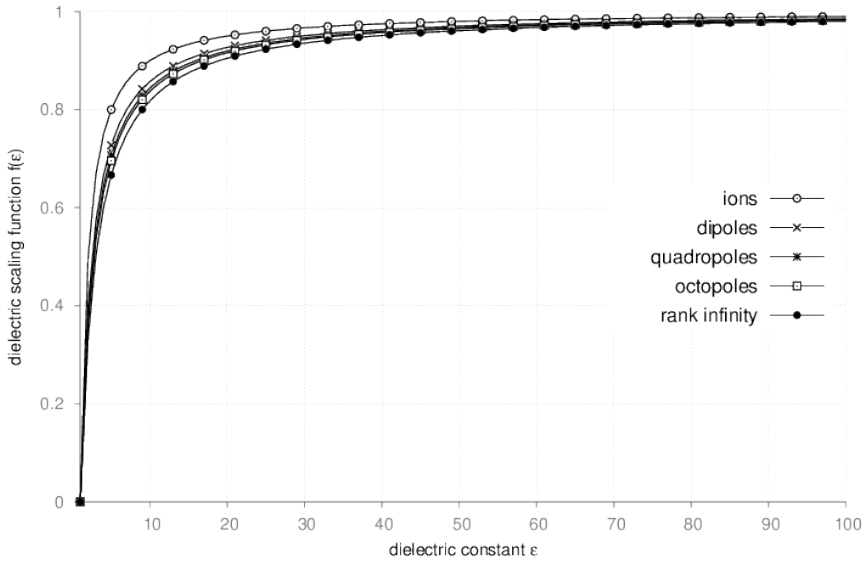


Figure 2: Dielectric scaling functions for different multipole orders [13].

will come back to this in the following sections.

In Onsager's method , equation 12, the multipole moments generated by the electrostatic field of the solvent are considered as fixed properties of solute molecules. Since the multipole moments are changed by the reaction field of the solvent, it is not possible to use the vacuum multipoles. Since they can not easily be measured they are sometimes uses as empirical parameters. The back-polarization not only causes the induced polarizations but can also perturb the geometrical configuration of solutes.
The effect of the back-polarization of the solute by the solvent can be treated in a self-consistent iterative manner which typically increases the solute-solvent interaction by $\sim$ 20% [13]. It is worth noting that in the calculation of solvatochromic effects, only the fast electronic polarizability is able to respond to a light-induced electronic excitation during the time scale of light absorption. In these cases it is required to replace the dielectric constant by the optical dielectric constant, ε_{opt}.

An inspection of the scaling function $f_l(\varepsilon_{\text{sol}})$ for dielectric constants in the regime of typical solvents reveals an interesting point about the continuum solvation models, figure 2 [13]. The figure shows that the scaling factors of higher-order multipole moments are very close to each other. Besides that, the dielectric constant of water at 20°C is 80.1: hence its scaling factor for every multipole moment
is higher than 0.975 and close to the scaling factor of an infinitely strong dielectic ($f(\varepsilon) = 1$), i.e. a conductor. Therefore one can imagine why a conductor-like solvation model might work for most of the solvents, in particular for water.

2.3.2 Polarizable continuum models

The so-called dielectric polarizable continuum model (D-PCM, previously called PCM) of solvation requires the solution of the classical electrostatic Poisson problem:

$$-\vec{\nabla} \times [\varepsilon(\vec{r})\vec{\nabla}V(\vec{r})] = \ 4\pi\rho_{\text{mol}}(\vec{r}), \tag{15}$$

where $\rho_{\text{mol}}(\vec{r})$ is the solute charge distribution and $\varepsilon_{\text{sol}}(\vec{r})$ is the general position-dependent permittivity. For a molecule embedded in a cavity of a homogeneous and isotropic solvent, $\varepsilon(\vec{r})$ can be approximated in a simple way [14]:

$$\varepsilon(\vec{r}) = \begin{cases} 1 & \vec{r} \in \Omega_{\text{cavity}} \\ \varepsilon_{\text{sol}} & \vec{r} \notin \Omega_{\text{cavity}} \end{cases} \tag{16}$$

where ε_{sol} is the dielectric constant of the solvent and Ω_{cavity} is the cavity surface.
Considering the solvent as an infinitely extended dielectric continuum, Equation 16, and the exact dielectric boundary conditions (EDBC), the electrostatic potential, $V(\vec{r})$, can be written as the sum of the solute potential and the surface potential of the cavity due to the apparent charge distribution $\sigma(\vec{s})$ on the boundary between solute and solvent,

$$V(\vec{r}) = V_{\text{mol}}(\vec{r}) + V_{\sigma}(\vec{r}) \tag{17}$$

where

$$V_{\sigma}(\vec{r}) = \int_{\Omega} \frac{\sigma(\vec{s})}{|\vec{r} - \vec{s}|} d^2s. \tag{18}$$

With neglecting the volume electrostatic potential (a consequence of defining an infinitely extended continuum) there is no need to define a cut-off for the continuum environment. An exact integration over a molecule-like shaped surface, $\vec{s}$, is computationally demanding and numerical solutions are based on discretization of the surface [13]. The electrostatic potential at segment s due to multipoles at the remaining segments ($s^{'}$) can be obtain upon a Taylor expansion of the electrostatic potential:

$$V_{\sigma}^{s} = \sum_{s^{'}} T_{0}^{ss^{'}} q^{s^{'}} \ - \sum_{s^{'}} T_{\gamma_1}^{ss^{'}} \mu_{\gamma_1}^{s^{'}} \ + \frac{1}{2} \sum_{s^{'}} T_{\gamma_1\gamma_2}^{ss^{'}} Q_{\gamma_1\gamma_2}^{s^{'}} - ... \tag{19}$$

where

$$T_{\gamma_1\gamma_2...\gamma_m}^{ss^{'}} = \ \nabla_{\gamma_1}^{s} \nabla_{\gamma_2}^{s} ... \nabla_{\gamma_m}^{s} \ \frac{1}{|\mathbf{r}^{s} - \mathbf{r}^{s^{'}}|} \quad \text{and} \quad \nabla_{\gamma_j}^{s} = \frac{\partial}{\partial r_{\gamma_j}^{s}}. \tag{20}$$

Here T_{γ_m} is an $m-$th order interaction tensor where the zeroth order tensor is defined as $T_{0}^{ss^{'}} = \frac{1}{|\mathbf{r}^{s} - \mathbf{r}^{s^{'}}|}$ [**?**]. The apparent surface charge $\sigma(\vec{s})$ can be calculated by an iterative self-consistent reaction field procedure. After convergence the solvation energy can be written as follows:

$$E_{\text{sol}}^{\text{mol}} = \ V_{\text{mol}} \mathbf{S}\boldsymbol{\sigma} = \ V_{\text{mol}} \mathbf{q}. \tag{21}$$

$\mathbf{q}$ is called the vector of screening charges that collects the expansion coefficients of the ASC in the piecewise constant basis (charges, dipoles, ...) and $\mathbf{S}$ is a diagonal matrix representing the segment areas. Considering the multipole expansion the interaction energies will be converged by incorporating octopoles while they are really needed to describe lone pairs [13].

However, coupling the integration of surface electrostatic potential, Eq. 18, with the self-consistent reaction field procedure is a hurdle– one has to calculate the derivatives of the electrostatic potential– to compute the interaction energies and geometry gradients with different quantum chemistry models via an efficient implementation [14].

The conductor-like screening model (COSMO) also attempts to solve the EDBC equation with an approximation on top of it: replacing the dielectric medium by a conductor medium, i.e. $\varepsilon = \infty$ – thus equation 22 – and afterwards scaling it with a scaling function, equation 23. It was shown in figure 2 that the latter is indeed a good approximation. Therefore, instead of solving the Poisson equation one has to solve the following simpler equation 23,

$$V(\vec{r}) = V_{\text{mol}}(\vec{r}) + V_{\sigma}(\vec{r}) = V_{\text{mol}} + \mathbf{Tq} = 0 \tag{22}$$

which results

$$\mathbf{Tq} = -f(\varepsilon)V_{\text{mol}}(\vec{s}). \tag{23}$$

V_{mol} presents the surface potential and $\mathbf{T}$ denotes the interaction matrix of the surface segments where

$$T_{i\neq j} = ||\mathbf{t}_i - \mathbf{t}_j||^{-1} \quad \text{and} \quad T_{ii} \cong 3.8\sqrt{S_i}. \tag{24}$$

The latter is an approximation that improves the convergence of the result with respect to number of segments – S_i is the area of segment i centered in $\mathbf{t}_i$. By scaling the solute potential with a dielectric scaling function (as it was shown the scaling function of all multipole moments are very close to each other and therefore in the COSMO model the scaling factor is equal to scaling factor of dipoles, $x_l = \frac{1}{2}$.), $f(\varepsilon)$, the next equation can be obtained, equation 25.

$$\mathbf{Tq} = -f(\varepsilon_{\text{sol}})V_{\text{mol}} \tag{25}$$

Since $\mathbf{T}$ is a symmetric positive-definite matrix, the solute-conductor interaction energy can be calculated easily as follows:

$$E_{\text{sol}}^{\text{mol}} = V_{\text{mol}}\mathbf{q} = \ -f(\varepsilon)V_{\text{mol}}\mathbf{T}^{-1}V_{\text{mol}} \tag{26}$$

Subsequently, one can introduce an operator $\mathbf{B}$ which generates the potential $\boldsymbol{V}_{\text{mol}}$ from the charge-density – this tool is available in all QM codes –, $\mathbf{Q}_{\text{mol}}$, thus we have

$$E_{\mathrm{sol}}^{\mathrm{mol}} = \ f(\varepsilon)\mathbf{Q}^{mol}\ \mathbf{B}^{t}\mathbf{T}^{-1}\mathbf{B}\mathbf{Q}^{mol} \equiv \ \mathbf{Q}^{mol}\mathbf{D}\mathbf{Q}^{mol}. \tag{27}$$

The last expression (Green's function of the dielectric continuum) is very similar to the expression for Coulomb interaction energy which presents in any QM code, equation 28.

$$E_{\mathrm{Coloumb}}^{\mathrm{mol}} = \frac{1}{2}\mathbf{Q}^{mol}\mathbf{C}\mathbf{Q}^{mol} \tag{28}$$

Therefore the dielectric continuum can be represented as an additional charge-charge interaction, which is just mediated by the dielectric (conductor-like) continuum. This expression also helps to implement the continuum solvation models with linking it to the routines that are already available in QM codes. In addition, the analytical expression for gradients of the dielectric continuum green's function is also relatively simple and agreeable with vacuum one [13].

One of the advantages of COSMO is that the outlying charge error is an order of magnitude smaller than for any DPCM model [15]. The reason for this is that the electrostatic potential is much less sensitive to the outlying charge than the electrostatic field (needed in the DPCM framework), while COSMO only uses the potential. In addition, a double layer scheme has been introduced by Klamt for the outlying charge correction (OCC) which takes into account the local outlying charges, while the DPCM and IEFPCM can only be corrected by scaling the results [16].

COSMO is compatible with the statistical thermodynamics of interacting surfaces (COSMO-RS approach) and can be used to approximate the chemical potentials and is therefore used a lot in chemical engineering [17].

In some QM codes, the COSMO approach has been implemented with a bit different scaling function ($x_l = 0$), which is called the conductor polarizable continuum model (C-PCM) [18]. To demonstrate the equivalence of both methods, one may consider a solvent with $\varepsilon = 50$. Equation 27 shows that the interaction energy has a linear relationship with the scaling function. For the chosen dielectric constant, the scaling factor of COSMO is 0.97 while the C-PCM scaling factor is 0.98. Although different implementations use different segmentation methods which probably makes more significant differences between them, specially for larger systems and gradient calculations. Nowadays, COSMO and/or C-PCM are the mostly applied continuum models. The TURBOMOLE [19] implementation uses the NEW-CAV scheme which is a modified type of the GEOPOL algorithm [13]. The alternative approach is the continuous surface charge (CSC) COSMO approach which has been proposed by York and Karplus [20] to rewrite the surface charge in terms of an expansion of surface element basis functions, which are typically simple spherical Gaussian functions. Furthermore applying the COSMO approach to large systems requires fair consideration to define the boundary condition in a proper way [21, 22, 23].

Further developments include combinations of exact dielectric and conductor-like models where the difference between scaled-COSMO and the exact dielectric is considered through a remaining surface polarization [13], i.e.,

$$\boldsymbol{\sigma}'_d = \boldsymbol{\sigma}(\varepsilon) - f(\varepsilon)\boldsymbol{\sigma}(\infty). \tag{29}$$

Other types of approaches were introduced to model solvents with low dielectric constants, particularly the integral equation formalism, IEFPCM [24, 25, 26] approach that considers the general formulation of continuum solvation models (CSMs). The SS(V)PE approach [27] takes into account the volume polarization and might be considered as the most accurate method [28]. However the gain of accuracy versus the additional computational cost makes them less interesting for applications.

In addition, such continuum solvation models can not describe dispersion and other non-classical interactions and essentially can not see the difference between two solvents which have the same or close dielectric constants, for instance Benzene and Cyclohexane. To describe these effects one has to go beyond the standard continuum solvation models.

2.3.3 Post Hartree-Fock continuum solvation models

Probably the most important remaining challenges for the CSMs are treating time-dependent and excited states problems or to deal with the electronic correlation problem beyond the DFT accuracy [29].

Adding the interaction with the apparent surface charges to the solute Hamiltonian causes the need to solve a non-linear Schrödinger equation, Eq. 30.

$$\left\{\widehat{H} + \langle \Psi_{\text{mol}}|\widehat{\mathbf{Q}}|\ \Psi_{\text{mol}}\rangle \cdot\ \widehat{\mathbf{V}}\right\}|\Psi_{\text{mol}}\rangle = E_{\text{mol}}|\Psi_{\text{mol}}\rangle \tag{30}$$

There are two perturbative approaches including PTE and PTD which neglect the coupling between solvation and electronic correlation. In the perturbation to density approach (PTD), the post-HF problem is solved in vacuum and then the solvation energy is computed directly for the solute density. The perturbation to energy approach (PTE) solves the SCF/CSM problem and then uses the orbitals to compute the amplitudes in the canonical way [29].

The perturbation to energy and density scheme (PTED) attempts to combine the PTE and the PTD approaches in an iterative manner. The PTE relaxed density is used to update the CSM operator. The CSM operator is applied to update the apparent surface charges that are utilized to solve the PTD equations and repeat the cycles until achieving the self-consistency [30].

In order to study the solvent effect on electronic excited states, there are two proposed approaches, Linear response CSM and the state-specific approaches. The LR-CSM formulation describes the solvent effect (non equilibrium) properly and has no capability to be used to study the excited state relaxations or the charge transfer states [31, 32]. While the state specific approach (SS) is designed to study the relaxation of reaction field with respect to electronically excitation of the solute in a PTED-like way. The excited state relaxed density is applied to compute a new reaction field which is added to the Fock operator as an external field to calculate the new orbitals and subsequent new relaxed density iteratively until self-consistency is achieved.
Lunkenheimer and Köhn reported a PTED implementation of COSMO-ADC(2) which uses the MP2-SCRF parameters of an electronic state (either can be a ground state or an excited state) to evaluate the excitation energies of other states by the ADC(2) model [33]. This approach is our method of choice to study the lowest excited state of Berenil in solutions.

3 Stability of Ground State Conformers

Berenil, the double protonated form of Diminazene, has 5 rotating bands (without counting rotation of Amino groups) and 32 spatial conformers in general. Before describing the computational details, it is convenient to introduce some contractions. Figure 3 shows atom numbers on a DFT-optimized conformer of Berenil that are used hereafter. In most of the geometry optimizations– details are discussing below– , the Triazenyl and Aryl rings stay in-plane while Amidine groups are rotated by approximately 30° around $N_{34}-C_{31}-C_{24}-C_{23}$ and $N_8-C_7-C_4-C_3$ dihedral angles.

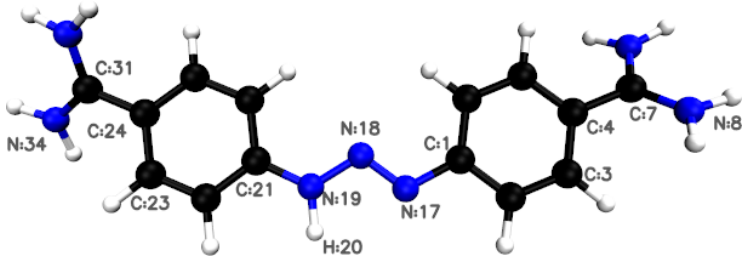


Figure 3: A typical 3-dimensional conformer of Berenil. Atom numbers are chosen by contraction.

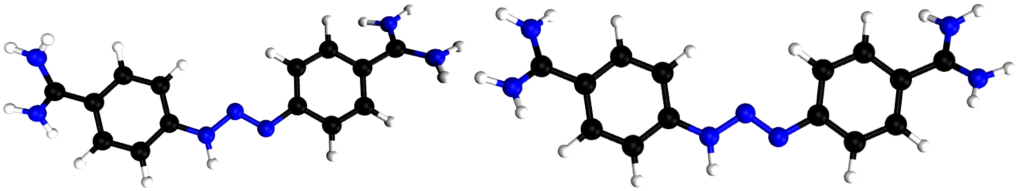
Consequently, the *E-azo*-isomer of Berenil comes in 8 conformers which occur in 4 pairs of enantiomers. Four conformers of Berenil, each from an enantiomeric pair, are considered in this study – figures 4 and 5 show optimized geometries with their respective contracted names.
The **trans-111** and **trans-112** are two spacial conformers of *E-azo,s-trans*$_{N_{18}-N_{19}}$-isomer. In the former, trans-111, Amidine groups are rotated in opposite direction while in the latter, Amidine groups are rotated in the same direction.
Considering spacial *E-azo,s-cis*$_{N_{18}-N_{19}}$-isomers, figure 5, the Amidine groups are rotated in the likewise and contrariwise directions for the **trans-212** and **trans-211**, respectively.
The *(Z-azo)*-isomer has only two minima (Fig. 6) based on our calculations, **cis-111** and **cis-122** which do not have in-plane Aryl rings.

in vacuo geometry optimizations are performed within the framework of density functional theory (DFT) using the B3-LYP-D3(BJ) exchange-correlation functional [34, 35, 36] and the Def2-TZVP basis set [37] (doubly polarized at C/N/O, singly polarized at H) without any geometrical constrain. All calculations were done with version 7.0 of TURBOMOLE program package [19]. [1] The relative energies, Boltzmann factors and relative probabilities of the six mentioned isomers are listed in table 1.

To evaluate the DFT results, energies of the DFT-optimized geometries are calculated at MP2 level with Def2-TZVP basis set. These calculations are done by employing the

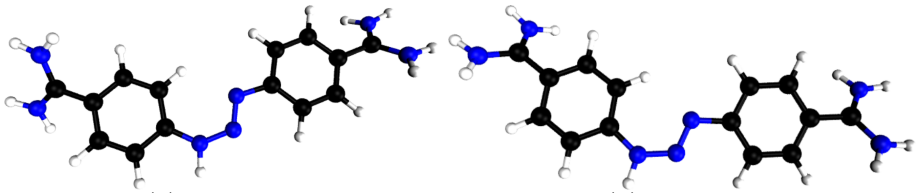
[1]The DFT optimized geometries of the 6 mentioned conformers are given to me by Prof. Hättig.

resolution-of-identity approximation– corresponding auxiliary basis set [38] are used– and the frozen-core approximation is applied on the optimized SCF orbitals.



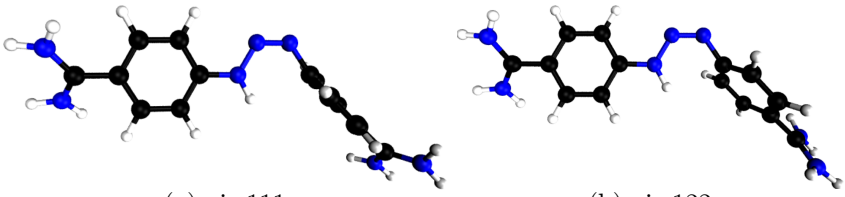
(a) trans-111 (b) trans-112

Figure 4: *E-azo,s-trans* conformers



(a) trans-211 (b) trans-212

Figure 5: *E-azo,s-cis* conformers



(a) cis-111 (b) cis-122

Figure 6: *Z-azo,s-trans* conformers

A comparison between MP2 and DFT energies shows consistency between the results which means that molecule is not a highly correlated system. Therefore the DFT-optimised geometries seem reliable enough (for our propose) and to have qualitatively accurate results, second order treatment of correlation energy should be sufficient *i.e.,* MP2, CC2, ADC(2) and etc. In addition, it was confirmed that the cis-111 and cis-122 conformers have much higher energy that could not be stabilized by solvent perturbation or using more costly methods or bigger basis sets. Therefore these conformers would not be considered in the further studies. Furthermore, having the similar Boltzmann's weights for trans-111 and trans-112 show that the relative orientation of Amidine groups is not a stabilizing geometrical factor.

With a set of single point calculation with the SVP basis set (which has lower number of basis functions) at MP2 level, the basis set dependency of the relative energies of conformers is investigated. The relative stability of results with respect to changing basis sets let us use the smaller basis set, SVP, for further investigations where having a qualitative picture is sufficient. Based on vacuum calculations, we conclude that in vacuum and thermodynamically equilibrated condition (T=298K), the most stable and dominantly occupied conformers of Berenil are *E-azo,s-trans* conformers with the same weights (W), trans-111 and trans-112. Furthermore, having the similar Boltzmann's weights for trans-111 and trans-112 show that the relative orientation of Amidine groups is not a stabilizing geometrical factor.

However these conclusions might be changed if we consider the role of solvent, *i.e.* the relative stability of trans-211 and trans-212 and/or the relative probability weights of trans-111 and trans-112 might be changed. In order to taking into account the solvent role, the conductor-like screening model (COSMO) is used to model aqueous solution ($\varepsilon = 50$, rsolv=1.3333). The COSMO-MP2 results listed in table 1 shows that the trans-111 and trans-112 conformers stay the most stable conformers in aqueous solutions with the same probability. And the relative orientation of the $-\mathrm{C(NH_2)_2}$ plays no role for the ground state energies even with inclusion of implicit solvent interactions. Other conformers also show the mentioned trend consistently at different levels of theory.

Table 1: The relative energies in kcal $\cdot$ mol^{-1}, Boltzmann factors, B_f, and relative probabilities (W) of different conformers with different methods and basis sets. Vacuum results are shown by "VAC" and the COSMO results are shown by "COSMO".

conformer	method	basis set	ΔE	B_f(298K)	W
cis-111	B3-LYP(VAC)	def2-TZVP	14.3	0.00	0.00
cis-122	B3-LYP(VAC)	def2-TZVP	14.3	0.00	0.00
trans-111	B3-LYP(VAC)	def2-TZVP	0.0	1.00	0.50
trans-112	B3-LYP(VAC)	def2-TZVP	0.0	1.00	0.50
trans-211	B3-LYP(VAC)	def2-TZVP	4.4	0.00	0.00
trans-212	B3-LYP(VAC)	def2-TZVP	4.4	0.00	0.00
cis-111	MP2-VAC	def2-TZVP	14.4	0.00	0.00
cis-122	MP2-VAC	def2-TZVP	14.4	0.00	0.00
trans-111	MP2-VAC	def2-TZVP	0.0	1.00	0.50
trans-112	MP2-VAC	def2-TZVP	0.0	0.99	0.50
trans-211	MP2-VAC	def2-TZVP	4.0	0.00	0.00
trans-212	MP2-VAC	def2-TZVP	4.0	0.00	0.00
cis-111	MP2-VAC	SVP	15.6	0.00	0.00
cis-122	MP2-VAC	SVP	15.6	0.00	0.00
trans-111	MP2-VAC	SVP	0.0	1.00	0.50
trans-112	MP2-VAC	SVP	0.0	1.00	0.50
trans-211	MP2-VAC	SVP	4.0	0.00	0.00
trans-212	MP2-VAC	SVP	4.0	0.00	0.00
trans-111	COSMO-MP2	SVP	0.0	1.00	0.50
trans-112	COSMO-MP2	SVP	0.0	0.99	0.50
trans-211	COSMO-MP2	SVP	3.3	0.00	0.00
trans-212	COSMO-MP2	SVP	3.2	0.00	0.00

4 Excitation properties

4.1 Simulation of UV/Vis Spectrum

The energy and one-photon absorption oscillator strength (OS) of electronically excited states (EESs) of Berenil are calculated at COSMO-ADC(2) level with different basis sets. The SVP basis set is used to calculate the vertical excitation energies (VEE) of the first 10 EESs for all trans- conformers (The results for the EESs with VEE up to the 5 eV are shown in table 2) – see the supporting informations for the results of 10 EESs. These calculations show that for all trans- conformers there are only two excitation levels which have lower 4 eV vertical excitation energies.

The COSMO and Vacuum results are presented in table 2. Interacting with continuum causes the decreasing of the lowest excitation energy by 0.11 eV. In general, the COSMO and Vacuum results are close and thus the vacuum excited state geometry optimisations should be close to COSMO ones as well. We will use the latter in the next chapter to find the S_1 geometries.

Getting higher excitation energies because of basis set incompleteness is a common problem. To control the efficiency of the SVP basis set, the same properties are calculated for electronically excited states which have the VEE smaller than 4 eV with aug-cc-pVDZ basis set for all conformers where the EESs of the trans-111 conformer are calculated up to the 5 eV. Although the VEE of the lowest electronically excited state are different (approximately by 0.3 eV) with respect to basis sets, SVP and aug-cc-pVDZ, the SVP results are qualitatively descriptive for the excited state properties. Therefore we will use both mentioned basis sets to take advantages of performance and accuracy in the following.

The results in table 2 reveal that the first electronic excited state has the highest oscillator strength among those that appear below 5ev. The table shows that neither the VEEs nor OSs are changed with respect to relative orientation of $-C(NH_2)_2$ groups. This fact is illustrated with simulation of UV/Vis spectra of four conformers at COSMO-ADC(2)/SVP level, figure 7.

Therefore, because of the much higher probability of having trans-111 and trans-112 at thermodynamically equilibrated condition (298K) – these conformers have the same probability, see table 1 – and the no-role playing by relative orientation of Amidine functional groups, it is sufficient to focus on one of those most stable conformers to study the excited state properties of Berenil. The latter is demonstrated with comparing the calculated spectrum of trans-111 at COSMO-ADC(2)/aug-cc-pVDZ level of theory and the experimental spectrum which shows a very good agreement, at least around the maximum of the peak.

Up to know, computational methods were compared only relatively. In order to compare the experimental and computational spectra, the first step is to reproduce a spectrum

Table 2: Electronic excited state properties of 4 trans- conformers calculated with COSMO-ADC(2) method with SVP and aug-cc-pVDZ basis sets.

Environment		Vacuum		COSMO			
Basis set		SVP		SVP		aug-cc-pVDZ	
conformer	EES	VEE	OS	VEE	OS	VEE	OS
trans-111	1	3.61	1.26	3.52	1.28	3.32	1.17
	2	3.91	0.00	3.87	0.00	3.84	0.00
	3	4.68	0.01	4.66	0.01	4.53	0.01
	4	4.78	0.02	4.77	0.02	4.58	0.08
	5	5.01	0.10	4.92	0.12	4.74	0.09
trans-112	1	3.54	1.30	3.52	1.27	3.32	1.17
	2	3.73	0.01	3.88	0.00	3.84	0.00
	3	4.61	0.00	4.66	0.01		
	4	4.65	0.09	4.77	0.02		
	5	4.75	0.02	4.92	0.11		
trans-211	1	3.60	0.29	3.64	1.14	3.49	1.17
	2	3.65	0.92	3.74	0.08	3.69	0.01
	3	4.64	0.02	4.71	0.02		
	4	4.73	0.04	4.78	0.00		
	5	4.77	0.05	5.00	0.09		
trans-212	1	3.59	0.39	3.64	1.04	3.49	1.15
	2	3.68	0.81	3.76	0.17	3.70	0.04
	3	4.65	0.02	4.71	0.02		
	4	4.73	0.04	4.79	0.00		
	5	4.77	0.05	5.01	0.09		

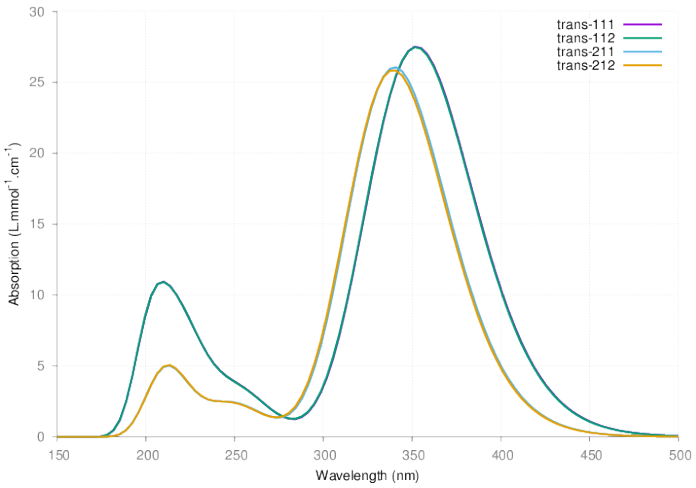


Figure 7: The calculated UV/Vis spectra of Berenil at COSMO-ADC(2)/SVP level of theory. For each conformer 8 excitation levels are included and the peaks are broaden with a Gaussian function with 0.3 eV standard deviation.

which being comparable with experimental results, *i.e.* the same physical quantities and units should be used and refuse to introduce arbitrary units and scaling factors as much as possible.

In order to simulate the UV/Vis spectrum, each peak is broaden by a Gaussian function with 0.3 eV standard deviation. Units and related physical quantities are calculated by following the reference [39] and a technical note from the Gaussian Inc. to convert oscillator strengths to extinction coefficients. [2] Afterwards a script to convert the TURBOMOLE output to a gnuplot script for plotting UV/Vis spectra is written.

A preliminary comparison between the experimental spectrum of Berenil and the simulated spectrum is done. An experimental spectrum of Berenil was scaled with extinction coefficient reported in reference [40] and the error bar of the same reference is shown in figure .
[3]

4.2 Natural transition orbital analysis

Calculation of excitation energies with a correlated method, usually ends up with a set of coefficients that describe the contribution of each pair of occupied and virtual orbitals. If there would be no dominant contribution (which is often the case), it is difficult to say something about the "nature" of the excitation. Natural transition orbitals (NTO) analysis

[2] A pre-factor is missed, in the recipe of Gaussian Inc.

[3] I am not an experimental chemist. Therefore it would be appreciated if experts reproduce the UV/Vis spectrum of reference [40] and help us to compare our theoretical spectrum with the standard one.

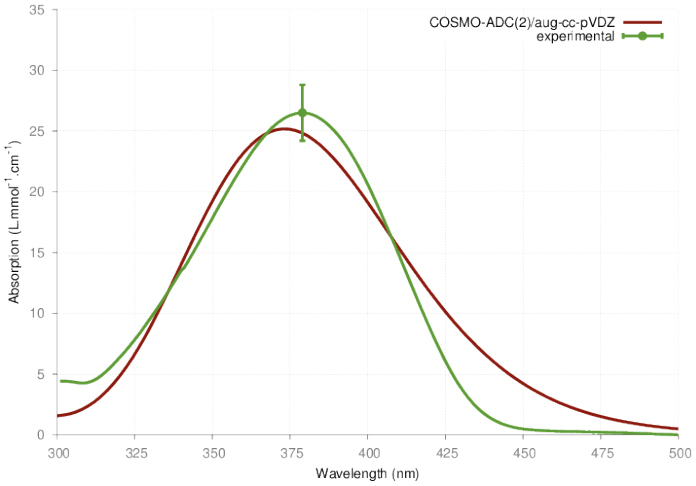


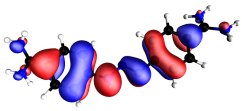
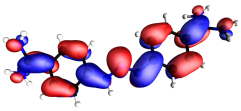
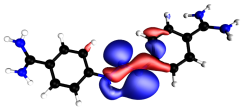
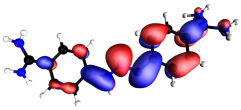
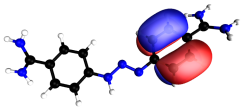
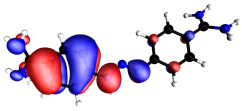
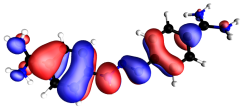
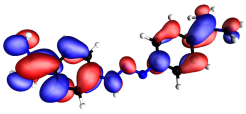
Figure 8: The experimental UV/Vis spectrum of Berenil in NaCl 250mM solution vs. Simulated spectrum of trans-111 conformer at COSMO-ADC(2)/aug-cc-pVDZ level of theory.

[41] intend to reproduce the particle-hole or occupied-to-virtual interpretation by giving a compact orbital representation for the electronic transition density matrix.
The NTO analysis is pursued to characterize the excited states characters as well as the changing of bonding pattern of electronically excited states, table 3. The lowest in energy electronically excitation occurs from an occupied NTO which has bonding character for the *azo-* , N=N, bond – no vertical node between these two Nitrogen atoms. In contrast, the virtual NTO has anti-bonding character for the N=N bond (vertical node). Consequently the π-bond between these N atoms is weakened or broken in this particular excited state.

In addition, the C−N bonds between the −NH−N=N− and the $C_6H_4(C(NH_2)_2)$ groups show a bonding character in the excited state. The C−C bonds between Benzene rings and Amidine groups, have bonding character in both occupied and virtual natural transition orbitals.

The influence of the mentioned points to the geometrical changes and electronic charge density distribution of the lowest excited state will be discussed in the next chapter. In general, one can characterize the lowest electronically excitation as a $\pi \rightarrow \pi^*$, and the second one as a $n \rightarrow \pi^*$ transitions.
The same NTO analysis is done for the other trans- conformers (see the SI).

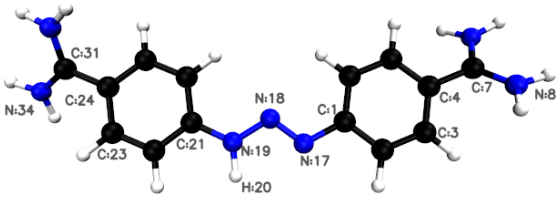
Table 3: Natural transition orbitals of trans-111, calculated at COSMO-ADC(2)/SVP level. The isosurfaces chosen to visualize orbitals are on +0.025 (blue) and -0.025 (red). Contributions of each pair is shown in percentage.

conformer	EEL	OCC	VIR	Contrib.
trans-111	1			96.2
trans-111	2			99.7
trans-111	3			70.6
trans-111	4			62.3
trans-111	5			86.2

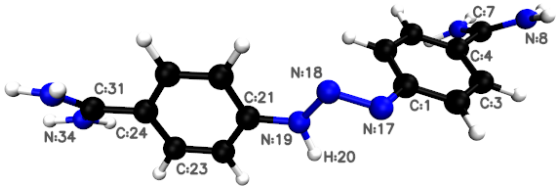
5 Berenil in the excited state

5.1 The relaxed structure of the excited state

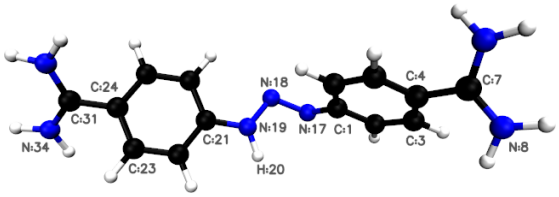
Geometry optimisations for the lowest EES of all trans- conformers are performed using the ground state optimised geometry as the initial geometry (Frank-Condon point of the EES) for at ADC(2) level with SVP basis set. [4] Afterwards, a set of single point calculations to evaluate energies at COSMO-ADC(2) level with SVP and aug-cc-pVDZ basis sets are pursued.



(a) S_0 structure of trans-111



(b) S_1 structure of trans-111



(c) R-S_1 structure of trans-111

Figure 9: Atom numbers are the same as atom numbers of table 12 and the contractions that have been introduced in chapter 3.

The main trends of the EES geometric relaxation for all trans- conformers are similar. The molecule rotates around the N=N *azo*- bond for $\approx \pm 90°$ to dispose itself of the anti-bonding unfavorable interaction of the N=N *azo*-bond– the anti-bonding character of the *azo*-bond is shown by NTO analysis. Hence two minima (figure 9) are found for each conformer, one for rotating clock-wise and one for rotating anti-clock-wise around the N=N

[4] a module to calculate COSMO-ADC(2) gradients is not implemented yet; so that was not possible to use this method for excited state geometry optimisation.

azo- bond. These conformers are named S_1 and R-S_1 (rotated S_1) for the lead conformer, trans-111, respectively (see table 12). Table 12 shows some selected internal coordinates which are changed (or unchanged), notably from Frank-Condon point (S_0 geometry) to the Semi-transition state ($TS^*_{S_1}$) and the Transition state (TS_{S_1}) of the first electronically excited level and the S_1 minima (S_1 and R-S_1)– to find a table which includes other selected internal coordinates, please see the SI.

Table 4: Geometrical details of different structures of Berenil trans-111 conformer. Distances are given in pm. Triangular and dihedral angles are given in degree.

Internal coord.	S_0	TS^*	TS_{S_1}	S_1	R-S_1
Selected torsion angles					
$N_{34}-C_{31}-C_{24}-C_{23}$	327	323	322	320	322
$N_8-C_7-C_4-C_3$	325	329	329	316	315
$C_2-C_1-N_{17}-N_{18}$	179	180	180	180	179
$C_{22}-C_{21}-N_{19}-N_{18}$	180	181	180	191	171
$C_{22}-C_{21}-N_{19}-H_{20}$	0	1	1	11	12
$C_{21}-N_{19}-N_{18}=N_{17}$	180	179	180	174	187
$C_1-N_{17}=N_{18}-N_{19}$	180	179	179	92	-92
Selected bond lengths					
C_1-N_{17}	140	135	135	137	137
$N_{17}=N_{18}$	125	134	134	138	138
$N_{18}-N_{19}$	133	137	137	135	135
$N_{19}-H_{20}$	102	103	103	103	103
$N_{19}-C_{21}$	138	135	135	138	138
Selected non-bonded distances					
C_7-C_{31}	1248	1251	1266	1185	1185
C_1-C_{21}	463	461	462	431	431
Selected triangular angles					
$C_1-N_{17}-N_{18}$	116	114	115	115	115
$N_{17}-N_{18}-N_{19}$	112	105	104	110	110
$N_{18}-N_{19}-C_{21}$	123	123	124	120	120

Neither the length of the C−N bonds nor the dihedral angles of Amidine groups change during the geometrical relaxation. The NTO analysis also does not show changes in bond-

ing character of these groups.
Except the N=N *azo-* bond, other rotating bonds roughly stay in-plane which is consistent with the fact that the N−Ar bonds find a binding character, based on NTO analysis. Considering the bond lengths, the *azo-* bond becomes longer while the C_1−N_{17} bond gets shorter. The mentioned non-bonded distances are shown that the terminal Amidine groups get closer to each other in the relaxed excited state geometry mainly due to rotation of the N_{17}=N_{18} bond. These changes could be important which are studied closely in the next chapter.
In conclusion, there are two characteristic changes associated by the geometrical relaxation, rotation by $\approx \pm 90^\circ$ of the C_1−N_{17}=N_{18}−N_{19} dihedral angle and increasing of the N_{17}=N_{18} bond length.

5.2 Energy levels at S_1 geometries

Interestingly, The lowest excitation energy of the S_1 and the R-S_1 geometries are negative (figure 10– please see the SI for other conformers). Thus an intersection between the ground and the excited states might be happened during the geometrical relaxations in the excited state. The latter could explain the experimental observation about the non-radiative relaxation and the weak fluorescence intensity of Berenil in aqueous solutions. S_1 state would not be populated due to the intersection and therefore the fluorescence intensity is low. Though because of the single-reference nature of ADC(2) method and the diabaticity of the intersection phenomenon, giving a quantitative description about the transition process needs further considerations which is beyond the scope of the current study.

Table 5 shows the gaps between electronic states at S_0 and S_1 geometries. For trans-111 and trans-112 conformers, the gaps between the lowest EES and the electronic ground state is 0.24 eV at S_1 geometry. It clears that the relative orientation of Amidine groups has no role to the energy gap between the electronic states at S_1 geometry as well.

Table 5: Relative energies of electronic ground state, Frank-Condon point and the energies of electronic ground state and lowest excited state of the S_1 structure at COSMO-ADC(2)/aug-cc-pVDZ level for trans- conformers. Numbers are given in eV.

Conformer	S_0 structure		S_1 structure	
	S_0	S_1	S_0	S_0
trans-111	0.00	3.32	2.02	1.78
trans-112	0.00	3.32	2.02	1.78
trans-211	0.00	3.48	2.35	1.85
trans-212	0.00	3.49	2.36	1.85

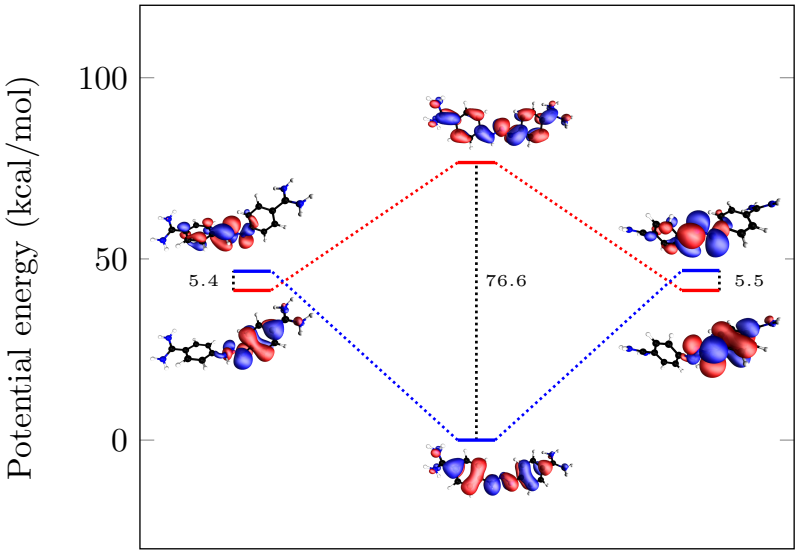


Figure 10: The schematic potential energy surface shows the electronic de-excitation process of trans-111 at the COSMO-ADC(2)/aug-cc-pVDZ level.

5.3 Intrinsic reaction coordinate analysis

Although two minima on the lowest EES (S_1 and R-S_1) are characterized, that is still questionable if there is any other path on the excited state or the ground state potential energy surfaces to connect the Frank-Condon point to other minima on the excited or the ground state. To investigate the connectivity of the minima and for studying the relaxation path in more details, intrinsic reaction coordinate (IRC) analysis are carried out for the trans-111 conformer.

The idea of the IRC analysis is to calculate nuclei gradients of a given geometry, calculating acceleration and velocities via the Newton's second law and let the nuclei to move for a given δt (= 2 ps in this case) and again calculating the gradients for the new coordinates, accelerations and velocities until reaching a minimum where the gradients are zero with a given threshold. In the IRC analysis, the kinetic energy (velocities) of the previous step is damped to zero to keep the molecule on the corresponding potential energy surface. The IRC analysis where done at ADC(2)/SVP level for the lowest EES and consistently at MP2/SVP level for the electronic ground state calculations.

Figure 11 presents the results of the IRC analysis which includes the results of more than 8500 single point calculations of energy and gradients. The energy of each single point calculation versus the N_{17}=N_{18} bond distance and the C_1−N_{17}=N_{18}−N_{19} dihedral angle (These are the characteristic measures of the de-excitation path) of each point are shown in the figure where the density of points is a descriptor for the size of gradients along the

chosen dimensions.

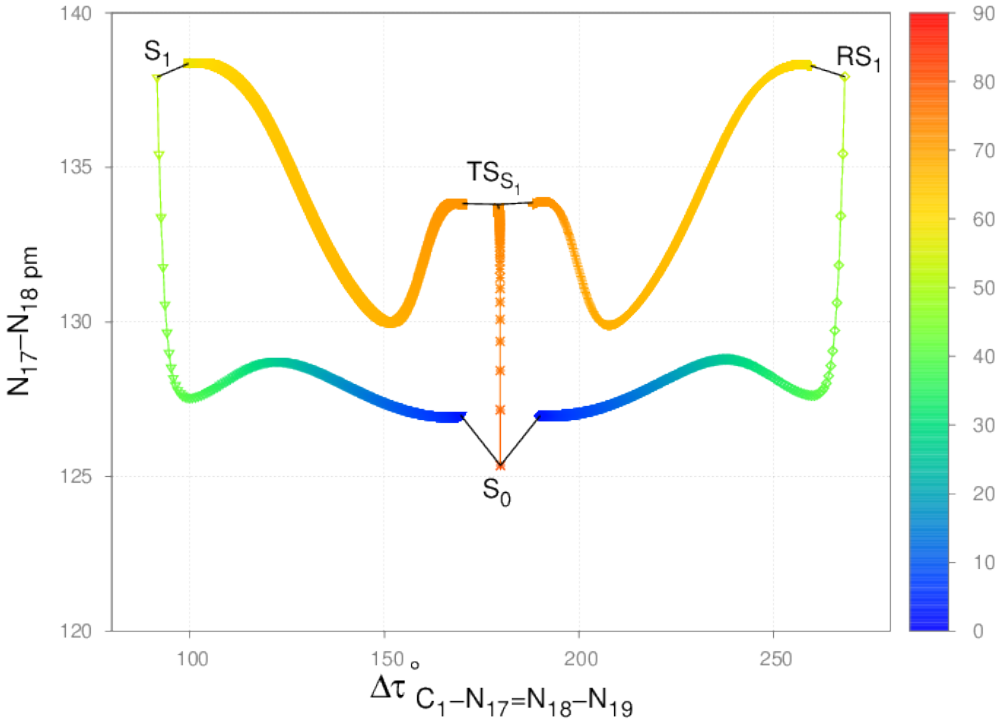


Figure 11: Intrinsic reaction coordinate analysis along the de-excitation path, from Frank-Condon point to the transition state of the lowest EES to the S_1 and RS_1 points at ADC(2)/SVP level and from S_1 and R-S_1 geometries to the ground state optimised geometry, S_0, for the ground electronic state at MP2/SVP level. Relative energies are shown by colors in kcal $\cdot$ mol^{-1}. Black lines connect the last structure of a part of the IRC calculations to the initial structure of the next part.

The IRC calculations can be interpreted as following:

1) After vertical excitation at S_0 geometry, as a first phase the N_{17}=N_{18} bond distance increases from 125 pm to 134 pm with a negligible changes of dihedral angle. Then the IRC calculations converged to a point which is named TS*. At this point, gradients become too small that pursuing the IRC calculations would not be feasible.

Therefore the IRC analysis is continued by a transition state optimisation on the lowest electronic excited state with choosing the TS* geometry as the initial geometry. The geometry optimisation converged fast to a point which is confirmed to be a transition state (TS_{S_1}) with calculating the Hessians and getting only one imaginary frequency at ADC(2)/SVP level.

The geometrical indices of the TS* and TS_{S_1} are listed in table 12. A short look at the table shows that these geometries are indeed very similar thus roughly speaking, we conclude that the first phase of the de-excitation path is to reach a transition state at the lowest EES where the potential energy surface becomes bifurcated to two paths that lead to S_1 and R-S_1 minima.

2) By perturbing the transition state geometry along the normal mode which gives imaginary frequency in both directions, two sets of IRC analysis are run at ADC(2)/SVP level to control if the TS_1 is connected to S_1 and R-S_1 geometries. With approximately 1900 single point calculations of energy and gradients in each direction, the molecule gets close to the characterized minima, though it was not possible to converge exactly to the minima because of very low gradients norm around the minima.

We have checked if geometry optimisations from the last structures of the second phase of IRC calculations converge to the mentioned minima at the ADC(2)/SVP level.
During the second phase the bond length increases where at the same time the dihedral angle rotated by $\approx \pm 90°$.

3) The last part of the IRC analysis belongs to the geometrical relaxation of the S_1 and the R-S_1 structures on the ground state. These structures are not minima of the ground state and their gradients are non-zero. This part of the analysis has an important message that following the gradient paths from the mentioned structures lead to the trans-111 conformer and not any other conformer– for instance *Z-azo*-conformers.

5.4 Electronic charge density analysis

Atomic partial charges are calculated by Mulliken definition of partial charges at MP2 and ADC(2) level with SVP basis set. This analysis gives us a qualitative picture about the changes of electronic structure during the electronic excitation and the de-excitation process.

Berenil is a double protonated cation, though the electronic charges are distributed over the molecule. Based on the calculations– please see table 6–, approximately 0.8 positive charge are localized on Amidine groups in both the electronic ground state, S_0, and the excited state, S_1, minima.

The Phenyl groups become partially negative by vertical excitation, but afterwards become more positively charged with roughly 0.3 at S_1 structure.

The total charge of Triazenyl group is 0.11 and -0.02 at the S_0 and the S_1 states, respectively. However the second part of table 6 reveals that positive and negative partial charges

are compensated with each other. The H_{20} atom holds about 0.15 where the partial charge of C_1 atom changes from 0.07 to -0.03. [5] From the S_0 to S_1, N_{17} looses some negative charges (-0.06) while N_{18} gains some, -0.04.

In conclusion, the electronic charge density analysis lets us know where are the most partially negative and positive sites of the molecule in different electronic states.

Table 6: Mulliken atomic partial charges are calculated with ADC(2)/SVP. Atoms which have less than ±0.1 charge for all structures are not shown. Amidine groups are named 1 and 2 based on their chemical priority.

Groups	S_0	FC	TS	S_1
Triazenyl	0.11	0.14	0.06	-0.02
1-Phenyl	0.19	0.15	0.14	0.30
1-Amidine	0.80	0.73	0.71	0.82
2-Phenyl	0.11	0.00	0.23	0.28
2-Amidine	0.79	0.75	0.80	0.79

	Triazenyl			
Atoms	S_0	FC	TS	S_1
C_1	0.07	0.04	0.02	-0.03
N_{17}	-0.16	-0.16	-0.19	-0.10
N_{18}	-0.02	-0.07	-0.07	-0.06
N_{19}	-0.08	0.08	0.04	-0.11
H_{20}	0.15	0.17	0.18	0.16
C_{21}	0.15	0.08	0.07	0.13

6 Geometrical flexibility of the excited state

6.1 Flexibility of Triazenyl group

The chosen approach to research the flexibility of the excited state is to rotate each torsion angle of the Triazenyl group of S_1 structure gradually and calculated the energies of these unoptimised geometries at COSMO-ADC(2)/SVP level of theory. The presumption of the approach is that the unoptimised geometries are closed to relax one and therefore the energies are qualitatively representative.

In addition, the obtained results show that the vacuum optimised structure are closed enough to COSMO-ADC(2) results since the S_1 structure of vacuum is quiet closed to the

[5] carbon atoms C_1 and C_{21} are formally atoms of Phenyl groups, however they are included in Triazenyl group.

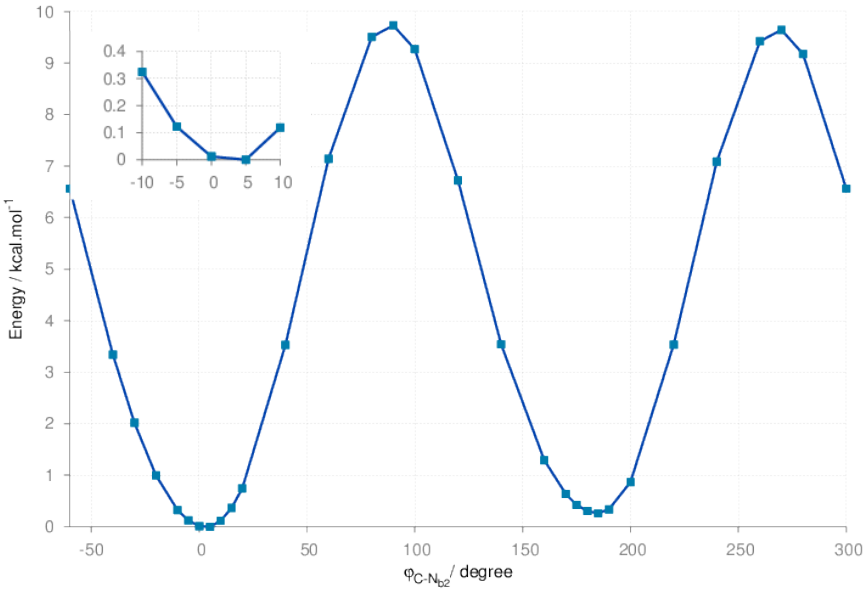


Figure 12: The barrier of rotation at kcal $\cdot$ mol^{-1} around the $C_{21}-N_{19}$ bond versus the $C_{22}-C_{21}-N_{19}-H_{20}$ torsion angle for the S_1 structure and electronic state, calculated at the COSMO-ADC(2)/SVP level without preliminary geometry optimization.

minima of each path.

The barrier for rotating around the $C_{21}-N_{19}$ bond is about 12 kcal $\cdot$ mol^{-1} which means that rotation around the band is possible at room temperature but not very fast. Passing the barrier would lead to give the mirror image of the trans-111 S_1 structure (see figure 12).
It is found that the energy of 5° rotated structure is slightly lower than the energy of vacuum optimised structure. Relaxing the geometry a COSMO-ADC(2) level will increase the difference, though the tiny difference (lower than 0.1 kcal $\cdot$ mol^{-1}) shows that the vacuum optimised geometry is a good approximation for the COSMO-ADC(2) optimised as well.

Rotation around the $N_{19}-N_{18}$ bond (figure 13) is possible up to ±40° degree. Beyond this regime, the repulsion of Aromatic groups becomes too strong – the SCF calculations for geometries out of the mentioned regime did not converged at all. Therefore one can conclude that there is no simple path for transformation between *s-cis* (trans-211,212) and *s-trans* (trans-111,112) in the excited state and such a isomerization reaction might be occurred with much higher barrier via a more complex trajectory. The latter conclusion is aligned with the results of the IRC analysis. There is also no experimental evidence for such a isomerization.

Figure 14 illustrates the unoptimised path of transformation of S_1 to R-S_1 structure. The

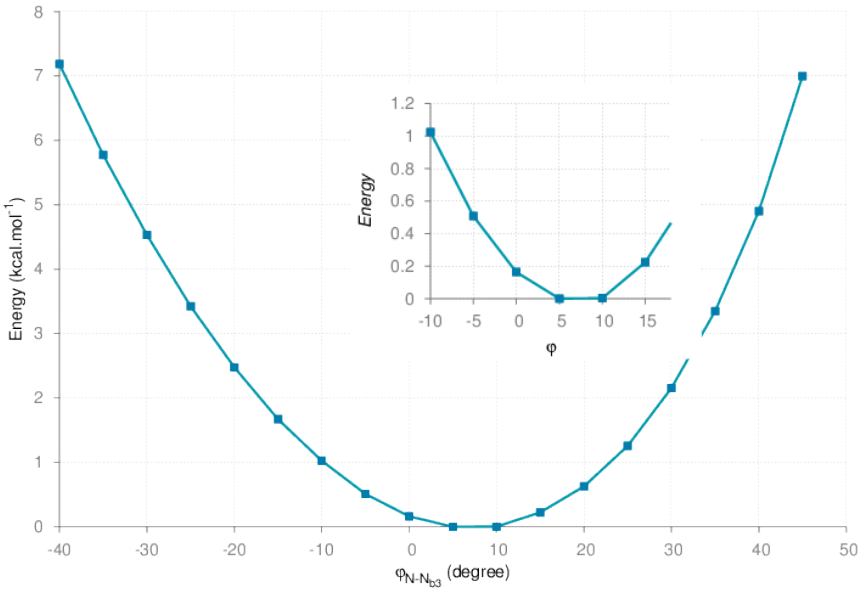


Figure 13: The barrier of rotation at kcal $\cdot$ mol^{-1} around the $N_{19}-N_{18}$ bond versus the $H_{20}-N_{19}-N_{18}-N_{17}$ torsion angle for the S_1 electronic state and structure, calculated at the COSMO-ADC(2)/SVP level without preliminary geometry optimization. The green line is fitted to the data points which are shown with moderate yellow.

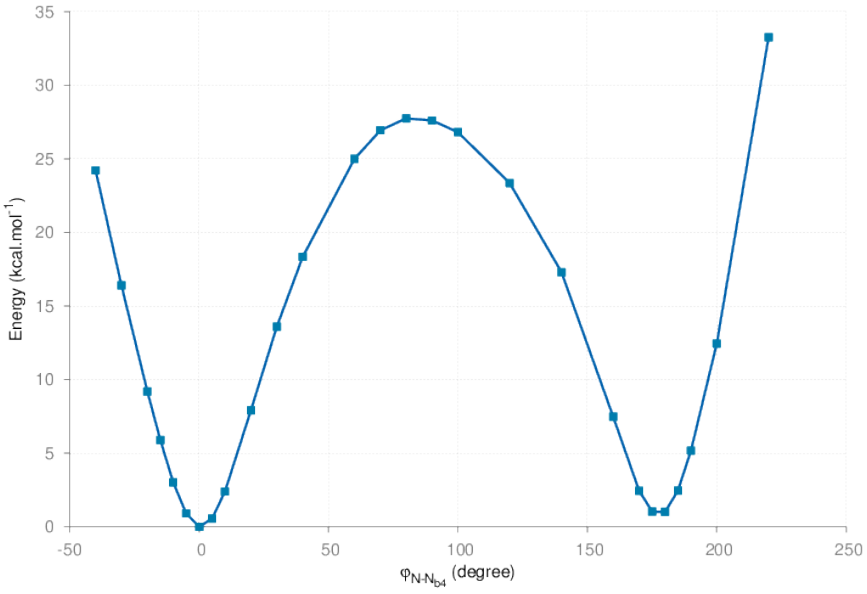


Figure 14: The barrier of rotation at kcal $\cdot$ mol^{-1} around the N_{17}=N_{18} bond versus the N_{19}−N_{18}=N_{17}−C_1 torsion angle for the S_1 electronic state and structure. Energies are evaluated with single point calculation at the COSMO-ADC(2)/SVP level without preliminary geometry optimization

barrier is about 25 kcal $\cdot$ mol^{-1} which is too high to let the molecule to pass over at the room temperature. Although the N_{17}=N_{18} bond does not have π binding character, the molecule can not rotate around this bond with large amplitude motions. Assuming that there is 10 kcal $\cdot$ mol^{-1} thermal energy available at the room temperature, the molecule can rotate around the bond up to $\pm 20°$.

The barrier of rotation around the N_{17}−C_1 single bond in the excited state, figure 15 , is similar to rotation around C_{21}−N_{19} bond except the fact that passing the barrier of this rotation will give one conformer of trans-112 S_1 structures. At the room temperature though, these transformations should be slow.

The barrier of rotation around the N_{17}−C_1 single bond in the excited state, figure 15 , is similar to rotation around C_{21}−N_{19} bond except the fact that passing the barrier of this rotation will give one conformer of trans-112 S_1 structures. At the rootm temperature though, these transformations should be slow.

The un-relaxed potential energy scans assist to scrutinize the flexibility of Triazenyl group of Berenil around the S_1 structure of the excited state of trans-111 conformer. Furthermore it confirms that the vacuum optimised geometry should be close enough to COSMO optimised one and reveals that probably there is no hidden minimum on the excited state.

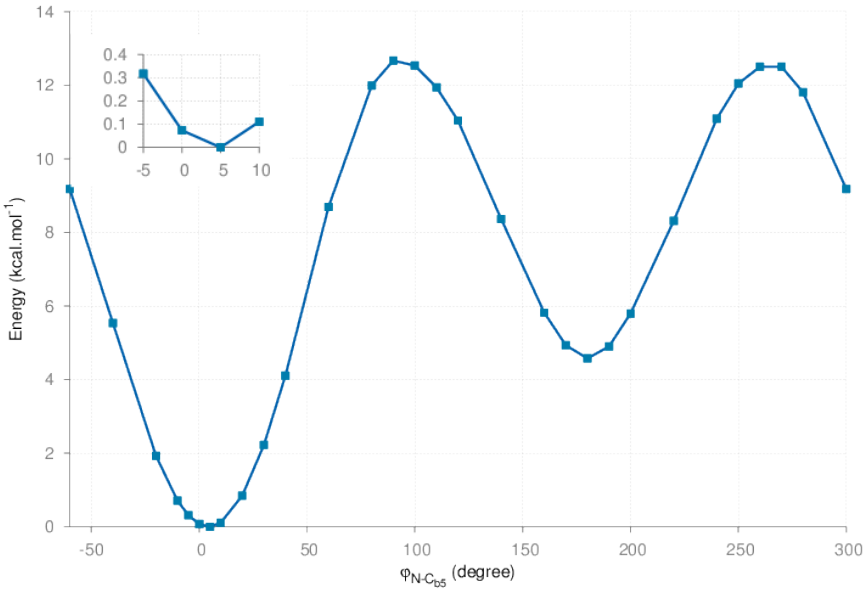


Figure 15: The barrier of rotation at kcal $\cdot$ mol^{-1} around the $N_{17}-C_1$ bond versus the $N_{18}-N_{17}-C_1-C_2$ torsion angle for the S_1 electronic state and structure. Energies are evaluated with single point calculation at the COSMO-ADC(2)/SVP level without preliminary geometry optimization

6.2 Flexibility of non-bonded distance of Amidine groups

In previous section, the rotating bonds of Triazenyl group of the molecule were screwed to study the flexibility of Berenil at a minimum of the lowest EES un-relaxed PES by an one-at-the-time approach. Here, we are going to study the flexibility of the molecule when two torsion angles (particularly $H_{20}-N_{19}-N_{18}=N_{17}$ and $N_{19}-N_{18}=N_{17}-N_{17}$) rotate spontaneously in the excited state via constructing a minimal additive force-field.

$$\Delta E = \sum_i \Delta E_i \tag{31}$$

Additive force-fields are built by assuming that degrees of freedom are decoupled and thus the changes of total (potential) energy of the system is equal to sum of changes of each degree of freedom, equation 31.

Then to have an analytical expression for the un-relaxed PES around a particular geometry, it is enough to intrapolate the un-relaxed potential energy curves to have an analytical expression for each term of the latter summation.
The intrapolation of un-relaxed energy curves– calculated at COSMO-ADC(2)/SVP level– of changes of $H_{20}-N_{19}-N_{18}=N_{17}$ and $N_{19}-N_{18}=N_{17}-N_{17}$ dihedral angles are done via a simple polynomial expansion around the S_1 structure, equation 32. The two dimensional

force-field parameters are listed in table 7.

$$\Delta E_{\tau_i}(\delta\tau_i) = a \cdot (\delta\tau_i - b)^2 + \ c \cdot (\delta\tau_i - b)^3 + d \cdot (\delta\tau_i - b)^4 \tag{32}$$

Table 7: The minimal force-field parameters. Please find the plots of fitted curves versus data points in the SI

parameters	$\tau_{\mathrm{H_{20}-N_{19}-N_{18}=N_{17}}}$	$\tau_{\mathrm{N_{19}-N_{18}=N_{17}-C_1}}$
a	0.003661	0.028256
b	7.76111	0.53892
c	$2.48389 \cdot 10^{-5}$	$-3.91249 \cdot 10^{-5}$
d	$3.01446 \cdot 10^{-7}$	$-1.75301 \cdot 10^{-5}$

The two dimensional calculated potential energy surface is presented in figure 16. Assuming that the molecule could gain maximally about 15 kcal $\cdot$ mol^{-1} potential energy due to interactions with environment, a range for rotation of dihedral angles is given by figure 16.

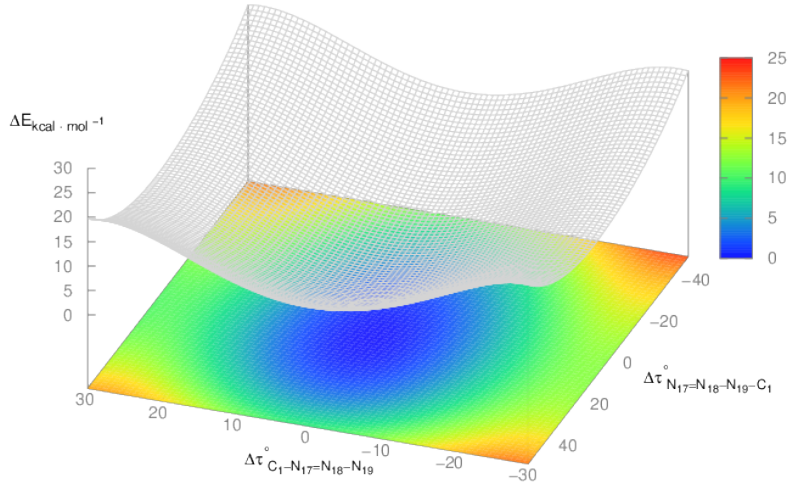
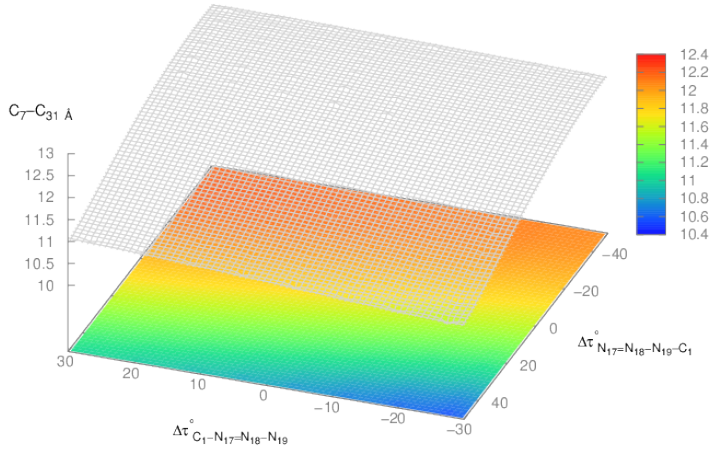


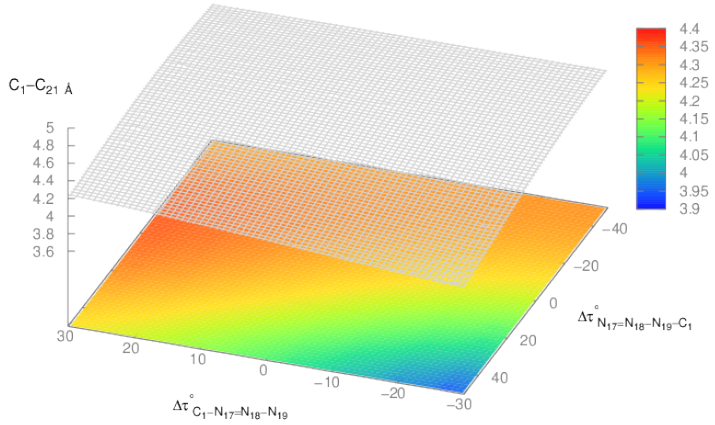
Figure 16: 2D-PES of dihedral angles rotations based on intrapolation of un-relaxed COSMO-ADC(2)/SVP.

Rotation around the N_{19}−N_{18} and N_{18}=N_{17} bonds can change the distance between terminal Amidine groups. These dihedral angles can rotate due to zero point vibrations and gaining thermal energy from solutions. and the amplitude of these rotations are explored qualitatively with the two dimensional force-field.

Figure ?? shows the distance between C_7 and C_{31} atoms as a measure for the distance of Amidine groups. It is approximated that the distance between Amidine groups can be changed in the range of $\approx 10.8 - 12$ Ångstrom due to rotations of the dihedral angles.



(a) Non-bonded distance between carbon atoms of Amidines



(b) Non-bonded distance between carbon atoms of internal group

Figure 17: Carbon-Carbon distances and dihedral angles are given in Ångstrom and angles, respectively.

This preliminary study shows that the Triazenyl group is flexible enough to react to the changes of the Amidines distance where ever necessary, in the given range.

7 Summary and conclusion

The first chapter of my master thesis is devoted to give a briefing about the Berenil. The molecule that is studied in this project. Interesting fluorescence properties of the molecule and it is potential to interact with some of the vital molecules (DNA and/or RNA) gives me passion to know more about this small molecule.

In the second chapter, the response theory and continuum solvation models are reviewed shortly to give a perspective about computational methods that are used in the thesis. I would be happy if I could show that developing theoretical methods, indeed can open doors to understand molecular chemistry better.

The chapter 3 is allotted to find the role of different conformers of Berenil in the electronic ground state where it is shown that the molecule has a relatively in-plane structure and there are couple of conformers, trans-111 and trans-112, that are thermodynamically much more stable than the other. Further, based on different calculations, we figure out that the relative orientation of Amidine groups of the molecule does not play a role to change the relative properties of trans- conformers.

In chapter 4, electronic vertical excitation and a related property, one-photon absorption, is studied where we have shown that the excitation to the lowest electronic excited state causes the main peak of UV/Vis spectrum of the molecule. With NTO analysis the changing of bonding patterns during the excitation is scrutinized.

In chapter 5, excited state geometrical analysis at ADC(2)/SVP level reveals that the rotation around $C_1{-}N_{17}{=}N_{18}{-}N_{19}$ angle occurs while the roughly planar structure of the molecule with respect to other torsion angles of internal group (Ar−NH−N=N−Ar) remains. The Amidine groups do not rotate remarkably as well. From S_0 to S_1, the *azo*-, $N_{17}{=}N_{18}$ bond becomes longer where the $N_{19}{-}N_{18}$ and $C_1{-}N_{17}$ bonds become shorter. The IRC analysis as well as the electronic charge density analysis assist us to characterize the de-excitation path and understand the changes of electronic structure during the de-excitation. Particularly, finding the partially positive and negative sites will help us to add explicit water molecules to our model to expand our study in close future.

Figure 10 shows a $\text{kcal} \cdot \text{mol}^{-1}$ gap between the electronic ground and lowest excited state at S_1 structure, calculated at COSMO-ADC(2)/aug-cc-pVDZ level. Explicit solvent effects, changing of solvent and/or geometrical restrictions (for instance where the Amidine groups are binded to some molecule and their distance has to stay fixed) and dynamics of the excited state could fill the gap or make it larger which is crucial to understand the fluorescence properties of molecule in different environments.
It was shown that Berenil binds to DNA by terminal Amidine groups and the selectivity of binding is connected to the distance of Amidine groups [2]. As mentioned in the introduction, Berenil has intensive fluorescence emission when binded to an active site of DNA

and/or RNA.
Assuming that the active site has a relatively fixed structure and thus the distance between Amidine groups remains fixed, it is interesting to investigate the relaxed structure of the excited state if the energy of it increases much enough to change the electronic states order due to the geometrical restrictions. The flexibility of the molecule is studied in chapter 6 with calculating several un-relaxed potential energy curves to get a preliminary perspective to the questions that we are going to answer.

8 Supporting Informations

8.1 Excited state properties for 10 excitation levels

Table 8: Excitations energies and one-photon absorption of trans- conformers calculated with COSMO-ADC(2)/SVP. Excitation energies are given in eV and oscillator strengths are given in atomic units.

Conformer	EES	VEE	OS	Conformer	EES	VEE	OS
trans-111	1	3.517	1.277	trans-211	1	3.639	1.135
	2	3.874	0.000		2	3.738	0.078
	3	4.662	0.007		3	4.708	0.022
	4	4.768	0.021		4	4.783	0.004
	5	4.918	0.115		5	5.003	0.085
	6	5.513	0.231		6	5.621	0.079
	7	5.717	0.043		7	5.820	0.031
	8	5.910	0.143		8	5.843	0.009
	9	6.041	0.051		9	5.917	0.109
	10	6.133	0.244		10	6.074	0.033
trans-112	1	3.519	1.274	trans-212	1	3.637	1.044
	2	3.877	0.003		2	3.756	0.167
	3	4.664	0.007		3	4.713	0.021
	4	4.768	0.021		4	4.785	0.004
	5	4.919	0.114		5	5.007	0.085
	6	5.514	0.232		6	5.625	0.080
	7	5.717	0.042		7	5.816	0.027
	8	5.911	0.146		8	5.864	0.018
	9	6.043	0.040		9	5.920	0.109
	10	6.134	0.253		10	6.080	0.028

Table 9: NTOs of trans-112 conformer calculated with COSMO-ADC(2)/SVP

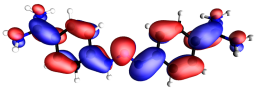
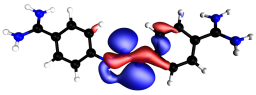
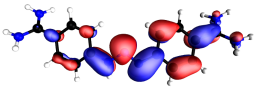
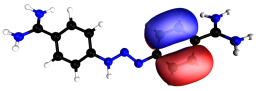
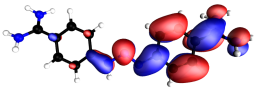
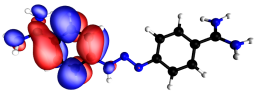
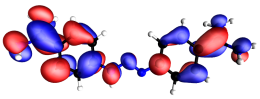
Conformer	EES	OCC	VIR	Contrib.
trans-112	1			96.2
trans-112	2			99.7
trans-112	3			70.5
trans-112	4			62.3
trans-112	5			86.2

Table 10: NTOs of trans-211 conformer, calculated with COSMO-ADC(2)/SVP

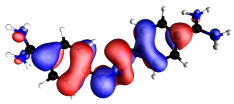
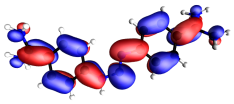
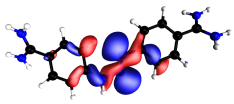
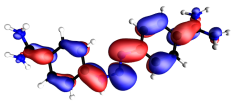
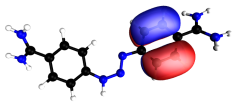
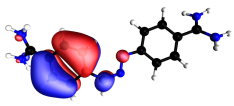
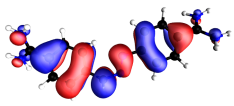
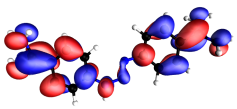
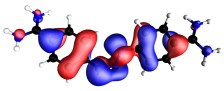
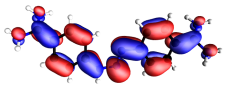
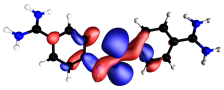
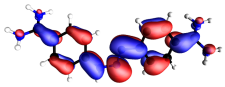
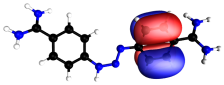
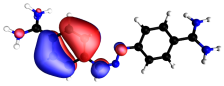
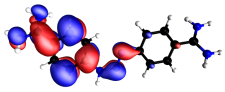
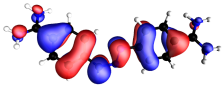
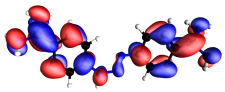
Conformer	EEL	OCC	VIR	Contrib.
trans-211	1			95.6
trans-211	2			98.8
trans-211	3			69.8
trans-211	4			55.5
trans-211	5			84.5

Table 11: NTOs of trans-212 conformer, calculated with COSMO-ADC(2)/SVP.

Conformer	EEL	OCC	VIR	Contrib.
trans-212	1			95.4
trans-212	2			98.0
trans-212	3			69.7
trans-212	4			55.4
trans-212	5			84.5

8.2 S_0-S_1 potential energies

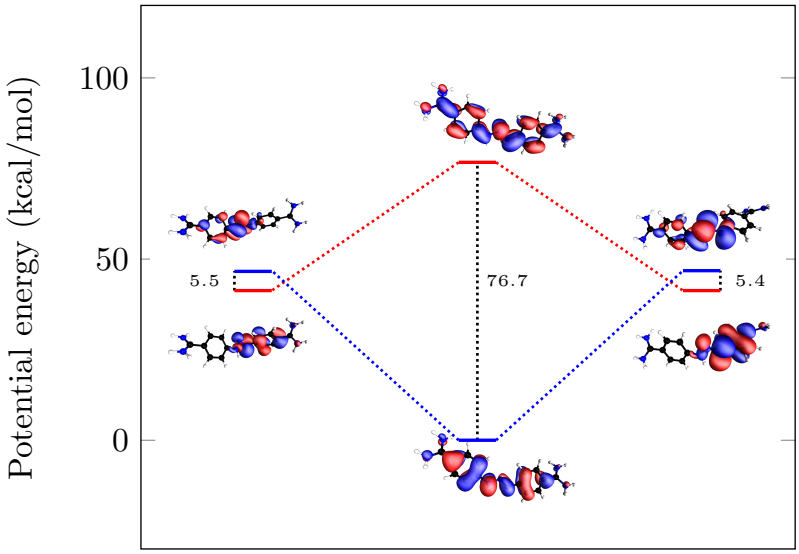


Figure 18: The schematic potential energy surface shows the electronic de-excitation process of trans-112 at the COSMO-ADC(2)/aug-cc-pVDZ level.

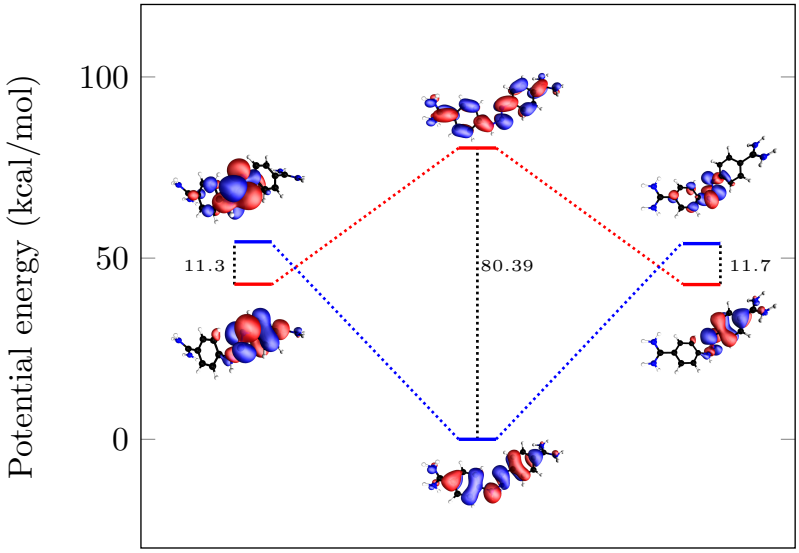


Figure 19: The schematic potential energy surface shows the electronic de-excitation process of trans-211 at the COSMO-ADC(2)/aug-cc-pVDZ level.

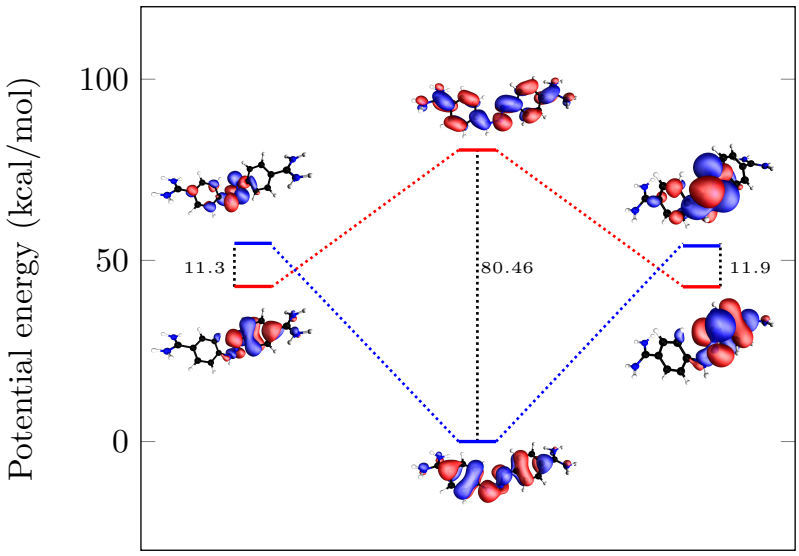


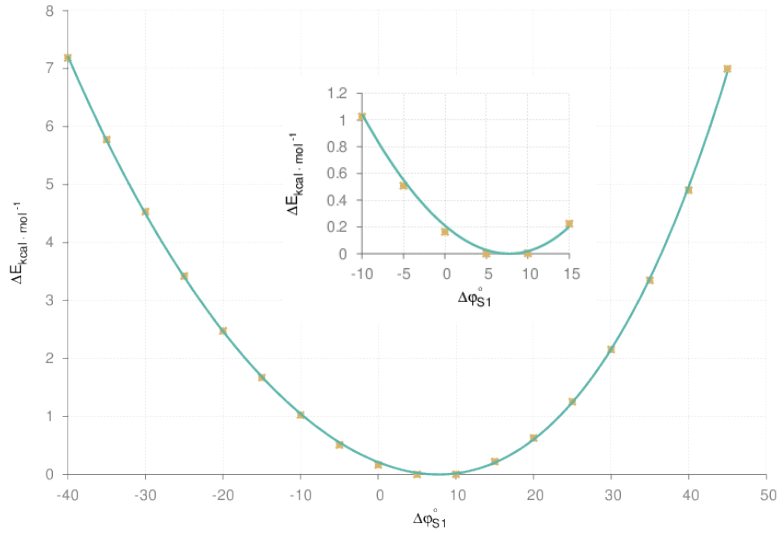
Figure 20: The schematic potential energy surface shows the electronic de-excitation process of trans-212 at the COSMO-ADC(2)/aug-cc-pVDZ level.

8.3 Geometrical analysis

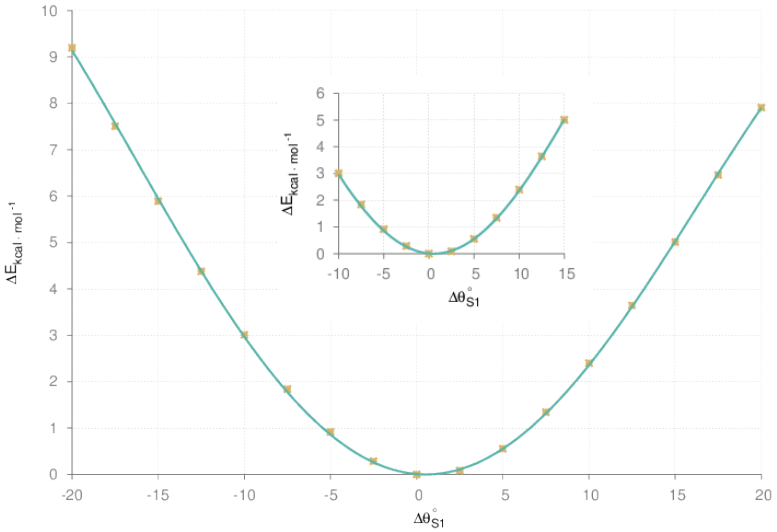
Table 12: Geometrical details of different structures of Berenil trans-111 conformer. Distances are given in pm. Triangular and dihedral angles are given in degree.

Internal coord.	S_0	TS^*	TS	S_1	R-S_1
Selected bond lengths					
$C_{24}-C_{31}$	146	147	147	146	146
$C_{21}-C_{22}$	140	142	142	141	141
$C_{22}-C_{23}$	138	139	139	139	139
$C_{31}-N_{34}$	132	133	133	133	133
C_7-N_8	132	133	133	132	132
C_4-N_7	146	145	145	147	147
C_1-C_2	140	143	143	142	142
C_2-C_3	138	138	138	139	139
Selected non-bonded distances					
$N_{17}-H_{20}$	226	218	215	234	234
Selected triangular angles					
$N_{32}-C_{31}-N_{34}$	120	121	121	121	121
$C_{23}-C_{24}-C_{31}$	120	120	120	120	120
$N_8-C_7-N_9$	120	120	119	122	122
$C_3-C_4-C_7$	120	121	120	120	119
$N_{19}-C_{21}-C_{22}$	118	120	119	119	119
$N_{17}-C_1-C_2$	115	117	116	115	115
$N_{18}-N_{19}-H_{20}$	116	115	114	118	118

8.4 Force field fitted curves



(a) Intrapolated curve (green) of un-relaxed PES with respect to rotation of $H_{20}-N_{19}-N_{18}-N_{17}$ dihedral angle, φ



(b) Intrapolated curve (green) of un-relaxed PES with respect to rotation of $N_{19}-N_{18}=N_{17}-C_1$ dihedral angle, θ

References

[1] C. S. Chow and F. M. Bogdan, Chemical Reviews. **97**(5), 1489 (1997).

[2] S. Neidle, Natural Product Reports. **18**, 291 (2001).

[3] J. R. Thomas and P. J. Hergenrother, Chemical Reviews. **108**(4), 1171 (2008).

[4] T. Helgaker, S. Coriani, P. J. and Kasper Kristensen, J. Olsen and K. Ruud, Chemical Reviews. **112**, 543 (2012).

[5] O. Christiansen, P. Jørgensen and C. Hättig, International Journal of Quantum Chemistry. **68**, 1 (1998).

[6] O. Christiansen, H. Koch, P. Jørgensen and T. Helgaker, Chemical Physics Letters. **263**, 530 (1996).

[7] M. Schreiber, M. R. Silva-Junior, S. P. A. Sauer and W. Thiel, Journal of Chemical Physics. **128**, 134110 (2008).

[8] C. Hättig and F. Weigend, Journal of Chemical Physics. **113**, 5154 (2000).

[9] F. Weigend, Physical Chemistry Chemical Physics. **8**, 1057 (2006).

[10] A. Hellweg, C. Hiättig, S. Höfener and W. Klopper, Theoretical Chemistry Accounts. **117**, 587 (2007).

[11] J. Schirmer, Physical review A. **26**, 2395 (1982).

[12] C. Hättig, 50 37 (2005).

[13] A. Klamt. *COSMO-RS From Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design*. ELSEVIER B.V., 2005.

[14] B. Mennucci, WIREs Computational Molecular Science. **2**, 386 (2012).

[15] A. Klamt, Journal of Physical Chemistry. **100**, 3349 (1996).

[16] A. Klamt and V. Jonas, Journal of Chemical Physics. **105**, 9972 (1996).

[17] A. Klamt, V. Jonas, T. Bürger and J. Lohrenz, Journal of Physical Chemistry A. (102), 5074 (1998).

[18] V. Barone and M. Cossi, Journal of Physical Chemistry A. **102**, 1995 (1998).

[19] F. Furche, R. Ahlrichs, C. Hättig, W. Klopper M. Sierka and F. Weigend, WIREs Computational Molecular Science. **4**, 91 (2014).

[20] F. Furche, R. Ahlrichs, C. Hättig, W. Klopper, M. Sierka and F. Weigend, WIREs Computational Molecular Science. **4**, 91 (2013).

[21] V. Weijo, M. Randrianarivony, H. Harbrecht and L. Frediani, Journal of Computational Chemistry. **103**, 11060 (1999).

[22] M. Bugeanu, R. D. Remigio, K. M. and S. Sommerfelt Reine, H. Harbrecht and L. Frediani.

[23] B. Stamm, E. Canés, F. Lipparini and Y. Maday, Journal of Chemical Physics. **144**, 054101 (2016).

[24] E. Cances, B. Mennucci and J. Tomasi, Journal of Chemical Physics. **107**, 3032 (1997).

[25] J. T. B. Mennucci, E. Cansés, Journal of Physical Chemistry B. **144**(054101) (2016).

[26] B. M. E. Cansés, Journal of Mathematical Chemistry. **23**, 309 (1998).

[27] D. Chipman, Journal of Chemical Physics. **116**(10129) (2002).

[28] D. Chipman, Journal of Chemical Physics. **107**, 80 (2002).

[29] F. Lipparani and B. Mennucci, Journal of Chemical Physics. **144**, 160901 (2016).

[30] F. O. del Valle, M. Aguilar and S. Tolosa, Journal of Molecular Structure. **279**, 223 (1993).

[31] O. Christiansen and K. V. Mikkelsen, Journal of Chemical Physics. **110**, 1365 (1991).

[32] O. Christiansen and K. V. Mikkelsen, Journal of Chemical Physics. **110**, 8348 (1999).

[33] B. Lunkenheimer and A. köhn. **9**, 977 (2013).

[34] A. D. Becke, The Journal of Chemical Physics. **98**(7), 5648 (1993).

[35] S. Grimme, J. Antony, S. Ehrlich and H. Krieg, Journal of Chemical Physics. **132**, 154104 (2010).

[36] S. Grimme, S. Ehrlich and L. Georigk, Journal of Computational Chemistry. **32**, 1456 (2011).

[37] F. Weigenda and R. Ahlrichs, Physical Chemistry Chemical Physics. **7**, 3297 (2005).

[38] F. Weigend, M. Häser, H. Patzelt and R. Ahlrichs, Chemical Physics Letters. **294**, 143 (1998).

[39] P. J. Stephens and N. Harada, Chirality. **22**, 229 (2010).

[40] A. F. Thünemann, D. S. and Robert Sachse, H. Schlaad and H. Möhwald, Langmuir. **22**, 2323 (2006).

[41] R. L. Martin, Journal of Chemical Physics. **118**, 4775 (2003).