

Polarizable Embedded ADC(2) Gradients and
DNA-Confinement Effects on the Electronically Excited-State
Relaxation Pathways of Berenil

by

Alireza Marefat Khah

Dissertation submitted in partial fulfillment
of the requirements for
the degree of a Doctor in Natural Sciences (Dr. rer. nat.)

RUHR-UNIVERSITÄT BOCHUM

FAKULTÄT FÜR CHEMIE UND BIOCHEMIE

LEHRSTUHL FÜR THEORETISCHE CHEMIE

Acknowledgment

First, I would like to express my greatest appreciations to my supervisor Prof. Dr. Christof Hättig for all the support and opportunities he has given me. I acknowledge all his contributions of time, ideas and funding in my PhD studies to make it a great and rich experience for me.

I gratefully thank Professor Patrick Nürnberger for proposing the berenil project to us. Having the opportunity to collaborate with his team was a great advantage for me. I also thank him for being my second supervisor, and a member of my PhD examination board. I am thankful to Professor Jacob Kongsted for his mentorship and all the insightful and friendly discussions during my internship in Odense, and also for accepting to be a member of my PhD examination board.

I am grateful to my collaborators, Lena Grimmelsmann and Peter Reinholdt, for their scientific contributions and discussions. Special thanks to Peter Reinholdt and Robert Treß for reading my thesis and giving me productive feedback.

I would also like to thank my colleagues in the research group of Christof Hättig, specially my coffee partner, Marius, and my previous and current office-mates Nora, Lisa and Robert. Finally, I am particularly grateful to my dear wife, Sarah, for all her supports, scientific inspirations and sacrifices. Without her accompany this would have not come true.

Alireza M. Khah

Ruhr Universität Bochum

October 2019

Contents

1	Introduction	2
2	Second-Order Correlated Wavefunction Methods	10
2.1	Coupled Cluster Singles and Approximated Doubles	11
2.2	Excited-State Energies and Gradients of Second-Order Methods	14
3	Quantum Mechanics/Molecular Mechanics Models	24
3.1	A Hierarchy of QM/MM Methods	25
3.2	Post-SCF Polarizable Embedding Reaction Field	29
4	Modelling of Excited-States Relaxation Pathways	36
4.1	Intramolecular Relaxation Pathways from Intrinsic Reaction Coordinates .	37
4.2	Volume-Conserving Torsions Relaxation Pathways	41
5	Ultrafast Dynamics of a Triazene: Excited-State Pathways and the Im-	
	act of Binding to the Minor Groove of DNA and Further Biomolecular	
	Systems	46
6	How a Linear Triazene Photoisomerizes in a Volume-Conserving Fashion	56
7	Analytic Excited State Gradients for the QM/MM Polarizable Embed-	
	ded Second-Order Algebraic Diagrammatic Construction for the Polar-	
	ization Propagator PE-ADC(2)	70

8	Avoiding Electron Spill Out in QM/MM Calculations on Excited States with Simple Pseudopotentials	84
9	Berenil in DNA-Minor-Groove Confinement	108
9.1	Computational Details	109
9.1.1	Molecular Dynamics Simulation	109
9.1.2	Calculation of Polarizable Potential	110
9.1.3	QM/MM Calculations	112
9.1.4	Models of berenil@DNA complex	112
9.2	Results and Discussions	113
9.3	Conclusion and Outlook on the Berenil@DNA Project	119
10	Summary and Conclusion	122
	Appendices	144

1 Introduction

Developing *ab initio* quantum chemistry methods, together with spectroscopic techniques, [1] provide molecular-level insights about the photochemistry of biosystems [2] and photophysical processes. [3] To advance photopharmacology, [4,5] improve environment-sensitive probes, [6] develop functional molecular switches [7,8] and to design novel molecular motors, [9,10] we need to understand the underlying molecular mechanisms of photophysical processes. For instance, based on the well-known *trans-cis* photoisomerization mechanism of azobenzene, [11–15] derivatives of it have been tailored to restore visual responses in blind Retinae. [16]

At the beginning of our research, very little was known about the excited-state relaxation and photoisomerization processes of aromatic triazenes. [17–21] Similar to the azobenzene, the triazene functional group of berenil (diminazene aceturate) can perform a *trans-cis* photoisomerization. [22,23] Berenil is a widely used drug for treating sleep sickness in

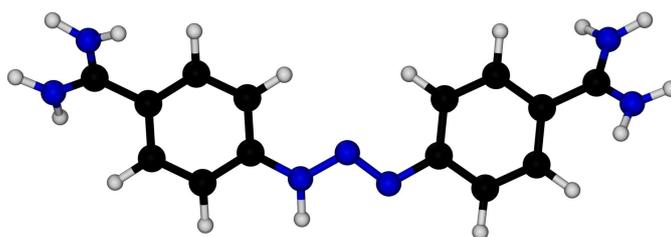


Figure 1: Berenil, the double protonated state of diminazene

animals and also an AT-rich DNA minor groove binder with a nanomolar dissociation constant (it can also bind to other biomolecular compounds). [24] During my PhD project,

the photo-induced relaxation processes of berenil has been studied as a model system for understanding the photophysics of aromatic triazenes, in solutions and also in the confinement of DNA minor groove with *ab initio* quantum chemistry methods, *vide infra*. The challenge of my PhD project was to develop a quantum chemistry method that can model electronic relaxation processes in a heterogeneous molecular environment, and exploit the potential of the method for the simulation of electronic excited-state relaxation processes of berenil when it is confined into the DNA minor groove.

The first objective of my PhD was to investigate the electronic ground- and excited-state relaxation pathways of berenil upon the photoexcitation through interpreting the measured time-resolved fluorescence up-conversion and transient absorption data by utilizing the second-order wavefunction-based method ADC(2). This project was conducted in collaboration with the research group of Professor Patrick Nürnbergger, which provided us with the experimental spectroscopic data. Former computational studies that are reported in my master thesis [25] indicated that after the photoexcitation of berenil to the excited state at the Franck-Condon point, the molecule relaxes to a conical intersection between the ground and the lowest excited state. As the first project of my PhD, reported in chapter 5, the intrinsic reaction coordinates for the relaxation of berenil in the lowest singlet electronic excited state have been analyzed. From this analysis, we could identify a two-phase relaxation mechanism (an N=N bond stretching followed by a bicycle-pedal type motion), which is volume conserving. These results are in very good agreement with the conclusions that are derived by our collaborators from the ultrafast time-resolved fluorescence spectroscopy.

This line of research has been pursued further to interpret the time-resolved transient absorption data obtained by the physical chemistry group. From the results of these studies, a comprehensive picture of the ground- and excited-state relaxation pathways of berenil has been emerged, which is described in chapter 6. In the electronic ground state, berenil has three pairs of isomers, cmp. Fig. 1 in chapter 6. Different isomerization pathways have been theoretically studied with an emphasize on comparing the ground and excited-state deexcitation processes with the corresponding experimental data measured by transient absorption spectroscopy. By studying the ground and excited-state absorption along the photoisomerization pathways, it was found that there are two volume conserving photoisomerization paths to the *E-azo-s-trans* and *Z-azo-s-trans* isomers, with a bicycle-pedal and a Hula-twist mechanisms, respectively. With these mechanisms, we could explain why the photoisomerization pathways exhibit low sensitivity to changes in the solvent viscosity and a moderate response to bio-molecular embedding.

In the next step, we aimed to implement a suitable *ab initio* quantum chemistry method actually to study the excited-state relaxation processes of berenil when it is embedded in the minor groove of DNA, chapter 9. For the computation of electronic excited states of single-reference molecules, such as berenil, wavefunction-based second-order methods including the coupled cluster singles and approximate doubles (CC2) and the algebraic diagrammatic construction through second order (ADC(2)) are usually more reliable than the (time dependent) mean-field methods *e.g.*, Hartree-Fock and Density Functional Theory, by accounting for the exact electronic exchange-correlation and approximate electronic coulomb correlation. [26–28] The following chapter 2 gives an overview of the background

and the capabilities of correlated second-order methods for excited states, and in particular the ADC(2) method, which is used for studying berenil excited-state relaxation processes. The computational cost of the second-order methods increased with the fifth power of the system size; hence, it is impossible to use the ADC(2) method for studying berenil@DNA without further approximations. A general idea to ease the calculations for systems that consist of an active (berenil in this study) and an inactive subsystem *e.g.*, the DNA and surrounding solvent molecules, is to apply the QM/MM strategy. [29] The QM/MM methods have been developed substantially in the last two decades for the simulation of chemical reactions in the electronic ground state. [30,31]

Compared to QM/MM methods that are restricted to an electrostatic and mechanical embeddings, the development of polarizable embedded QM/MM methods, in particular for electronic excited states [1, 31, 32], is more challenging, which is a subject of current research. [32–34] Towards developing a method to study the excited-state dynamics of berenil, we have developed the theory for analytical excited-state gradients for the QM/MM polarizable embedded ADC(2) and implemented them in the `ricc2` module of the TURBOMOLE package. [35] In the introductory chapter 3, more detail about the hierarchy of QM/MM methods and the specific features of our implementation of PE-ADC(2) are given. To the best of my knowledge, this is the first implementation of excited-state gradients for a correlated method within a polarizable embedding.

The newly implemented PE-ADC(2) code was first tested for the host-guest complex of 2,3-diazabicyclo-[2.2.2]oct-2-ene (DBO) inside cucurbit[7]uril (CB[7]) in chapter 7. For this system, an artificially low excitation energy was observed when a basis set with diffuse

functions is used in the calculations. This problem is attributed to the so-called electron spill out, which occurs due to the lack of Pauli repulsion. [31] As a remedy for this issue, the effective core potentials (ECPs) [36] are used to mimic the effect of Pauli repulsion. The usual available ECPs in the literature are only optimized to model the inner electrons and not for use in the MM part of QM/MM setups. Therefore, we parametrized simple ECPs that model all electrons of an atom. For further details, see chapter 8. These simple all-electron ECPs have been used in a benchmark study that has been carried out together with the quantum chemistry group of Professor Jacob Kongsted. Also, the PE- and PE(ECP)-ADC(2) methods were tested against the full-QM PNO-ADC(2) for the calculation of excitation energies. The results of our benchmark study that are reported in chapter 8 show that PE(ECP)-ADC(2) is more accurate than the original PE-ADC(2) and is much more stable against an unphysical electron spill-out.

Furthermore, choosing the PNO-ADC(2) as the reference method provides us a unique opportunity to analyze the performance of polarizable QM/MM methods even further. In contrast to DFT methods, at the correlated ADC(2) level, the dispersion interaction is included intrinsically through the electron correlation. Based on the localization of molecular orbitals, an excitation energy decomposition scheme has been implemented in the `pnoccsd` code of TURBOMOLE by Professor Hättig, which can isolate the dispersion contribution to the excitation energy. By carrying out the latter energy decomposition analysis, I clarified that the next major source of error in the polarizable embedded calculations is the lack of dispersion interactions. This is the first time, to the best of my knowledge, that the importance of dispersion interactions on the electronic wavefunction-based QM/MM

calculations has been assessed quantitatively.

The developed PE(ECP)-ADC(2) method has been used for studying the excited-state relaxation pathways of berenil in the DNA-minor-groove confinement with an atomistic representation of solvent molecules. In chapter 9, it is shown that the excitation energies of berenil@DNA can be calculated with PE(ECP)-ADC(2) using basis set that have diffuse functions (aug-cc-pVDZ), which is in excellent agreement with experimental UV/Vis data. Furthermore, the excited state relaxation pathways of berenil in the confinement of DNA minor groove has been studied in detail. We found that the excited state relaxation mechanism of berenil@DNA has a bicycle-pedal type character. Moreover, it is illustrated that the latter volume-conserving relaxation mechanism can be influenced by the water molecules of the first solvation shell of berenil@DNA. At the moment, the influence of hydrogen-bonded water molecules on the *trans-to-cis* photoisomerization of berenil@DNA and the collective excited state relaxation mechanism of berenil and hydrogen-bonded water molecules at the DNA minor groove confinement is under the study.

Outline of the Thesis

In chapter 2, a brief overview of single-reference second-order methods that are suitable for the calculation of analytic excited-state gradients and excited-state properties is given. At the end of the chapter, it is explained why ADC(2) is our method of choice for the entire project. In chapter 3, the general framework of QM/MM methods and the post-SCF reaction field scheme for coupling a correlated QM method with the polarizable environ-

ment that is used in the implementation of the PE-ADC(2) excited-state gradients will be explained. In our implementation of PE-ADC(2) excited-state gradients, the coupling between the correlated part of the wavefunction and polarization of the environment has been taken into account via the post-SCF reaction field scheme. In chapter 3, an overview about the general framework of the dipoles polarizable QM/MM methods and the post-SCF reaction field scheme is given.

In the last introductory chapter 4, the assessment of intrinsic reaction coordinate analysis for studying the intramolecular excited state relaxation pathways and the possible torsional relaxation pathways of berenil will be discussed. Chapters 5 and 6 explain the results of our studies on the excited-state relaxation processes and photoisomerization pathways of berenil. In chapter 7, the theory, the implementation aspects and the application of PE-ADC(2) excited-state gradients to study the excited state geometries and fluorescence emission energies of the host-guest systems are presented. Chapter 8 consists of the results of our benchmark study on the performance of the PE(ECP)-ADC(2) method and the discussion about the role of excited-state dispersion to the excitation energies. The relaxation mechanism of berenil@DNA are reported in chapter 9 where the effects of DNA-minor-groove confinement and the solvent effects to the volume-conserving excited-state relaxation process of berenil are discussed in detail. Finally, the most important findings and advancements are highlighted in a summary and conclusion chapter.

2 Second-Order Correlated Wavefunction Methods

One of the well-known single-reference post-Hartree-Fock methods is Møller-Plesset perturbation theory, MPPT. [37] In the spirit of Rayleigh-Schrödinger perturbation theory, in MPPT theory, the Hamiltonian is partitioned into the zeroth-order Fock operator \hat{F} , and the first-order fluctuation potential \hat{W} . The nuclear-nuclear repulsion contribution h_{nuc} is a part of the zeroth-order Hamiltonian with no importance in our discussion.

$$\hat{H} = \hat{F} + \hat{W} + h_{\text{nuc}} . \quad (1)$$

In MPPT, the sum of the zeroth- and first-order energy corrections $E^{(1)}$ is identical with the Hartree-Fock energy E^{HF} and the first improvement on top of the HF energy is at the second-order. Thus, a second-order Møller-Plesset perturbation theory MP2 [38] calculation comprises two steps: first an iterative Hartree-Fock (HF) step and then a posterior non-iterative second-order energy correction step to calculate ΔE_{MP2} . A standard implementation of MP2 has an $\mathcal{O}(\mathcal{N}^5)$ computational complexity with the system size \mathcal{N} . The MPPT structure is incompatible with response theory, which is a systematic framework to calculate excited state properties. [39] In consequence, the calculation of excitation energies and transition moments is not meaningful at this level of theory. Therefore, there is a demand to develop second-order methods that determine variables of the response equations by a single set of equations [40]. The iterative variant of the doubles correction to configuration interaction singles CIS(D_∞) [41], the algebraic diagrammatic construction through the second-order ADC(2) [28] and the coupled cluster singles and approximated doubles

CC2 [42, 43] are three examples of the single-reference, second-order wavefunction-based methods that one can use for calculation of excitation energies. They are all implemented in the `ricc2` module of TURBOMOLE program package [35]. In the following, I give a short description of the second-order methods CC2, ADC(2), CIS(D_∞) and their relations to highlight the pros and cons of our method of choice, ADC(2), for the berenil project.

2.1 Coupled Cluster Singles and Approximated Doubles

The coupled cluster (CC) model corresponds to a nonlinear exponential parametrization of the wavefunction. Choosing the Hartree-Fock state as the reference state, the CC wavefunction reads as [38]

$$|\text{CC}\rangle = \exp(\hat{T})|\text{HF}\rangle, \quad (2)$$

with $\hat{T} = \hat{T}_1 + \hat{T}_2 + \dots + \hat{T}_n = \sum_i \sum_{\mu_i} \hat{T}_{\mu_i}$ as the cluster operator and μ as an excitation manifold. For instance, the \hat{T}_2 operators generate doubly excited determinants out of the reference state $|\text{HF}\rangle$,

$$\hat{T}_2|\text{HF}\rangle = \sum_{\substack{i>j \\ a>b}} t_{ij}^{ab} \hat{\tau}_{ij}^{ab}|\text{HF}\rangle = \sum_{\substack{i>j \\ a>b}} t_{ij}^{ab} |ij\rangle^{ab}. \quad (3)$$

The t_{ij}^{ab} is a cluster amplitude and $\hat{\tau}_{ij}^{ab}$ is an excitation operator that generates the doubly excited Slater determinants. In the above equation and in the following, a, b, c, \dots are indices for unoccupied and i, j, k, \dots are indices for occupied molecular orbitals of the reference determinant $|\text{HF}\rangle$, respectively. The CC hierarchy (CCS, CCSD, CCSDT, ...) is obtained by truncation of \hat{T} after the n -th excitation level. For $n \rightarrow N_e$ it converges to the

full configuration interaction limit while the computational complexity scales as $\mathcal{O}(\mathcal{N}^{2n+2})$.

To obtain the CC energy

$$E^{\text{CC}} = \langle \text{HF} | H \exp(\hat{T}_1 + \hat{T}_2) | \text{HF} \rangle , \quad (4)$$

one has to first solve the amplitude equations that are determined by projection of the coupled cluster wavefunction onto the corresponding excitation manifold

$$\langle \mu_i | \exp(-\hat{T}) H \exp(\hat{T}) | \text{HF} \rangle = 0 . \quad (5)$$

For the CCSD model, using the \hat{T}_1 -transformed formulation, $\hat{O} = \exp(-\hat{T}_1) \hat{O} \exp(\hat{T}_1)$, the CCSD singles and doubles amplitude equations, in the order Ω_1 and Ω_2 , read as

$$\Omega_1 = \langle \mu_1 | \hat{H} + [\hat{H}, \hat{T}_2] | \text{HF} \rangle = 0 , \quad (6)$$

$$\Omega_2 = \langle \mu_2 | \hat{H} + [\hat{H}, \hat{T}_2] + \frac{1}{2} [[\hat{H}, \hat{T}_2], \hat{T}_2] | \text{HF} \rangle = 0 . \quad (7)$$

A rough idea about the quality of coupled cluster energies and wavefunctions can be obtained based on a perturbation analysis. With the MPPT partitioning of the Hamiltonian, *cmp.* Eq. 1, the excitation operators and cluster amplitudes can be expanded in orders of perturbation

$$\hat{T} = \hat{T}^{(0)} + \hat{T}^{(1)} + \hat{T}^{(2)} + \dots \quad (8)$$

$$t_\mu = t_\mu^{(0)} + t_\mu^{(1)} + t_\mu^{(2)} + \dots , \quad (9)$$

and so can be the energy. In the CCSD model, the \hat{T}_1 and \hat{T}_2 cluster amplitudes are correct through second-order. Thus, the E^{CCSD} is correct up to the third-order in the

fluctuation potential, while for the CCSDT model, E^{CCSDT} is correct through the fourth-order. [38] This analysis also reveals that in order to have a CC method with the energy correct up to the second-order, it is sufficient to calculate \hat{T} correctly up to the first-order in the fluctuation potential. The latter point and the CCSD unfavorable computational-cost scaling, $\mathcal{O}(\mathcal{N}^6)$, were the ground to simplify the doubles equation of CCSD, Eq. 7, and introduce the coupled cluster singles and approximated doubles method CC2 with an approximated doubles equation, [42]

$$\Omega_2 = \langle \mu_2 | \hat{H} + [\hat{F}, \hat{T}_2] | \text{HF} \rangle = 0 . \quad (10)$$

CC2 has a lower computational-cost scaling, $\mathcal{O}(\mathcal{N}^5)$, than CCSD with an energy correct through second-order (like MP2) in the fluctuation potential. In contrast to MP2, single excitations of CC2 are relaxed via a single set of coupled equations. [43] Owing to the latter, CC2 becomes a viable choice for calculation of excitation energies, excited-state gradients [44] and transition moments [45], which in combination with the resolution-of-the-identity approximation [27] can be calculated for molecules of the size of berenil very efficiently.

2.2 Excited-State Energies and Gradients of Second-Order

Methods

For the CC wavefunction ansatz, excitation energies ω^f are determined as the eigenvalues of the asymmetric Jacobian matrix \mathbf{A} where E^f and \bar{E}^f are right and left eigenvectors.

$$\bar{E}^f \mathbf{A} E^f = \omega^f \quad , \quad \mathbf{A}_{\mu_i \nu_j} = \frac{\partial \Omega_{\mu_i}}{\partial t_{\nu_j}} . \quad (11)$$

The CC2 Jacobian matrix reads as

$$\mathbf{A}_{\mu_i \nu_j}^{\text{CC2}}(T_1, T_2) = \begin{pmatrix} \langle \mu_1 | [\hat{H} + [\hat{H}, \hat{T}_2], \tau_{\nu_1}] | \text{HF} \rangle & \langle \mu_1 | [\hat{H}, \tau_{\nu_2}] | \text{HF} \rangle \\ \langle \mu_2 | [\hat{H}^{\text{QM}}, \tau_{\nu_1}] | \text{HF} \rangle & \langle \mu_2 | [\hat{F}, \tau_{\nu_2}] | \text{HF} \rangle \end{pmatrix} . \quad (12)$$

In the CC2 model, the singles equations are identical with those of the CCSD model, Eq. 6. However, if the only goal is to have a method that is just correct through second-order, some of the terms in the CC2 model exceed the criteria. In the CC2 Jacobian, Eq. 12, terms that are introduced via the similarity transformation with $\exp(-\hat{T}_1)$ contribute only in third and higher orders to the excitation energies. Looking back to the MPPT theory and assuming that the Brillouin condition is fulfilled, the only required information to calculate the second-order ground-state energy $E^{(2)}$ is obtained from the first-order wavefunction amplitudes [38]

$$|\text{MP1}\rangle = \hat{T}_2^{(1)} |\text{HF}\rangle \quad (13)$$

where \hat{T}_1 amplitudes are zero for the optimized reference state due to the Brillouin condition. The Jacobian of the CIS(D $_{\infty}$) model [41] is obtained by substituting the CC2 ground

state amplitudes with the MP1 amplitudes in Eq. 12

$$\mathbf{A}^{\text{CIS}(\text{D}_\infty)} = \mathbf{A}^{\text{CC2}}(0, T_2^{(\text{MP1})}) . \quad (14)$$

The excitation energies of the CIS(D_∞) model is also correct through second order of the fluctuation potential. The accuracy of CC2 is higher than CIS(D_∞), while the computational cost of CIS(D_∞) is lower than for CC2. The computational saving comes from the fact that for CIS(D_∞) there is no need to solve iterative equations for Lagrange multipliers. A more approximated second-order method is CIS(D), [46] which uses a perturbative treatment of electron correlation. The computational cost of CIS(D) is lower than CIS(D_∞), although it still scales with the fifth power of the system size. CIS(D) does not provide a systematic improvement over the CIS method for calculation of excited-state stationary geometries and behaves erratically in the case of near degenerate states. [47] Therefore, CIS(D) can not be a method of choice for berenil.

ADC(2) is another second-order method that is derived from the polarization propagator theory. [28] In the ADC approach, the exact excitation energies are obtained through solving a Hermitian eigenvalue problem [48]

$$\mathbf{M}\mathbf{X} = \Theta\mathbf{X}, \quad \mathbf{X}^\dagger\mathbf{X} = \mathbf{1} \quad (15)$$

where the exact excited state eigenfunction $|\Psi_n\rangle$ that is expanded in terms of a complete set of intermediate states $|\tilde{\Psi}_J\rangle$ [49]

$$|\Psi_n\rangle = \sum_J X_{Jn} |\tilde{\Psi}_J\rangle . \quad (16)$$

In Eq. 15, $M_{IJ} = \langle \tilde{\Psi}_I | \hat{H} - E_0 | \tilde{\Psi}_J \rangle$ is an element of the secular matrix \mathbf{M} , \mathbf{X} is the matrix of eigenvectors X_{Jn} and Θ_n is the excitation energy of excited state n . The ADC equation can be expanded based on the Møller-Plesset partitioning of the Hamiltonian,

$$\mathbf{M} = \mathbf{M}^{(0)} + \mathbf{M}^{(1)} + \mathbf{M}^{(2)} + \dots + \mathbf{M}^{(n)} , \quad (17)$$

and therefore, the eigenvalues of $\mathbf{M}^{(2)}$ give the second-order correction of excitation energies. The ADC(2) secular matrix $\mathbf{M}^{(2)}$ is found to be identical with the symmetrized CIS(D_∞) Jacobian, [50]

$$\mathbf{M}^{(2)} = \mathbf{A}^{\text{ADC}(2)} = \frac{1}{2}(\mathbf{A}^{\text{CIS}(D_\infty)} + (\mathbf{A}^{\text{CIS}(D_\infty)})^\dagger) . \quad (18)$$

Equations 14 and 18 show the mathematical relations between CC2, CIS(D_∞) and ADC(2) that proved to be useful in the development of new functionalities for these methods, which are also employed in the development of PE-ADC(2), cmp. chapter 7. In our PE-ADC(2) implementation, we used the available implemented routines of PERI-CC2 [34,51] as much as possible.

In comparison to CC2, the ADC(2) approach is simpler and due to its symmetric structure, it leads to some computational saving if both right and left eigenvectors are needed, Eq. 11. Similar to the CIS(D_∞) method, the iterative solution of Lagrange multiplier equations are not needed for ADC(2). These computational savings can be significant, in particular for the berenil project, where we require to evaluate excited-state molecular gradients for thousands of times in the calculation of intrinsic reaction coordinate pathways (*vide infra*). In contrast to CC2, for ADC(2) and CIS(D_∞) a gradient calculation has no iterative $\mathcal{O}(\mathcal{N}^5)$ scaling-cost step after identifying the eigenvalues, [50] which overall gives

ADC(2) a factor of 4 lower computational cost than CC2 for the calculation of excited-state gradients. Moreover, the non-symmetric structure of the CC2 Jacobian has a significant impact on the shape of the potential energy surfaces in the vicinity of the intersection between the same-symmetry states *e.g.*, for a conical intersection between the states with the same symmetry. Refs. 50, 52 clarify that for methods with non-symmetric Jacobian matrix *e.g.*, for coupled cluster methods, the potential energy surfaces in the vicinity of the same-symmetry intersections are unphysical. It is worth noting that this limitation is not universal for all coupled cluster models and there are Hermitian coupled cluster models that are potentially applicable for near excited-state potential energy crossings. [53] For instance, the similarity constrained coupled cluster model that is recently introduced by Kjøenstad and Koch, [54, 55] lifted this limitation for CCSD. However, the excited-state gradients of this method are not yet implemented and the applicability of the method still needs to be tested. Based on the above discussion, we selected ADC(2) as our method of choice to study the non-radiative deexcitation processes of berenil. In general, for studying the photophysical properties of single-reference organic molecules with a similar system size *i.e.*, for molecules with 20 to 50 atoms, ADC(2) is a competitive approach and in this thesis we have capitalized on that to study the excited-state processes and implemented the PE-ADC(2) excited-state gradients for studying the photophysics of this family of chromophores in complex molecular environments, *vide infra*.

In the following, few points have been highlighted that are related to the design of approximations that are essential for efficient implementation of polarization-correlation cross terms for the post-Hartree-Fock methods such as PE-ADC(2). In short, the resolution-of-

identity RI approximation and using of the doubles-amplitude-direct formulation are two important efficiency bringing factors that we tried to retain in the implementation of PE-ADC(2) excited-state gradients without significant lost in accuracy. A major bottleneck in large-scale MP2, CC2 and ADC(2), etc. calculations is the evaluation of four-index two-electron integrals. With the resolution-of-the-identity approximation, these integrals are approximated by

$$(\alpha\beta|\gamma\delta) \approx \sum_{QP} (\alpha\beta|Q)[V^{-1}]_{PQ}(P|\gamma\delta) \quad (19)$$

where $\alpha, \beta \dots$ refers to atomic orbitals (AOs) basis and matrix V contains two-center integrals and P and Q are referring to an auxiliary basis set. [40] Thus, a demand to save 4-index integrals on disk or memory has been removed and in return, instead of four-index integrals, simpler 3-index integrals have to be evaluated. The RI approximation does not reduce the fifth power scaling, but it reduces the computational cost by (depending on the size of the basis set and the number of electrons) one or two orders of magnitudes, while the introduced error is more than two orders of magnitudes smaller than the basis set error, if an optimized auxiliary basis set with the same quality as the basis set is used. [27] The machinery of implying RI- approximation is perfectly compatible with PE-ADC(2) without any programming effort. Therefore, in our implementations and calculations on berenil this well-established approximation has always been used.

The next important strategy to lift an I/O (input/output) bottleneck for making the code more efficient is to exploit the doubles amplitude-direct formulation of second-order methods. The doubles-doubles block of the ADC(2) secular matrix (and CC2 Jacobian),

Eq. 12 has a simple structure

$$\mathbf{A}_{\mu_2\nu_2} = \delta_{\mu_2\nu_2}\epsilon_{\mu_2} , \quad (20)$$

which is used to define an effective Jacobian matrix for CC2 (or an effective secular matrix for ADC(2))

$$\mathbf{A}_{\mu_1\nu_1}^{\text{eff}} = \mathbf{A}_{\mu_1\nu_1} - \sum_{\xi_2} \frac{\mathbf{A}_{\mu_1\xi_2}\mathbf{A}_{\xi_2\nu_1}}{\epsilon_{\xi_2} - \omega^f} . \quad (21)$$

This simple reformulation is used to avoid the storage of doubles amplitudes, which for large systems can create an I/O bottleneck [45]. In the efficient implementation of the QM/MM polarizable embedded second-order methods, [34] like PERI-CC2 and PE-ADC(2) (cmp. chapter 7), retaining the simple structure of the doubles-doubles block and the effective Jacobian formulation, Eq. 21, is essential and therefore few approximations that are inspired by MPPT have been made, [56] which in combination with the post-SCF reaction field scheme [34,57] provide an accurate, yet efficient, and simple framework for the implementation of excited-state gradients and molecular properties. For further details see chapters 3 and 7.

The excited-state (f) relaxed properties and gradients at the ADC(2) level are calculated from the orbital relaxed Lagrangian L^f

$$L^f = E^{\text{MP2}} + \bar{t}_{\mu_2}^f \Omega_{\mu_2} + \sum_{i,j=1}^2 \sum_{\mu_i\nu_j} E_{\mu_i}^f \mathbf{A}_{\mu_i\nu_j} E_{\nu_j}^f + \bar{\omega}(1 - \sum_{i=1}^2 \sum_{\mu_i} E_{\mu_i}^f E_{\mu_i}^f) + \sum_{\mu_0} \bar{k}_{\mu_0}^f F_{\mu_0} . \quad (22)$$

The first term in Eq. 22 is the ground state MP2 energy. The second term determines the Lagrange multipliers, the third term gives the ADC(2) excitation energy and the fourth term ensures the orthonormality of the eigenvector E^f . Here we remind that for ADC(2)

the right and left eigenvectors are identical. Since the molecular orbitals are the linear combination of atom-centered atomic orbitals, the derivatives of the orbitals with respect to nuclei positions have to be included in the Lagrangian, cmp. Eq. 22, to make sure that the orbitals, the integrals and also the Fock matrix F continue to be consistent. In Eq. 22, the orbital rotation constraint in the subspace $\bar{\kappa}_0$ is imposed by the $\bar{\kappa}_{\mu 0}^f$ Lagrange multipliers which are determined by solving the so-called coupled-perturbed Hartree-Fock equations. After solving the equations for Lagrangian multipliers \bar{t} and $\bar{\kappa}$, the excited state gradients can be evaluated in atomic orbital basis as

$$\begin{aligned} \left(\frac{dL^{\text{ADC}(2),f}}{dx}\right)_{x=0} &= \sum_{\alpha\beta} D_{\alpha\beta}^{\text{eff},ao} h_{\alpha\beta}^{[x]} + \frac{1}{2} \sum_{\alpha\beta\gamma\delta} d_{\alpha\beta\gamma\delta}^{\text{sep},ao} (\alpha\beta|\gamma\delta)^{[x]} - \sum_{\alpha\beta} F_{\alpha\beta}^{\text{eff},ao} S_{\alpha\beta}^{[x]} \\ &+ \sum_{\alpha\beta P} \Delta_{\alpha\beta}^{\text{ao},P} (\alpha\beta|P)^{[x]} - \sum_{PQ} \gamma_{PQ} V_{PQ}^{[x]} \end{aligned} \quad (23)$$

where x is the x -component of a nuclear coordinate. Detailed expressions of the AO and auxiliary densities are given in Ref. 50.

On the evaluation of the performance of the second-order methods, few words worth to be mentioned. First of all, the methods that are considered here are single-reference methods, which implies that the single-determinant Hartree-Fock wavefunction should already be an accurate zeroth-order approximation. [38] Most of the organic chromophores are single-reference systems; however, in some cases the single-determinant Hartree-Fock wavefunction is not adequate and thus multi-reference methods are needed. [58] In particular importance for the berenil project, the electronic structure of molecules at the conical intersection of the singlet ground and lowest electronic excited state, S_0 and S_1 , have a multi-reference character. Therefore, to locate the S_0/S_1 intersection seam one has to use

a multi-reference method. [59] Hence, for the molecules with a S_0/S_1 state crossing *e.g.*, benzenil, the ADC(2) relaxed geometries of the S_1 state are much more qualitative and thus, we can only determine regions that are close to conical intersections, *cmp.* chapter 6.

Furthermore, the perturbation analysis argument itself is not always conclusive. [38] For instance, it is not true that a third-order method *e.g.*, CCSD is always more accurate than a second-order method. The accuracy of CC2, CCSD, and CASPT2 have been benchmarked against CC3. [60,61] It turns out that in most of the cases, for $\pi\pi^*$ and $n\pi^*$ excitations, the CC2 excitation energies are closer to the reference values than the CCSD and the CASPT2 methods (due to a favourable error cancelation of higher-orders); but in the case of Rydberg excitations and double-excitations CCSD performs better than CC2. [62,63] Also, a recent benchmark study clarifies that the ADC(3) overcorrects ADC(2), although ADC(3) computational cost scales with $\mathcal{O}(N^6)$. [64] Additionally, there are interesting trends between the second-order methods themselves. For ground state geometries, the MP2 method performs slightly better than CC2, although in the CC2 model the single excitations are allowed to be relaxed. [65] Therefore, the main application of CC2 is the calculation of excitation energies and excited-state properties. [27,42,65] Moreover, various benchmark studies for small and medium-sized chromophores indicate that the performance of CC2 and the more approximated ADC(2) method on the calculation of excitation and emission energies are about the same. [50,66]

An attractive alternative for post-Hartree-Fock methods to tackle the electron correlation problem is Density Functional Theory. [67–71] In particular, the Time-Dependent Density Functional Theory, TD-DFT, gives another choice to obtain results that are more accu-

rate than those of Time-Dependent Hartree-Fock, although most of the benchmark studies show that TD-DFT is not as accurate as CC2 or ADC(2). [66, 72–74]. Thanks to its lower computational cost, TD-DFT is the method of choice in many computational studies of electronically excited states. However, among the hybrid TD-DFT functionals, only those that have 100% exact exchange (for the long-range interaction) are able to describe the asymptotic behavior of charge-transfer excitations correctly. [75, 76] For the excited-state dynamics, TD-DFT with different functionals have been compared with second-order methods and it reveals that in particular for small or medium sized single-reference molecules like berenil, ADC(2) should be the method of choice to study excited state dynamics if it is possible. [77] For ADC(2), different mean absolute errors (between 0.08 to 0.23 eV) were reported for the calculation of excitation energies and/or adiabatic emission energies in various molecular systems. [64, 66, 72, 73] Therefore, for the applications that require highly accurate computational predictions, ADC(2) may not be the method of choice due to its finite accuracy.

In this chapter, the basic theoretical framework for the calculation of excitation energies and excited-state gradients of the second-order methods and some of the practical considerations to have an efficient implementation of them are mentioned. In the next chapter, the basic principles of QM/MM methods and the main considerations about the implementation of PE-ADC(2) excited-state gradients will be given.

3 Quantum Mechanics/Molecular Mechanics Models

In hybrid QM/MM methods, [29] the total system is partitioned into distinguished active and inactive subsystems: solute and solvent, ligand and receptor, guest and host, to name a few. The idea is to describe the active part with a proper quantum mechanical method and treat the inactive part with a more approximated method. An important ingredient is the coupling between the subsystems $E^{\text{QM/MM}}$. Technically, for ground state, it is possible to start from the full quantum mechanical treatment of the subsystem interactions and derive an expression for the QM/MM interaction energy by applying a series of well-defined approximations. [78] However, for excited states, as we discussed early and as pointed out in the paper by Köhn and Lunkerheimer [79], as well as List, Norman, Kongsted and Aagaard [80], there are additional aspects and problems, as soon as more states come into play. The physical background of the QM/MM partitioning scheme has been recently reviewed. [32] Here, I mention only one of the fundamental assumptions of the QM/MM calculations. On the way to derive the QM/MM approximation from the full-QM treatment of subsystem interactions, it is assumed that the density of two quantum subsystems are not overlapping; therefore, the localized density of one of the subsystems can be approximated by a multipole expansion, *vide infra*. In consequence, when two subsystems are getting closer, the QM/MM approximation deviates more from reference supermolecular calculation. [81] Furthermore, due to the lack of Pauli repulsion, the electrons can spill out from the QM subsystem and leak to the MM subsystem and, thus, the

QM/MM approximation starts to break. For the same reason in the excited state calculations where electrons are approaching the outer electronic shells, keeping the QM/MM approximations under control requires special treatments. In chapter 8, we introduce the PE(ECP) approach to model the Pauli repulsion effect of the MM subsystem.

3.1 A Hierarchy of QM/MM Methods

The interaction energy $E^{\text{QM/MM}}$ of subsystems can be decomposed into physically well-defined contributions including the interaction between the static molecular densities that are usually replaced by a multipole expansion E^{es} , the polarization interaction between induced and total moments E^{pol} and the non-classical van der Waals dispersion and Pauli repulsion interactions, which are usually treated together in E^{vdW}

$$E^{\text{QM/MM}} = E^{\text{es}} + E^{\text{pol}} + E^{\text{vdW}} . \quad (24)$$

In the simplest QM/MM calculations, the interaction energy between QM and MM subsystems (QMsub and MMsub in Eq. 25) is included into the total energy of the system $E_{\text{Sys}}^{\text{Tot}}$ at the same level as the MM subsystem via a subtractive scheme

$$E_{\text{Sys}}^{\text{Tot}} = E_{\text{QMsub}}^{\text{QM}} + E_{\text{Sys}}^{\text{MM}} - E_{\text{MMsub}}^{\text{MM}} . \quad (25)$$

Within this framework, one can readily combine every quantum chemistry method with any other QM or MM approach, which can be the method of choice for many applications. [30] This approach has been expanded to incorporated arbitrary number of QM or MM layers within a general subtractive scheme, which is called ONIOM (Our own N-layered Integrated

molecular Orbital and molecular Mechanics). [82] The disadvantage of this scheme is that the densities of the subsystems do not interact with each other and the coupling is included only via the energy expression (Eq. 25) *i.e.*, at a mechanical level.

In the additive QM/MM scheme,

$$E_{\text{Sys}}^{\text{Tot}} = E_{\text{QMsub}}^{\text{QM}} + E_{\text{MMsub}}^{\text{MM}} + E^{\text{QM/MM}} , \quad (26)$$

the interaction energy of the subsystems is calculated in the quantum mechanical calculation. Using the additive scheme, and at the level of electrostatic embedding, the interaction between the electrons of the QM subsystem and the permanent electrostatic potential of the MM subsystem (environment) is coupled to the quantum mechanical calculations via an one-electron coupling operator \hat{G}^{es} , *vide infra*. At this level, molecules in the environment are usually approximated by a (distributed) multipole expansion, $\sum_{m=1}^M \sum_{k=0} \mathbf{Q}_m^k$, where M is the number of MM sites and k is the multipole rank. [83] Most of the QM/MM calculations are done at this level with truncating the multipole expansion immediately after $k = 0$, which is called point-charge embedding. [30, 78] The integrals that are needed to describe the interaction of the QM subsystem electrons and the point-charges are identical with the nuclei-electron integrals that are available in almost all quantum chemistry programs. However, point-charge embedding is only accurate for the long-range interactions and have limited resolution to capture the anisotropic features of the electron density. [84] At the multipole-embedding level, the point multipole expansion is truncated after a desired multipole rank *e.g.*, after $k = 1$ for dipole embedding and $k = 2$ for quadrupole embedding. In practice, the higher-order multipole interaction integrals can be calculated

recursively. However, the extra computational cost to calculate higher-order distributed multipoles and the limited corresponding gain in accuracy has been found discouraging in some cases. [85] In most of the polarizable embedding calculations, which is a more complete level of embedding, *vide infra*, the point multipole expansion is truncated after quadrupoles. [32]

Based on the long-range perturbation analysis, after the permanent electrostatic interaction, polarization E^{pol} is the most important contribution that should be included in the calculation of $E^{\text{QM/MM}}$, Eq. 24. [86] The polarizable QM/MM level is a more complete QM/MM level of embedding, which includes the polarization interaction of the QM and MM subsystems into account. Fluctuating charge, Drude oscillator and induced point dipoles are commonly used models to include polarization. Our PE-ADC(2) method is a polarizable embedding QM/MM approach, which uses the induced point dipole model. Redistribution of charges to equalize the electronegativity of atoms in a molecule and oscillation of some off-center partial charges are principles of fluctuating charge and Drude oscillator models, respectively. [87] Recently, the mathematical relation between these models have been discussed by different groups. [88–90]

In the dipole polarizable embedded QM/MM approach [91–93], a dipole polarizability α is allotted to each MM atom and the QM/MM polarization energy can be calculated as

$$E^{\text{pol}} = -\frac{1}{2} \sum_{uv} \left(F_u^{\text{ne}} + \sum_{pq} \varepsilon_{pq}^u D_{pq} \right) R^{uv} \left(F_v^{\text{ne}} + \sum_{pq} \varepsilon_{pq}^v D_{pq} \right) \quad (27)$$

where $\mathbf{R} = (\alpha^{-1} - \mathbf{T}_2)^{-1}$ is the symmetric response matrix. In Eq. 27, p and q are molecular orbital indices, D_{pq} is the electron density of the QM subsystem, u and v are multipole sites

of the MM subsystem, ε_{pq}^u is the multipole interaction integral and F^{ne} is the non-electronic static field. Accordingly, the polarization contribution to the reaction field operator \hat{G}^{pol} , which explicitly depends on the density of the QM subsystem, \mathbf{D} , is

$$\hat{G}^{pol} = \hat{G}^{pol}(\mathbf{D}) = - \sum_{uv} \left(F_u^{ne} + \sum_{pq} \varepsilon_{pq}^u D_{pq} \right) R^{uv} \hat{\varepsilon}^v \quad (28)$$

where $\hat{\varepsilon}_{pq}^u$ is the multipole interaction operator. Finally, the sum of static and polarization contribution can be named the "reaction field operator", $\hat{G}(\mathbf{D}) = \hat{G}^{es} + \hat{G}^{pol}(\mathbf{D})$. [33, 94]

In the QM/MM calculations with a correlated method, the correlation-polarization cross terms should be included in the QM/MM reaction field for both ground and excited states. Accurate calculation of the polarization contribution and, in general, the QM/MM interaction energy, usually pays off since a QM/MM calculation with a small QM subsystem can provide an accuracy close to supermolecular calculations and, thus, unnecessary large QM subsystems can be avoided. [95] In my PhD project, the goal was to investigate the relaxation mechanism of berenil@DNA, which requires an efficient implementation of QM/MM ADC(2) excitation energies and excited-state gradients. The latter implied that the evaluation of QM/MM correlation-polarization coupling terms should not introduce any new computational bottleneck compared to an efficient implementation of ADC(2). In other words, computation of the new terms should not introduce an extra loop over the most time-consuming steps of the calculations and do not dismantle the efficiency-bringing factors of the second-order methods that are discussed in chapter 2. In our implementation of PE-ADC(2) gradients, the post-self-consistent field (post-SCF) scheme has been used and the polarization-correlation cross terms are approximated through second-order in the

fluctuation potential. The post-SCF scheme, together with the same approximations for polarization-correlation cross terms has been employed in Ref. 34 for the implementation of PERI-CC2 excitation energies, one-photon absorption oscillator strengths and in Ref. 51 for two-photon absorption cross sections. Another closely related approach has been introduced by Caricato for the calculation of CCSD correlation-polarization coupling terms within a polarizable continuum solvation model. [96,97] In the following, the post-SCF reaction field scheme is compared with other related schemes that are used in the implementation of PE-CC methods. [94]

3.2 Post-SCF Polarizable Embedding Reaction Field

For any post-SCF (*i.e.*, post-HF and post-KS) method, there are different possibilities to calculate the coupling of the polarization and wavefunction parameters, which differ in costs, accuracy and complexity of the implementation. [98–100] For the mean-field method PE-HF, the fully self-consistent polarized molecular orbitals are obtained as solution of the Hartree-Fock equations through the SCF algorithm for the polarizable embedded Fock matrix F^{PE} [83]

$$F_{ai}^{\text{PE}} = F_{ai}^{\text{QM}} + G_{ai}(\mathbf{D}^{\text{HF}}) = 0 \quad (29)$$

where $G(\mathbf{D}^{\text{HF}})$ is the reaction field operator calculated from the HF density. My starting point for the PE-ADC(2) gradients was to implement the PE-HF gradients in TURBO-MOLE by following the implementation of Ref. 101. At this level, since F^{QM} depends also on the density, the reaction field $G(\mathbf{D}^{\text{HF}})$ and the density \mathbf{D}^{HF} can be determined with-

out much complications in a fully self-consistent manner. Furthermore, for PE-DFT the effective Kohn-Sham operator is identical to the F^{PE} in PE-HF. [102] Since in TURBO-MOLE all the relevant routines are shared, the PE-DFT gradients were also automatically implemented. The polarized orbitals and the polarization consistent Fock matrix can be used in a vacuum-like post-Hartree-Fock calculation in which the correlated part of the wavefunction is only coupled to the environment by the ground state mean field. [83] PTE, perturbation in energy, is a general term used for the latter self-consistent reaction field scheme originally for polarizable continuum method, [99] which also cover QM/MM and QM/QM methods. The advantage of the PTE scheme is its simplicity. It can easily be coupled to various correlated wave-function-based methods with a minor programming efforts. However, at the PTE level the induced moments are kept fixed in the post-HF step. In consequence, in the calculation of excitation energies, the environment can not adapt, which causes a significant loss in accuracy. [103] Very recently, a perturbative reaction field scheme is advocated by the Dreuw's group, which includes perturbative corrections to the PTE excitation energies. [104, 105] The second implementation of PE-ADC(2) excitation energies that is published after our work is implemented within the latter scheme. [106] However, this perturbative treatment can not be used for potential energy surfaces since a non-degenerate perturbation theory gives inconsistent, qualitatively wrong potential energy surfaces in the region of avoided crossings, which are omnipresent for polyatomic molecules . [57] A more complete but also costly reaction field scheme is the perturbation in energy and density method, PTED. In this scheme, the induced dipoles, molecular orbitals and the density of the QM system at the given excited state are determined self-consistently in

a macro self-consistent loop. The PTED-COSMO-ADC(2) (COSMO stands for conductor-like screening model [107]) method have been implemented by Lunkenheimer and Köhn for excitation energies and one-photon transition matrix elements. [79] Although the PTED scheme is considered to be the most complete coupling scheme, it introduces a loop over the correlated calculations, which increases the computational cost significantly. Furthermore, calculation of gradients within this scheme is complicated due to the complex loop structure.

The post-SCF reaction field scheme was formulated in Ref. 34 and combined for the second-order method CC2 with an approximation for the density in the correlation-polarization cross terms, which discards some higher-order contributions to these terms. These approximations avoid the necessity of storing double excitation amplitudes, which as explained in the chapter 2 is essential to retain as much as possible the computational efficiency of second-order methods. With these approximations, the contributions of correlation-polarization cross terms are included to an order consistent with the order of QM approach while maintaining the QM/MM methods as efficient as possible. [34, 56] In the post-SCF reaction field scheme, a post-SCF calculation *e.g.*, PE-CC2, PE-ADC(2) etc. starts from a fully self-consistent PE-HF calculation and adds the reaction field contributions $G(\mathbf{D}^\Delta)$ from the post-SCF contributions of the density ($\mathbf{D}^\Delta = \mathbf{D}^{\text{CC}} - \mathbf{D}^{\text{HF}}$); but in contrast to PTED, without a feed-back loop to the PE-HF calculation. Furthermore, the post-SCF contributions to the reaction field are approximated based on a perturbation expansion of the coupling operator through the orders of the fluctuation potential. As it was mentioned in the previous chapter, these approximations are essential for an efficient implementation

of polarizable embedded RI-CC2 and RI-ADC(2) etc., [34, 56] which for MP2 reduce to the PTE scheme. To realize the latter approximations, by introducing the normal order operators $\hat{O}_N = \hat{O} - \langle \text{HF} | \hat{O} | \text{HF} \rangle$, the non-approximated PE-CC coupling operator \hat{G}_N^Δ can be expanded with respect to the correlated density of the QM subsystem as

$$\hat{G}_N^\Delta = \hat{G}^\Delta(\mathbf{D}^{(1)}) + \hat{G}^\Delta(\mathbf{D}^{(2)}) + \dots \quad (30)$$

where $\mathbf{D}^{(1)}, \mathbf{D}^{(2)} \dots$ are the MPPT corrections to the density. To arrive at a consistent implementation for n th-order energy, the interaction energy should also be correct up to the n th-order and all the terms that include \hat{G}_N^Δ components and correcting energy in a higher order than n can be approximated. Prior to my work, an efficient implementation of the PERI-CC2 excitation energies, linear and quadratic response properties *e.g.*, one- and two- photon absorption cross-sections were reported. [34, 51] Also in parallel to this project, the Post-SCF coupling scheme has been used in the implementation of COSMO-ADC(2) excited-state gradients [57] and the magnetic circular dichorism spectra at the COSMO-CC2 level. [108] In summary, the post-SCF reaction field scheme can be considered as an applicable choice for efficient calculation of energy, gradients and various response properties of ground and excited states.

In the current work, the post-SCF reaction field scheme is used with the same approximation for the correlation-polarizable cross terms for the implementation of the ground state PE-MP2 gradients and PE-ADC(2) excited-state energies and gradients as is explained in detail in chapter 7. For MP2, which is the ground-state method that is consistent with ADC(2), the post-SCF reaction field scheme with these approximations for the correlation-

polarization cross terms becomes identical with PTE-MP2. This is in agreement with Ángyan's analysis, who showed that the differences between PTE-MP2 and PTED-MP2 enter only in fourth-order in the correlation energy and, thus, PTE-MP2 is consistent with the theory behind MP2. [109]

In chapter 2, we have discussed that the ADC(2) excited-state gradients (cmp. Eq. 23) are calculated by differentiation of the ADC(2) relaxed Lagrangian. For the calculation of PE-ADC(2) excited-state gradients the explicit PE contribution to PE-ADC(2) have been identified (cmp. chapter 7), which after solving the equations for ground state Lagrange multipliers and the coupled perturbed Hartree-Fock equations reads as

$$\begin{aligned} \Delta L^{\text{PE},f} = & -\frac{1}{2} \sum_{uv} \left(F_u^{ne} + 2F_u^{elec}(\mathbf{D}^{eff}) - F_u^{elec}(\mathbf{D}^{\text{HF}}) \right) R_{uv} \left(F_v^{ne} + F_v^{elec}(\mathbf{D}^{\text{HF}}) \right) \\ & - \sum_{uv} F_u(\mathbf{D}^\xi) R_{uv} F_v(\mathbf{D}^\eta) + \sum_{pq} \mathbf{D}_{pq}^{eff} \sum_u Q_u^{es} \varepsilon_{pq}^u + E^{ne} \end{aligned} \quad (31)$$

where \mathbf{D}^{eff} is the MP2 effective density. \mathbf{D}^η and \mathbf{D}^ξ are two auxiliary densities that include the effects of excitonic coupling term.

$$\mathbf{D}^\eta = \sum_{\nu_1} \frac{\partial \mathbf{D}^\Delta}{\partial t_{\nu_1}} E_{\nu_1}^f, \quad \mathbf{D}^\xi = \sum_{\mu_1} E_{\mu_1}^f \frac{\partial \mathbf{D}^\Delta}{\partial t_{\mu_1}}. \quad (32)$$

Thus, for PE-ADC(2) gradients the terms that originate from $\Delta L^{\text{PE},f}$ are:

$$\begin{aligned} \frac{d\Delta L^{\text{PE},f}}{dx} = & \sum_{pq} D_{pq}^{eff} \sum_u Q_u^{\text{HF}} \frac{d\varepsilon_{pq}^u}{dx} + \sum_{pq} D_{pq}^{\text{HF}} \sum_u \left(Q_u^{eff} - Q_u^{\text{HF}} \right) \frac{d\varepsilon_{pq}^u}{dx} \\ & + \sum_u Q_u^{eff} \frac{d\varepsilon_{\text{nuc}}^u}{dx} + \frac{dE^{\text{LJ}}}{dx} + \sum_{pq} D_{pq}^\eta \sum_u Q_u^\xi \frac{d\varepsilon_{pq}^u}{dx} + \sum_{pq} D_{pq}^\xi \sum_u Q_u^\eta \frac{d\varepsilon_{pq}^u}{dx}, \end{aligned} \quad (33)$$

which has to be added to the vacuum-like contributions (Eq. 23). These terms have been implemented by me in the `ricc2` code of TURBOMOLE package. For the implementation

of the final expression (Eq. 33), I had to generalize the integral routines to calculate the derivatives of the multipoles interaction integral, which were in the beginning only available for point charges, add the contraction of the densities and multipole moments with the derivative integrals and implement the Lennard-Jones contribution. For further details about the intermediate steps see chapter 7.

In this chapter, the hierarchy of QM/MM approaches, the role of polarization in the calculation of excited-state energies and gradients and different reaction field schemes that are designed for wavefunction-based methods (cmp. section 3.2) have been briefly reviewed. In the implementation of polarizable embedded post-SCF methods, it is important to consider the main efficiency and accuracy factors of the parent quantum chemistry method and care must be taken not to introduce a computational bottleneck or an accuracy destroyer. Our implementation of PE-ADC(2) excited-state gradients that is reported in chapter 7 is in line with this way of thought. In the same chapter, the performance of polarizable embedded QM/MM PE-ADC(2) in the calculation of excited-state optimized geometries and fluorescent emission energies are investigated against the full-QM ADC(2) calculations for two exemplary supramolecular systems.

4 Modelling of Excited-States Relaxation Pathways

Ultrafast time-resolved spectroscopic techniques provide a gentle way to probe dynamics of transient molecular species, transition states and reactive intermediates with a femtoseconds time-resolution. [110] The excited-state ultrafast relaxation of berenil in the electronically excited states have been studied in the group of Professor Nürnberger by means of the time-resolved fluorescence up-conversion spectroscopy, cmp. chapter 5. In the fluorescence up-conversion technique, the relaxation processes are monitored through the changes of the emitted light intensity. [111] The kinetics of an excited-state relaxation process can be explained by depletion of molar concentration of an electronically excited molecule at a transient conformation $[M^*]$ as a function of time,

$$[M^*]_t = [M^*]_0 \cdot \exp\left(-\frac{t}{\tau_F}\right) \quad (34)$$

where τ_F is defined as the life-time of the process. The time-resolved spectra of a multi-process relaxation can be obtained with a more complicated delay function to model the time-dependent signal intensity, [111]

$$F(t) = \sum_{i=1}^n A_i \exp\left(-\frac{t}{\tau_i}\right) . \quad (35)$$

The number of involved processes n in the excited-state relaxation of berenil and their relative life-times τ_i for each process in different molecular environments are the spectroscopic data that we aimed to understand in our study. In Eq. 35, A_i is the amplitude of process i . Fluorescence life-times depend on the mechanisms of the relaxation processes and

also the possible interactions with the molecular environment *e.g.*, [110,112] polarization and reorganization of solvent molecules, [113,114] (de)formation of hydrogen bonds, [115] biomolecular binding and molecular confinement. [116] Therefore, a single exponential decay may represent many processes with close life-times that could possibly be decomposed into several processes if high resolution data are available. [111] For berenil a two-phase relaxation process has been observed by our collaborators both in aqueous solution and in complexation to biomolecules. Both processes are ultrafast with the time delays that are in sub-pico second regime. In addition, both relaxation processes show insignificant response to viscosity of solvent and to the molecular structure of the environment. In the following, the assessment of intrinsic reaction coordinates calculation, which is used in chapters 5,6 and 9 for studying berenil excited-state relaxation processes and the possible intramolecular processes for *trans-cis* isomerization, are discussed. Furthermore, the possible *trans-cis* torsion rotation mechanisms are discussed for berenil.

4.1 Intramolecular Relaxation Pathways from Intrinsic Reaction Coordinates

Within the Born-Oppenheimer approximation, [26] reactions are large amplitude motions (in comparison to vibrations) of nuclei on the potential energy surface of the ground or excited states between local minima. In the Born-Oppenheimer molecular dynamics simulations, reaction mechanisms can be extracted from the real-time propagation of the system and statistical averaging of the trajectories ensemble. [117] Unfortunately, *ab initio*

molecular dynamics calculations beyond the DFT level are computationally expensive and thus it is difficult to obtain at this level a statistically converged ensemble of trajectories, which restricts the application of these techniques to reactions with low activation energies of few kcal·mol⁻¹. For this reason, many techniques have been developed to accelerate molecular dynamics simulations of reactions — sometimes with the help of additional assumptions. [118, 119] For instance, in the meta-dynamics technique [119], it is expected to describe a reaction by a few already known collective variables that represent the reaction coordinates. [120] However, selection of optimal collective variables is not a trivial task. [121] Therefore, for studying excited-state relaxation pathways, which is costly for berenil, we decided to calculate and analyze *ab initio* intrinsic reaction coordinates. Intrinsic reaction coordinates can be used also in the analysis of molecular dynamics and also for enhanced calculations of dynamic reaction coordinates. [122, 123]

In 1970, [124] Fukui proved that the gradient of the potential energy is extremal along the mass-weighted steepest descent path s . Thus, the coordinates $\mathbf{x}(s)$ of the so-called intrinsic reaction coordinates (IRC) are the solution of the following differential equation

$$\frac{d\mathbf{x}}{ds} = -\frac{\mathbf{g}(\mathbf{x})}{|\mathbf{g}(\mathbf{x})|} \quad (36)$$

where $\mathbf{g}(\mathbf{x})$ is the gradient at position x . It can be shown that the vibrational and rotational motions are completely excluded from the IRC pathway. [125] Therefore, the essence of atomistic details of a reaction mechanism can be extracted from IRCs without a need to analyze a large ensemble of trajectories. [126] Usually, an IRC calculation starts from a slightly distorted transition-state geometry (along the normal mode with imaginary eigen-

value) or from the Franck-Condon point of an excited state by evaluating the classical equation of motion for a given small time step. [127] In this algorithm, the initial velocities are set to zero at every step. The first IRC calculations have been done in a pioneer work of Morukama in 1977 with a two-step algorithm to reduce the number of gradient calculations. [128] Other algorithms were also introduced to reduce the number of steps and/or to provide an accurate reaction path, which is used in modeling of reactions kinetics. [129,130] The intrinsic reaction coordinate (IRC) analysis is carried out extensively in quantum chemistry calculations to confirm that a transition state belongs to a certain reaction. [126] For the berenil project, the excited-state IRC calculations were used to characterize different phases of excited-state relaxation process, which are also identified based on time-resolved fluorescence spectroscopy. [25] Also, by projecting the obtained IRC pathway to the internal coordinates *e.g.*, bonds, torsions and angles, we clarified the detailed mechanism of the excited-state relaxation, *vide infra*. Analysing of IRCs in terms of linearly-independent internal coordinates helped us to understand the mechanism for ground- and excited-state (photo)isomerization of berenil, cmp. chapter 6. Furthermore, in chapter 6, the calculated IRC pathways were employed to interpret transient absorption spectra. In the same chapter different geometrical indices, excitation energies and one-photon transition matrices have been calculated along the ground- and excited-state IRC pathways, which are used to characterize the time-resolved transient absorption signals. Moreover, calculation of excited-state IRC pathways for berenil@DNA with our PE-ADC(2) method was employed to understand the role of DNA-confinement and surrounding water molecules into the volume-conserving relaxation of berenil@DNA, cmp. chapter 9.

However, in practice, there are several pitfalls in the IRC approach. The optimal free energy pathways might be different from the IRC paths considerably due to the impact of molecular collisions or the extra thermal kinetic energies. [126] This can happen in flexible systems *e.g.*, in the case of supermolecular systems and in reactions with curvy reaction paths. In addition, for cases like our studies of the relaxation pathways of berenil, the so-called valley-ridge inflection may happen. [131,132] The IRC calculations that do not start from (the vicinity of) a stationary point *e.g.*, a ground- or excited-state Franck-Condon point, can lead to a transition state that connects two minima. In these cases, it is possible that two IRC calculations with a slightly different setup (different starting geometries or different time-steps, etc.) lead to completely different results due to the choice of the starting geometries or through contributions of molecular vibrations that are perpendicular to the IRC tangent vector. Typically, it is possible that an IRC calculation converges to a local minima without meeting the transition state, at all. [126] The IRC calculations that are reported in chapter 9 for S_1 excited state of berenil@DNA converged to one of the S_1 minimum geometries directly. Whether the latter is an example of valley-ridge inflection or a physical effect of the DNA-minor-groove confinement is under investigation.

In essence, the entropic contributions are neglected in the IRC analysis approach. This is the price that must be paid to avoid expensive ensemble averaging. However, one should consider that our goal is not to reproduce the time-dependent fluorescence up-conversion and/or transient absorption spectra directly. Rather, we are interested in gaining a qualitative understanding about the time-resolved data, which are obtained by the global fitting of the time-dependent spectra with assuming that the excited-state deexcitation comprises of

a finite number of processes, which is also bound to the resolution of the instrument. [111] Therefore, and based on the results that are discussed in the next coming chapters for berenil, the IRC analysis approach can be considered as a cost-effective approach for complementary computational studying of time-resolved fluorescent and transient absorption data.

4.2 Volume-Conserving Torsions Relaxation Pathways

Many organic chromophores consist of π -systems, which upon photoexcitation can form a $\pi\pi^*$ excitation. For these systems, the excited-state relaxation process often comprises a torsion twist around the broken π -bond to diminish the unfavorable π^* antibonding interaction, which is the general mechanism of rather well-known light-induced *trans-cis* isomerization. [19] This photo-induced pure rotation mechanism *i.e.*, torsional motion around a single bond, is the design principle of some of the molecular motors. [133] The pure torsional rotation around a bond between two sizeable fragments is accompanied by spatially large amplitude motions of the chromophore and is then in solution coupled to solvent relaxation, and thus, its life-time depends on the viscosity, polarity and other properties of the solvent. [134] In a (bio)molecular confinement or in a viscous solvent, the pure torsion relaxation pathway is hindered; and hence, can be characterized by a delay in the relaxation. [116] In my master thesis, we observed that the N=N double bond of berenil is halfway (clockwise or anti-clockwise) twisted in the two minima of the lowest electronically excited state S_1 . [25] Our colleagues in the group of Professor Nürnbergger also observed

that the excited-state relaxation process has very little sensitivity to the viscosity of the solvent and also to embedding by biomolecular binding molecules, *e.g.* in binding to the minor-groove of AT-rich DNA. This secular response to solvent viscosity discards pure rotation for the excited-state relaxation mechanism of berenil and thus we should consider other excited-state relaxation mechanisms that are possible for berenil.

The ultrafast deexcitation dynamics of berenil in complexation to biomolecular environment and its low sensitivity to solvent viscosity hint that berenil has a volume-conserving relaxation mechanism. In a volume-conserving mechanism, the nuclei motions are spatially localized and thus the relaxation process can take place without large amplitude motions of bulky groups on the excited- or ground-state potential energy surface. These properties have also been observed for the photoswitching molecule azobenzene, [135] the green fluorescent protein chromophore, [136] the photoactive yellow protein [137,138], rhodopsin [139] that is involved in the vision process and its mutants [140], to name a few. The excited-state relaxation processes of all the latter mentioned cases have been called volume-conserving, despite the structural and the mechanistic differences. As it is apparent from the above examples, here we emphasize that having a volume-conserving relaxation mechanism does not necessarily go with non-radiative deexcitation.

For the chromophores that exhibit an excited-state torsion twist (usually via a photo-induced $\pi\pi^*$ excitation), the volume-conserving relaxation motions compete with the pure torsion mechanism. Some of these volume-conserving motions are called torsional motions, [141] which comprise synchronous rotation around multiple bonds. The Hula-twist [142] and the bicycle-pedal rotation [139] comprise synchronous rotations around two

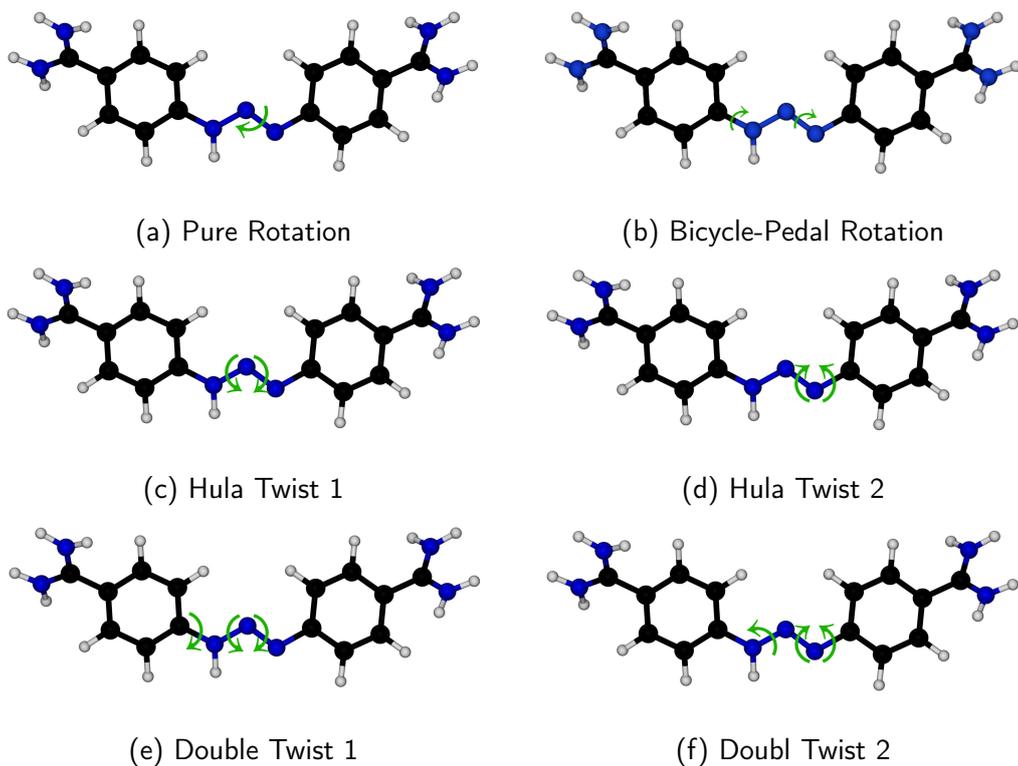


Figure 2: Possible, volume-conserving torsional motions for berenil

adjacent or decussate bonds, [141] respectively. Double twist [14] and folding table [143] mechanisms are other volume-conserving isomerization mechanisms that incorporate synchronous rotations of three, in the given order, adjacent or decussate bonds. For berenil, in addition to pure rotation, bicycle-pedal mechanism with clock- or anti-clockwise rotation around the double N=N bond can lead to two S_1 minima that are connected via a S_1 transition state. Two Hula twist and two double twist mechanisms are other possible relaxation pathways for berenil. These mechanisms are given in figure 2.

The Hula-twist motion can be explained as an out-of-plane or into-the-plane flip of one atom, which is identical with disrotatory rotation around two adjacent bonds, figure 2c and

2d. The bicycle-pedal mechanism is somewhat more complicated. Imagine a bicycle gear stays fixed on the bond that connects (connecting bond) two decussate rotating bonds, figure 2b. Each pedal of the bicycle models one of the rotating bonds, which rotates around the bicycle gear axis *i.e.*, when one of the decussate rotating bonds is moving toward the paper plane the other bond synchronously moves upward with respect to the paper plane. In figure 2b, the arrays refer to the rotation of bicycle pedals and not to the rotation of molecular torsions. In a perfect bicycle pedal type mechanism, the dihedral angle of connecting bond remains unchanged. Projection of bicycle-pedal motion to internal coordinates gives a disrotatory rotation around the decussate bonds, which in the IRC calculations, have been observed for berenil, cmp. chapters 5 and 6.

For berenil the two-phase excited-state relaxation process consists of a N=N bond elongation and a bicycle-pedal motion. Both of the excited-state relaxations are volume-conserving and thus berenil exhibits an ultrafast relaxation in solutions and also in biomolecular confinement, cmp. chapter 5. Interestingly, both bicycle-pedal and Hula-twist relaxation mechanisms are involved in the photoisomerization process of berenil. In the excited state, the relaxation process starts from the Franck-Condon point of the ground-state dominant conformer *E-azo-s-trans* and converges to a S_1 minimum, which continues on the ground state potential energy surface to the *Z-azo-s-cis* conformer via a Hula-twist mechanism, cmp. chapter 6. It should also be noted that retaining a volume-conserving mechanism does not mean that the relaxation process is insensitive to molecular environment. Using the picosecond X-ray crystallography technique, Jung et. al. [137] uncovered that the volume-conserving bicycle-pedal isomerization pathway of the photoactive yellow

protein can be controlled via modification of hydrogen-bonded residues of the embedding protein of the chromophore. In line with the latter study, our PE-ADC(2) calculations reveal that the S_1 relaxation process of berenil in DNA-minor-groove confinement retains a bicycle-pedal character, which is also influenced by molecular environment via hydrogen-bonding interactions with water molecules of the first solvation shell, cmp. chapter 9.

5 Ultrafast Dynamics of a Triazene: Excited-State Pathways and the Impact of Binding to the Minor Groove of DNA and Further Biomolecular Systems

Reprinted with permission from *J. Phys. Chem. Lett.* 2017, 8, 1986–1992.
Copyright 2019 American Chemical Society.

This paper reports the results of our collaborative project with the Physical Chemistry Group of Professor Patrick Nürnbergger. The research question has been asked by the physical chemistry group who have carried out the experiments, spectroscopy measurements and the related data analysis and data representations.^{1 2}

My master project was also about the excited state relaxation of berenil. Most of the calculations that were reported in the paper have been done during my master studies and this part does not belong to my PhD. In the master thesis [25], it is reported that the relaxation mechanism has two phases, which is one of the experimental observations. However, the crucial data analysis to identify the relaxation mechanism and to understand the viscosity independent relaxation processes, and writing the paper was done during my PhD. In the writing process, two separated computational and experimental drafts were transferred to the first full version by Lena Grimmelsmann. She has done the introduction, the graphical abstract, and the preparation of the final manuscript. The main line of the

¹Article: <https://pubs.acs.org/doi/abs/10.1021/acs.jpcllett.7b00472>

²SI: https://pubs.acs.org/doi/suppl/10.1021/acs.jpcllett.7b00472/suppl_file/jz7b00472_si_001.pdf

paper was discussed between all the co-authors and they all contribute in the processes of preparing the final draft. Professor Nürnberger and Professor Hättig supervised us to shape the final discussion. The submission process has been done in the experimental group.



The screenshot shows the RightsLink interface. At the top left is the Copyright Clearance Center logo. To its right is the RightsLink logo. Further right are navigation buttons for Home, Create Account, and Help, along with an email icon. Below the Copyright Clearance Center logo is the ACS Publications logo with the tagline "Most Trusted. Most Cited. Most Read." The main content area displays the following information:

Title: Ultrafast Dynamics of a Triazene: Excited-State Pathways and the Impact of Binding to the Minor Groove of DNA and Further Biomolecular Systems

Author: Lena Grimmelsmann, Alireza Marefat Khah, Christian Spies, et al

Publication: Journal of Physical Chemistry Letters

Publisher: American Chemical Society

Date: May 1, 2017

Copyright © 2017, American Chemical Society

On the right side, there is a LOGIN button and a text box that reads: "If you're a copyright.com user, you can login to RightsLink using your copyright.com credentials. Already a RightsLink user or want to learn more?"

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

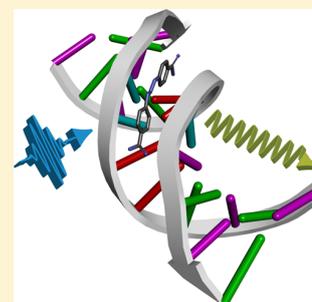
Ultrafast Dynamics of a Triazene: Excited-State Pathways and the Impact of Binding to the Minor Groove of DNA and Further Biomolecular Systems

Lena Grimmelsmann,[†] Alireza Marefat Khah,[‡] Christian Spies,[†] Christof Hättig,^{*,‡} and Patrick Nuernberger^{*,†}

[†]Physikalische Chemie II and [‡]Theoretische Chemie, Ruhr-Universität Bochum, 44780 Bochum, Germany

Supporting Information

ABSTRACT: Many synthetic DNA minor groove binders exhibit a strong increase in fluorescence when bound to DNA. The pharmaceutical-relevant berenil (diminazene aceturate) is an exception with an extremely low fluorescence quantum yield (on the order of 10^{-4}). We investigate the ultrafast excited-state dynamics of this triazene by femtosecond time-resolved fluorescence experiments in water, ethylene glycol, and buffer and bound to the enzyme β -trypsin, the minor groove of AT-rich DNA, and G-quadruplex DNA. Ab initio calculations provide additional mechanistic insight. The complementing studies unveil that the excited-state motion initiated by $\pi\pi^*$ excitation occurs in two phases: a subpicosecond phase associated with the lengthening of the central N=N double bond, followed by a bicycle-pedal-type motion of the triazene bridge, which is almost volume-conserving and can proceed efficiently within only a few picoseconds even under spatially confined conditions. Our results elucidate the excited-state relaxation mechanism of aromatic triazenes and explain the modest sensitivity of the fluorescence quantum yield of berenil even when it is bound to various biomolecules.



Triazene compounds are employed for a variety of medical and pharmaceutical applications. For instance, berenil is the most widely used drug for treating sleeping sickness in animals, and it also acts as babesiacide.^{1–6} However, little is known about the underlying mechanism and how the properties of berenil change upon interaction with biomolecules. Berenil shows a high binding affinity to enzymes like β -trypsin^{7,8} or diamine oxidase,⁹ it is an efficient minor-groove binder of AT-rich DNA,^{1,10–20} and it also strongly binds to G-quadruplex DNA.^{21,22} A characteristic blue fluorescence emitted by berenil after UV excitation is observable when it is added to cell cultures that are in the stationary or lag phase of a growth cycle, yet it is absent when berenil is added to cells in the log phase.^{1,23} Until now, no explanation could be found for this exceptional behavior, and, in general, emission from berenil is not reported but the molecule is rather described as nonfluorescent.^{16,24} While these observations suggest that the deactivation channel of berenil after photoexcitation is sensitive to the surrounding, the virtual absence of fluorescence hints to very fast excited-state dynamics comprising photoisomerization and internal conversion.

Berenil (diminazene aceturate) is composed of two amidinophenyl units linked by a triazene bridge (Figure 1). It has some similarities to azobenzene in which two phenyl rings are linked by an azo group. The abilities to bind to biomolecules are very different for azobenzene and berenil, the latter being more flexible and exhibiting more possibilities for hydrogen-bond and electrostatic interactions due to the charged terminal groups. Regarding the photodynamics, despite numerous papers on the behavior of azobenzene after

excitation, there is still an ongoing debate considering its photochemistry.^{25–30} The reaction mechanisms discussed for its photoisomerization include inversion, rotation, the hula-twist,³⁰ and the concerted double twist.^{31,32} By contrast, the triazene bridge in berenil contains an amino (NH) group in addition to the azo bond, which increases the number of torsional angles that can be involved in the excited-state relaxation. This opens alternative pathways for deactivation.

To unravel the mechanism of the ultrafast excited-state dynamics of berenil, we combine ultrafast fluorescence upconversion measurements of berenil bound to large biomolecules or in solvents with different viscosity with correlated wave function calculations for the potential energy surface of the excited state. In this way, the ultrafast deactivation and the extremely low fluorescence quantum yield found for all of these conditions can be rationalized.

Steady-state absorption and emission spectra of berenil in buffer (50 mM Tris-HCl, 250 mM KCl) are shown in Figure 1. The absorption maximum is located at 368 nm with an extinction coefficient of $(26\,500 \pm 2300) \text{ M}^{-1} \text{ cm}^{-1}$.³³ The fluorescence exhibits a large Stokes shift and peaks at 476 nm and is extremely weak, as confirmed by the low quantum yield of 0.9×10^{-4} that we derived in a Strickler–Berg analysis³⁴ with several reference substances for consistency.

Received: February 27, 2017

Accepted: April 13, 2017

Published: April 20, 2017

48

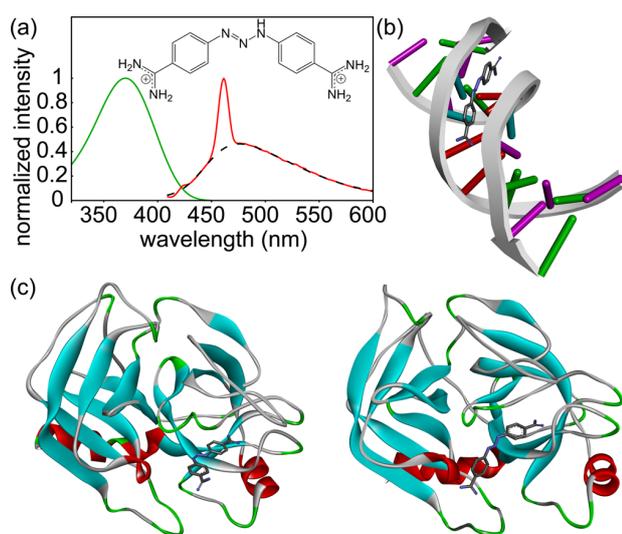


Figure 1. (a) Chemical structure of berenil and its absorption (green) and emission spectrum (red, excited at 400 nm) in buffer solution. Because of the low fluorescence quantum yield, Raman scattering in the solvent contributes as a pronounced peak at 462 nm. The dashed line is the Raman-corrected emission, as employed for the Strickler–Berg analysis. Cartoon representation of berenil bound to (b) AT-rich DNA (PDB accession code 2DBE)¹³ and to (c) β -trypsin (left: 3GY6; right: 3GY2).⁸

As biomolecular binding partners, we studied β -trypsin, AT-rich DNA (sequence 5' to 3': CGAATTTCAAAAGAAATT-CG), and G-quadruplex DNA (AGGGAGGGCGCTGGG-AGGAGGG) because of the reported high binding affinity of berenil to these kinds of biomolecules.^{7,8,13,21} The binding between the particular system and berenil is reflected in characteristic shifts¹⁰ in the absorption spectra (Figure S2 in the Supporting Information). Time-resolved fluorescence spectroscopy (see Supporting Information for experimental details) provides insight into the decay behavior, as shown in Figure 2 for berenil in buffer solution and bound to the different biomolecules as well as in neat water and ethylene glycol, respectively. Only results for an emission wavelength of 480 nm are shown for a direct comparison (best-fit parameter values with error limits are summarized in Table 1); data and fits for other wavelengths are presented in the SI. In each case, the best global deconvolution fit for all measured samples yields two lifetimes. For the buffer solution (blue curve), global decay times of 0.24 and 0.71 ps are found, and the amplitude A_2 of the second lifetime increases with increasing fluorescence wavelength (amplitude ratio 70:30 at 480 nm, 65:35 at 540 nm, inset of Figure 2).

When berenil is bound to β -trypsin, two main structures are present with berenil in different conformations (Figure 1c), and several hydrogen bonds between one of the amidinium groups and the protein are formed, while further interactions are possible.⁸ Nevertheless, the two fitted lifetimes are only slightly increased to 0.28 and 1.41 ps. The amplitude of the latter again increases with wavelength (83:17 at 480 nm, 76:24 at 540 nm, inset of Figure 2).

For berenil bound to AT-rich DNA (Figure 1b), one amidinium moiety hydrogen-bonds directly to an adenine base, whereas the other one interacts with the DNA via a mediating water molecule.¹³ This binding to the DNA minor groove restricts the volume available for geometrical changes of berenil,

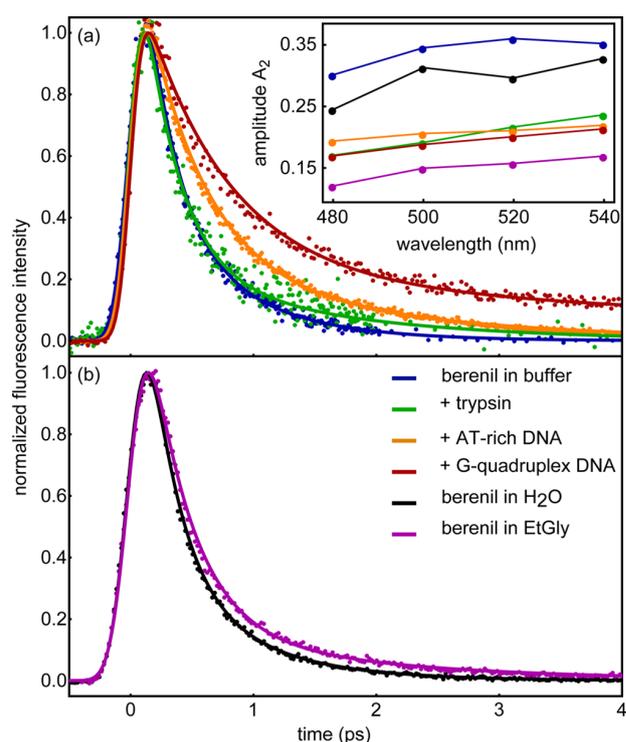


Figure 2. Transient fluorescence signal of berenil at an emission wavelength of 480 nm. (a) In Tris-HCl and the corresponding salt (in the case of unbound berenil and bound to the DNAs: KCl; in the case of berenil bound to trypsin: CaCl₂): unbound (blue), bound to trypsin (green), bound to AT-rich DNA (orange), and bound to G-quadruplex DNA (red). (b) In water (black) and ethylene glycol (magenta). Dots show data points, while lines are results from global fitting. The inset at the top-right shows the relative amplitude A_2 of the long lifetime in these fits as a function of fluorescence wavelength.

which should be even more confined when berenil is bound to the G-quadruplex DNA. Looking at the decay traces of berenil bound to the DNAs, the traces clearly deviate from the one of berenil in buffer only. The corresponding lifetimes are 0.48 and 1.77 ps for berenil bound to AT-rich DNA (orange line) and 0.69 and 6.55 ps for G-quadruplex DNA (red line), respectively. By analogy to the cases with buffer only and binding to β -trypsin, the amplitude A_2 increases with increasing wavelength (81:19 at 480 nm, 78:22 for 540 nm and 83:17 at 480 nm, 79:21 for 540 nm, respectively, inset of Figure 2). When calculating amplitude-weighted lifetimes τ_{ave} (Table 1), the general trend is the same as for the individual time constants: While τ_{ave} is 0.38 ps for berenil in buffer, it is larger for berenil bound to trypsin (0.47 ps) and to AT-rich DNA (0.73 ps). With 1.69 ps, the largest τ_{ave} value is found for berenil bound to G-quadruplex DNA.

These results are in opposition to the behavior of another efficient DNA binder, DAPI, for which the fluorescence lifetime is enhanced from 100 ps for DAPI dissolved in buffer to several nanoseconds when DAPI is bound to the minor groove of a DNA dodecamer.³⁵ Compared with this, the lifetime prolongation we observe for berenil is rather small. To further explore if the reason for this is the expendability of large geometrical changes after photoexcitation, we studied berenil in two solvents with different viscosity. The lifetimes for berenil in neat water (0.25 and 0.73 ps) are basically identical to the ones

Table 1. Time Constants Obtained by Global Fitting of the Fluorescence Upconversion Data^a

sample	A_1	τ_1 (ps)	A_2	τ_2 (ps)	τ_{ave} (ps)
berenil in buffer	0.70 ± 0.07	0.24 ± 0.03	0.30 ± 0.07	0.71 ± 0.08	0.38 ± 0.06
berenil + trypsin	0.83 ± 0.04	0.28 ± 0.02	0.17 ± 0.04	1.41 ± 0.24	0.47 ± 0.07
berenil + AT-rich DNA	0.81 ± 0.01	0.48 ± 0.01	0.19 ± 0.01	1.77 ± 0.08	0.73 ± 0.03
berenil + G-quad. DNA	0.83 ± 0.01	0.69 ± 0.02	0.17 ± 0.01	6.55 ± 0.52	1.69 ± 0.11
berenil in water	0.75 ± 0.03	0.25 ± 0.01	0.25 ± 0.03	0.73 ± 0.03	0.37 ± 0.03
berenil in ethylene glycol	0.88 ± 0.01	0.34 ± 0.01	0.12 ± 0.01	1.57 ± 0.08	0.49 ± 0.02

^aRelative amplitudes and amplitude-weighted average fluorescence lifetimes correspond to an emission wavelength of 480 nm (i.e., the data displayed in Figure 2; see Table S2 in the Supporting Information for other emission wavelengths).

ethylene glycol, the lifetimes are only slightly extended (0.34 and 1.57 ps), and, remarkably, the amplitude A_2 takes the lowest values with the same overall trend of an increase with wavelength (88:12 at 480 nm, 83:17 at 540 nm, inset of Figure 2), which is reflected in the rather small increase in τ_{ave} from 0.37 to 0.49 ps (Table 1). The viscosity of ethylene glycol (20.8 mPa·s at 20 °C)³⁶ is ~20 times higher than the viscosity of water (1.00 mPa·s at 20 °C).³⁶ In numerous studies where deactivation follows a *cis*–*trans* isomerization by rotation around a C=C bond, changing the solvent viscosity leads to a significant increase in the fluorescence lifetime.^{37–41} The rather weak dependence of the fluorescence lifetime of berenil on the solvent viscosity leads to analogous implications as the studies with berenil bound to biomolecules, namely, that a very fast nonradiative deactivation mechanism must be involved that cannot be associated with pronounced intramolecular motion.

If the deactivation proceeded via rotation, the fluorescence lifetimes should be significantly enhanced when changing the viscosity of the solvent. Therefore, we exclude the rotation as a possible deactivation channel. An in-plane motion of the phenyl rings as for the inversion mechanism is expected to be less affected by a higher viscosity. When berenil is bound to trypsin or DNA the volume accessible for the molecule should be substantially smaller than that for unbound berenil. Therefore, the time needed for deactivating structural dynamics along the inversion mechanism would still be expected to be longer for bound berenil. The associated enhancement of the emission lifetimes is deducible from the decay traces. However, the time-resolved fluorescence studies disclose that there is only a slight prolongation, which argues against deactivation pathways with predominant character of rotation and inversion.^{31,42–44} Consequently, we also exclude inversion as deactivation mechanism.

To get further insight into the absorption and fluorescence spectra, we used ab initio electronic structure calculations. Under the experimental and also physiological conditions^{18,20} diminazene is doubly protonated. Using dispersion-corrected density functional theory (DFT-D3) for geometry optimizations, we could characterize six local minima on the electronic ground-state potential energy surface of the doubly protonated diminazene as pairs with “*E*-*azo*, *s*-*trans*” (ET), “*E*-*azo*, *s*-*cis*” (EC), and “*Z*-*azo*, *s*-*trans*” (ZT) structure. The details of the computational procedure as well as the molecular structures of all conformers are available in the Supporting Information. Second-order Møller–Plesset perturbation theory (RI-MP2) has been used to obtain accurate relative energies. According to these results, at room temperature in the thermodynamic equilibrium only the ET minima are significantly occupied. Characteristic for the *E*-*azo* conformers is that the triazene bridge and the phenyl rings are coplanar and the amidinium

groups are twisted by ~30° either dis- or conrotatory out of this plane.

The energies and oscillator strengths for the five lowest excited states of the ET conformers have been calculated with the algebraic diagrammatic construction through second order,^{45,46} ADC(2), in vacuum and with the COSMO solvation model^{47,48} for aqueous solution ($\epsilon = 50$) with the SVP and aug-cc-pVDZ basis sets (see Supporting Information). The results show that the relative orientation (dis- or conrotatory) of the amidinium groups has no influence on the spectrum. The excitation into S_1 is the only bright transition below 5 eV. The vertical excitation energy of 3.32 eV for the transition into S_1 computed at the COSMO-ADC(2)/aug-cc-pVDZ level agrees well with the first maximum in the experimental UV/vis spectrum at ~3.35 eV (Figure 1a).^{10,33} At the ground-state structure the transition to S_1 is dominated by a $\pi\pi^*$ excitation from an occupied natural transition orbital⁴⁹ (NTO), which has bonding character for the N=N azo bond (no nodal plane between these two nitrogen atoms) to a virtual NTO with antibonding character for the N=N bond (see Figure S6 in the Supporting Information). Consequently, the π bond between these N atoms is broken in the S_1 state. The C–N π bonds between the NH–N=N moiety and the benzene rings are strengthened in the S_1 state.

The results for the absorption spectrum in vacuum and in aqueous solution show only minor differences, that is, for the transition into S_1 , a red shift by 0.02 eV and a decrease in the oscillator strength by ~2%. Furthermore, in the smaller and computationally cheaper SVP basis set the excitation energies are only slightly blue-shifted, but the same orbitals are involved in the excitation to S_1 , and thus it also describes the electronic structure of S_1 correctly. Because neither the excitation energy nor the oscillator strengths depend on the orientation of the –C(NH₂)₂ groups, we have chosen one of the ET conformers (the ET_{1S0} structure in Figure 3) and the ADC(2)/SVP level of theory without COSMO to investigate the relaxation of the photoactivated molecule.

For the S_1 state, we found two local minima denoted as ET_{1S1} and ET_{1RS1}, which can be reached from the Franck–Condon point after excitation. Their structures, which differ from the ground-state geometry by clockwise (ET_{1RS1}) and anticlockwise (ET_{1S1}) torsion around the azo bond, are shown in Figure 3. The two minima on the S_1 surface are connected by a transition state, ET_{1TS–S1}, which differs from the ground-state structure mainly by a stretch of the azo bond. At the ET_{1S1} and ET_{1RS1} geometries the COSMO-ADC(2)/aug-cc-pVDZ energy of the S_1 state is slightly below the COSMO-RI-MP2/aug-cc-pVDZ energy of S_0 . This indicates that in these regions the two states might be connected by conical intersections.

To investigate the deactivation pathways, we carried out 50 intrinsic reaction coordinate (IRC) calculations at the

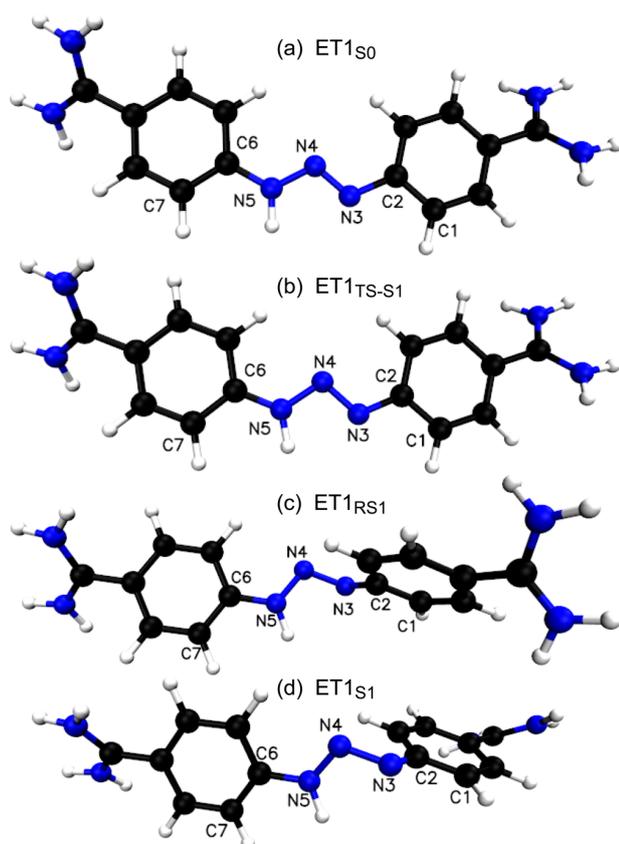


Figure 3. Optimized stationary points for the ground state and the lowest excited state at the ADC(2)/SVP level of theory. At the transition state $ET1_{TS-S1}$, the phenyl groups are almost coplanar while at the minima $ET1_{S1}$ and $ET1_{RS1}$ they are twisted by $\sim 90^\circ$ around the $N_3=N_4$ bond anticlockwise and clockwise, respectively.

ADC(2)/SVP level on the S_1 surface starting from the Franck–Condon point. The idea of the IRC analysis is to calculate nuclear gradients at a given geometry, let the nuclei move for a certain time step δt ($= 2$ fs in this case) according to Newton's second law, and again calculate the gradients and a geometry update for the new coordinates until a stationary point is reached where the gradients are zero within a given threshold. In the IRC analysis,^{50,51} the kinetic energy is at each step damped to zero to follow the minimum energy reaction pathway.⁵²

Figure 4a shows how the two most important internal coordinates, the length of the $N_3=N_4$ azo bond and the $C_2-N_3=N_4-N_5$ dihedral angle, change along the IRC trajectories. The density of the points can be interpreted as the inverse of the gradient (force) norm along the internal coordinates. This Figure shows that the first phase (IRC₁) of the de-excitation is a fast departure from the Franck–Condon region, which is characterized by a $N_3=N_4$ bond stretch and brings the molecules to a flat region on the S_1 surface. An optimization starting from the end point of the IRC₁ path led to the transition state $ET1_{TS-S1}$, which is very similar to the last snapshot of the IRC₁ calculation. To continue the IRC calculations, the TS_{S1} geometry has been screwed along the normal mode with negative eigenvalue for $\sim \pm 10^\circ$. The obtained paths (IRC₂ and IRC₃) show that the second phase of the de-excitation involves a torsion of the $C_2-N_3=N_4-N_5$

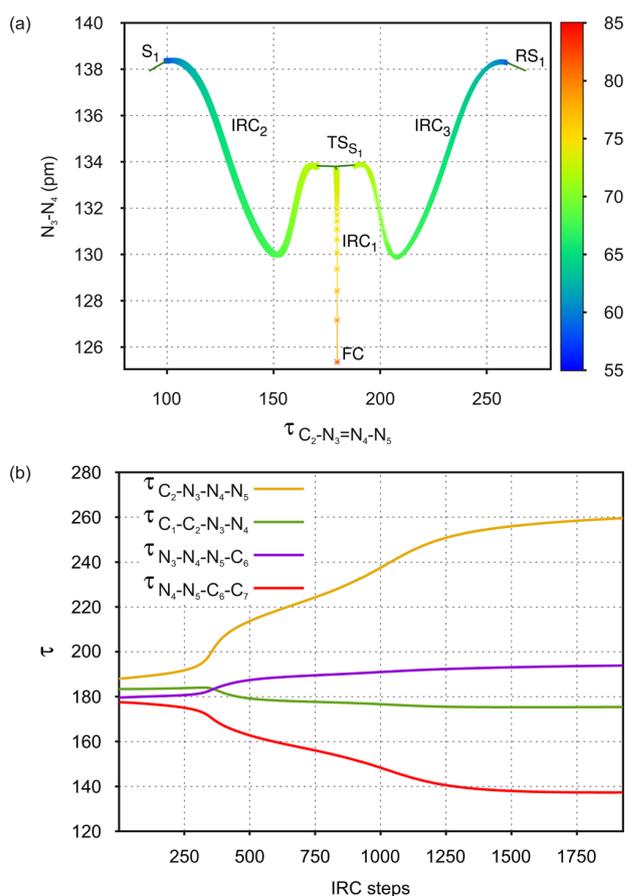


Figure 4. (a) Calculated (ADC(2)/RI-MP2)/SVP intrinsic reaction coordinates are mapped to the $N_3=N_4$ double bond and the $C_2-N_3-N_4-N_5$ dihedral angle. The Figure presents the results of three independent IRC calculations. Colors indicate for each point the energy relative to the ground state at $ET1_{S1}$ in $\text{kcal}\cdot\text{mol}^{-1}$. The IRC₁ starts from the FC point. The IRC₂ starts from the $\sim -10^\circ$ screwed TS_{S1} geometry along the normal mode with imaginary frequency and the IRC₃ from the $\sim +10^\circ$ rotated TS_{S1} geometry. The green lines connect the starting and ending point of the IRC calculations to the corresponding stationary points. (b) Evolution of triazene dihedral angles along the IRC₃ trajectory.

dihedral angle. The two pathways lead to the two minima $ET1_{S1}$ and $ET1_{RS1}$ on the S_1 surface described above (Figure 3).

The two clockwise and anticlockwise rotation paths are similar to each other and involve, in addition to changes of the length and torsion angle of the azo bond, torsions around other bonds at the triazene group. To further unravel the mechanism of the second de-excitation phase, Figure 4b shows the four dihedral angles at the triazene group along the IRC₃ path. Several mechanisms have been discussed in the literature for the (photo)isomerization of chromophores with double bonds. The most important ones are the pure torsion rotation, the Hula-twist mechanism,^{43,53} in-plane inversion, concerted double-twist,³¹ and the bicycle-pedal type mechanism.^{44,54} The moderate sensitivity of the deactivation dynamics to the viscosity of the solvent observed in the experiments already ruled out inversion and rotation as relaxation mechanisms.

The Hula-twist mechanism would for berenil imply simultaneous torsions around the N_3N_4 and N_4N_5 (or C_2N_3) bonds with about the same amplitude, whereas the concerted

double-twist mechanism is identified with concerted out-of-plane motion of $N_3=N_4$ nitrogens—here the simultaneous rotation around C_2N_3 , N_3N_4 and N_4N_5 bonds. According to Figure 4b, the molecule rotates by $\sim 70^\circ$ around the N_3N_4 bond but only by $\sim 20^\circ$ around the N_4N_5 and C_2N_3 bonds. Consequently, neither the Hula-twist nor the concerted double twist are the mechanisms of the second phase. The torsion around the azo bond is accompanied by a simultaneous rotation around the N_5C_6 bond by $\sim 40^\circ$. Although the amplitude of the latter motion is somewhat smaller than for the torsion around the azo bond, it is by far the most important component of the torsional motions next to the torsion around the azo bond. Thus we conclude that the relaxation mechanism is predominantly a bicycle-pedal type motion, which is almost volume-conserving.^{44,55} This explains the fast radiationless de-excitation in solution and with various biomolecular binding partners because it does not require large amplitude motions of the terminal phenyl and amidinium groups.

While no indication for an intersystem crossing is evident from our data or from the literature, the ultrafast fluorescence decay is attributed to ultrafast geometrical changes, rapidly leading to internal conversion back to the electronic ground state. From several possibilities for this structural motion, the combined information from experiment and theory allows for an unambiguous identification. A rotation of the phenol's sidegroup,⁵⁶ that is, the amidinium groups, after photo-excitation can be ruled out as decisive process for two reasons: On the one hand, this motion would be hindered when berenil is hydrogen-bonded to the biomolecules, yet the dynamics do not change drastically, in contrast with the minor-groove binder DAPI where this mechanism is known to contribute to the relaxation³⁵ and hence causes an enhancement of emission lifetimes by more than an order of magnitude; on the other hand, our simulations show that these motions are not the favorable mechanism for the relaxation from the Franck–Condon point on the excited-state potential energy surface.

Hence, the triazene bridge is crucially involved in the de-excitation but in a way that is different from a cis–trans isomerization found for compounds with $C=C$ double bonds,^{37–41} as can be deduced from the experimentally observed viscosity dependence. To clarify the structural dynamics of the triazene bridge, we can first address similarities to the dynamics of azobenzene, for which also two lifetimes were found in the fluorescence decay traces. Quick et al.³⁰ assigned the 0.3 ps lifetime of *trans*-azobenzene to a rapid motion away from the Franck–Condon region to a local minimum, while the longer lifetime of 3 ps was attributed to a population relaxation. This rather slow population relaxation in the excited state is associated with a barrier of 8 kJ/mol.

In the case of berenil, the short lifetime is in the range between 0.24 ps for unbound berenil and 0.69 ps for berenil bound to G-quadruplex DNA. They match well the first fluorescence lifetime of azobenzene and can be interpreted accordingly. As is evident from the simulations and the IRC₁ path in Figure 4a, the initial motion setting in after photoexcitation and concurrent cleavage of the $N=N$ π -bond in the S_1 state is a lengthening of the $N=N$ bond distance, which is accompanied by a change of the oscillator strength. These dynamics are experimentally monitored by the rapidly decaying fluorescence contribution; that is, the first lifetime reflects the motion out of the Franck–Condon region. Albeit slight prolongations are observed, the first lifetime is much less affected by the different confining conditions than is the second

one, which spans an order of magnitude from 0.71 ps for unbound berenil to 6.55 ps for berenil bound to G-quadruplex DNA. Hence the experiments imply that relaxation occurs in two phases associated with two different types of motion.

This finding is supported by the calculations at the ADC(2) level. The computational investigations of the relaxation from the Franck–Condon point to the minima on the S_1 surface show the two relaxation phases. Initially, the molecule approaches a transition state where the azo bond is stretched but the triazene and phenyl groups are still coplanar. Through this first phase, $16.1 \text{ kcal}\cdot\text{mol}^{-1}$ of the absorbed energy is released. The second phase is dominated by a bicycle-pedal type motion around the $N_3=N_4$ and N_5-C_6 bonds that links the transition state TS_{S_1} to the two minima (S_1 and $R-S_1$) on the S_1 surface, depending on the direction of the torsional motion. Also, the second phase of the relaxation is strongly exothermic and releases $\sim 19.4 \text{ kcal}\cdot\text{mol}^{-1}$ energy. Similar to the first phase, this step requires no large amplitude motions of the bulky groups. The fluorescence exhibits an initial red shift. We do not observe (or possibly resolve) a dynamic Stokes shift starting at the excitation wavelength with a delayed onset of long-wavelength emission, as may be caused by solvent reorganization or intramolecular vibrational relaxation. Nevertheless, the spectral composition of the fluorescence contains valuable information. Around the S_1 and $R-S_1$ regions, the energy gap between S_0 and S_1 is small, and the two states are probably connected via conical intersections. Hence, the central fluorescence wavelength should be different for the two relaxation phases as a consequence of the shrinking energy gap when following the IRC₂ and IRC₃ reaction paths. Hence, the associated release of energy during relaxation is manifested in the relative amplitudes of the fluorescence decays. The inset of Figure 2 displays an increasing relative amplitude of the longer lifetime with increasing emission wavelength; that is, the emission corresponding to the second relaxation phase is spectrally centered at lower energy photons.

In summary, the joint experimental and theoretical study shows that the efficient minor-groove binder berenil performs an ultrafast motion in the electronically excited state after $\pi\pi^*$ excitation. The dynamics comprise two phases: first, a contribution out of the Franck–Condon region associated with the lengthening of the initial $N=N$ double bond and a subsequent bicycle-pedal type motion of the central triazene unit. Time-resolved fluorescence studies of berenil in water, buffer, or ethylene glycol and bound to β -trypsin, AT-rich DNA, or G-quadruplex DNA revealed that the first phase proceeds on a subpicosecond time scale, whereas the second phase still occurs within only a few picoseconds for all of the investigated conditions despite several hydrogen bonds or different spatial restrictions. The volume-conserving nature of the bicycle-pedal type motion renders possible extremely fast excited-state dynamics in berenil and thus explains the very low fluorescence quantum yield and the frequently encountered classification of berenil as being “nonfluorescent”.

■ ASSOCIATED CONTENT

📄 Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcllett.7b00472.

Details of sample preparation, experimental procedures, absorption/emission spectra, and transient fluorescence

signals for all samples, computational details, optimized structures.(PDF)

AUTHOR INFORMATION

Corresponding Authors

*C.H.: E-mail: christof.haettig@rub.de.

*P.N.: E-mail: patrick.nuernberger@rub.de.

ORCID

Patrick Nuernberger: 0000-0002-4690-0229

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We are grateful to Prof. Christian Herrmann and Prof. Lars Schäfer for helpful discussions and to Bastian Geißler for assistance with global fitting. This work has been supported by the Deutsche Forschungsgemeinschaft within the Cluster of Excellence RESOLV (EXC 1069) and the Emmy-Noether program (PN).

REFERENCES

- (1) Newton, B. A. Berenil: A trypanocide with selective activity against extranuclear DNA. In *Mechanism of Action of Antimicrobial and Antitumor Agents*; Corcoran, J. W., Hahn, F. E., Snell, J. F., Arora, K. L., Eds.; Springer: Berlin, Germany, 1975; pp 34–47.
- (2) Peregrine, A. S.; Mamman, M. Pharmacology of diminazene. *Acta Trop.* **1993**, *54*, 185–203.
- (3) Khaw, M.; Panosian, C. B. Human antiprotozoal therapy: Past, present, and future. *Clin. Microbiol. Rev.* **1995**, *8*, 427–439.
- (4) Kuriakose, S.; Muleme, H. M.; Onyilagha, C.; Singh, R.; Jia, P.; Uzonna, J. E. Diminazene aceturate (berenil) modulates the host cellular and inflammatory responses to trypanosoma congolense infection. *PLoS One* **2012**, *7*, e48696.
- (5) Kuriakose, S.; Uzonna, J. E. Diminazene acetate (berenil), a new use for an old compound? *Int. Immunopharmacol.* **2014**, *21*, 342–345.
- (6) da Silva Oliveira, G. L.; de Freitas, R. M. Diminazene acetate—An antiparasitic drug of antiquity: Advances in pharmacology & therapeutics. *Pharmacol. Res.* **2015**, *102*, 138–157.
- (7) Pereira, M. T.; Silva-Alves, J. M.; Martins-Jose, A.; Lopes, J. C. D.; Santoro, M. M. Thermodynamic evaluation and modeling of proton and water exchange associated with benzamidine and berenil binding to β -trypsin. *Braz. J. Med. Biol. Res.* **2005**, *38*, 1593–1601.
- (8) Perilo, C. S.; Pereira, M. T.; Santoro, M. M.; Nagem, R. A. P. Structural binding evidence of the trypanocidal drugs Berenil and Pentacarinat active principles to a serine protease model. *Int. J. Biol. Macromol.* **2010**, *46*, 502–511.
- (9) McGrath, A. P.; Hilmer, K. M.; Collyer, C. A.; Shepard, E. M.; Elmore, B. O.; Brown, D. E.; Dooley, D. M.; Guss, J. M. Structure and inhibition of human diamine oxidase. *Biochemistry* **2009**, *48*, 9810–9822.
- (10) Festy, B.; Lallemand, A.-M.; Riou, G.; Brack, C.; Delain, E. Mécanisme d'action des diamidines trypanocides. Importance de la composition en bases dans l'association berenil-polynucléotides. *C. R. Hebd. Séances Acad. Sci. D* **1970**, *271*, 684–687.
- (11) Zimmer, C.; Wähner, U. Nonintercalating DNA-binding ligands: Specificity of the interaction and their use as tools in biophysical, biochemical and biological investigations of the genetic material. *Prog. Biophys. Mol. Biol.* **1986**, *47*, 31–112.
- (12) Yoshida, M.; Banville, D. L.; Shafer, R. H. Structural analysis of d(GCAATTGC)₂ and its complex with berenil by nuclear magnetic resonance spectroscopy. *Biochemistry* **1990**, *29*, 6585–6592.
- (13) Brown, D. G.; Sanderson, M. R.; Skelly, J. V.; Jenkins, T. C.; Brown, T.; Garman, E.; Stuart, D. I.; Neidle, S. Crystal structure of a berenil-dodecanucleotide complex: the role of water in sequence-specific ligand binding. *EMBO J.* **1990**, *9*, 1329–1334.
- (14) Jenkins, T. C.; Lane, A. N.; Neidle, S.; Brown, D. G. NMR and molecular modeling studies of the interaction of berenil and pentamidine with d(CGCAAATTTGCG)₂. *Eur. J. Biochem.* **1993**, *213*, 1175–1184.
- (15) Barceló, F.; Portugal, J. Berenil recognizes and changes the characteristics of adenine and thymine polynucleotide structures. *Biophys. Chem.* **1993**, *47*, 251–260.
- (16) Pilch, D. S.; Kirolos, M. A.; Liu, X.; Plum, G. E.; Breslauer, K. J. Berenil [1,3-bis(4'-amidinophenyl)triazene] binding to DNA duplexes and to a RNA duplex: Evidence for both intercalative and minor groove binding properties. *Biochemistry* **1995**, *34*, 9962–9976.
- (17) Taberner, L.; Bella, J.; Alemán, C. Hydrogen bond geometry in DNA-minor groove binding drug complexes. *Nucleic Acids Res.* **1996**, *24*, 3458–3466.
- (18) Reddy, B. S. P.; Sondhi, S. M.; Lown, J. W. Synthetic DNA minor groove-binding drugs. *Pharmacol. Ther.* **1999**, *84*, 1–111.
- (19) Mallena, S.; Lee, M. P. H.; Bailly, C.; Neidle, S.; Kumar, A.; Boykin, D. W.; Wilson, W. D. Thiophene-based diamidine forms a “super” AT binding minor groove agent. *J. Am. Chem. Soc.* **2004**, *126*, 13659–13669.
- (20) Zsila, F. Glycosaminoglycans are potential pharmacological targets for classic DNA minor groove binder drugs berenil and pentamidine. *Phys. Chem. Chem. Phys.* **2015**, *17*, 24560–24565.
- (21) Zhou, J.; Le, V.; Kalia, D.; Nakayama, S.; Mikek, C.; Lewis, E. A.; Sintim, H. O. Diminazene or berenil, a classic duplex minor groove binder, binds to G-quadruplexes with low nanomolar dissociation constants and the amidine groups are also critical for G-quadruplex binding. *Mol. Biosyst.* **2014**, *10*, 2724–2734.
- (22) Wang, C.; Carter-Cooper, B.; Du, Y.; Zhou, J.; Saeed, M. A.; Liu, J.; Guo, M.; Roembke, B.; Mikek, C.; Lewis, E. A.; Lapidus, R. G.; Sintim, H. O. Alkyne-substituted diminazene as G-quadruplex binders with anticancer activities. *Eur. J. Med. Chem.* **2016**, *118*, 266–275.
- (23) Stockert, J. C.; Del Castillo, P.; Llorente, A. R.; Rasskin, D. M.; Romero, J. B.; Gómez, A. New fluorescence reactions in DNA cytochemistry. 1. Microscopic and spectroscopic studies on nonrigid fluorochromes. *Anal. Quant. Cytol. Histol.* **1990**, *12*, 1–10.
- (24) Gabelica, V.; Rosu, F.; De Pauw, E.; Antoine, R.; Tabarin, T.; Broyer, M.; Dugourd, P. Electron photodetachment dissociation of DNA anions with covalently or noncovalently bound chromophores. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 1990–2000.
- (25) Schultz, T.; Quenneville, J.; Levine, B.; Toniolo, A.; Martinez, T. J.; Lochbrunner, S.; Schmitt, M.; Shaffer, J. P.; Zgierski, M. Z.; Stolow, A. Mechanism and dynamics of azobenzene photoisomerization. *J. Am. Chem. Soc.* **2003**, *125*, 8098–8099.
- (26) Satzger, H.; Spörlein, S.; Root, C.; Wachtveitl, J.; Zinth, W.; Gilch, P. Fluorescence spectra of trans- and cis-azobenzene-emission from the Franck-Condon State. *Chem. Phys. Lett.* **2003**, *372*, 216–223.
- (27) Pancur, T.; Renth, F.; Temps, F.; Harbaum, B.; Krüger, A.; Herges, R.; Näther, C. Femtosecond fluorescence up-conversion spectroscopy of a rotation-restricted azobenzene after excitation to the S₁ state. *Phys. Chem. Chem. Phys.* **2005**, *7*, 1985–1989.
- (28) Lu, Y.-C.; Diao, E. W.-G.; Rau, H. Femtosecond fluorescence dynamics of rotation-restricted azobenzenophanes: New evidence on the mechanism of trans→cis photoisomerization of azobenzene. *J. Phys. Chem. A* **2005**, *109*, 2090–2099.
- (29) Bandara, H. M. D.; Burdette, S. C. Photoisomerization in different classes of azobenzene. *Chem. Soc. Rev.* **2012**, *41*, 1809–1825.
- (30) Quick, M.; Dobryakov, A. L.; Gerecke, M.; Richter, C.; Berndt, F.; Ioffe, I. N.; Granovsky, A. A.; Mahrwald, R.; Ernsting, N. P.; Kovalenko, S. A. Photoisomerization dynamics and pathways of trans- and cis-azobenzenes in solution from broadband femtosecond spectroscopies and calculations. *J. Phys. Chem. B* **2014**, *118*, 8756–8771.
- (31) Böckmann, M.; Doltsinis, N. L.; Marx, D. Enhanced photoswitching of bridged azobenzene studied by nonadiabatic ab initio simulation. *J. Chem. Phys.* **2012**, *137*, 22A505.
- (32) Böckmann, M.; Braun, S.; Doltsinis, N. L.; Marx, D. Mimicking photoisomerisation of azo-materials by a force field switch derived from nonadiabatic ab initio simulations: Application to photo-

switchable helical foldamers in solution. *J. Chem. Phys.* **2013**, *139*, 084108.

(33) Thünemann, A. F.; Schütt, D.; Sachse, R.; Schlaad, H.; Möhwald, H. Complexes of poly(ethylene oxide)-block-poly(L-glutamate) and diminazene. *Langmuir* **2006**, *22*, 2323–2328.

(34) Strickler, S. J.; Berg, R. A. Relationship between absorption intensity and fluorescence lifetime of molecules. *J. Chem. Phys.* **1962**, *37*, 814–822.

(35) Banerjee, D.; Pal, S. K. Dynamics in the DNA recognition by DAPI: Exploration of the various binding modes. *J. Phys. Chem. B* **2008**, *112*, 1016–1021.

(36) Tsierekzos, N. G.; Molinou, I. E. Thermodynamic properties of water + ethylene glycol at 283.15, 293.15, 303.15, and 313.15 K. *J. Chem. Eng. Data* **1998**, *43*, 989–993.

(37) Taylor, J. R.; Adams, M. C.; Sibbett, W. Investigation of viscosity dependent fluorescence lifetime using a synchronously operated picosecond streak camera. *Appl. Phys.* **1980**, *21*, 13–17.

(38) Bagchi, B.; Fleming, G. R.; Oxtoby, D. W. Theory of electronic relaxation in solution in the absence of an activation barrier. *J. Chem. Phys.* **1983**, *78*, 7375–7385.

(39) Åkesson, E.; Hakkarainen, A.; Laitinen, E.; Helenius, V.; Gillbro, T.; Korppi-Tommola, J.; Sundström, V. Analysis of microviscosity and reaction coordinate concepts in isomerization dynamics described by Kramers' Theory. *J. Chem. Phys.* **1991**, *95*, 6508–6523.

(40) Vauthey, E. Isomerisation dynamics of a thiocarbocyanine dye in different electronic states and in different classes of solvents. *Chem. Phys.* **1995**, *196*, 569–582.

(41) Vogt, G.; Nuernberger, P.; Gerber, G.; Improta, R.; Santoro, F. Femtosecond study on the isomerization dynamics of NK88. II. Excited-state dynamics. *J. Chem. Phys.* **2006**, *125*, 044513.

(42) Liu, R. S. H.; Asato, A. E. The primary process of vision and the structure of bathorhodopsin: A mechanism for photoisomerization of polyenes. *Proc. Natl. Acad. Sci. U. S. A.* **1985**, *82*, 259–263.

(43) Liu, R. S. H. Photoisomerization by Hula-Twist: A fundamental supramolecular photochemical reaction. *Acc. Chem. Res.* **2001**, *34*, 555–562.

(44) Schapiro, I.; Weingart, O.; Buss, V. Bicycle-pedal isomerization in a rhodopsin chromophore model. *J. Am. Chem. Soc.* **2009**, *131*, 16–17.

(45) Schirmer, J. Beyond the random-phase approximation: A new approximation scheme for the polarization propagator. *Phys. Rev. A: At., Mol., Opt. Phys.* **1982**, *26*, 2395–2416.

(46) Hättig, C. Structure optimizations for excited states with correlated second-order methods: CC2 and ADC(2). *Adv. Quantum Chem.* **2005**, *50*, 37–60.

(47) Klamt, A.; Schüürmann, G. COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc., Perkin Trans. 2* **1993**, *2*, 799–805.

(48) 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.

(49) Martin, R. L. Natural transition orbitals. *J. Chem. Phys.* **2003**, *118*, 4775–4777.

(50) Fukui, K. A formulation of the reaction coordinate. *J. Phys. Chem.* **1970**, *74*, 4161–4163.

(51) Fukui, K. The path of chemical reactions - the IRC approach. *Acc. Chem. Res.* **1981**, *14*, 363–368.

(52) Hellweg, A. Heuristic control of kinetic energy in dynamic reaction coordinate calculations. *J. Comput. Chem.* **2013**, *34*, 1835–1841.

(53) Imamoto, Y.; Kuroda, T.; Kataoka, M.; Shevyakov, S.; Krishnamoorthy, G.; Liu, R. S. H. Photoisomerization by Hula Twist: 2,2'-dimethylstilbene and a ring-fused analogue. *Angew. Chem., Int. Ed.* **2003**, *42*, 3630–3633.

(54) Warshel, A. Bicycle-pedal model for the first step in the vision process. *Nature* **1976**, *260*, 679–683.

(55) Jung, Y. O.; Lee, J. H.; Kim, J.; Schmidt, M.; Moffat, K.; Šrajer, V.; Ihee, H. Volume-conserving trans-cis isomerization pathways in photoactive yellow protein visualized by picosecond X-ray crystallography. *Nat. Chem.* **2013**, *5*, 212–220.

(56) Kovalenko, S. A.; Schanz, R.; Farztdinov, V. M.; Hennig, H.; Ernsting, N. P. Femtosecond relaxation of photoexcited parinitroaniline: solvation, charge transfer, internal conversion and cooling. *Chem. Phys. Lett.* **2000**, *323*, 312–322.

6 How a Linear Triazene Photoisomerizes in a Volume-Conserving Fashion

In this full paper, we have reported the results of our computational and experimental studies on the photoisomerization pathways of berenil. All the experimental works have been done in the physical chemistry group of Professor Nürnberger.^{3 4}

Most of the calculations were done during my PhD and only a few ones were taken from the previous letter and the master thesis. The main theoretical discussion was designed by me under the supervision of Professor Hättig and Professor Nürnberger. The discussion of the article has been completed and improved in a close discussion with my collaborators: Lena Grimmelsmann, Professor Nuernberger and Professor Hättig.

All the co-authors decided together about the strategy of presenting the data. Lena Grimmelsmann prepared the first half of the introduction, the discussion of experimental data and figure 3 of the paper. I provided the theoretical part, the second half of the introduction and all the other figures. All the co-authors participated in the process of editing the first draft. I have done the finalization of the manuscript and communication with the journal. The referees' responses and questions were discussed between all of us.

³Article: <https://pubs.rsc.org/en/content/articlelanding/2018/cp/c8cp05208e#divAbstract>

⁴ESI: <http://www.rsc.org/suppdata/c8/cp/c8cp05208e/c8cp05208e1.pdf>



Cite this: *Phys. Chem. Chem. Phys.*,
2018, 20, 28075

How a linear triazene photoisomerizes in a volume-conserving fashion†

Alireza Marefat Khah,^a Lena Grimmelsmann,^b Johannes Knorr,^b
Patrick Nuernberger^{*b} and Christof Hättig^{*a}

Understanding deactivation mechanisms of functional groups is a key step to design novel photo-active devices and molecular imaging agents. Here, we elucidate the photochemistry of linear triazenes, an extended analogue of the photo-switchable azo group, exemplarily for the widely used DNA-minor-groove binder berenil. Combining ultrafast spectroscopy and *ab initio* calculations unveils that the *E*-azo,*s-trans* structure of berenil predominates in the gas phase and in aqueous solution, and ADC(2) intrinsic reaction coordinate calculations disclose that the excited-state relaxation to the S_1 minima/conical intersections follows a two-step mechanism: N=N bond stretching followed by a bicycle-pedal rotation in the triazene bridge. Furthermore, studying the ground-state pathways shows that a fraction of the molecules relaxes back to the *E*-azo,*s-trans* isomer while the other part photoisomerizes to the *Z*-azo,*s-trans* via a hula-twist motion, as evidenced by experimental quantum yields of $\Phi \approx 0.5$ found for berenil in water, ethylene glycol, or bound to β -trypsin. Moreover, our studies show that while the excited-state relaxation is insensitive to the environment, the ground-state dynamics depend on biomolecular binding partners.

Received 15th August 2018,
Accepted 24th October 2018

DOI: 10.1039/c8cp05208e

rsc.li/pccp

1 Introduction

Light-induced *cis-trans* isomerizations are rather simple processes for the conversion of electronic energy into kinetic energy on a molecular scale. Despite the notation “simple”, these processes are extensively studied because of the large variety of applications in biology, chemistry and technology. They are for example involved in the primary step of vision,^{1,2} optical data storage devices^{3,4} and switchable molecular motors.^{5–7} Many photoisomerization processes occur on a (sub-)picosecond timescale. With the development of ultrafast laser sources the investigation of these phenomena in real time became feasible. From a theoretical point of view, conical intersections serve as efficient funnels from an upper to a lower electronic state. Substantial structural changes can lead excited molecules to and “through” a conical intersection and thus are often the reason for an ultrashort excited-state lifetime. Azobenzene and stilbene present two of the most extensively studied systems featuring rapid photoisomerization processes which involve a conical intersection. Despite the similarities in their molecular structure, the excited-state dynamics differ significantly.

Azobenzene exhibits ultrafast excited-state dynamics even under strong constraints,⁸ whereas, because of the different isomerization mechanism, for stilbene the dynamics are slowed down. For azobenzene, a hula-twist mechanism is proposed,⁹ while stilbene isomerizes via rotation around the C=C double bond.¹⁰ Moreover, in the literature further mechanisms are described for these and other organic chromophores which can be classified as either large-amplitude or volume-conserving motions. Inversion and rotation belong to the former class,⁹ whereas the hula-twist, the NN-twist, the bicycle-pedal motion and related torsional mechanisms are characterized as volume-conserving.^{11–16}

The investigation of the *cis-trans* isomerization of triazenes started in 1972 with the discovery of silyltriazenes by Wiberg and Pracht.^{17,18} However, for many triazenes, photogenerated isomers are thermally unstable and reform the initial configuration so that time-resolved spectroscopy techniques are necessary to allow an observation. For this reason, initially only those triazenes forming rather stable isomers were investigated,¹⁹ later insightful work with various spectroscopy approaches on the ground-state isomerization processes after photoexcitation was reported for 1,3-diphenyltriazene.^{20–23} The triazene compound diminazene, which is of interest in the present study, is in combination with its counterion aceturate well-known as berenil, a widely used drug for treating the sleeping sickness in animals.²⁴ Depending on the environment, the fluorescence intensity of berenil differs significantly. While a characteristic blue fluorescence can be detected when berenil is added to cell cultures

^a Theoretische Chemie, Ruhr-Universität Bochum, 44780 Bochum, Germany.
E-mail: christof.haettig@rub.de

^b Physikalische Chemie II, Ruhr-Universität Bochum, 44780 Bochum, Germany.
E-mail: patrick.nuernberger@rub.de

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c8cp05208e

which are in the stationary or lag phase of a growth cycle, the fluorescence is absent when berenil is added to cells in the log phase.²⁴ The virtual absence of fluorescence is also characteristic for solutions of berenil in organic solvents,^{25,26} which indicates rapid deactivation dynamics. In a recent communication,²⁷ the excited-state relaxation pathway of berenil was explored by means of femtosecond fluorescence upconversion spectroscopy and wavefunction-based *ab initio* calculations at the level of the algebraic diagrammatic construction through second order, ADC(2). It was found that ultrafast excited-state dynamics proceed on a sub-picosecond timescale even under strong constraints, showing similarities but also differences with regard to azobenzene.⁸ This unexpected insensitivity to changes in the molecular environment and the ultrafast excited-state relaxation of the ET1 conformer (*vide infra*) were rationalized through an analysis of the intrinsic reaction coordinates (IRC). It was illustrated that the deexcitation mechanism consists of two phases: an elongation of the N=N double bond followed by a bicycle-pedal-type motion of conjugated bonds which minimizes the energy to a conical intersection between the lowest electronically excited and the ground state. This volume-conserving motion can explain the lack of sensitivity to viscous environments (as recently confirmed in high-pressure studies²⁸) and the binding to large biomolecules as well as the extremely short emission lifetime. The study concentrated solely on excited-state dynamics, and the employed technique of fluorescence upconversion only allowed the pursuit of emissive reaction pathways. Following absorption changes on the femtosecond timescale would further open the possibility to observe spectral signatures associated with the formation and dynamics of ground-state species.

In the current study, the ultrafast photodynamics of berenil are investigated with transient absorption spectroscopy. This complements extensive *ab initio* MP2 and ADC(2) calculations and the analysis of the IRC pathways which are used to provide molecular insight into the involved ground- and excited-state processes. Dissecting the role of different ground-state conformers and the (photo-)isomerization pathways allows us to draw a comprehensive picture which connects the steady-state and transient absorption spectra of berenil with light-induced geometrical changes as elucidated by the calculations.

2 Material and methods

2.1 Sample preparation

Berenil, β -trypsin and ethylene glycol were purchased from Sigma Aldrich. For the experiment with trypsin, a solution of 4.5 mg berenil and 15 mL 50 mM Tris-HCl and 25 mM CaCl₂ was prepared. Afterwards 244.7 mg of β -trypsin were added. In the case of pure buffer, 28.8 mg berenil were dissolved in 25 mL buffer (for the UV white-light probe studies) or 10.3 mg berenil were dissolved in the same volume of buffer (for the visible white-light probe studies). For the ethylene glycol studies, 10.4 mg berenil were dissolved in 25 mL of ethylene glycol for the probe with visible white-light and 30.6 mg berenil for the

probe with UV white-light, respectively. For all solutions, the final optical density of the sample, prepared in fused-silica cuvettes (Hellma) with a path length of 0.02 cm, was verified using an absorption spectrometer (Jasco V-770). The absorption was again measured after the time-resolved measurements to exclude the possibility of pronounced photodegradation.

2.2 Transient absorption

The femtosecond transient absorption (TA) setup comprises an amplified laser system (Spitfire Ace, Spectra Physics) delivering near-infrared pulses (central wavelength 800 nm, pulse energy 4.2 mJ, pulse duration 100 fs, repetition rate 1 kHz). Frequency-doubling to 400 nm in a β -barium borate (BBO) crystal yields the pump pulses employed in the TA experiments. Every second pump pulse is blocked by a chopper driven at 500 Hz. A small fraction of the 800 nm pulses are used to pump a linearly moving calcium fluoride (CaF₂) disk to generate the white-light continuum in the wavelength range between 320 and 610 nm, which serves as probe pulses. The integration of another BBO crystal in front of the CaF₂ plate and the use of adequate filters leads to a UV white-light continuum between 280 and 390 nm. Magic angle configuration is used to eliminate anisotropy effects.^{29,30} Both pump and probe beams overlap in the sample solution, which is circulated through a 0.02 cm path length flow cell. The pump pulse is blocked behind the sample, whereas the white-light signal is guided to a spectrometer (Acton SpectraPro SP2500i, grating 150 grids per mm) with an attached CCD camera (Roper Scientific, binned to 1024 pixels at 1 kHz readout). A sketch of the setup is shown in the ESI.†

TA spectra were recorded at 400 time delays from -1 ps to 3.8 ns. 250 steps are equidistant from -1 to 15 ps and 150 time steps are spaced on a logarithmic time scale. At each delay, 3000 spectra were recorded and the TA spectra were calculated according to

$$\Delta A(\lambda_{\text{probe}}, t) = -\lg \left[\frac{I_{\text{pumped}}(\lambda_{\text{probe}}, t)}{I_{\text{unpumped}}(\lambda_{\text{probe}})} \right], \quad (1)$$

where $I_{\text{pumped}}(\lambda_{\text{probe}}, t)$ and $I_{\text{unpumped}}(\lambda_{\text{probe}})$ represent the probe intensity in presence and absence of a pump pulse, respectively. Afterwards, the 1500 TA spectra per delay time were averaged. Altogether, at least 8 TA maps are averaged per sample. The presented TA maps are corrected for background light and for the chirp of the probe beam,^{30,31} for which the temporal center of the coherent artifact is extracted and fitted with a third-order polynomial. The analysis of the TA maps is performed using the software GloTarAn³² which also accounts for the instrument response function and the coherent artifact.

2.3 Computational methods

The *ab initio* calculations were carried out with the TURBOMOLE³³ program package. For the ground-state geometry optimizations the B3LYP-D3(bj)/def2-TZVP³⁴⁻³⁷ functional and basis set are used. They are combined with MP2/def2-TZVP single point calculations to obtain the relative energies of the conformers. The MP2 and ADC(2) calculations used the resolution-of-the-identity approximation with the corresponding optimized

auxiliary basis sets³⁸ and the frozen-core approximation on the optimized SCF 1s core orbitals of the C, N and O atoms. The polarizable conductor-like screening model (COSMO)³⁹ is used to calculate the MP2 ground-state energies (COSMO-MP2) in aqueous solution ($\epsilon = 50$, $r_{\text{solv}} = 1.3333 \text{ \AA}$). The solvation effects are also included in the excitation energy and transition moment calculations with PTED-COSMO-ADC(2)⁴⁰ including a perturbative correction for the state specific contribution to excitation energies (hereafter COSMO-ADC(2)) reported in Table 2 and the simulated absorption spectra shown in Fig. 2b.

The vertical excitation energies and one-photon oscillator strengths are calculated with the algebraic diagrammatic construction through second⁴¹ order, ADC(2), a propagator method which is based on the same ground-state wavefunction as MP2. From a practical point of view, ADC(2) can also be seen as a modification⁴² of the coupled-cluster method CC2,⁴³ which in difference to the latter has a hermitian secular matrix and uses, like MP2, a strict first-order truncation for the ground-state wavefunction. ADC(2) provides the same accuracy⁴⁴ for excitation energies as CC2, but has about four times lower computational costs for the calculation of gradients,⁴² *i.e.* excited-state structure optimization and intrinsic reaction coordinate⁴⁵ (IRC) calculations.

The ground- and excited-state relaxation of the ET conformer are studied by analyzing the IRC paths calculated with MP2/SVP (ground state) and ADC(2)/SVP (excited state). The calculations have been started from the Frank-Condon (FC) point or (slightly distorted) transition state or bifurcation point (second-order saddle point) structures. The initial velocities of the atoms were set to zero. Then the molecule is moved on the Born-Oppenheimer potential energy surface (PES) with respect to the forces that are obtained from classical Newtonian mechanics with a fixed time interval. In the next step, the atomic velocities are again set to zero and the last step is repeated until the components of the Cartesian molecular gradients became smaller than a defined threshold. With a properly small time interval, this procedure allows to smoothly probe the PESs and provides a way to map the connectivity network between the stationary points. The so obtained IRC paths are used in the current work to determine relaxation and isomerization mechanisms. Furthermore, the last point of each IRC calculation was used for a numerical (or an analytical) frequency calculation at the same level of theory and as an initial structure for a transition state or a geometry optimization calculation, which is used to confirm that two stationary points are indeed connected. The reaction pathways for the conversion of the *Z*-azo,*s*-*trans* and *E*-azo,*s*-*cis* conformers into the *E*-azo,*s*-*trans* conformer on the ground-state PES were obtained with the chain-of-states method⁴⁶ which is implemented in woelfling program of the TURBOMOLE package. Finally, the excited-state absorption (ESA) spectra⁴⁷ and relative energies of the ground- and excited-state stationary points of the most stable conformers are calculated with ADC(2)/aug-cc-pVTZ using structures optimized, respectively, at the MP2/cc-pVTZ (ground state) and ADC(2)/cc-pVTZ (excited state) level of theory.

3 Results and discussion

3.1 Computational assessment of ground-state conformers

The computational investigation starts with searching for the most stable ground-state conformers. Starting from different initial structures of berenil, we could identify six distinct conformers, which appear as three pairs of spatial conformers of the triazene functional group as shown in Fig. 1 based on the B3LYP-D3(bj)/def2-TZVP geometry optimizations: *E*-azo,*s*-*trans* (ET1,ET2), *E*-azo,*s*-*cis* (EC1,EC2) and *Z*-azo,*s*-*trans* (ZT1,ZT2) pairs where 1 and 2 stand for gauge and eclipsed relative orientation of the terminal amidinium functional groups that are twisted by ≈ 30 degrees out of the molecular plane. Cartesian coordinates of all conformers are available in ESI.†

The DFT-optimized geometries were used for single point calculations with MP2/SVP and MP2/def2-TZVP. The close agreement of the energy differences between the conformers at the MP2/SVP, MP2/def2-TZVP and DFT/B3-LYP-D3(bj)/def2-TZVP levels of theory shows that this molecule is not a strongly correlated system and that the DFT optimized geometries are reliable. Therefore, the second-order treatment of the correlation energy in the MP2 and ADC(2) methods should be sufficient. Both DFT and MP2 agree on the fact that the ZT1 and ZT2 conformers have high relative energies. Furthermore, similar Boltzmann probabilities for ET1 and ET2 show that the relative orientation of the amidinium groups is not important for the energy. With a set of single point energy calculations with the SVP basis set (which has less basis functions) at the MP2 level, the basis set dependence of the relative energies of conformers is investigated. The insensitivity of the results to this change indicates that the smaller SVP basis set should be sufficient for the relaxation and isomerization pathway calculations. Finally, based on these calculations, we conclude that in vacuum and

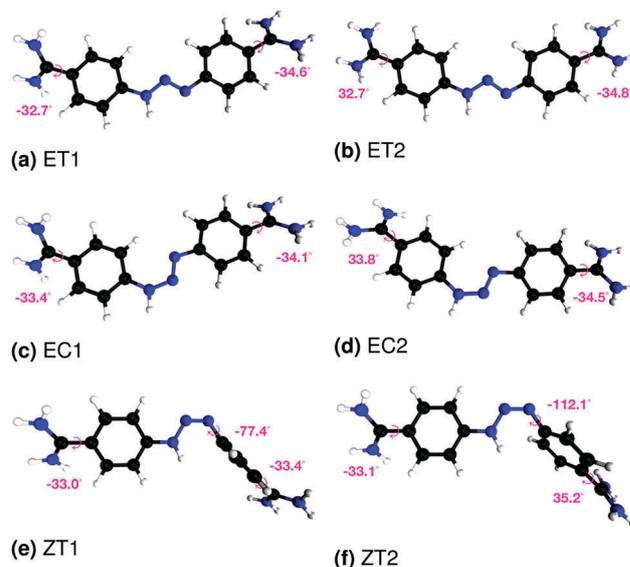


Fig. 1 The structure of berenil ground-state conformers computed at B3LYP-D3(bj)/def2-TZVP level of theory. All of the conformers are approximately planar with respect to torsion angles that are not specified.

Table 1 The relative energies in kcal mol⁻¹ and relative Boltzmann probabilities (W_B) at $T = 298$ K of different conformers with different methods and basis sets. Vacuum results are indicated by "VAC" and the COSMO results by "COSMO"

Conformer	Method	Basis set	ΔE	W_B
ZT1	B3-LYP(VAC)	def2-TZVP	14.3	0.00
ZT2	B3-LYP(VAC)	def2-TZVP	14.3	0.00
ET1	B3-LYP(VAC)	def2-TZVP	0.0	0.50
ET2	B3-LYP(VAC)	def2-TZVP	0.0	0.50
EC1	B3-LYP(VAC)	def2-TZVP	4.4	0.00
EC2	B3-LYP(VAC)	def2-TZVP	4.4	0.00
ZT1	MP2-VAC	def2-TZVP	14.4	0.00
ZT2	MP2-VAC	def2-TZVP	14.4	0.00
ET1	MP2-VAC	def2-TZVP	0.0	0.50
ET2	MP2-VAC	def2-TZVP	0.0	0.50
EC1	MP2-VAC	def2-TZVP	4.0	0.00
EC2	MP2-VAC	def2-TZVP	4.0	0.00
ZT1	MP2-VAC	SVP	15.6	0.00
ZT2	MP2-VAC	SVP	15.6	0.00
ET1	MP2-VAC	SVP	0.0	0.50
ET2	MP2-VAC	SVP	0.0	0.50
EC1	MP2-VAC	SVP	4.0	0.00
EC2	MP2-VAC	SVP	4.0	0.00
ZT1	COSMO-MP2	SVP	16.3	0.00
ZT2	COSMO-MP2	SVP	16.3	0.00
ET1	COSMO-MP2	SVP	0.0	0.50
ET2	COSMO-MP2	SVP	0.0	0.50
EC1	COSMO-MP2	SVP	3.3	0.00
EC2	COSMO-MP2	SVP	3.2	0.00

thermodynamically equilibrated conditions at $T = 298$ K, the most stable and dominantly occupied conformers of berenil are the *E*-azo,*s-trans* conformers (ET1 and ET2) with the same Boltzmann probabilities (W_B). The COSMO-MP2 results listed in Table 1 show that the ET1 and ET2 conformers stay the most stable conformers in aqueous solutions with the same probability. The relative orientation of the amidinium groups plays again no role for the ground-state energies even with inclusion of implicit solvent interactions. Also, the other conformers show this trend consistently at different levels of theory.

3.2 Simulation of steady-state absorption

The vertical excitation energies and oscillator strength for the 5 lowest electronically excited states of berenil are calculated at the COSMO-ADC(2) level with the aug-cc-pVDZ (aDZ) and SVP basis sets using the DFT-optimized ground-state geometries. The results are shown in Table 2. In general, for the most stable ground-state conformers ET1 (Fig. 1a) and ET2 (Fig. 1b), the COSMO and vacuum results are close. Furthermore, the minima of the unrelaxed torsion potentials, calculated at COSMO-ADC(2)/SVP level (available in ESI,† Fig. S1) are close to the fully optimized vacuum ADC(2)/SVP structure and thus the optimized excited-state geometries found in the vacuum calculations can be expected to be close to those that would be found with COSMO. Although in COSMO-ADC(2) calculations the vertical excitation energies of the lowest excited states of the ET conformer change by approximately 0.2 eV from the SVP to the aDZ basis set, a natural transition orbital (NTO) analysis indicates that the character of the excitation of the first two

Table 2 Electronic transitions of the ET and EC conformers calculated at the ADC(2) and COSMO-ADC(2) levels with the SVP and aDZ basis sets. Vertical excitation energies (ΔE in eV) and oscillator strengths (f) are shown

State	Vacuum		COSMO			
	SVP		SVP		aDZ	
	ΔE	f	ΔE	f	ΔE	f
ET1 _{S1}	3.54	1.30	3.52	1.28	3.32	1.17
ET1 _{S2}	3.73	0.01	3.87	0.00	3.84	0.00
ET1 _{S3}	4.61	0.00	4.66	0.01	4.53	0.01
ET1 _{S4}	4.65	0.09	4.77	0.02	4.58	0.08
ET1 _{S5}	4.75	0.02	4.92	0.12	4.74	0.09
ET2 _{S1}	3.54	1.30	3.52	1.27	3.32	1.17
ET2 _{S2}	3.73	0.01	3.88	0.00	3.84	0.00
ET2 _{S3}	4.61	0.00	4.66	0.01	—	—
ET2 _{S4}	4.65	0.09	4.77	0.02	—	—
ET2 _{S5}	4.75	0.02	4.92	0.11	—	—
EC1 _{S1}	3.60	0.29	3.64	1.14	3.49	1.17
EC1 _{S2}	3.65	0.92	3.74	0.08	3.69	0.01
EC1 _{S3}	4.64	0.02	4.71	0.02	4.58	0.02
EC1 _{S4}	4.73	0.04	4.78	0.00	4.64	0.01
EC1 _{S5}	4.77	0.05	5.00	0.09	4.78	0.07
EC2 _{S1}	3.59	0.39	3.64	1.04	3.49	1.15
EC2 _{S2}	3.68	0.81	3.76	0.17	3.70	0.04
EC2 _{S3}	4.65	0.02	4.71	0.02	—	—
EC2 _{S4}	4.73	0.04	4.79	0.00	—	—
EC2 _{S5}	4.77	0.05	5.01	0.09	—	—

excited states are similar for the SVP and aDZ basis sets (see ESI,† Tables S1–S4) and the SVP results are still accurate enough for qualitative purposes regarding the stationary points and the dynamics while the ADC(2)/aDZ calculations can reproduce the experimental results almost quantitatively. Moreover, it is seen that solvent effects have a minimal impact on both the vertical excitation energies (a 0.08 eV solvatochromic red-shift with aDZ basis set) and the oscillator strength of the lowest excited state (a bright state with $N=N_{\pi\pi^*}$ character) of the ET conformer while they cause a 0.15 eV blue-shift of the second excited state, a dark state with $N=N_{n\pi^*}$ character. The aDZ results are available in the ESI,† Tables S3 and S4. The observed solvatochromic red-shift of the bright state can be understood by the interaction between the transition density of the solute and the solvent which stabilizes excited states typically the more the larger the transition moments are. On the other hand, $n\pi^*$ excitations of organic chromophores are in polar solvent usually blue-shifted because of the reduced interaction between the lone-pair (n) and the solvent in the excited state. This opposite response of $\pi\pi^*$ and $n\pi^*$ excited states can cause a switching between the states when they are close in energy and, indeed, in the case of the EC conformer, the NTO analysis of the vacuum ADC(2)/SVP results shows an $N=N_{n\pi^*}$ character for the lowest excited state while at the COSMO-ADC(2)/SVP level, the lowest excited state has $N=N_{\pi\pi^*}$ character. With the larger aDZ basis, however, the $N=N_{\pi\pi^*}$ state is already in vacuum the lowest excited state and stays the lowest in solution. The energy gaps between the first and second excited state are with the aDZ basis 0.1 eV in vacuum and 0.2 eV in aqueous solution. NTOs from the (COSMO)-ADC(2)

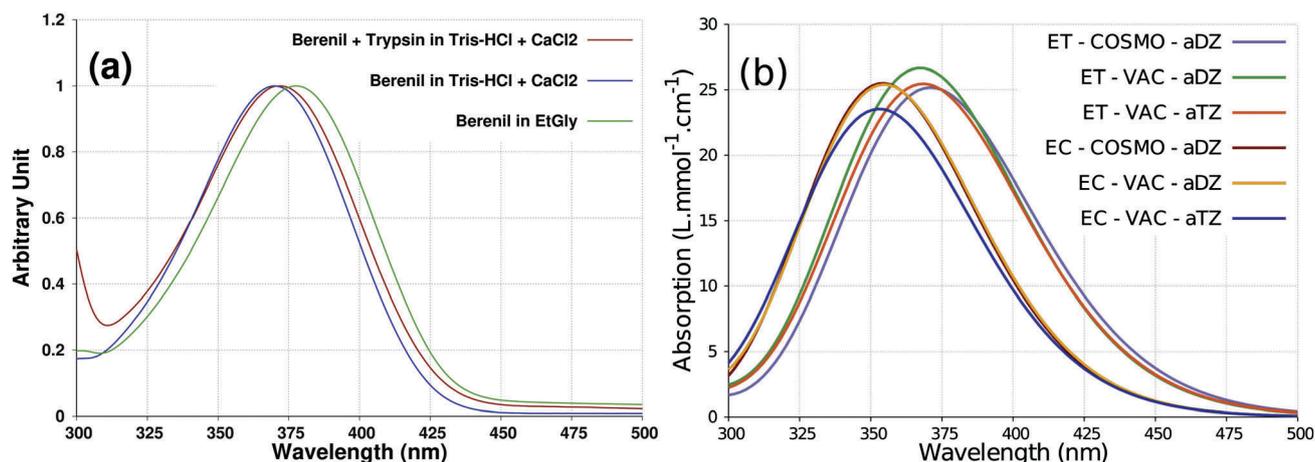


Fig. 2 (a) Normalized experimental steady-state absorption spectra of berenil in different environments. (b) Steady-state absorption spectra calculated at COSMO-ADC(2)/ and ADC(2)/aug-cc-pVXZ (X = D,T) levels with 0.3 eV broadening factor.

calculations with the SVP and aDZ basis sets are available in Tables S5–S8 of the ESI.† Because of the small energy gap between the $N=N_{\pi\pi^*}$ and $N=N_{n\pi^*}$ states one can expect relatively complex excited-state dynamics for the meta-stable EC conformer.

To further validate the stability of the results with respect to the basis-set size and to compare the theoretical and experimental spectra, the steady-state absorption spectra of berenil under different experimental conditions and the simulated spectra from COSMO-ADC(2)/aDZ, ADC(2)/aDZ and ADC(2) in the aug-cc-pVTZ (aTZ) basis are shown in Fig. 2. The experimental extinction coefficient of berenil in distilled water is reported to be $\epsilon_{\text{exp}} = (26.5 \pm 2.3) \text{ mM}^{-1} \text{ cm}^{-1}$ at a wavelength of 368 nm,⁴⁸ which is very close to the theoretical values (373 and 368 nm) with the aDZ and aTZ basis sets, respectively. Also the SVP results are only shifted by 0.2 eV. Therefore, we will use the both types of basis sets to take advantage of their performance and accuracy, *i.e.* using the aug-cc-pVXZ (X = D,T) basis sets whenever a quantitative comparison with the experimental results is needed and the smaller SVP basis set for the calculation of deexcitation and reaction pathways. Moreover, the results in Table 2 reveal that the first electronically excited state has the highest oscillator strength among those that appear below 5 eV. Table 2 also shows that neither the vertical excitation energies nor the oscillator strengths change with the relative orientation of the amidinium groups for both the EC and the ET conformers. Therefore, hereafter, we will only consider the ET1, EC1 and ZT1 conformers and will omit the “1” from the abbreviated names.

The above conclusions corroborate our previously reported study²⁷ of the excited-state pathways of berenil where the lead conformer, ET, was analyzed to understand the measured fluorescence upconversion data and to clarify the deexcitation mechanism. Albeit our calculations predict a probability close to 0% for the conformers other than ET, we want to point out that the vertical excitation energies of the lowest singlet excitation of the ET and EC conformers are as close as ≈ 0.1 eV and therefore it would be difficult to identify putative trace amounts

of the less stable conformer EC based on the steady-state absorption spectrum. Thus, in the calculations further below, we will concentrate on the most stable conformer ET, as the lead conformer of berenil, but also investigate the less stable EC and ZT conformers to examine if these conformers play a role in the photochemistry of berenil.

Due to the hydrogen atom bound to one of the nitrogens of the triazene bridge, linear triazenes have in principle the possibility of a double-bond migration by a hydrogen transfer, which could involve solvent molecules or even comprise tunneling contributions. For 1,3-diphenyltriazene, the ground-state backreaction from a *Z* to an *E* isomer proceeds significantly faster in protic solvents than in aprotic ones (ms *versus* 100 ms range), which was explained by a solvent-mediated hydrogen transfer.^{20,22} We note that the ET conformers of berenil would stay the same because of the symmetry of the molecule, whereas EC and ZT could, in principle, interconvert and both conformers, ZT and EC can then isomerize to the ET structure which explains why the ET isomer is always reformed in our samples even when photoisomers are generated (see below).

3.3 Transient absorption measurements

The TA maps of berenil in different environments are displayed in Fig. 3 (left columns). At first glance, all maps look very similar and show a negative signal centered around 370 nm, a positive transient signal around 480 nm, and a weak positive signal centered at 300 nm. The most pronounced differences arise at the long-wavelength region of the maps.

A global analysis of the TA data recorded using the visible white-light was performed, yielding three time constants for the TA map of the buffer and the ethylene glycol solutions, and four time constants when berenil is bound to trypsin. The decay-associated spectra (DAS) are shown in Fig. 3 (right panels). The DAS component associated with the shortest decay time (red) of 0.36 ps for the buffer sample, 0.40 ps for bound berenil and 0.34 ps for the ethylene glycol sample shows negative features in the region between 320 nm and 420 nm and in the range

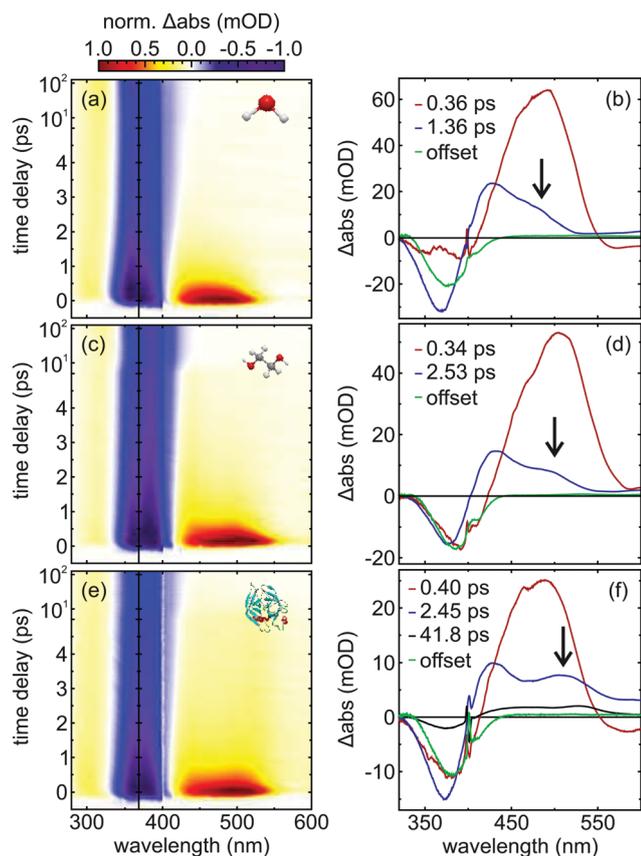


Fig. 3 Left column: Transient absorption maps for berenil pumped at 400 nm and probed in the region between 280 and 368 nm with UV white-light (narrow panel) and in the visible white-light region between 368 and 600 nm (wide panel). At 368 nm the data is normalized to -1 for comparison (see color coding). At 400 nm scattered pump light is visible. Right column: Corresponding decay associated spectra (DAS) and relaxation time scales resulting from global fitting using GloTarAn. The data is recorded for berenil in 50 mM Tris-HCl and 25 mM CaCl₂ aqueous buffer solution (a and b), for berenil in ethylene glycol (c and d), and for berenil bound to trypsin in 50 mM Tris-HCl and 25 mM CaCl₂ aqueous buffer solution (e and f). The arrows indicate the transient feature discussed in the text.

between 550 nm and 600 nm; however, this latter feature is only clearly visible for the buffer and trypsin sample but not for the ethylene glycol solution due to a general slight solvatochromic red-shift of the spectra. A large positive feature is present for all samples in the region above 410 nm. Our previous study has shown a broad fluorescence peak between 440 and 600 nm,²⁷ hence most of the corresponding stimulated emission (SE) signal overlaps with the positive signal, yet the negative signal above 550 nm confirms the SE contribution. The negative signal in the spectral region below 420 nm can be assigned to partial ground-state bleach (GSB) recovery, since the wavelength region matches well the absorption spectra shown in Fig. 2. The large positive signal red-shifted compared to the GSB can be assigned to excited-state absorption (ESA); we note that due to the parallel fit model of DAS, the absorption at early delay times is actually the sum of all DAS and not just represented by the red line. Thus, the first time constant describes berenil in the electronic excited state, in accordance

with the average excited-state fluorescence decay time,²⁷ but with the difference that the overlapping of SE, GSB, and ESA does not allow an explicit decomposition of these excited-state dynamics into two sub-ps processes as in the emission studies.

The DAS offset component (green) shows a negative feature for all samples in the region between 330 and 440 nm. This component directly reveals a remaining GSB signal still observable after some ns. However, an absorption measurement right before and after the transient absorption experiment reveals no permanent change of the absorption. Therefore, the offset is not associated with the formation of a photoproduct that remains intact for infinite times, but it rather originates from a photoproduct that reforms to the reactant on a timescale of nanoseconds to seconds. The positive absorption at 300 nm observable in the UV white-light studies (narrow panels in Fig. 3a, c and e) which persists to ns delays can be assigned to this photoproduct (see Fig. S3 in ESI† for a DAS of berenil in buffer for the UV white-light study showing contributions from ESA in the first ps and the formation of the photoproduct). Since the steady-state absorption spectra reveal no permanent changes, it is likely to assume structural changes like a *cis-trans* photoisomerization. Our previous combined theoretical and experimental study already implied the bicycle-pedal motion as isomerization mechanism,²⁷ thus a different isomer can lead to the product absorption at 300 nm. The quantum yield for this photoisomerization is determined to be $\Phi = 0.48$ for the buffer sample, $\Phi = 0.46$ for the ethylene glycol sample and $\Phi = 0.47$ when berenil is bound to trypsin. This is in good agreement with the quantum yield of $\Phi = 0.47$ reported for the *E-Z* photoisomerization with respect to the N=N bond of 1,3-diphenyltriazene in methanol,²² a system for which also the thermal ground-state backrelaxation was measured and occurs in the ms regime.^{20,22} The *Z* isomer of 1,3-diphenyltriazene absorbs blue-shifted relative to the *E* form which is predominant in solution,²¹ and generally the photogenerated *Z* isomers of triazenes exhibit an absorption peak at shorter wavelengths compared to the *E* isomers.¹⁹ Hence, our data shows that a new isomer is formed after photoexcitation, and the existing literature supports the assumption of a *Z* isomer absorbing around 300 nm. As will be shown further below, the structure of the photogenerated isomer is elucidated in the quantum-chemical calculations.

The DAS components associated with the second time constant (blue) exhibit a derivative-like shape, with negative features below 400 nm and positive features above 400 nm. However, a significant difference is observable in the region around 500 nm, as indicated by arrows in Fig. 3. For berenil dissolved in buffer, a tiny shoulder around 485 nm appears. In the ethylene glycol solution the shoulder gets more pronounced and shifts to 500 nm. When berenil is bound to trypsin, a peak is clearly visible at 510 nm. Furthermore, the ratio of the positive and negative areas yields 1.4 for the buffer sample and 2 for the ethylene glycol and trypsin sample, indicating the emergence of a new peak. The associated time constants increase from 1.36 ps for the buffer sample, to 2.45 ps when bound to trypsin and 2.53 ps when dissolved in ethylene glycol. The derivative-like shape is typical for the description of a

shift^{49,50} due to solvation and/or cooling processes.^{51–53} The typical solvation time of water is on the order of 1 ps,⁵⁴ which would approximately match with τ_2^{buffer} , however the solvation time for the highly viscous solvent ethylene glycol is 30 times higher.^{55,56} Therefore, solvation is likely not the process giving rise to τ_2 . Cooling processes are mainly observable in systems with fast excited-state dynamics,⁵⁷ because in this case most of the excitation energy is not released by a radiative process. Instead, the energy remains in the molecule and has to be removed by vibrational cooling. The ultrafast excited-state dynamics observed for berenil hint to internal conversion leading to the population of highly excited vibrational states of the electronic ground state. This vibrational excess energy results in a red-shift of the ground-state absorption and a shifting to shorter wavelength with increasing delay time.^{57–61} Since water has a larger number of low-frequency modes,⁶² it can accept the excess energy more efficiently than ethylene glycol,⁶³ reflected in a prolongation of τ_2^{EtGly} . Hence, τ_2 is assigned to hot ground-state absorption, with τ_2^{trypsin} showing a value close to τ_2^{EtGly} implying cooling as fast as in ethylene glycol.

The assignment of the new peak at 500 nm allows various interpretations that will be discussed in the following. The red-shift in comparison to the ground-state absorption indicates an increase of the π -system of berenil. The deprotonation of the triazene bridge would lead to such an increased π -system, as is corroborated by the shifting of the absorption peak of berenil dissolved in NaOH buffer (pH 13.5) to 425 nm (see Fig. S5 in ESI†). In order to exclude that the deprotonation occurs at the amidinium groups, we performed test measurements with another triazene with different terminal groups (thiazole yellow, see ESI† for the molecular structure and Fig. S5 for the data) dissolved in water and in basic solution, which exhibit the same behavior implying that strong basic media lead to the deprotonation of the triazene bridge. Despite the well-matching spectral regions for the absorption of deprotonated berenil and the new transient peak indicated by the arrows in Fig. 3, the latter can barely be explained by deprotonation, because the time scale of 2.5 ps seems too fast for a deprotonation and subsequent reprotonation, also when considering diffusion of a released proton.⁶⁴ In a former study of 1,3-diphenyltriazene, the formation of cyclic dimers is proposed²⁰ which might lead to a red-shifted absorption signal in comparison to the monomer. However, the observed feature is most pronounced when berenil is bound to trypsin, where dimer formation would be sterically hindered. Another possibility is that the feature originates from a different isomer of berenil absorbing red-shifted in comparison to the ET isomer. We tend to discard this scenario, because the calculations of Section 3.1 have disclosed that only the ET isomer is stable at room temperature, and our sample volume is exchanged continuously so that also no photoproducts can accumulate. Regarding the ratio of positive and negative areas of the DAS exhibiting the peak around 500 nm, the relative increase from 1.4 (in buffer) to 2 (bound to trypsin) might indicate that the new peak is not related to the hot ground-state absorption (which then would look similar in all cases) but rather an additional process occurring on the

same time scale. The fluorescence upconversion study²⁷ has shown that there is an emissive contribution with 1.4 ps in the case of berenil bound to trypsin, so at least some population survives for more than 1 ps in the ES. Since the peak is more pronounced for more restricted environments, it seems to be likely that a deactivating motion of berenil is hindered, as might also be reflected in the DAS associated with τ_3 only found for trypsin. The calculations hint towards a further possible origin related to distorted ground-state structures transiently present during the relaxation process (see Discussion in Section 3.5).

In the case of trypsin, the global fitting leads to this third DAS contribution (black component in Fig. 3f) with a lifetime of 41.8 ps, revealing also a derivative-like shape with a negative feature in the region between 330 and 400 nm and a positive feature for wavelengths larger than 400 nm. This positive feature is characterized by an additional peak at 525 nm. For strong solute–solvent interactions *via* hydrogen bonding, biexponential cooling dynamics are expected, because of the local heating of the first solvation shell.⁶⁵ The large binding constant ($59.6 \times 10^4 \text{ M}$)⁶⁶ points towards a strong interaction between berenil and trypsin, suggesting a release of the excess energy in two different ways: in addition to the fast component of 2.45 ps, the slower component of 41.8 ps might reflect how the strongly-bound trypsin, which also gets some of the excess energy, cools and thereby influences the berenil absorption features. This slowed-down cooling of berenil bound to trypsin could further be connected to the restricted volume available for the solvent molecules in the pocket where berenil binds.

Whereas the possibility of double-bond migration by hydrogen transfer in the ground state was discussed in Section 3.2, we also want to mention the scenario of an excited-state intramolecular hydrogen transfer (ESIHT), which (*e.g.* for keto–enol tautomerism) occurs on a timescale of 100 fs or even faster.^{67–73} In this respect, the calculations of the ground-state conformers tells us that only the ET conformers are initially excited, which upon ESIHT would again yield ET isomers. That is, even if there was an extremely fast initial ESIHT contribution, the observed formation of a photoproduct must originate from another molecular motion, as substantiated by calculations in the following.

3.4 Deexcitations, from the excited state to the ground state

Fig. 4a shows the IRC pathways for the deexcitation after vertical excitation of the ET conformers (IRC₁–IRC₅) from the FC point *via* a transition state (ET_{TSS1}) to the minima ET_{S1} and ET_{RS1} and after deexcitation to the S₀ electronic state back to the ET ground-state equilibrium structure. The last point of each of these IRC calculations is used for a numerical frequency calculation at the same level of theory and for a transition state or geometry optimization calculation, which is used to confirm that the two stationary points are indeed connected. From Fig. 4a, it is apparent that in the first phase of the deexcitation process the N=N azo-bond is stretched which leads to a planar transition state, ET_{TSS1}. In the second phase of the deexcitation, the N=N double bond is rotated by $\pm 90^\circ$ which brings the structure close to a conical intersection between S₀ and S₁. In the earlier stage of the second phase of the S₁ state deexcitation,

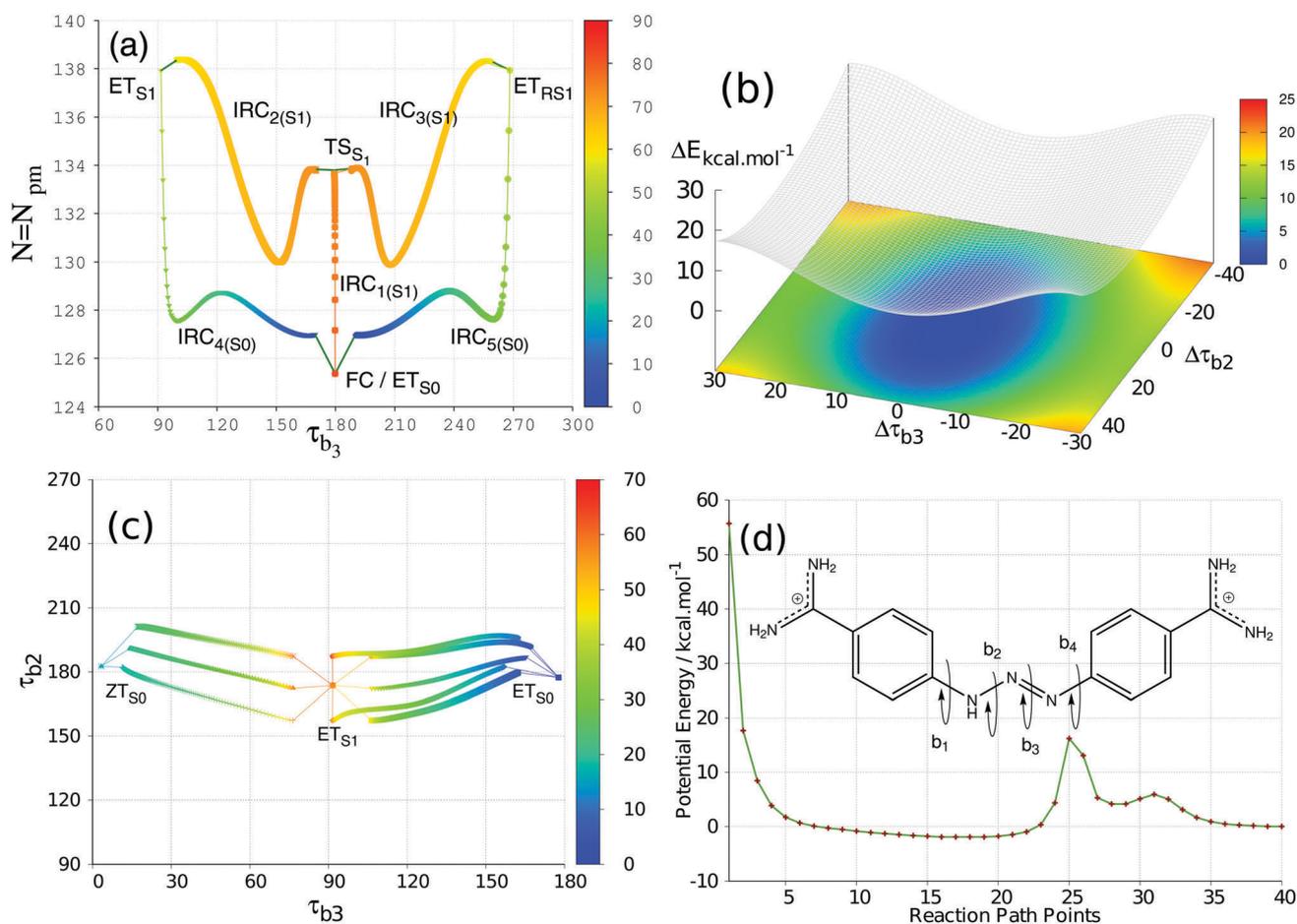


Fig. 4 (a) Evaluation of the N=N bond length and the C-N=N-N torsion angle (τ_{b3}) along the IRC₁, IRC₂ and IRC₃ paths calculated at ADC(2)/SVP level at the lowest electronically excited state and the IRC₄ and IRC₅ ground-state paths calculated at MP2/SVP level of theory. (b) Interpolated unrelaxed 2D torsion potential in the vicinity of the ET lowest electronically excited-state minimum, ET_{S1} calculated at COSMO-ADC(2)/SVP level. (c) Evaluation of the τ_{b2} and τ_{b3} torsion angles along the MP2/SVP IRC paths at the ground state starting from the screwed ($\pm 15^\circ$ around τ_{b2} and τ_{b3}) ET_{S1} geometry. (d) MP2/SVP optimized reaction path starting from the ET_{S1} to EC_{S0} with 40 points and the contracted names for different torsion angles.

IRC₂ and IRC₃ in Fig. 4a, the N=N bond length decreases from 134 pm to 130 pm since the N=N bond rotation diminishes the N=N π -antibonding interaction. In the next stage of the process, the N=N bond elongates to compensate the steric and Coulomb interactions between two positively charged amidino-phenyl units. The clockwise and anti-clockwise rotation around the N=N bond is energetically almost degenerate. They differ from each other's mirror image only in the torsion angles of the amidinium groups. Further details are given in the ESI.† The proposed two-phase deexcitation mechanism is confirmed by the reported fluorescence upconversion measurements.²⁷ The ultrafast non-radiative deexcitation of the molecule can be explained by the possibility that the molecule is funneled to the ground state *via* a conical intersection between S₀ and S₁ in a non-radiative process. Monitoring the central torsion angles along the IRC₃ trajectory, Fig. 5a reveals that the second phase of the deexcitation mechanism follows a bicycle-pedal type motion, which is a volume-conserving deexcitation path and, thus, the ultrafast deexcitation of berenil even in viscous and mechanically confined environments can be understood.

The ground-state structures of the ET and ZT conformers (ET_{S0} and ZT_{S0}) and the structure of the lowest excited state minimum ET_{S1} are compared in Fig. 6. For further details about the two-phase deexcitation process and the definition of different relevant deexcitation mechanisms we refer to ref. 27.

To investigate the ground-state isomerization processes, IRC calculations (MP2/SVP) are pursued from the ET_{S1} and ET_{RS1} geometries on the ground-state PES. Both paths end in the ET_{S0} minimum, which shows that the excited molecules can be deexcited to the ground-state minimum without isomerization, IRC₄ and IRC₅ in Fig. 4a. In the ET region the relaxation on the S₁ PES funnels molecules through a conical intersection radiationless to the ground state. On the S₀ PES the relaxation proceeds through a two-step relaxation mechanism: an initial N=N bond contraction (IRC₄ and IRC₅ in Fig. 4a) which reduces the potential energy substantially and a bicycle-pedal type motion to the ET_{S0} minimum, Table 3.

However, triazene is a flexible functional group and the deexcitation does not necessarily happen at the S₁ minima or minima of the intersection seam. To examine the flexibility of

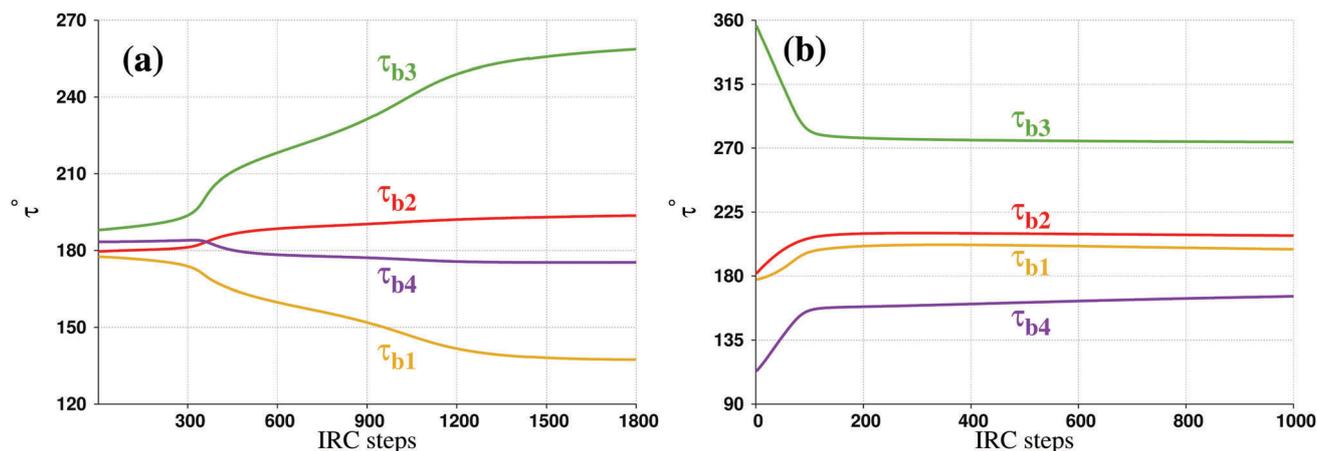


Fig. 5 Evaluations of (a) the ET conformer's triazene torsion angles along the IRC_{3(S₁)} path calculated at ADC(2)/SVP level (b) the ZT conformer's torsion angles along the S₁ deexcitation IRC pathway calculated at ADC(2)/SVP level, starting from the Franck–Condon point, ZT₁ geometry in Fig. 1e. Geometry optimizations, starting from the final structure of both IRC paths, (a) and (b), are converged to the same minimum on the S₁ excited state, ET_{RS1}.

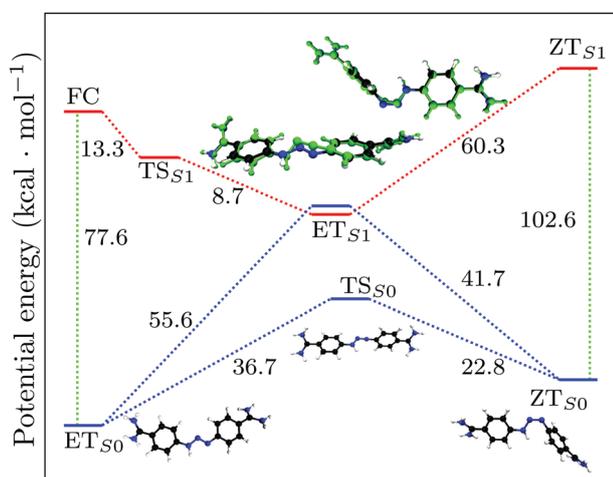


Fig. 6 Ground-state and excited-state isomerization paths of ET to ZT calculated at ADC(2)/cc-pVTZ/aug-cc-pVTZ level of theory. Color codes: (solid line) stationary and Franck–Condon point, (dotted line) geometrical changes, (blue) electronic ground-state PES, (red) lowest electronic excited state, (green) bright electronic transitions. The ground state geometries of ET and ZT conformers (atoms colors) are superimposed to the geometry of the lowest excited-state minimum, ET_{S1} (green).

the molecule on the excited-state PES, the unrelaxed torsion potentials of τ_{b2} and τ_{b3} calculated at COSMO-ADC(2)/SVP level are interpolated with a polynomial for developing a minimal additive forcefield (see Table S9 in ESI[†]) to construct a two-dimensional PES around the ET_{S1} geometry as a function of τ_{b2} and τ_{b3} , Fig. 4b. The latter shows that the ET_{S1} structure is flexible enough to tolerate up to 15 degrees rotation around the τ_{b2} and τ_{b3} torsion angles. Based on this result, 8 MP2/SVP IRC calculations are run with $\pm 15^\circ$ distorted ET_{S1} geometries for 500 steps, followed by MP2/SVP geometry optimizations. Fig. 4c shows the τ_{b2} and τ_{b3} torsion angles along these 8 IRC paths. It is shown that, in addition to the deexcitation path to the ET_{S0} ground-state geometry, photoisomerization to ZT_{S0} is also possible. Results in Table 3 show that on the ground-state

Table 3 The classifications of deexcitation mechanisms are based on the maximum changes of the triazene torsion angles along the ground- and excited-state IRC pathways of the ET and ZT isomers. The Hula-twist and bicycle-pedal type mechanisms are shown with HT and BP, respectively

Transition	$\Delta\tau_{b1}$	$\Delta\tau_{b2}$	$\Delta\tau_{b3}$	$\Delta\tau_{b4}$	Mechanism
ET _{TSS1} –ET _{S1}	40.53	13.66	70.51	10.77	BP
ET _{S1} –ET _{S0}	56.80	16.55	78.03	6.13	BP
ET _{S1} –ZT _{S0}	19.58	28.04	74.27	38.42	HT
ZT _{S1} –ET _{S1}	24.56	28.55	81.92	67.21	HT

surface, the relaxation path from ET_{S1} to ET_{S0} follows a bicycle-pedal type motion while the isomerization path to ZT_{S0} involves a hula-twist motion.

The calculated wavelength for the vertical excitation of the ZT isomer (318 nm) agrees reasonably with the experimental positive product signal of the TA maps at ≈ 300 nm. If the ZT isomer is excited from the ground state to the S₁ state, it relaxes on the S₁ surface from the Franck–Condon (FC) point to the same S₁ minima as reached when the ET isomer is excited. The deexcitation consists of a N=N bond stretch and a hula-twist⁷⁴ motion as shown in Fig. 5b (note the changes of τ_{b3} and τ_{b4}). The missing (pseudo-) planar symmetry in the case of ZT in comparison to ET and EC is the reason for the different deexcitation mechanisms. Therefore, the IRC calculation starting from the Franck–Condon point leads to the S₁ minimum without meeting a transition state located between rotatory and disrotatory twisted S₁ minima. In Fig. 6 a schematic PES of the ET and ZT isomers and their thermal and photochemical isomerization pathways are given. Cartesian coordinates are available in ESI[†]. It is shown that following the photoisomerization of ET to ZT, ZT can isomerize to ET on the ground state *via* a transition state and an in-plane N=N bond inversion mechanism. Note that in addition to this mechanism, solvent-mediated double-bond migration might also contribute in the ET reformation in the ground state (Section 3.2).

As discussed above, the ET conformer can be deexcited to the ZT conformer *via* the S₁ minima. But none of the IRC paths

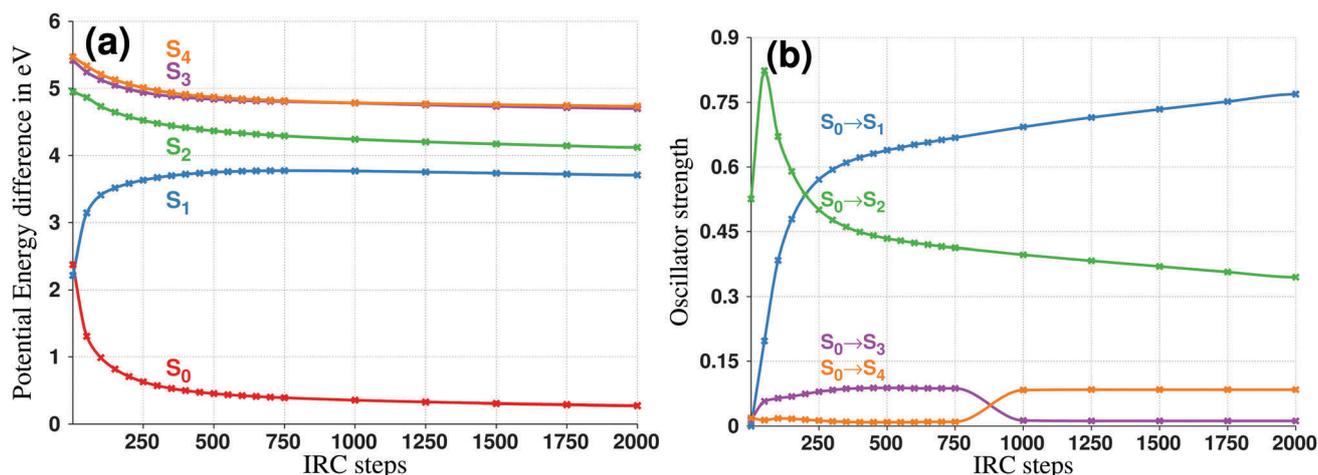


Fig. 8 (a) ADC(2)/aTZ energies of the S_0 – S_4 states along the IRC₅ (Fig. 4a) pathway in eV. (b) Evaluation of oscillator strengths for excitations from ground state (S_0) to higher singlet ESs of the IRC₅ pathway calculated at ADC(2)/aTZ level. The calculations are performed for 50-step intervals up to step 750 and 250-step intervals up to step 2000 of the IRC path.

excited state (S_1) is also around 370 nm, thus in this spectral region the calculated ESA signal might overlap with the GSB contributions. In the vicinity of the ET_{S₁}, molecular population of the S_1 excited state can be transferred to the ground state *via* the conical intersection. Unlike the ground-state minimum, the ET_{S₁} geometry is not planar, therefore, it is possible that electronic transitions that have very low oscillator strength in the steady-state absorption (by reason of symmetry) exhibit higher oscillator strength in the transient absorption. The vertical excitation energies and oscillator strengths of excitation from the S_0 state to the 4 lowest singlet ESs at the ET_{S₁} geometry (optimized at ADC(2)/cc-pVTZ level) are calculated with ADC(2)/aTZ, Table 4. At this geometry, the S_0 to S_2 is a bright excitation with a 496 nm excitation energy. The ground-state relaxation process continues on the ground-state PES with a two-step-like relaxation mechanism that is discussed in the previous section (Fig. 4a, IRC₄ and IRC₅). Fig. 8 shows the evaluation of vertical excitation energy and oscillator strength of the ground-state relaxation process along the IRC₅ path calculated at ADC(2)/aTZ level. The latter calculations and NTO analysis show that two bright excitations are involved in the ground-state relaxation process, the $S_0 \rightarrow S_1$ with a $\pi\pi^*$ character which is the dominant steady-state absorption signal and the $S_0 \rightarrow S_2$ with an $n\pi^*$ character which is transiently bright at the earlier stage of the ground-state relaxation process (blue and green curves in Fig. 8b, respectively). NTOs are given in Tables S10–S17 in ESI.† The ground-state relaxation process reestablishes the (pseudo-)planar structure of the ET_{S₁} conformer which causes the decline in the oscillator strength of the $n\pi^*$ excitation and the increase in the oscillator strength of the $\pi\pi^*$ excitations due to symmetry reasons while the excitation energies only change moderately, Fig. 8a. Based on these calculations, it is expected that in the transient absorption spectra, a signal with a lifetime close to the ET_{S₁} excited-state absorption peak (the 400 nm maximum of the second component of DAS spectra) appears around 496 nm (the 510 nm maximum of the second DAS

component in Fig. 3f) due to the induced oscillator strength of the $n\pi^*$ excitation caused by the molecular distortion on the ground state. While the molecule relaxes to its ground state (pseudo-) planar geometry, the oscillator strength of the $S_0 \rightarrow S_2$ excitation and the energy gap of S_1 and S_2 decrease, which enhance the $S_0 \rightarrow S_1$ and $S_0 \rightarrow S_2$ signals overlap. Therefore, there is no contribution from $S_0 \rightarrow S_2$ transient absorption at longer times and the steady-state absorption has only one peak. Moreover, for the more viscous solvents or confined environments, a stronger $S_0 \rightarrow S_2$ transient absorption signal is reasonable due to the slower ground-state relaxation process, which is consistent with the experimental observation (indicated by arrows in Fig. 3b, d and f).

Based on the above discussion, either the second DAS component is caused by a contribution from the molecules that are still in the excited state and exhibit an ESA around 375 nm or the geometrically induced $S_0 \rightarrow S_2$ hot ground-state absorption with a signal at 496 nm, or both. Note that the DAS signals result from a global fit of the TA data and especially the early time dynamics must be thought of as a sum of the DAS, thereby not representing isolated but basically all participating species. Therefore, for a quantitative comparison between the computational result and the experimental data, one would have to simulate the whole dynamic processes and, for example, consider the hydrogen-transfer process as well. Nevertheless, the static approach, that is used here, is useful for acquiring a qualitative insight into the deexcitation process.

4 Conclusions

In this joint experimental and computational study of the deexcitation dynamics of the triazine compound berenil in three different environments, it could be shown that the main spectroscopic features and deexcitation mechanisms are related to the excited-state (N=N stretching, bicycle pedal motion)

dynamics of the ET conformer occurring on a sub-ps time scale. By funneling through a conical intersection, the molecules return to the ground-state PES where one half of them reform the ET isomer and the other half isomerizes to the ZT isomer, which absorbs blue-shifted with regard to the ET isomer. Furthermore, IRC calculations of various relaxation pathways from the ET_{S1} region to the ground state and the reaction path optimization confirms that there is no barrier-less minimum energy photoisomerization pathway from the ET conformer to the EC conformer, although on the ground state there is a thermal isomerization pathway between ET and EC conformers. Through analysis of the transient absorption spectra and the calculations, a connection between experimental excited-state absorption signals and calculated absorption signatures of the transition state and the minimum of the S₁ state as well as of the geometrically induced S₀ → S₂ transition, which is environment-dependent and is more pronounced in the case of berenil bound to trypsin, could be made. It is worth noting that other processes, e.g. ground- and excited-state proton transfers and solvent- and protein-induced relaxation can happen on the same timescale, and thus compete with the discussed mechanisms and require additional efforts to be included in the follow-up simulations. The calculations have further unveiled that other conformers of berenil, although not present in the studied solutions, exhibit a rich photochemistry. Future studies might aim at (photo)-generating these conformers and then verify whether these new reaction channels are pursued.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This work has been supported by the Deutsche Forschungsgemeinschaft within the Cluster of Excellence RESOLV (EXC 1069).

References

- R. W. Schoenlein, L. A. Peteanu, R. A. Mathies and C. V. Shank, *Science*, 1991, **254**, 412–415.
- Q. Wang, R. W. Schoenlein, L. A. Peteanu, R. A. Mathies and C. V. Shank, *Science*, 1994, **266**, 422–424.
- Z. F. Liu, K. Hashimoto and A. Fujishima, *Nature*, 1990, **347**, 658–660.
- T. Ikeda and O. Tsutsumi, *Science*, 1995, **268**, 1873–1875.
- N. Koumura, R. W. Zijlstra, R. A. van Delden, N. Harada and B. L. Feringa, *Nature*, 1999, **401**, 152–155.
- C. A. Schalley, K. Beizai and F. Vögtle, *Acc. Chem. Res.*, 2001, **34**, 465–476.
- B. L. Feringa, *Acc. Chem. Res.*, 2001, **34**, 504–513.
- F. Renth, J. Bahrenburg and F. Temps, in *Photon-Working Switches*, ed. Y. Yokoyama and K. Nakatani, Springer, Japan, Tokyo, 2017, pp. 237–259.
- M. Quick, A. L. Dobryakov, M. Gerecke, C. Richter, F. Berndt, I. N. Ioffe, A. A. Granovsky, R. Mahrwald, N. P. Ernsting and S. A. Kovalenko, *J. Phys. Chem.*, 2014, **118**, 8756–8771.
- G. R. Fleming, S. H. Courtney and M. W. Balk, *J. Stat. Phys.*, 1986, **42**, 83–104.
- M. Böckmann, N. L. Doltsinis and D. Marx, *J. Chem. Phys.*, 2012, **137**, 22A505.
- M. Böckmann, S. Braun, N. L. Doltsinis and D. Marx, *J. Chem. Phys.*, 2013, **139**, 084108.
- I. Schapiro, O. Weingart and V. Buss, *J. Am. Chem. Soc.*, 2009, **131**, 16–17.
- A. Warshel, *Nature*, 1976, **260**, 679–683.
- D. Polli, P. Altoè, O. Weingart, K. M. Spillane, C. Manzoni, D. Brida, G. Tomasello, G. Orlandi, P. Kukura, R. A. Mathies, M. Garavelli and G. Cerullo, *Nature*, 2010, **467**, 440–443.
- J. Conyard, I. A. Heisler, Y. Chan, P. C. Bulman Page, S. R. Meech and L. Blancafort, *Chem. Sci.*, 2018, **9**, 1803–1812.
- N. Wiberg and H. J. Pracht, *Chem. Ber.*, 1972, **105**, 1377–1387.
- N. Wiberg and H. J. Pracht, *Chem. Ber.*, 1972, **105**, 1392–1398.
- D. Fanghänel, G. Timpe and V. Orthman, in *Organic Photochromes*, ed. A. V. El'tsov, Consultants Bureau (Division of Plenum Publishing Corp.), New York, 1990, pp. 136–153.
- J. Baro, D. Dudek, K. Luther and J. Troe, *Ber. Bunsenges. Phys. Chem.*, 1983, **87**, 1155–1161.
- J. Baro, D. Dudek, K. Luther and J. Troe, *Ber. Bunsenges. Phys. Chem.*, 1983, **87**, 1161–1164.
- J. C. Scaiano, C. Chen and P. McGarry, *J. Photochem. Photobiol., A*, 1991, **62**, 75–81.
- M. Barra and N. Chen, *J. Org. Chem.*, 2000, **65**, 5739–5744.
- B. A. Newton, in *Antibiotics III – Mechanism of Action of Antimicrobial and Antitumor Agents*, ed. J. W. Corcoran, F. E. Hahn, J. F. Snell, and K. L. Arora, Springer, Heidelberg Berlin, 1975, pp. 34–47.
- D. S. Pilch, M. A. Kirolos, X. Liu, E. Plum and K. J. Breslauer, *Biochemistry*, 1995, **34**, 9962–9976.
- V. Gabelica, F. Rosu, E. D. Pauw, R. Antoine, T. Tabarin, M. Broyer and P. Dugourd, *J. Am. Soc. Mass Spectrom.*, 2007, **18**, 1990–2000.
- L. Grimmelsmann, A. Marefat Khah, C. Spies, C. Hättig and P. Nuernberger, *J. Phys. Chem. Lett.*, 2017, **8**, 1986–1992.
- L. Grimmelsmann, V. Schuabb, B. Tekin, R. Winter and P. Nuernberger, *Phys. Chem. Chem. Phys.*, 2018, **20**, 18169–18175.
- S. Schott, A. Steinbacher, J. Buback, P. Nuernberger and T. Brixner, *J. Phys. B: At., Mol. Opt. Phys.*, 2014, **47**, 124014.
- U. Megerle, I. Pugliesi, C. Schriever, C. F. Sailer and E. Riedle, *Appl. Phys. B: Lasers Opt.*, 2009, **96**, 215–231.
- S. Schott, L. Röss, J. Hrusak, P. Nuernberger and T. Brixner, *Phys. Chem. Chem. Phys.*, 2016, **18**, 33287–33302.
- J. J. Snellenburg, S. P. Laptanok, R. Seger, K. M. Mullen and I. H. M. van Stokkum, *J. Stat. Softw.*, 2012, **49**, 1–22.
- TURBOMOLE V7.1 2016 and V7.2.1 2017, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989–2007, TURBOMOLE GmbH, since 2007; available from <http://www.turbomole.com>.
- A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648–5652.

- 35 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, 2010, **132**, 154104.
- 36 S. Grimme, S. Ehrlich and L. Georigk, *J. Comput. Chem.*, 2011, **32**, 1456.
- 37 F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.*, 2005, **7**, 3297.
- 38 F. Weigend, M. Häser, H. Patzelt and R. Ahlrichs, *Chem. Phys. Lett.*, 1998, **294**, 143.
- 39 A. Klamt and G. Schüürmann, *J. Chem. Soc., Perkin Trans. 2*, 1993, 386.
- 40 B. Lunkenheimer and A. Köhn, *J. Chem. Theory Comput.*, 2013, **9**, 977.
- 41 J. Schirmer, *Phys. Rev. A: At., Mol., Opt. Phys.*, 1981, **26**, 2395–2416.
- 42 C. Hättig, *Adv. Quantum Chem.*, 2005, **50**, 37–60.
- 43 C. Hättig and F. Weigend, *J. Chem. Phys.*, 2000, **113**, 5154–5161.
- 44 N. O. C. Winter, N. K. Graf, S. Leutwyler and C. Hättig, *Phys. Chem. Chem. Phys.*, 2013, **15**, 6623–6630.
- 45 K. Fukui, *Acc. Chem. Res.*, 1981, **14**, 363.
- 46 P. Plessow, *J. Chem. Theory Comput.*, 2013, **9**, 1305–1310.
- 47 M. Pabst and A. Köhn, *J. Chem. Phys.*, 2008, **129**, 214101.
- 48 A. F. Thünemann, D. Schütt, R. Sachse, H. Schlaad and H. Möhwald, *Langmuir*, 2006, **22**, 2323.
- 49 R. Mundt, C. T. Ziegenbein, S. Fröbel, O. Weingart and P. Gilch, *J. Phys. Chem. B*, 2016, **120**, 9376–9386.
- 50 A. J. Reuss, C. Grünwald, M. Braun, J. W. Engels and J. Wachtveitl, *ChemPhysChem*, 2016, **17**, 1369–1376.
- 51 H. Marciniak and S. Lochbrunner, *Chem. Phys. Lett.*, 2014, **609**, 184–188.
- 52 C. Consani, M. Berberich, F. Würthner and T. Brixner, *Chem. Phys. Lett.*, 2017, **683**, 83–90.
- 53 T. Bolze, J.-L. Wree, F. Kanal, D. Schleier and P. Nuernberger, *ChemPhysChem*, 2015, **19**, 138–147.
- 54 S. K. Pal, J. Peon, B. Bagchi and A. H. Zewail, *J. Phys. Chem. B*, 2002, **106**, 12376–12395.
- 55 P. van der Meulen, H. Zhang, A. M. Jonkman and M. Glasbeek, *J. Phys. Chem.*, 1996, **100**, 5367–5373.
- 56 R. Adhikary, P. Mukherjee, T. W. Kee and J. W. Petrich, *J. Phys. Chem. B*, 2009, **113**, 5255–5261.
- 57 I. Delfino, C. Manzoni, K. Sato, C. Dennison, G. Cerullo and S. Cannistraro, *J. Phys. Chem.*, 2006, **110**, 17252–17259.
- 58 J.-M. L. Pecourt, J. Peon and B. Kohler, *J. Am. Chem. Soc.*, 2001, **123**, 10370–10378.
- 59 J. Grilj, P. Buchgraber and E. Vauthey, *J. Phys. Chem. A*, 2012, **116**, 7516–7522.
- 60 J. Grilj, C. Zonca, a. M. L. Daku and E. Vauthey, *Phys. Chem. Chem. Phys.*, 2012, **14**, 6352–6358.
- 61 F. Renth, M. Foca, A. Petter and F. Temps, *Chem. Phys. Lett.*, 2006, **428**, 62–67.
- 62 D. M. Carey and G. M. Korenowski, *J. Chem. Phys.*, 1998, **108**, 2669–2675.
- 63 W. Sawodny, K. Niedenzu and J. W. Dawson, *Spectrochim. Acta, Part A*, 1967, **28**, 799–806.
- 64 E. Pines, D. Huppert and N. Agmon, *J. Chem. Phys.*, 1988, **88**, 5620–5630.
- 65 S. A. Kovalenko, R. Schanz, H. Hennig and N. P. Ernsting, *J. Chem. Phys.*, 2001, **115**, 3256–3273.
- 66 M. T. Pereira, J. M. Silva-Alves, A. Martins-Jose, J. C. D. Lopes and M. M. Santoro, *Braz. J. Med. Biol. Res.*, 2005, **38**, 1593–1601.
- 67 N. P. Ernsting, S. A. Kovalenko, T. Senyushkina, J. Saam and V. Farztdinov, *J. Phys. Chem. A*, 2001, **105**, 3443–3453.
- 68 R. Ghosh and D. K. Palit, *Photochem. Photobiol. Sci.*, 2013, **12**, 987–995.
- 69 P. K. Verma, F. Koch, A. Steinbacher, P. Nuernberger and T. Brixner, *J. Am. Chem. Soc.*, 2014, **136**, 14981–14989.
- 70 T. N. V. Karsili, B. Marchetti, M. N. R. Ashfold and W. Domcke, *J. Phys. Chem. A*, 2014, **118**, 11999–12010.
- 71 O. F. Mohammed, D. Xiao, V. S. Batista and E. T. J. Nibbering, *J. Phys. Chem. A*, 2014, **118**, 3090–3099.
- 72 P. K. Verma, A. Steinbacher, F. Koch, P. Nuernberger and T. Brixner, *Phys. Chem. Chem. Phys.*, 2015, **17**, 8459–8466.
- 73 P. K. Verma, A. Steinbacher, A. Schmiedel, P. Nuernberger and T. Brixner, *Struct. Dyn.*, 2016, **3**, 023606.
- 74 J. J. Szymczak, M. Barbatti and H. Lischka, *J. Phys. Chem. A*, 2009, **113**, 11907–11918.
- 75 P. Plessow, *J. Chem. Theory Comput.*, 2013, **9**, 1305–1310.

7 Analytic Excited State Gradients for the QM/MM Polarizable Embedded Second-Order Algebraic Diagrammatic Construction for the Polarization Propagator PE-ADC(2)

Reprinted with permission from *J. Chem Theory Comput.* 2018, 14, 4640-4650.

Copyright 2019 American Chemical Society.

In this paper, an implementation of polarizable embedded ADC(2) excited state gradients is reported. It is demonstrated that the quality of PE-MP2 ground state and PE-ADC(2) excited-state geometries are comparable with the supermolecular calculations. A saviour electron spill out in the non-covalent host-guest complex of DBO@CB[7] has been reported and supplementing the MM system with ECPs is proposed as a solution to the latter issue.^{5,6}

My contribution to this project comprises the major portion of theory development and implementation under the supervision of Christof Hättig and in case of applications, I have performed all the calculations and compiled the numbers, tables and figures. Since the polarizable embedding (PE) theory and COSMO share some subroutines in TURBOMOLE program, Sarah Karbalaei Khani contributed partially to the implementation where the both approaches have a common structure. The first draft of the manuscript was written

⁵Article: <https://pubs.acs.org/doi/pdf/10.1021/acs.jctc.8b00396>

⁶SI: https://pubs.acs.org/doi/suppl/10.1021/acs.jctc.8b00396/suppl_file/ct8b00396_si_001.pdf

by me and further developed together with my co-authors.

 **Copyright Clearance Center**  **Home** **Create Account** **Help** 

 **ACS Publications** Most Trusted. Most Cited. Most Read. **Title:** Analytic Excited State Gradients for the QM/MM Polarizable Embedded Second-Order Algebraic Diagrammatic Construction for the Polarization Propagator PE-ADC(2)

Author: Alireza Marefat Khah, Sarah Karbalaeei Khani, Christof Hättig

Publication: Journal of Chemical Theory and Computation

Publisher: American Chemical Society

Date: Sep 1, 2018

Copyright © 2018, American Chemical Society

LOGIN

If you're a copyright.com user, you can login to RightsLink using your copyright.com credentials. Already a **RightsLink user** or want to [learn more?](#)

PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

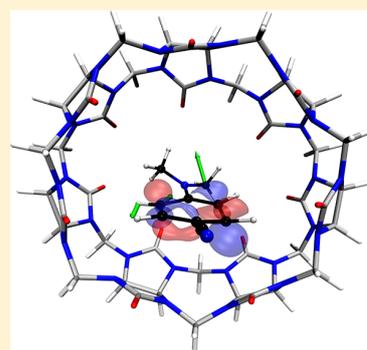
Analytic Excited State Gradients for the QM/MM Polarizable Embedded Second-Order Algebraic Diagrammatic Construction for the Polarization Propagator PE-ADC(2)

Alireza Marefat Khah,^{*†} Sarah Karbalaei Khani, and Christof Hättig[†]

Arbeitsgruppe Quantenchemie, Ruhr-Universität Bochum, Bochum D-44801, Germany

 Supporting Information

ABSTRACT: An implementation of a QM/MM embedding in a polarizable environment is presented for second-order Møller–Plesset perturbation theory, MP2, for ground state energies and molecular gradients and for the second-order Algebraic Diagrammatic Construction, ADC(2), for excitation energies and excited state molecular gradients. In this implementation of PE-MP2 and PE-ADC(2), the polarizable embedded Hartree–Fock wave function is used as uncorrelated reference state. The polarization-correlation cross terms for the ground and excited states are included in this model via an approximate coupling density. A Lagrangian formulation is used to derive the relaxed electron densities and molecular gradients. The resolution-of-the-identity approximation speeds up the calculation of four-index electron repulsion integrals in the molecular orbital basis. As a first application, the method is used to study the photophysical properties of host–guest complexes where the accuracy and weaknesses of the model are also critically examined. It is demonstrated that the ground state geometries of the full quantum mechanical calculation for the supermolecule can be well reproduced. For excited state geometries, the deviations from the supermolecular calculation are slightly larger, but still the environment effects are captured qualitatively correctly, and energy gaps between the ground and excited states are obtained with sufficient accuracy.



INTRODUCTION

Most photochemistry happens in solution, on surfaces, or in biomolecular matrices where the inhomogeneous structure of the molecular environment can disturb the electronic and molecular structure of the chromophore and thereby have non-negligible effects on various molecular properties.¹ This leads to a demand for multiscale methods that can be applied to systems with many atoms but only treat an active site, which has much fewer atoms, with high accuracy, so that the computational costs are still manageable. Often specific interactions between the system of interest and the environment are significant and require that the atomistic structure of the environment is explicitly taken into account. For these systems, an efficient way to include the effects of a complex molecular environment in a quantum chemical calculation is to use QM/MM methods,^{2,3} which describe the environment at an atomistic level using classical mechanics and electrostatics.

In the QM/MM approach, the energy of the total system can be divided into the energies of the QM and the MM subsystems and an interaction term, $E^{\text{QM/MM}} = E^{\text{QM}} + E^{\text{MM}} + E^{\text{QM-MM}}$. The interaction between the QM and MM subsystems is often approximated by mechanical contributions (which do not affect the electrons) plus an electrostatic interaction between the electronic density of the QM subsystem with a set of point charges and also static multipole moments (dipoles, quadrupoles, etc.). In such an electrostatic multipole embedding, the MM subsystem shows no direct response to the changes in the electronic density of the QM

subsystem that happen at the time scale of electronic excitations. It can only respond indirectly through changes in the structure, which happen at much slower time scales. Therefore, a polarizable embedding is important for a realistic description of environment effects on electronic excitation processes.⁴

The QM/MM approach has been used for a long time with several QM methods for ground state energies and structures where (nonpolarizable) electrostatic multipole interactions are often sufficient. But when it comes to electronic excitations and excited state structures for which the inclusion of the polarizability of the environments is essential, the QM/MM approach has only rarely been used with wave-function-based methods that go beyond the time-dependent density functional theory (TDDFT).

In recent years, the QM/MM polarizable embedding method has been developed to describe different excited state properties^{5,6} in combination with TDDFT⁷ and coupled cluster wave functions^{8–10} as well as multireference methods.^{11,12} Implementations for analytic molecular gradients for polarizable MM embeddings have been reported for ground state structures at the DFT^{13,14} and MP2¹⁵ levels and for excited states at the TDDFT level.^{16,17}

The correct description of charge-transfer states with TDDFT is already problematic in the gas phase. If the

Received: April 25, 2018

Published: July 24, 2018

quantum system is placed in a polarizable environment even functionals that were designed for electronic excitations like CAM-B3LYP¹⁸ can fail for systems with charge transfer states.¹⁹ Since wave-function-based methods do not suffer from this problem, the second-order methods CC2²⁰ and ADC(2)²¹ are often the methods of choice for medium-sized quantum systems where they have only moderately higher computational cost.^{22–24} For the calculation of molecular gradients for excited-state potential energy surfaces, ADC(2) is of particular interest as for this method the computational costs for the gradients are up to a factor of 4 lower than those for CC2, while the accuracy of the resulting structures is almost the same for both methods.²⁵ Furthermore, ADC(2) has a Hermitian secular matrix, which provides, in contrast to CC2, a physically correct description of the potential energy surface in the vicinity of conical intersection.^{22,26}

In the current contribution, we report for PE-ADC(2) the first implementation of analytic excited state gradients for a correlated wave-function-based method in an atomistic polarizable embedding. Applications of the PE-ADC(2) gradients are presented for host–guest systems where supermolecular reference calculations are still possible so that they can be used to evaluate and discuss the strengths and also weaknesses of the approach.

THEORY

Polarizable Embedded Hartree–Fock. For the sake of brevity, we will only recapitulate the essential equations for the theory of the polarizable embedding.^{5,7} For a more detailed discussion of the physical background of polarizable embeddings and the definition and implementation of linear response theory in a polarizable embedding, we refer the reader to a recent review by List et al.²⁷ We start with the free energy expression of the polarizable embedded Hartree–Fock (PE-HF) model:

$$\begin{aligned} \mathcal{G}^{\text{PE-HF}} = & \langle \text{HF} | \hat{H}^{\text{QM}} + \hat{G}^{\text{es}} | \text{HF} \rangle \\ & - \frac{1}{2} \sum_{uv} (F_u^{\text{ne}} + F_u^{\text{elec}}(\mathbf{D}^{\text{HF}})) \\ & \times R_{uv} (F_v^{\text{ne}} + F_v^{\text{elec}}(\mathbf{D}^{\text{HF}})) + E^{\text{ne}} \end{aligned} \quad (1)$$

In this equation, \hat{H}^{QM} is the vacuum-like Hamiltonian for the QM system. The interaction of the QM electrons with the static multipole moments of the MM system are taken into account through the one-electron operator:

$$\hat{G}^{\text{es}} = \sum_u Q_u^{\text{es}} \sum_{pq} \epsilon_{pq}^u \hat{E}_{pq} \quad (2)$$

where Q_u^{es} and ϵ_{pq}^u are, respectively, the static (permanent) multipole moments and the multipole interaction integrals. The index u runs here over all multipole components at all MM sites. Finally, \hat{E}_{pq} is the second quantization singlet replacement operator.²⁸ Here and in the following, the indices p, q, r , and s are used for general molecular orbitals. R_{uv} is the response matrix with the indices u and v running in general over all components of induced multipole moments at polarizable MM sites. In the current work, we only include distributed dipole–dipole polarizabilities. F_v^{ne} are then the components of the electric fields at the polarizable MM sites due to the nuclei in the QM subsystem and the static multipoles of all other MM sites, and $F_v^{\text{elec}}(\mathbf{D})$ are the contributions due to electron density \mathbf{D} . For further details,

we refer to ref 7. The last term in eq 1 contains the nonelectronic contributions to the free energy:

$$E^{\text{ne}} = E^{\text{nuc-es}} + E^{\text{MM}} + E_{6-12}^{\text{LJ}} \quad (3)$$

where $E^{\text{nuc-es}}$ is the electrostatic interaction of the QM nuclei with the MM sites, and E^{MM} is the internal energy of the MM system. To account for the London dispersion interaction and the short-range repulsion, we include a Lennard–Jones 6-12 term (E_{6-12}^{LJ}) between the atoms in the QM and the atoms in the MM subsystem,

$$E_{6-12}^{\text{LJ}} = \sum_{s=1}^S \sum_{m \in \text{QM}} \epsilon^{sm} \left[\left(\frac{r_{sm}}{|\mathbf{R}_m - \mathbf{R}_s|} \right)^{12} - 2 \left(\frac{r_{sm}}{|\mathbf{R}_m - \mathbf{R}_s|} \right)^6 \right] \quad (4)$$

The pair interaction parameters r_{sm} and ϵ^{sm} are derived from atomic parameters using the Lorentz–Berthelot combination rules $r_{sm} = \frac{1}{2}(r_s + r_m)$ and $\epsilon^{sm} = \sqrt{\epsilon^s \epsilon^m}$ where r_s and ϵ^s are, respectively, atomic radii and well depths.

The minimization of the free energy for the polarizable embedded Hartree–Fock model, $\mathcal{G}^{\text{PE-HF}}$, with respect to variations of the orbitals leads to the condition

$$\langle \text{HF} | [\hat{H}^{\text{QM}} + \hat{G}^{\text{es}} + \hat{G}^{\text{pol,HF}}, \hat{E}_{pq}^-] | \text{HF} \rangle = 0 \quad (5)$$

where we introduced the antisymmetrized operators $\hat{E}_{pq}^- = \hat{E}_{pq} - \hat{E}_{qp}$. The polarization contribution to the reaction field potential is given by

$$\begin{aligned} \hat{G}^{\text{pol,HF}} = & - \sum_{uv} (F_v^{\text{ne}} + F_v^{\text{elec}}(\mathbf{D}^{\text{HF}})) R_{vu} \sum_{pq} \epsilon_{pq}^u \hat{E}_{pq} \\ = & \sum_u Q_u^{\text{ind,HF}} \sum_{pq} \epsilon_{pq}^u \hat{E}_{pq} \end{aligned} \quad (6)$$

Since within the PE model the interaction with the environment only affects the electrostatic potential experienced by the electrons and the nonelectronic contributions, its formulation and implementation for Kohn–Sham density functional theory (DFT) are analogous to the above equations. Working expression for analytic nuclear gradients at the PE-DFT and PE-HF level have been reported by List et al. in ref 13 and for conciseness will not be repeated here. As a part of the current work, we have implemented the gradients for PE-HF and PE-DFT in the HF and DFT gradient programs `grad` and `rdgrad` of a development version of the TURBOMOLE package.²⁹ The self-consistently optimized PE-HF determinant is used as reference wave function for the correlation treatment at the PE-MP2 and the calculation of excitation energies at the PE-ADC(2) level in accordance with the recent implementation of PERI-CC2 for excitation energies and one- and two-photon transition moments.^{9,30}

Polarizable Embedded Second-Order Møller–Plesset Perturbation Theory. The reaction field scheme that has been used in previous work on PERI-CC2 reduces at the level of MP2, that is, when the CC2 cluster amplitudes are replaced by those from first-order perturbation theory, to a scheme that in the literature is known³¹ as perturbation on the energy (PTE). The PTE scheme is also the reaction field scheme that is most consistent with second-order Møller–Plesset perturbation theory.³² Within this scheme, the PE-MP2 correlation energy is calculated fully analogously to the vacuum case just using the molecular orbitals and the Fock operator

$$\hat{F}^{\text{PE}} = \hat{F}^{\text{QM}} + \hat{G}^{\text{HF}} \quad (7)$$

from a PE-HF calculation where $\hat{G}^{\text{HF}} = \hat{G}^{\text{es}} + \hat{G}^{\text{pol,HF}}$. The working expressions for the PE-MP2 molecular gradients can be obtained by differentiating the orbital relaxed Lagrangian:

$$\begin{aligned} L^{\text{PE-MP2}} = & \langle \text{HF} | \hat{H}^{\text{QM}} + \hat{G}^{\text{es}} + [\hat{W}^{\text{QM}}, \hat{T}_2^{(1)}] | \text{HF} \rangle \\ & + \langle \bar{t}_2^{(1)} | [\hat{F}^{\text{PE}}, \hat{T}_2^{(1)}] + \hat{W}^{\text{QM}} | \text{HF} \rangle \\ & + \langle \text{HF} | [\hat{H}^{\text{QM}} + \hat{G}^{\text{HF}}, \hat{\kappa}] | \text{HF} \rangle \\ & - \frac{1}{2} \sum_{uv} (F_u^{\text{ne}} + F_u^{\text{elec}}(\mathbf{D}^{\text{HF}})) \\ & \times R_{uv} (F_v^{\text{ne}} + F_v^{\text{elec}}(\mathbf{D}^{\text{HF}})) + E^{\text{ne}} \end{aligned} \quad (8)$$

In the above equation, $\hat{T}_2^{(1)} = \sum_{\mu_2} t_{\mu_2}^{(1)} \tau_{\mu_2}$ is the double excitation operator built from the first-order amplitudes, $t_{\mu_2}^{(1)}$, with μ_2 running over all double replacements and $\langle \bar{t}_2^{(1)} | = \sum_{\mu_2} \bar{t}_{\mu_2}^{(1)} \langle \mu |$ is the first-order wave function correction built from the first-order Lagrange multipliers. \hat{W}^{QM} denotes the fluctuation potential operator and $\hat{\kappa}$ is an operator that contains the orbital rotation Lagrange multipliers:

$$\hat{\kappa} = \sum_{pq} \bar{\kappa}_{pq} \hat{E}_{pq}^- \quad (9)$$

The values of the orbital-rotation Lagrange multipliers $\bar{\kappa}_{pq}$ are determined by the CPHF (also known as Z-vector) equations, which are obtained by requiring $L^{\text{PE-MP2}}$ to be stationary with respect to orbital variations:

$$\sum_{rs} \langle \langle \text{HF} | [\hat{F}^{\text{PE}}, \hat{E}_{pq}^-, \hat{E}_{rs}^-] | \text{HF} \rangle + A_{pq,rs}^{\kappa, \text{MP2}} \bar{\kappa}_{rs} = -\eta_{pq}^{\kappa, \text{MP2}} \quad (10)$$

The right-hand-side vector η_{pq}^{κ} is given by

$$\begin{aligned} \eta_{pq}^{\kappa, \text{MP2}} = & \langle \text{HF} | [\hat{F}^{\text{PE}} + \hat{W}^{\text{QM}}, \hat{E}_{pq}^-, [\hat{W}^{\text{QM}}, \hat{E}_{pq}^-], \hat{T}_2^{(1)}] | \text{HF} \rangle \\ & + \langle \bar{t}_2^{(1)} | [\hat{F}^{\text{PE}}, \hat{E}_{pq}^-, \hat{T}_2^{(1)}] + [\hat{W}^{\text{QM}}, \hat{E}_{pq}^-] | \text{HF} \rangle \\ & + \langle \bar{t}_2^{(1)} | \left[\frac{d\hat{F}^{\text{PE}}}{d\kappa_{pq}} - [\hat{F}^{\text{PE}}, \hat{E}_{pq}^-] \right], \hat{T}_2^{(1)} | \text{HF} \rangle \end{aligned} \quad (11)$$

and, thus, includes environment contributions only via \hat{F}^{PE} and the orbitals. The CPHF matrix $A_{pq,rs}^{\kappa}$ for the PE model is defined as

$$\begin{aligned} A_{pq,rs}^{\kappa} = & \langle \text{HF} | [[\hat{W}^{\text{QM}}, \hat{E}_{pq}^-], \hat{E}_{rs}^-] | \text{HF} \rangle \\ & + \sum_{uv} \langle \text{HF} | [\hat{E}_{pq}^-, \hat{E}_{rs}^-] | \text{HF} \rangle R_{uv} \langle \text{HF} | [\hat{E}_{rs}^-, \hat{E}_{rs}^-] | \text{HF} \rangle \end{aligned} \quad (12)$$

After solving the CPHF eqs 10, the effective one-electron density for MP2 can be calculated as in the vacuum case:^{33,34}

$$D_{pq}^{\text{eff}} = D_{pq}^{\text{HF}} + \langle \bar{t}_2^{(1)} | [\hat{E}_{pq}^-, \hat{T}_2^{(1)}] | \text{HF} \rangle + \frac{1}{2} (\bar{\kappa}_{pq} + \bar{\kappa}_{qp}) \quad (13)$$

For the derivation of the expressions for the PE-MP2 gradients, we split the PE-MP2 Lagrangian into $L^{\text{PE-MP2}} = L^{\text{MP2,QM}} + \Delta L^{\text{PE}}$ where $L^{\text{MP2,QM}}$ contains the vacuum-like contributions and ΔL^{PE} the explicit PE terms, which are not present in the vacuum case:

$$\begin{aligned} \Delta L^{\text{PE}} = & \langle \text{HF} | \hat{G}^{\text{es}} | \text{HF} \rangle + \langle \bar{t}_2^{(1)} | [\hat{G}^{\text{HF}}, \hat{T}_2^{(1)}] | \text{HF} \rangle \\ & + \langle \text{HF} | [\hat{G}^{\text{HF}}, \hat{\kappa}] | \text{HF} \rangle \\ & - \frac{1}{2} \sum_{uv} (F_u^{\text{ne}} + F_u^{\text{elec}}(\mathbf{D}^{\text{HF}})) \\ & \times R_{uv} (F_v^{\text{ne}} + F_v^{\text{elec}}(\mathbf{D}^{\text{HF}})) + E^{\text{ne}} \\ = & -\frac{1}{2} \sum_{uv} (F_u^{\text{ne}} + 2F_u^{\text{elec}}(\mathbf{D}^{\text{eff}}) - F_u^{\text{elec}}(\mathbf{D}^{\text{HF}})) \\ & \times R_{uv} (F_v^{\text{ne}} + F_v^{\text{elec}}(\mathbf{D}^{\text{HF}})) \\ & + \sum_{pq} D_{pq}^{\text{eff}} \sum_u Q_u^{\text{es},u} \epsilon_{pq}^u + E^{\text{ne}} \end{aligned} \quad (14)$$

The explicit embedding contributions to the gradients for the QM atoms are computed as

$$\begin{aligned} \frac{d\Delta L^{\text{PE}}}{dx} = & \sum_{pq} D_{pq}^{\text{eff}} \sum_u Q_u^{\text{HF}} \frac{d\epsilon_{pq}^u}{dx} + \sum_{pq} D_{pq}^{\text{HF}} \\ & \times \sum_u (Q_u^{\text{eff}} - Q_u^{\text{HF}}) \frac{d\epsilon_{pq}^u}{dx} + \sum_u Q_u^{\text{eff}} \frac{d\epsilon_{\text{nuc}}^u}{dx} \\ & + \frac{dE^{\text{LJ}}}{dx} \end{aligned} \quad (16)$$

The multipole moments Q_u^{eff} are the sum of the static and the induced moments computed from the MP2 density \mathbf{D}^{eff} and Q_u^{HF} are the respective Hartree–Fock counterparts. ϵ_{nuc}^u is the sum of the multipole-charge interaction tensors for the interaction with the nuclei, contracted with the nuclear charges.

The PE-MP2 model has been implemented for energies and molecular gradients in the `ricc2` program of a development version of the TURBOMOLE package.²⁹ The MPI-parallel implementation of the MP2 energy and gradients for the vacuum case have been described in ref 34. The contributions from $L^{\text{MP2,QM}}$ to the molecular gradients are evaluated as described in the latter reference. We stress that our implementation of PE-MP2 uses, as the vacuum implementation, the resolution-of-the-identity (RI) approximation for the electron repulsion integrals and is parallelized with OpenMP and MPI, which makes it applicable to relatively large molecular systems.

PE-ADC(2) for Excited States Energy and Gradients.

Although the underlying physical background of ADC(2) and CC2 are different, the kernel of the following discussion lies in the mathematical similarity between these methods, which facilitates easy development and implementation of the equations for PE-ADC(2) excitation energies and gradients as a simplification of PE-CC2.²² In the framework of response theory,³⁵ excitation energies are obtained for nonvariational methods like CC2 as eigenvalues of the Jacobi matrix \mathbf{A} :

$$(\mathbf{A} - \omega^f \mathbf{1}) \mathbf{E}^f = 0 \quad (17)$$

The CC2 Jacobian includes terms that only contribute in third or higher order to the excitation energies. These terms can be excluded by replacing in the CC2 Jacobian, $\mathbf{A}^{\text{CC2}}(T_1, T_2)$, the ground state cluster amplitudes by those from first-order Møller–Plesset perturbation theory (MP1). This substitution leads to the CIS(D_∞) model³⁶ for excitation energies. If the Jacobian is in addition symmetrized, one arrives at the ADC(2) 74 secular matrix:²²

$$\mathbf{A}^{\text{ADC}(2)} = \frac{1}{2}(\mathbf{A}^{\text{CC2}}(0, T_2^{\text{(MP1)}}) + \mathbf{A}^{\text{CC2}}(0, T_2^{\text{(MP1)}})^\dagger) \quad (18)$$

The PE-ADC(2) secular matrix can be obtained in the same way from the PERI-CC2 Jacobian,⁹

$$\mathbf{A}_{\mu\nu}^{\text{PE-CC2}} = \begin{pmatrix} \langle \mu_1 | [\hat{F}^{\text{PE}} + \hat{W}^{\text{QM}} + \hat{G}^\Delta + [\hat{W}^{\text{QM}}, \hat{T}_2], \tau_{\nu_1}] + \hat{G}^{\nu_1} | \text{HF} \rangle & \langle \mu_1 | [\hat{W}^{\text{QM}}, \tau_{\nu_2}] | \text{HF} \rangle \\ \langle \mu_2 | [\hat{W}^{\text{QM}}, \tau_{\nu_1}] | \text{HF} \rangle & \langle \mu_2 | [\hat{F}^{\text{PE}}, \tau_{\nu_2}] | \text{HF} \rangle \end{pmatrix} \quad (19)$$

In the last equation, we used T_1 -transformed operators defined as $\hat{O} = \exp(-\hat{T}_1)\hat{O}\exp(\hat{T}_1)$ and the derivative of the correlation contribution to the reaction field operator proposed in ref 9 as

$$\hat{G}^{\nu_1} = \frac{\partial \hat{G}^\Delta}{\partial t_{\nu_1}} = \sum_{uv} \sum_{pq} \varepsilon_{pq}^u \langle \text{HF} + \bar{t}_1 | [\hat{E}_{pq}, \tau_{\nu_1}] | \text{HF} \rangle R_{uv} \times \sum_{pq} \varepsilon_{rs}^v \hat{E}_{rs} \quad (20)$$

For PE-CC2, the operator \hat{G}^Δ is computed only from the density contribution that originates from the singles cluster amplitudes and singles (ground state) Lagrange multipliers:⁹

$$\hat{G}^\Delta = -\sum_{uv} F_v^{\text{elec}}(\mathbf{D}^\Delta) R_{vu} \sum_{pq} \varepsilon_{pq}^u \hat{E}_{pq} = \sum_u \mu_u^\Delta \sum_{pq} \varepsilon_{pq}^u \hat{E}_{pq} \quad (21)$$

with

$$D_{pq}^\Delta = \langle \text{HF} + \bar{t}_1 | e^{-\hat{T}_1} \hat{E}_{pq} e^{\hat{T}_1} | \text{HF} \rangle - \langle \text{HF} | \hat{E}_{pq} | \text{HF} \rangle \quad (22)$$

The secular matrix for PE-ADC(2) can be obtained from the PE-CC2 Jacobian in the same way as the vacuum ADC(2) secular matrix from the vacuum CC2 Jacobian. For an easier discussion, we split the PE-ADC(2) secular matrix into a vacuum-like and an embedding contribution, $\mathbf{A}^{\text{PE-ADC}(2)} = \mathbf{A}^{\text{ADC}(2),\text{QM}} + \mathbf{A}^{\text{ADC}(2),\text{PE}}$. Regarding the above approximation for the reaction field, the polarizable embedding contributes only to the singles–singles and the doubles–doubles blocks:

$$A_{\mu_1, \nu_1}^{\text{ADC}(2),\text{PE}} = \langle \mu_1 | [\hat{G}^{\text{HF}}, \tau_{\nu_1}] + \hat{G}^{\nu_1} | \text{HF} \rangle \quad (23)$$

$$A_{\mu_2, \nu_2}^{\text{ADC}(2),\text{PE}} = \langle \mu_2 | [\hat{G}^{\text{HF}}, \tau_{\nu_2}] | \text{HF} \rangle \quad (24)$$

Note that for ADC(2) the density contribution \mathbf{D}^Δ vanishes since the ground state singles amplitudes and Lagrange multipliers vanish. \hat{G}^{HF} is included through the Fock matrix \hat{F}^{PE} . Its contributions are already present in the PTE reaction field scheme. The excitonic coupling term³⁷ that arises from \hat{G}^{ν_1} originates from the derivative of \hat{G}^Δ with respect to t_{ν_1} , compare eq 20, which are nonvanishing. Its contribution to the linear transformations with the secular matrix during the iterative solution of the eigenvalue equations is implemented via two auxiliary densities:

$$\mathbf{D}^\eta = \sum_{\nu_1} \frac{\partial \mathbf{D}^\Delta}{\partial t_{\nu_1}} E_{\nu_1}^f \quad \text{and} \quad \mathbf{D}^\xi = \sum_{\mu_1} E_{\mu_1}^f \frac{\partial \mathbf{D}^\Delta}{\partial \bar{t}_{\mu_1}} \quad (25)$$

With these densities, the contributions from \hat{G}^{ν_1} to the contractions of the CC2 Jacobian or the ADC(2) secular

matrix with a vector from the right or the left can be evaluated as

$$\sum_{\nu_1} \langle \mu_1 | \hat{G}^{\nu_1} | \text{HF} \rangle E_{\nu_1}^f = \langle \mu_1 | \hat{G}(\mathbf{D}^\eta) | \text{HF} \rangle \quad (26)$$

$$\sum_{\mu_1} E_{\mu_1}^f \langle \mu_1 | \hat{G}^{\nu_1} | \text{HF} \rangle = \langle \text{HF} | [\hat{G}(\mathbf{D}^\xi), \tau_{\mu_1}] | \text{HF} \rangle \quad (27)$$

Considering that in ADC(2) theory the correlation effects are included up to second-order in the fluctuation potential, using the first-order Møller–Plesset ground state wave function parameters, the canonical choice²² for defining excited state total energies is to augment the excitation energies with the ground state energy from second-order Møller–Plesset perturbation theory. We use the corresponding approach for PE-ADC(2) and define the total energies of excited states as the sum of the PE-MP2 ground state and the PE-ADC(2) excitation energies. This leads to the following Lagrange function for the total energy of an excited state f :

$$\begin{aligned} L^{\text{PE-ADC}(2),f} &= \langle \text{HF} | \hat{F}^{\text{QM}} + \hat{G}^{\text{es}} + \hat{W}^{\text{QM}} + [\hat{W}^{\text{QM}}, \hat{T}_2] | \text{HF} \rangle \\ &+ \sum_{i,j=1}^2 \sum_{\mu_i, \nu_j} E_{\mu_i}^f A_{\mu_i, \nu_j}^{\text{PE-ADC}(2)} E_{\nu_j}^f \\ &+ \bar{\omega}^f \left(1 - \sum_{i=1}^2 \sum_{\mu_i} E_{\mu_i}^f E_{\mu_i}^f \right) \\ &+ \sum_{\mu_2} \bar{t}_{\mu_2}^f \langle \mu_2 | [\hat{F}^{\text{PE}}, \hat{T}_2] + \hat{W}^{\text{QM}} | \text{HF} \rangle \\ &+ \langle \text{HF} | [\hat{H}^{\text{QM}} + \hat{G}^{\text{HF}}, \hat{\kappa}] | \text{HF} \rangle \\ &- \frac{1}{2} \sum_{uv} (F_u^{\text{ne}} + F_u^{\text{elec}}(\mathbf{D}^{\text{HF}})) \\ &\times R_{uv} (F_v^{\text{ne}} + F_v^{\text{elec}}(\mathbf{D}^{\text{HF}})) + E^{\text{ne}} \end{aligned} \quad (28)$$

The first term represents the vacuum-like contributions to the PE-MP2 ground state energy plus the interaction of the electron with the static moments at the Hartree–Fock level. The PE-ADC(2) excitation energies are accounted for by the second term where the right and left eigenvectors (E^f) are identical since the PE-ADC(2) secular matrix is Hermitian. The third term implements the orthonormality condition for the eigenvectors, and the fourth and fifth terms account for the MP1 amplitude and the Hartree–Fock equations, similar to that in the PE-MP2 Lagrangian. Finally, the last two terms account again for the Hartree–Fock polarization energy and the nonelectronic energy contributions.

Requiring the above Lagrangian to be stationary with respect to the variations of the doubles amplitudes leads to the equations for the Lagrange multipliers $\bar{t}_{\mu_2}^f$:

$$\begin{aligned} \frac{dL^{\text{PE-ADC}(2),f}}{dt_{\mu_2}^f} &= \langle \text{HF} | [\hat{W}^{\text{QM}}, \tau_{\mu_2}] | \text{HF} \rangle \\ &+ \langle E_1 | [\hat{W}^{\text{QM}}, E_1], \tau_{\mu_2} | \text{HF} \rangle + \langle \bar{t}_2^f | [\hat{F}^{\text{PE}}, \tau_{\mu_2}] | \text{HF} \rangle = 0 \end{aligned} \quad (29)$$

As at the MP2 level, the Lagrange multiplier equations contain only PTE-like embedding contributions, which are automatically included by using the Fock matrix and the orbitals from a 75 PE-HF calculation. With canonical MOs, that is, if F_{pq}^{PE} is

diagonal, the above equations decouple and can be directly inverted, which is important for the efficiency of the ADC(2) implementation.²² The reaction field scheme with the approximate coupling densities that we use here conserves this important feature of ADC(2).

Requiring the excited state Lagrangian to be stationary with respect to orbital variations gives the CPHF equations that determine the Lagrange multipliers $\bar{\kappa}_{pq}$:

$$\sum_{rs} (\langle \text{HF} | [\hat{F}^{\text{PE}}, \hat{E}_{pq}^-, \hat{E}_{rs}^-] | \text{HF} \rangle + A_{pq,rs}^{\kappa} \bar{\kappa}_{rs}) = -\eta_{pq}^{\kappa, \text{PE-ADC}(2)} \quad (30)$$

For easier discussion, we split the right-hand-side vector as $\eta_{pq}^{\kappa, \text{PE-ADC}(2)} = \eta_{pq}^{\kappa, \text{PTE-ADC}(2)} + \eta_{pq}^{\kappa, \text{pol}}$ into a PTE part, $\eta_{pq}^{\kappa, \text{PTE-ADC}(2)}$, which is computed as in the vacuum case but with the MOs and the Fock matrix replaced by their PE counterparts, and the polarization contributions that arise from the excitonic coupling term:

$$\eta_{pq}^{\kappa, \text{pol}} = 2 \sum_t [G_{pt}(\mathbf{D}^\eta)(\mathbf{D}_{iq}^\xi + \mathbf{D}_{qt}^\xi) - G_{qt}(\mathbf{D}^\eta) \times (\mathbf{D}_{ip}^\xi + \mathbf{D}_{pt}^\xi)] \quad (31)$$

Once the CPHF equations are solved, the PE-ADC(2) gradients can be calculated as

$$\left(\frac{dL^{\text{PE-ADC}(2),f}}{dx} \right)_{x=0} = \sum_{\alpha\beta} D_{\alpha\beta}^{\text{eff},\text{ao}} h_{\alpha\beta}^{[x]} + \frac{1}{2} \sum_{\alpha\beta\gamma\delta} d_{\alpha\beta\gamma\delta}^{\text{sep},\text{ao}} (\alpha\beta|\gamma\delta)^{[x]} - \sum_{\alpha\beta} F_{\alpha\beta}^{\text{eff},\text{ao}} S_{\alpha\beta}^{[x]} + \sum_{\alpha\beta P} \Delta_{\alpha\beta}^{\text{ao},P} (\alpha\beta|P)^{[x]} - \sum_{PQ} \gamma_{PQ} (P|Q)^{[x]} + \frac{d\Delta L^{\text{PE},f}}{dx} \quad (32)$$

In the above equation, $D_{\alpha\beta}^{\text{eff},\text{ao}}$ is the relaxed one-electron density in the atomic orbital (AO) basis where $d_{\alpha\beta\gamma\delta}^{\text{sep},\text{ao}}$ and $F_{\alpha\beta}^{\text{eff},\text{ao}}$ are, respectively, the separable part of the two-electron density and the effective Fock matrix in the AO basis. In eq 32, the fourth and fifth terms originate from the RI approximation. For detailed expressions of the auxiliary three- and two-index densities and integrals, we refer to ref 33. The explicit PE contributions have been collected in the last term, where

$$\Delta L^{\text{PE},f} = -\frac{1}{2} \sum_{uv} (F_u^{\text{ne}} + 2F_u^{\text{elec}}(\mathbf{D}^{\text{eff}}) - F_u^{\text{elec}}(\mathbf{D}^{\text{HF}})) \times R_{uv}(F_v^{\text{ne}} + F_v^{\text{elec}}(\mathbf{D}^{\text{HF}})) - \sum_{uv} F_u(\mathbf{D}^\xi) R_{uv} F_v(\mathbf{D}^\eta) + \sum_{pq} D_{pq}^{\text{eff}} \sum_u Q_u^{\text{es}} \epsilon_{pq}^u + E^{\text{ne}} \quad (33)$$

Differentiation of the PE contribution with respect to a nuclear coordinate x leads to

$$\begin{aligned} \frac{d\Delta L^{\text{PE},f}}{dx} &= \sum_{pq} D_{pq}^{\text{eff}} \sum_u Q_u^{\text{HF}} \frac{d\epsilon_{pq}^u}{dx} + \sum_{pq} D_{pq}^{\text{HF}} \\ &\times \sum_u (Q_u^{\text{eff}} - Q_u^{\text{HF}}) \frac{d\epsilon_{pq}^u}{dx} + \sum_u Q_u^{\text{eff}} \frac{d\epsilon_{\text{nuc}}^u}{dx} \\ &+ \frac{dE^{\text{LJ}}}{dx} + \sum_{pq} D_{pq}^\eta \sum_u Q_u^\xi \frac{d\epsilon_{pq}^u}{dx} \\ &+ \sum_{pq} D_{pq}^\xi \sum_u Q_u^\eta \frac{d\epsilon_{pq}^u}{dx} \end{aligned} \quad (34)$$

The last two terms arise from the excitonic coupling contribution, which goes beyond the PTE approximation. In these terms, Q_u^ξ and Q_u^η are the induced multipole moments calculated with, respectively, the densities \mathbf{D}^ξ and \mathbf{D}^η without including nonelectronic terms (compare eq 21). For ADC(2), these two terms are equivalent so that only one of them needs to be evaluated if it is multiplied by a factor of 2. The first four terms have the same form as for PE-MP2 but are now calculated with the effective density for PE-ADC(2).

The PE-ADC(2) model has been implemented for excitation energies and excited state orbital-relaxed first-order properties and molecular gradients in the development version of the `ricc2` program of the TURBOMOLE package.^{29,38}

FIRST APPLICATIONS TO SUPRAMOLECULAR CHEMISTRY

To demonstrate the performance of the PE-ADC(2) model for excited state potential energy surfaces, we have selected two test cases from supramolecular chemistry where it was still possible to do full supermolecular reference calculations. The two test cases are the host-guest complexes of 4-(dimethylamino)benzonitrile, DMABN, and 2,3-diazabicyclo[2.2.2]oct-2-ene, DBO (see Figure 1), inside cucurbit[7]uril,

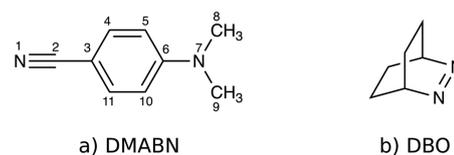


Figure 1. Lewis structures of the guest molecules.

CB[7], as the host molecule, which have interesting photo-physical properties.^{39–41} The DMABN molecule is a prototype case for a solvent-dependent dual fluorescence that results from emissions from two different electronically excited states: one with a locally excited (LE) character and the other one with an intramolecular charge transfer (ICT) character. The simplicity of its molecular structure as a typical push-pull chromophore made it a subject for many theoretical and experimental studies over the last decades.^{42–46} Recently, the UV-vis and fluorescence emission spectra of a DMABN-CB[7] complex in aqueous solution have been studied. They show one absorption and two emission peaks.⁴¹ In contrast to DMABN, DBO has a long but also solvent dependent fluorescence lifetime (up to 1 μs).³⁹ DBO has been used as a fluorescent probe to measure the polarizability of the inner cavity of CB[7], which is different from solvents.⁴⁷

In addition to the above-mentioned photophysical properties, these host-guest systems have some other helpful

Table 1. Representative Intramolecular Geometrical Parameters of DMABN in the Ground and the $S_1(\text{LE})$ Excited States from PE-MP2 and PE-ADC(2) in Comparison with the Results from Supermolecular (SM) and Vacuum (vac) Calculations^a

Internal coord.	ground state			locally excited state			
	SM	PE	vacuum	SM	PE	vacuum	
		Selected Torsion Angles (deg)					
$C_8-N_7-C_6-C_{10}$	20.5	20.0	16.2	28.2	15.4	20.3	
		Selected Bond Lengths (pm)					
C_3-C_4	141.3	140.8	141.0	140.4	141.0	142.2	
C_3-C_{11}	141.2	140.7	141.0	144.8	144.4	142.0	
C_4-C_5	140.1	139.8	139.8	142.5	141.3	145.2	
$C_{10}-C_{11}$	139.9	139.3	139.8	147.7	146.0	144.9	
C_5-C_6	141.6	142.0	142.1	143.3	144.7	142.0	
C_6-C_{10}	142.4	141.6	142.1	139.1	140.7	142.5	
C_6-N_7	138.6	140.2	139.2	143.6	138.9	140.0	
		Selected Out-of-Plane Angles (deg)					
$C_{11}-C_{10}-C_3-H$	1.0	0.0	0.2	-13.1	-10.1	-0.7	
$C_6-C_5-N_{10}-C_7$	0.2	-0.2	-0.3	-1.2	-1.2	0.1	
		Selected Bond Angles (deg)					
$C_8-N_7-C_9$	115.0	113.4	116.0	116.3	114.9	115.8	
$C_2-C_3-C_6$	175.8	178.8	179.9	167.6	170.4	179.9	
$N_7-C_6-C_3$	178.4	178.6	178.2	170.7	172.4	179.9	
		Intersubsystem Coordinates					
COM displacement (pm)		11			35		
tilting (deg)		2.3			1.0		

^aThe tilting angle is calculated as the angle between the normal vectors of the planes that include the positions of the carbon atoms C_3 , C_5 , and C_{10} (compare Figure 1).

properties for testing the performance of PE-MP2 and PE-ADC(2) for potential energy surfaces: the host system has only one structural isomer with a relatively rigid structure. Thus, the implemented gradients for PE-MP2 and PE-ADC(2) can be used to optimize the geometries of the guest molecules while keeping the geometry of the host molecule frozen. This excludes evaluation of the results uncertainties from the force field parametrization for the host structure and the effect of temperature fluctuations, which are beyond the scope of the current work. We should emphasize that in other applications these effects may play an important role.

Computational Details: Full Supermolecular Calculations. In the supermolecular MP2 geometry optimization,³³ the structure of the whole DMABN·CB[7] complex was fully optimized. These calculations are done with the cc-pVDZ basis set⁴⁸ and the corresponding optimized auxiliary basis set.⁴⁹ In the correlated calculations, the frozen core approximation was used for all electrons in 1s orbitals of non-hydrogen atoms. These calculations were carried out with the OMP/MPI parallel RI-MP2 program³⁴ of the TURBOMOLE package,^{29,38} version 7.2, on 8 computer nodes, each with two Intel Xeon ES-2670 CPUs (in total 16 cores) and 64 GB RAM. Under these conditions, one cycle of the MP2 geometry optimization takes about 113 min wall time. Then the full supermolecular QM geometry optimizations for the excited state with local excitation (LE) character of the DMABN·CB[7] complex at the ADC(2)/cc-pVDZ level were done with the OMP/MPI-parallel version of the `ricc2` module of TURBOMOLE on 3 nodes, each with two Intel Xeon ES-2640 CPUs with totally 20 cores and 256 GB memory. These calculations took approximately 10 h per cycle of the geometry optimization. The starting structure for the supermolecular ADC(2) geometry optimization was constructed by positioning the gas phase $S_1(\text{LE})$ structure of DMABN optimized at the ADC(2)/cc-pVDZ level at the center of the CB[7] cavity. The

supermolecular ADC(2) geometry optimization required from this starting structure more than 70 steps to converge to a geometry with a Cartesian gradient norm of 0.00018 and with the changes of the total energy smaller than 10^{-6} au.

Computational Details: PE-QM Calculations. For the PE calculations, one needs in addition to the coordinates for QM and MM systems static distributed multipole moments and dipole–dipole polarizabilities for the MM system and, for geometry optimizations, also Lennard-Jones parameters for the interaction between the MM and the QM atoms. For our calculations, we used the atomic charges and dipole and quadrupole moments from an intrinsic atomic orbital⁵⁰ (IAO) analysis of the electron density from a DFT/B3LYP-D3(bj) calculation^{51–53} on the host structure in the cc-pVDZ basis set for the corresponding host geometry of each PE calculation, *vide infra*. The isotropic polarizabilities were determined as proposed in ref 54, and the anisotropic dipole polarizabilities are set to zero. Finally, the Lennard-Jones parameters for the atoms of the host molecule were taken from ref 55, and the parameters of the guest molecules, DMABN and DBO, were taken from refs 56 and 57, respectively, if an atom type was not defined in ref 55.

Also the PE-MP2 and PE-ADC(2) calculations used the RI and the frozen core approximations. All electrons in the 1s orbitals of non-hydrogen atoms were excluded from the MP2 and ADC(2) calculations. Again, the cc-pVDZ orbital and auxiliary basis sets were used for the geometry optimizations to allow a direct comparison with the supermolecular calculations. A typical PE-MP2 step for one cycle of the geometry optimization of the DMABN·CB[7] complex was finished on one node with Intel Xeon X5670 CPUs (in total 12 cores with 42 GB RAM) after 24 s, and the calculation on the excited state with PE-ADC(2) took 113 s on the same machine. The vacuum MP2/cc-pVDZ and ADC(2)/cc-pVDZ calculations 77 for the DMABN molecule took 18 and 108 s, respectively, on

the same machine. Therefore, one can conclude that the computational cost of PE-ADC(2) and PE-MP2 calculations are close to the computational cost of vacuum calculations where the environmental effects are completely neglected.

Geometry Optimizations with PE-MP2 and PE-ADC(2). To study the error in the geometries that is introduced by describing the host–guest interaction within the PE approximations, we started for the DMABN molecule from the structures obtained in the full supermolecular MP2 and ADC(2) calculations, kept the coordinates for the CB[7] host frozen, and reoptimized the coordinates of the guest with PE-MP2 and PE-ADC(2). To check if the host molecule has a significant impact on the intramolecular geometry of DMABN, we compare in addition with results from vacuum MP2 and ADC(2) calculations. The results are summarized in Table 1. A comparison between the optimized ground state geometries from PE-MP2 and supermolecular MP2 is shown in Figure 2.

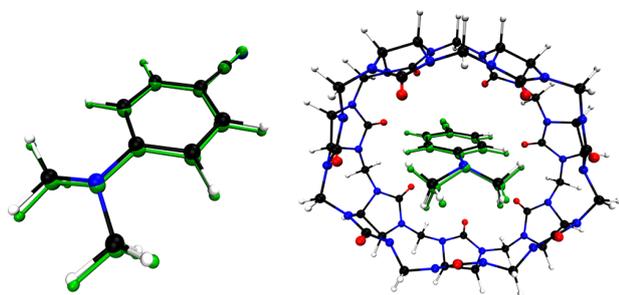


Figure 2. Comparison of the intra- and inter-subsystem coordinates from PE-MP2 and supermolecular MP2 for the ground state of DMABN·CB[7]. (left) Supermolecular MP2 structure (atom colors) superimposed on the PE-MP2 structure of DMABN (green). The host molecule is not shown for clarity. (right) Supermolecular MP2 geometry (atom colors) for the complex superimposed on the structure from PE-MP2 (green).

Both the intermolecular (between the host and guest system) and the intramolecular coordinates of the guest molecule (DMABN) are in good agreement. The mean absolute deviations are 0.36 pm for the bond lengths and 0.26° for the bond angles in the DMABN guest molecule, but in general the influence of the host on the internal coordinates of DMABN is small in the ground state. The DMABN center-of-mass for the PE-MP2 ground state structure is shifted by 11 pm and the molecular plane is tilted by $\sim 2^\circ$, comparing to the supermolecular MP2 structure. Compared with the intermolecular distances in the DMABN·CB[7] complex, this is a relatively small displacement, as can be seen in Figure 2.

The host effects are stronger in the excited state. In Figure 3, the fully optimized structure of DMABN·CB[7] for the $S_1(\text{LE})$ minimum at the ADC(2)/cc-pVDZ level is compared with the structure obtained with PE-ADC(2)/cc-pVDZ with a frozen host structure. In vacuum, the benzene ring in the $S_1(\text{LE})$ minimum exhibits a prominent antiquinoidal distortion. The interaction with the CB[7] host leads to a rotation of the nodes in the virtual natural transition orbital (NTO) shown in Table 2. As a consequence, the benzene ring shows a markedly different bond length distortion: the C_3-C_{11} , $C_{10}-C_{11}$, and C_5-C_6 bonds are 1–3 pm longer and the C_3-C_4 , C_4-C_5 , and C_6-C_{11} bonds are 1–4 pm shorter than those in the vacuum structure. This distortion is qualitatively correct in the PE-ADC(2) calculation, although not quantitatively reproduced.

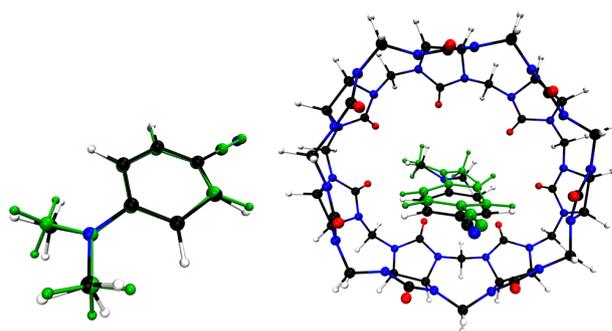


Figure 3. Comparison of the intra- and inter-subsystem coordinates from PE-ADC(2) and supermolecular ADC(2) for the $S_1(\text{LE})$ minimum of DMABN·CB[7]. (left) Supermolecular ADC(2) structure (atom colors) superimposed on the PE-ADC(2) structure of DMABN (green). The host molecule is not shown for clarity. (right) Supermolecular MP2 geometry (atom colors) for the complex superimposed on the structure from PE-MP2 (green).

Also the changes in the bending angles of the $-\text{CN}$ and $-\text{N}(\text{CH}_3)_2$ functional groups, the $C_2-C_3-C_6$ and $\text{N}_7-C_6-C_3$ angles, respectively, are qualitatively reproduced by the PE-ADC(2) model (see Table 1). The mean absolute deviation in the intramolecular bond lengths at 0.75 pm is about a factor of 2 larger than that in the ground state. Also the shift of the center of mass of DMABN between the supermolecular and the PE calculation for the $S_1(\text{LE})$ minimum is two times larger than that for the ground state. But these deviations are still small considering that the Lennard-Jones parameters have not been optimized for the PE-MP2 and PE-ADC(2) calculations. The main reason for the larger errors in the $S_1(\text{LE})$ state is probably that the binding energy is significantly higher than in the ground state. The supermolecular calculation gives a binding energy of $143 \text{ kJ}\cdot\text{mol}^{-1}$ for the ground state at the MP2/cc-pVDZ level and $200 \text{ kJ}\cdot\text{mol}^{-1}$ for the $S_1(\text{LE})$ state at the ADC(2)/cc-pVDZ level (in both cases not counterpoise corrected). One should also be aware that the supermolecular calculation has a larger basis set superposition error for the excited state than for the ground state because the populated virtual NTO is more diffuse than the depopulated occupied NTO. The PE-MP2 and PE-ADC(2) calculations, however, do not suffer from basis set superposition errors for the host–guest interaction.

Based on the reported timings, it is clear that pursuing supermolecular MP2 or ADC(2) calculations for systems larger than the studied complexes, for example, a chromophore bound to a protein, is not feasible.

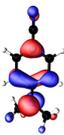
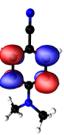
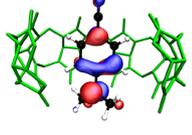
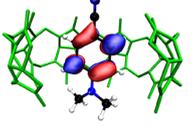
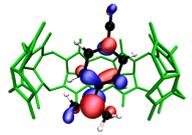
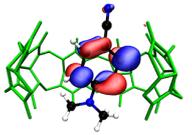
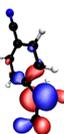
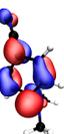
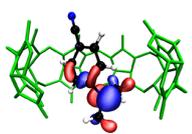
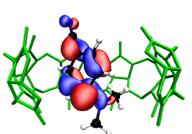
Complexation Effect on the Fluorescence Energies.

To investigate the effect of the supramolecular complexation on fluorescence energies, we studied two examples, DMABN·CB[7] and DBO·CB[7], and applied a slightly cheaper computational protocol than in the previous section. The structure of the CB[7] host was kept fixed in the MP2/cc-pVDZ ground state geometry for both complexes, while the coordinates of the DMABN and DBO guest molecules were optimized at the PE-ADC(2)/cc-pVDZ level.

For DMABN, we calculated the fluorescence emission energies out of the two minima of the lowest excited singlet state, $S_1(\text{LE})$ and $S_1(\text{ICT})$. The fluorescence energies of the LE and the twisted ICT states of DMABN·CB[7] complex have been measured experimentally in aqueous solutions.⁴¹

The experimental fluorescence energy for the LE state is 3.50

Table 2. Character of Lowest Electronic Excitation and Its Molecular Environment, Occupied (occ) and Virtual (vir) Natural Transition Orbitals, and Relative Contribution of the NTO Pair in the Excitations^a

	occ.	vir.	Contrib.
LE vacuum			92.7%
LE CB7			95.5%
LE SM			98.0%
ICT vacuum			99.8%
ICT CB7			99.6%

^aThe isosurface values for the visualization were set to ± 0.05 for all NTOs.

eV, which is in a very good agreement with the PE-ADC(2) result of 3.68 eV. The PE-ADC(2) result for the fluorescence energy of 2.05 eV for the twisted ICT state is 0.45 eV below the band maximum in the experimental fluorescence spectrum at 2.5 eV. This is in line with the CC2 results for the vacuum case from ref 44, which also is ~ 0.3 – 0.7 eV below the experimental band maxima in nonpolar solvents.

As another example we studied the $n\pi^*$ transition of DBO in the supramolecular complex with CB[7]. For this case, we also investigated the basis set dependence of the vertical excitation and fluorescence energy with different basis sets, see Table 3. The vertical excitation and fluorescence energies show a stable convergence for the cc-pVXZ (X = D, T, Q) basis sets, but they both collapse when diffuse functions are added as in the aug-cc-pVDZ basis. This issue is known in the literature as the charge penetration error or electrostatic catastrophe and occurs when the electron density of the QM region leaks significantly out into the MM region.^{17,58,59}

The natural transition orbitals for the DBO-CB[7] complex calculated at the PE-ADC/aug-cc-pVDZ level of theory with different embedding potentials are shown in Table 4. In the vacuum case, the lowest electronic excitation has $n\pi^*$ character, and the vertical excitation energy is 3.47 eV, which is in good agreement with the experimental band maximum λ_{\max} of the UV-vis spectrum at 3.31 eV.⁴⁰ For the cc-pVTZ and cc-pVQZ basis sets also the vacuum to CB[7] shift in the vertical excitation and fluorescence emission energies fit with,

Table 3. Vertical Excitation (VEE) and Fluorescence Emission Energies (FEE) of DBO-CB[7] Calculated at (ECP)/PE-ADC(2) with Different Basis Sets^a

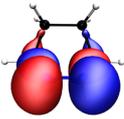
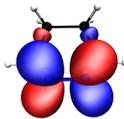
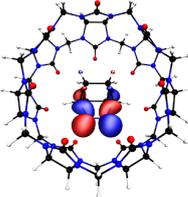
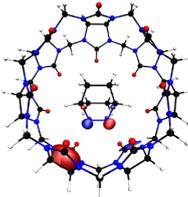
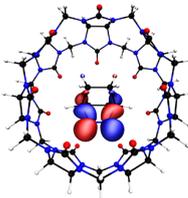
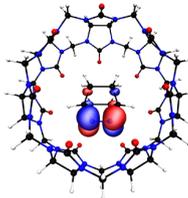
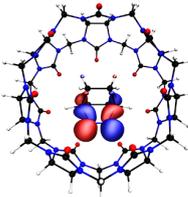
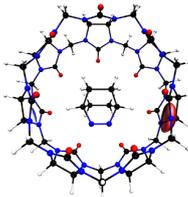
basis	gas phase		inside CB[7]		inside CB[7] (ECP)	
	VEE	FEE	VEE	FEE	VEE	FEE
cc-pVDZ	3.54	2.72	3.55	2.74	3.55	2.74
cc-pVTZ	3.46	2.65	3.47	2.67	3.47	2.67
cc-pVQZ	3.43	2.62	3.44	2.65	3.44	2.65
aug-cc-pVDZ	3.47	2.65	0.64	0.20	3.48	2.67
d-aug-cc-pVDZ	3.47	2.64	n.c.	n.c.	3.48	2.67
exptl ^b	3.31	2.79	3.32	2.82	3.32	2.82

^aAll VEEs and FEEs are calculated at PE-MP2/cc-pVDZ and PE-ADC(2)/cc-pVDZ optimized geometries. See the text for further details. ^bMeasured in an aqueous solution, ref 40.

respectively, -0.01 and $+0.03$ eV to the experimental results reported in ref 40.

However, in the PE-ADC(2) calculations with the aug-cc-pVDZ basis set, the excitation and fluorescence energies collapse to unphysical low values. The virtual natural transition orbital shown in Table 4 is in this case an orbital that is localized in the host structure and only accessible for the excitation process because the embedding potential does not account for the Pauli repulsion by the hosts electrons. We have analyzed this problem in more detail to identify to which 79 components of the electrostatic embedding (static charges,

Table 4. Environment Potential Used in the Calculation, Vertical Excitation Energies (VEE, given in eV) Calculated at PE-ADC(2)/aug-cc-pVDZ Level, Occupied (occ) and Virtual (vir) Natural Transition Orbitals, and Their Contributions to the Eigenvector

Potential	VEE	occ.	vir.	Contrib.
vacuum	3.47			99.8%
charges dipoles quadrupoles polarizability	0.64			99.4%
charges dipoles polarizability	3.55			99.7%
quadrupoles	0.85			99.9%

dipoles, quadrupoles or induced dipoles, that is, polarizabilities) the excitation energies are most sensitive. It turned out that in this case the vertical excitation energy collapses only when the static quadrupoles are included in the embedding potential (see Table 4). The expected $n\pi^*$ transition is recovered if we exclude quadrupoles from the embedding potential. The importance of the quadrupole moments for accurate embedding potentials has been discussed before in the context of various electrostatic models.^{60–63} In agreement with earlier observations, the exclusion of the quadrupoles has a significant effect on the excitation energies. Excluding quadrupoles in DBO-CB[7] shifts up the S_1 vertical excitation energy by 0.1 eV. To address this problem appropriately a robust embedding method must account somehow for the Pauli repulsion interaction. In this regard, we carried out pilot calculations where we added effective core potentials (ECP) to each non-hydrogen MM site to mimic at least partially the Pauli repulsion. In this test, we used the full PE potential (partial charges, dipoles, quadrupoles, and dipole polarizabilities) and, because pseudopotentials optimized for this purpose are not yet available, relativistic 2 electron SEFIT ECPs⁶⁴ for the MM subsystem. The results for VEEs and FEEs are presented in Table 3. With adding the nonelectrostatic contributions of the pseudopotentials, the charge leaking issue can be avoided and the VEE and FEE can be calculated with the same accuracy for the diffuse basis sets as for the other basis sets. Moreover the results of the ECP/PE embedding

model are very close to those from the PE embedding model without pseudopotentials for the basis sets without diffuse functions, which indicates that there are no artificial effects from the ECPs.

CONCLUSION

Analytic molecular gradients have been implemented for ground states at the PE-MP2 and for excited states at the PE-ADC(2) level into the `ricc2` program of the development version of the TURBOMOLE package.³⁸ In this implementation, the post-HF contributions to the reaction field potential is included in the ADC(2) response calculations via an approximate coupling density, which is calculated only from the single excitation parts of the excited state eigenvectors. This approach was used before successfully for the coupled cluster model CC2 and has the advantage that it conserves the high efficiency of the RI-CC2 or RI-ADC(2) implementations.⁹

The performance of the geometry optimizations with the QM/MM PE-ADC(2) gradients has been validated by a comparison to supermolecular MP2 and ADC(2) calculations for the DMABN-CB[7] host-guest complex. The QM/MM geometries can reproduce both the internal coordinates of the QM subsystem and its position and orientation in its environment reasonably provided that the interaction with the environment is small enough so that the assumptions that

underlie the QM/MM partitioning are valid. The use of basis sets with diffuse functions is problematic, however, if the embedding potential does not include any terms that model the Pauli repulsion, as it is the case for the original PE model that is used here. The diffuse functions enhance the charge leaking problem and can lead to low lying unphysical QM to MM charge transfer excitations, as demonstrated in the example of the DBO-CB[7] complex. Preliminary calculations show that pseudopotentials can be used as a remedy for the latter problem, although to exploit the full potential of this approach, the effective potentials should be optimized for hydrogens and also the valence electrons of the other atoms. Research in this direction is continued in our lab. The computational costs for PE-ADC(2) gradient calculations are only marginally higher than that for a corresponding gas phase calculation. Because of the low computational costs, PE-MP2 and PE-ADC(2) gradient calculations can be done many times for medium-sized systems (e.g., $\lesssim 50$ atoms) and thus be used to study the (short-time) dynamics after photoexcitation or to calculate intrinsic reaction coordinate pathways for relaxation processes at similar computational costs as for isolated systems. If the accuracy of the QM/MM approximation is controlled properly, this method can thus be used to study excited state potential energy surfaces and the short time dynamics after photoexcitation of medium size systems in complex molecular environments.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.8b00396.

Detailed information on the obtained reaction and interaction energies for both test sets (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: Alireza.MarefatKhah@ruhr-uni-bochum.de.

ORCID

Alireza Marefat Khah: 0000-0002-9593-0344

Christof Hättig: 0000-0002-5752-2710

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We are grateful to Prof. Jacob Kongsted and Dr. Jógvan Magnús Haugaard Olsen for providing us test cases of their PE-DFT implementation and for useful discussions. Financial support by Deutsche Forschungsgemeinschaft (DFG) through the cluster of excellence RESOLV (EXC 1069) and projects Ha 2588/8-1 and 2588/10-1 is gratefully acknowledged.

■ REFERENCES

- (1) Klymchenko, A. S. Solvatochromic and fluorogenic dyes as environment-sensitive probes: design and biological applications. *Acc. Chem. Res.* **2017**, *50*, 366–375.
- (2) Warshel, A.; Karplus, M. Calculation of ground and excited state potential surfaces of conjugated molecules. I. Formulation and parametrization. *J. Am. Chem. Soc.* **1972**, *94*, 5612–5625.
- (3) Warshel, A.; Levitt, M. Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* **1976**, *103*, 227–249.

- (4) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198–1229.
- (5) Olsen, J. M. H.; Kongsted, J. Molecular properties through polarizable embedding. *Adv. Quantum Chem.* **2011**, *61*, 107–143.
- (6) List, N. H.; Olsen, J. M. H.; Kongsted, J. Excited states in large molecular systems through polarizable embedding. *Phys. Chem. Chem. Phys.* **2016**, *18*, 20234–20250.
- (7) Olsen, J. M. H.; Aidas, K.; Kongsted, J. Excited States in Solution through Polarizable Embedding. *J. Chem. Theory Comput.* **2010**, *6*, 3721–3734.
- (8) Sneskov, K.; Schwabe, T.; Kongsted, J.; Christiansen, O. The polarizable embedding coupled cluster method. *J. Chem. Phys.* **2011**, *134*, 104108.
- (9) Schwabe, T.; Sneskov, K.; Haugaard Olsen, J. M.; Kongsted, J.; Christiansen, O.; Hättig, C. PERI-CC2: A Polarizable Embedded RI-CC2 Method. *J. Chem. Theory Comput.* **2012**, *8*, 3274.
- (10) List, N. H.; Coriani, S.; Kongsted, J.; Christiansen, O. Lanczos-driven coupled-cluster damped linear response theory for molecules in polarizable environment. *J. Chem. Phys.* **2014**, *141*, 244107.
- (11) Hedegård, E. D.; List, N. H.; Jensen, H. J. A.; Kongsted, J. The multi-configuration self-consistent field method within a polarizable embedded framework. *J. Chem. Phys.* **2013**, *139*, 044101.
- (12) Hedegård, E. D.; Jensen, H. J. A.; Kongsted, J. Polarizable Embedding Based on Multiconfigurational Methods: Current Developments and the Road Ahead. *Int. J. Quantum Chem.* **2014**, *114*, 1102–1107.
- (13) List, N. H.; Beerepoot, M. T. P.; Olsen, J. M. H.; Gao, B.; Ruud, K.; Jensen, H. J. A.; Kongsted, J. Molecular quantum mechanical gradients within the polarizable embedding approach—Application to the internal vibrational Stark shift of acetophenone. *J. Chem. Phys.* **2015**, *142*, 034119.
- (14) Caprasecca, S.; Jurinovich, S.; Viani, L.; Curutchet, C.; Mennucci, B. Geometry Optimization in Polarizable QM/MM Models: The Induced Dipole Formulation. *J. Chem. Theory Comput.* **2014**, *10*, 1588–1598.
- (15) Li, H. Analytic Energy Gradient in Combined Second-Order Møller-Plesset Perturbation Theory and Polarizable Force Field Calculation. *J. Phys. Chem. A* **2011**, *115*, 11824–11831.
- (16) Zeng, Q.; Liang, W. Analytic energy gradient of excited electronic state within TDDFT/MMpol framework: Benchmark tests and parallel implementation. *J. Chem. Phys.* **2015**, *143*, 134104.
- (17) Menger, M. F. S. J.; Caprasecca, S.; Mennucci, B. Excited-State Gradients in Polarizable QM/MM Models: An Induced Dipole Formulation. *J. Chem. Theory Comput.* **2017**, *13*, 3778–3786.
- (18) Yanai, T.; Tew, D.; Handy, N. C. A new hybrid exchange-correlation functional using the Coulomb-attenuating method (CAM-B3LYP). *Chem. Phys. Lett.* **2004**, *393*, 51–57.
- (19) Eriksen, J. J.; Sauer, S. P.; Mikkelsen, K. V.; Christiansen, O.; Jensen, H. J. A.; Kongsted, J. Failures of TDDFT in describing the lowest intramolecular charge-transfer excitation in para-nitroaniline. *Mol. Phys.* **2013**, *111*, 1235–1248.
- (20) Christiansen, O.; Koch, H.; Jørgensen, P.; Helgaker, T. The second-order approximate coupled cluster singles and doubles model CC2. *Chem. Phys. Lett.* **1996**, *263*, 530–539.
- (21) Schirmer, J. Beyond the random-phase approximation: A new approximation scheme for the polarization propagator. *Phys. Rev. A: At., Mol., Opt. Phys.* **1982**, *26*, 2395.
- (22) Hättig, C. Structure optimizations for excited states with correlated second-order methods: CC2 and ADC(2). *Adv. Quantum Chem.* **2005**, *50*, 37.
- (23) List, N. H.; Olsen, J. M. H.; Rocha-Rinza, T.; Christiansen, O.; Kongsted, J. Performance of Popular XC-Functionals for the Description of Excitation Energies in GFP-Like Chromophore Models. *Int. J. Quantum Chem.* **2012**, *112*, 789–800.
- (24) Jacquemin, D.; Duchemin, I.; Blase, X. 0–0 Energies Using Hybrid Schemes: Benchmarks of TD-DFT, CIS(D), ADC(2), CC2 and BSE/GW formalism for 80 Real-Life Compounds. *J. Chem. Theory Comput.* **2015**, *11*, 5340.

- (25) Winter, N. O.; Graf, N. K.; Leutwyler, S.; Hättig, C. Benchmarks for 0–0 transitions of aromatic organic molecules: DFT/B3LYP, ADC(2), CC2, SOS-CC2 and SCS-CC2 compared to high-resolution gas-phase data. *Phys. Chem. Chem. Phys.* **2013**, *15*, 6623–6630.
- (26) Tuna, D.; Lefrançois, D.; Wolański, Ł.; Gozem, S.; Schapiro, I.; Andruniów, T.; Dreuw, A.; Olivucci, M. Assessment of Approximate Coupled-Cluster and Algebraic-Diagrammatic-Construction Methods for Ground- and Excited-State Reaction Paths and the Conical-Intersection Seam of a Retinal-Chromophore Model. *J. Chem. Theory Comput.* **2015**, *11*, 5758–5781.
- (27) List, N. H.; Norman, P.; Kongsted, J.; Jensen, H. J. A. A quantum-mechanical perspective on linear response theory within polarizable embedding. *J. Chem. Phys.* **2017**, *146*, 234101.
- (28) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-structure Theory*; John Wiley & Sons, Ltd, 2000.
- (29) 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>.
- (30) Hršak, D.; Marefat Khah, A.; Christiansen, O.; Hättig, C. Polarizable Embedded RI-CC2 Method for Two-Photon Absorption Calculations. *J. Chem. Theory Comput.* **2015**, *11*, 3669–3678.
- (31) del Valle, F. J. O.; Tomasi, J. Electron correlation and solvation effects. I. Basic formulation and preliminary attempt to include the electron correlation in the quantum mechanical polarizable continuum model so as to study solvation phenomena. *Chem. Phys.* **1991**, *150*, 139–150.
- (32) Angyán, J. G. Choosing between alternative MP2 algorithms in the self-consistent reaction field theory of solvent effects. *Chem. Phys. Lett.* **1995**, *241*, 51–56.
- (33) Weigend, F.; Häser, M. RI-MP2: first derivatives and global consistency. *Theor. Chem. Acc.* **1997**, *97*, 331–340.
- (34) Hättig, C.; Hellweg, A.; Köhn, A. Distributed memory parallel implementation of energies and gradients for second-order Møller-Plesset perturbation theory with the resolution-of-the-identity approximation. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1159–1169.
- (35) 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**, *68*, 1–52.
- (36) Head-Gordon, M.; Oumi, M.; Maurice, D. Quasidegenerate second-order perturbation corrections to single-excitation configuration interaction. *Mol. Phys.* **1999**, *96*, 593–602.
- (37) 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.
- (38) Furche, F.; Ahlrichs, R.; Hättig, C.; Klopper, W.; Sierka, M.; Weigend, F. Turbomole. *WIREs Comput. Mol. Sci.* **2014**, *4*, 91.
- (39) Nau, W. M.; Greiner, G.; Rau, H.; Wall, J.; Olivucci, M.; Scaiano, J. C. Fluorescence of 2,3-Diazabicyclo[2.2.2]oct-2-ene Revisited: Solvent-Induced Quenching of the $n\pi^*$ Excited State by an Aborted Hydrogen Atom Transfer. *J. Phys. Chem. A* **1999**, *103*, 1579–1584.
- (40) Marquez, C.; Nau, W. M. Polarizabilities Inside Molecular Containers. *Angew. Chem., Int. Ed.* **2001**, *40*, 4387–4390.
- (41) Sayed, M.; Biedermann, F.; Uzunova, V. D.; Assaf, K. I.; Bhasikuttan, A. C.; Pal, H.; Nau, W. M.; Mohanty, J. Triple Emission from p-Dimethylaminobenzonitrile-Cucurbit[8]uril Triggers the Elusive Excimer Emission. *Chem. - Eur. J.* **2015**, *21*, 691–696.
- (42) Georgieva, I.; Aquino, A. J. A.; Plasser, F.; Trendafilova, N.; Köhn, A.; Lischka, H. Intramolecular Charge-Transfer Excited-State Processes in 4-(N,N-Dimethylamino)benzonitrile: The Role of Twisting and the $\pi\sigma^*$ State. *J. Phys. Chem. A* **2015**, *119*, 6232–6243.
- (43) Georgieva, I.; Aquino, A. J. A.; Plasser, F.; Trendafilova, N.; Köhn, A.; Lischka, H. Intramolecular Charge-Transfer Excited-State Processes in 4-(N,N-Dimethylamino)benzonitrile: The Role of Twisting and the $\pi\sigma^*$ State. *J. Phys. Chem. A* **2015**, *119*, 6232–6243.
- (44) Köhn, A.; Hättig, C. On the nature of the low-lying singlet states of dmabn. *J. Am. Chem. Soc.* **2004**, *126*, 7399–7410.
- (45) Hättig, C.; Hellweg, A.; Köhn, A. Intramolecular charge transfer mechanism in quinolidines: the role of the amino twist angle. *J. Am. Chem. Soc.* **2006**, *128*, 15672–15682.
- (46) Curchod, B. F. E.; Sisto, A.; Martínez, T. J. Ab initio multiple spawning photochemical dynamics of DMABN using GPUs. *J. Phys. Chem. A* **2017**, *121*, 265–276.
- (47) Assaf, K. I.; Nau, W. M. Curcubiturils: from synthesis to high-affinity binding and catalysis. *Chem. Soc. Rev.* **2015**, *44*, 394–418.
- (48) Dunning, 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.
- (49) 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.
- (50) Knizia, G. Intrinsic Atomic Orbitals: An Unbiased Bridge between Quantum Theory and Chemical Concepts. *J. Chem. Theory Comput.* **2013**, *9*, 4834–4843.
- (51) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- (52) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.
- (53) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32*, 1456.
- (54) Schröder, H.; Schwabe, T. Efficient Determination of Accurate Atomic Polarizabilities for Polarizable Embedding Calculations. *J. Comput. Chem.* **2016**, *37*, 2052–2059.
- (55) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.
- (56) Jørgensen, W. L.; Laird, E. R.; Nguyen, T. B.; Tirado-Rives, J. Monte Carlo Simulations of Pure Liquid Substituted Benzenes with OPLS Potential Functions. *J. Comput. Chem.* **1993**, *14*, 206–215.
- (57) Schäfer, L. V.; Müller, E. M.; Gaub, H. E.; Grubmüller, H. Elastic Properties of Photoswitchable Azobenzene Polymers from Molecular Dynamics Simulations. *Angew. Chem., Int. Ed.* **2007**, *46*, 2232–2237.
- (58) Reinholdt, P.; Olsen, J. M. H.; Kongsted, J. Polarizable Density Embedding: A Solution to the Electron Spill-Out Problem in Multiscale Modeling. *J. Phys. Chem. Lett.* **2017**, *8*, 5949–5958.
- (59) Demerdash, O.; Mao, Y.; Liu, T.; Head-Gordon, M.; Head-Gordon, T. Assessing many-body contributions to intermolecular interactions of the AMOEBA force field using energy decomposition analysis of electronic structure calculations. *J. Chem. Phys.* **2017**, *147*, 161721.
- (60) Narth, C.; Lagardère, L.; Polack, E.; Gresh, N.; Wang, Q.; Bell, D. R.; Rackers, J. A.; Ponder, J. W.; Ren, P. Y.; Piquemal, J.-P. Scalable Improvement of SPME Multipolar Electrostatics in Anisotropic Polarizable Molecular Mechanics Using a General Short-Range Penetration Correction up to Quadrupoles. *J. Comput. Chem.* **2016**, *37*, 494–505.
- (61) Misquitta, A. J.; Stone, A. J. Ab initio Atom-Atom Potentials Using CAMCASP: Theory and Application to Many-Body Models for Pyridine Dimer. *J. Chem. Theory Comput.* **2016**, *12*, 4184–4208.
- (62) Rackers, J. A.; Wang, Q.; Liu, C.; Piquemal, J.-P.; Ren, P.; Ponder, J. W. An optimized charge penetration model for use with the AMOEBA force field. *Phys. Chem. Chem. Phys.* **2017**, *19*, 276–291.
- (63) Elking, D. M.; Cisneros, G. A.; Piquemal, J.-P.; Darden, T. A.; Pedersen, L. G. Gaussian Multipole Model (GMM). *J. Chem. Theory Comput.* **2010**, *6*, 190–202.
- (64) Igel-Mann, G.; Stoll, H.; Preuss, H. Pseudopotentials for main group elements (IIIa through VIIa). *Mol. Phys.* **1988**, *65*, 1321–1328.

8 Avoiding Electron Spill Out in QM/MM Calculations on Excited States with Simple Pseudopotentials

In this manuscript, the PE(ECP) model has been introduced to model Pauli repulsion effects in QM/MM calculations. A simple set of all-electron ECPs have been optimized and tested against the fully quantum mechanical reference numbers that are obtained for excitation energies of various chromophores in different molecular environments.

The general research question has been put forward by Professor Hättig and Professor Kongsted and the work has been done under their supervision. Peter Reinholdt and I have designed the main framework of the study. This framework has been discussed between all the co-authors. I designed the benchmark set, and Peter Reinholdt carried out dynamics for two systems and collecting the configurations from the literature. I have prepared the final systems. Professor Hättig and I optimized the ECPs. I did all the PNO-ADC(2) and PE-ADC(2) calculations. Peter Reinholdt did all the TDHF and PE-TDHF calculations. The obtained results have been discussed between all the co-authors. I wrote the first full draft. I made tables 1-3 and figures 1 and 2. This draft was developed further by all the co-authors and Peter Reinholdt included the results and discussion for TDHF and PE-TDHF calculations.

Avoiding Electron Spill Out in QM/MM Calculations on Excited States with Simple Pseudopotentials

Alireza Marefat Khah,^{*,†} Peter Reinholdt,^{*,‡} Jógvan Magnus Haugaard Olsen,[¶]
Jacob Kongsted,^{*,‡} and Christof Hättig^{*,†}

[†]*Quantum Chemistry Group, Ruhr University of Bochum, Germany*

[‡]*Department of Physics, Chemistry and Pharmacy, University of Southern Denmark,
Odense, Denmark*

[¶]*Hylleraas Centre for Quantum Molecular Sciences, Department of Chemistry, UiT The
Arctic University of Norway, N-9037 Tromsø, Norway*

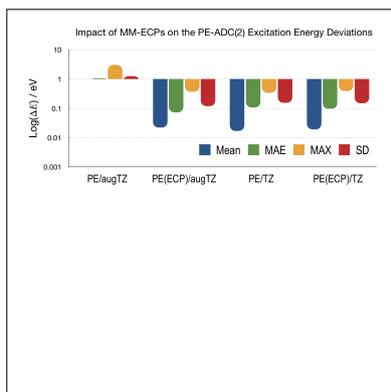
E-mail: alireza.marefatkhah@rub.de; reinholdt@sdu.dk; kongsted@sdu.dk;

christof.haettig@rub.de

Abstract

QM/MM calculations of electronic excitations with diffuse basis set have often large errors due to spilling of electrons from the quantum (QM) subsystem into the molecular mechanics (MM) environment. To avoid this, the Pauli repulsion effect by the environment on the electronic wavefunction has to be included. We propose transferable atomic all-electron pseudopotentials that can readily be combined with most molecular mechanics (MM) force-fields. To validate them, QM/MM excitation energies at the level of time-dependent Hartree-Fock theory and the algebraic diagrammatic construction through second-order are benchmarked against full quantum mechanical calculations. The QM/MM calculations with pseudopotentials give accurate results, stable with respect to augmentation of the basis with diffuse functions. An energy decomposition analysis shows that the largest contribution to the residual deviations from the full QM calculations is caused by the missing London dispersion interaction between the environment and the electrons in the QM subsystem.

Graphical TOC Entry



QM/MM methods¹ have become a standard tool for multiscale simulations of molecular and spectroscopic properties in complex environments.²⁻⁵ Many quantum chemistry packages are equipped with a QM/MM module or have been interfaced to molecular mechanics programs. For a fast convergence with the size of the quantum mechanical (QM) region, it is advantageous if the molecular mechanics (MM) part is polarizable.^{6,7} Different polarizable QM/MM schemes (QM/MMpol) have been developed that account for the polarization of the MM region through, e.g., point dipole-dipole polarizabilities,⁸⁻¹³ fluctuating charges,¹⁴⁻¹⁶ or Drude oscillators.¹⁷⁻¹⁹ Up to today, QM/MMpol approaches have been combined with (time-dependent) Hartree-Fock and Density Functional Theory^{8,10,12,13} and some of the most widely used correlated wavefunction methods as, e.g., coupled-cluster^{20,21} (CC), the algebraic-diagrammatic construction^{22,23} (ADC), multiconfigurational self-consistent field²⁴ (MCSCF), complete active space perturbation theory²⁵ (CASPT2), the second-order polarization propagator approach SOPPA,²⁶ and the density matrix renormalization group²⁷ (DMRG) approach.

In additive QM/MM schemes,² the missing effects of Pauli repulsion and dispersion interactions on the electrons in the QM subsystem are one of the pitfalls that require to go beyond a purely classical electrostatic and mechanical description. In the following, we focus on polarizable additive QM/MM methods for which the energy can be written as the sum of the internal energies of the QM and MM subsystems and an interaction term:

$$E^{\text{full}} = E^{\text{QM}} + E^{\text{MM}} + E^{\text{QM/MM}} . \quad (1)$$

The latter consists usually of a contribution from the permanent electrostatic moments of the particles in the MM region, $E_{\text{es}}^{\text{QM/MM}}$, a polarization contribution $E_{\text{pol}}^{\text{QM/MM}}$, and contributions from non-classical or van der Waals interactions $E_{\text{vdW}}^{\text{QM/MM}}$:

$$E^{\text{QM/MM}} = E_{\text{es}}^{\text{QM/MM}} + E_{\text{pol}}^{\text{QM/MM}} + E_{\text{vdW}}^{\text{QM/MM}} . \quad (2)$$

$E_{vdW}^{QM/MM}$ is usually approximated by mechanical potentials acting only on the nuclei, e.g., pair-wise 6-12 Lennard-Jones potentials. The missing effect of the Pauli repulsion on the electronic wavefunction leads to the so-called electron spill-out (ESO) problem, an unphysical leaking of the electron density into the MM environment. The ESO diminishes the accuracy of the calculations and in some cases causes completely unphysical results; in particular, when diffuse basis functions are required to capture the physical character of an excited state or to achieve the required accuracy. Several remedies have been proposed to avoid it. The modified Coulombic interaction approach²⁸ is often used in combination with standard MM force-fields. Unfortunately, this approach leads to a poor description of the electrostatic interactions across the QM/MM boundary. It has also been suggested that replacing point multipole moments with a Gaussian-smeared charge density can cure the issue.^{29,30} This can, in effect, lift the electrostatic singularities that occur at MM sites with point multipoles and provide a more accurate description of the continuous charge distribution in the short-range. However, it does not capture the effects of the Pauli repulsion on the electronic wavefunction. Moreover, modifying the electrostatic parameters of the MM force field makes the force field inconsistent and may require a laborious re-parametrization. In recent years, several extensions over the standard QM/MMpol approach have been proposed to substitute the purely mechanical description of $E_{vdW}^{QM/MM}$ by a density-dependent description of Pauli repulsion.^{29,31} For some of these more accurate but also more involved and costly approaches,²⁹ molecular gradients are not yet available and for some of them, preceding QM calculations for the MM subsystem or a re-parametrization of the electrostatics³¹ is required, which limit their applicability for large systems.

Therefore, there is a demand for tools to avoid the electron spill-out in QM/MM calculations, which satisfy the following criteria: **i)** efficient with computational costs that do not scale steeply with the size of the MM region, **ii)** accurate enough to be used with small QM regions and high-level QM methods, **iii)** compatible with the calculation of molecular gradients and response properties, and **iv)** easy to set up and parameterize.

Using pseudo-potential operators to model the effect of chemically inactive (core) electrons on the explicitly treated electrons is a well-known approach,³² first introduced by Hellmann³³ and Gombas³⁴ to model Pauli repulsion in metallic systems. Notably, for heavy elements, effective core potentials (ECPs) are frequently employed to reduce the computational costs and to include the most important relativistic effects on the valence electrons³⁵ in otherwise non-relativistic calculations. ECPs are also widely used in embedded cluster calculations for defects in ionic crystals and on their surfaces to avoid ESO as for these systems, the MM system contains only few atom- or ion-types and well-parametrized ECPs are available.³⁶ Especially parameterized ECPs have also been used in ground- and excited-state QM/MM calculations to model link atoms and functional groups at the QM/MM boundary.^{37–40} In the current work, we investigate the possibility of using atomic pseudo-potentials to avoid the electron spill-out problem in QM/MM calculations on electronically excited states. For simplicity of the parameterization, we assume that all atoms in the MM region fulfill the octet rule, i.e., have no dangling bonds that might serve as electron acceptors, and that their contribution to the Pauli repulsion can be described by a superposition of transferable one-center pseudopotentials that mimic approximately the core and valence electrons. We use the same ansatz as for the usual effective core potentials⁴¹ since for them the integrals are already available in most quantum chemistry packages:

$$\hat{V}_\lambda^{\text{ECP}}(i) = \sum_{l=0}^{l_{\text{max}}} \sum_k A_{l,k}^\lambda |\mathbf{r}_\lambda - \mathbf{r}_i|^{n_{l,k}^\lambda} \exp(-a_{l,k}^\lambda (\mathbf{r}_\lambda - \mathbf{r}_i)^2) \hat{P}_l^\lambda(i) \quad (3)$$

with

$$P_l^\lambda(i) = \sum_{m_l=-l}^l |\lambda l m_l(i)\rangle \langle \lambda l m_l(i)|. \quad (4)$$

In Eq. (3), $P_l^\lambda(i)$ projects the orbital occupied by electron i on the spherical harmonic function with angular momentum l around the position \mathbf{r}_λ of the MM site λ . In the current work, l_{max} is set to the highest angular momentum occupied in the ground state configuration

of the atom plus one, the exponents $n_{l,k}^\lambda$ are set to zero, and only one Gaussian function ($k = 0$) is used per angular momentum. For simplicity, we assume that the ECPs only depend on the number of atomic orbitals that are (partially) filled after fulfillment of the octet rule, i.e. we use for H the same ECP as for He, for B–F the same as for Ne, and for Al–Cl the same as for Ar. We thus optimized for the current work only ECP parameter for the three rare gas atoms He, Ne, and Ar. The coefficients $A_{l,0}^\lambda$ and exponents $a_{l,0}^\lambda$ for He, Ne, and Ar were determined by minimizing the penalty function:

$$L = \sum_{l=0}^{l_{max}} \sum_{a=1}^2 (\epsilon_{l,a}^{ECP} - \epsilon_{l,a}^{all-el})^2 . \quad (5)$$

where $\epsilon_{l,a}^{ECP}$ and $\epsilon_{l,a}^{all-el}$ are the canonical Hartree-Fock orbital energies of the lowest unoccupied orbitals a with angular momentum l in, respectively, ECP and all-electron calculations on the atoms in a large basis set, see SI. Adding such minimal ECPs to all MM sites increases the computational cost only moderately (see SI) and since the ECP integrals are more short-ranged than the Coulombic charge and multipole interaction integrals they will not affect the scaling of the overall costs with the sizes of the QM and MM regions. For the following, we focus on the polarizable embedding^{9,10,21,22} (PE) variant of QM/MM and, to distinguish between PE calculations with and without ECPs at the MM centers, we denote the ECP-augmented polarizable embedding as PE(ECP).

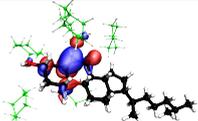
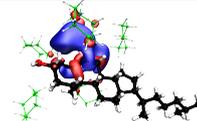
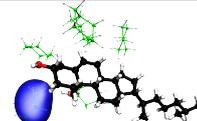
To assess the performance of PE(ECP), we compare below electronic excitation energies calculated within QM/MM set-ups with and without ECPs with full QM calculations at the time-dependent Hartree-Fock (TDHF) and the ADC(2) level. For the correlated wavefunction method ADC(2), the London dispersion interaction between the QM system and the MM environment is in addition to the Pauli repulsion another source of discrepancy between the full QM and the QM/MM calculations. To separate this error from the others, we used the PNO-ADC(2) code^{42,43} that is part of the development version⁴⁴ of TURBOMOLE. It is based on localized molecular orbitals (LMOs) for the occupied and pair natural orbitals for

the virtual space which allows to split the correlation contribution to the excitation energy originating from the double excitations, $T_2 = \sum_{aibj} t_{ij}^{ab} \tau_{ab}^{ij}$, according to the localization of the LMOs i and j into a major contribution where both LMOs are localized on the chromophore in the QM region, a smaller intermolecular contribution where one LMO is localized on the chromophore and one in the environment, which is identified with the dispersion interaction, and a minor contribution where both LMOs are localized in the environment. By subtracting the dispersion contribution from the reference full QM PNO-ADC(2) excitation energies (denoted below as NoD-PNO-ADC(2) results), we can isolate the electrostatic and Pauli repulsion contributions, which are the subject of the current study. In all the PNO-ADC(2) calculations, we used a tight PNO selection threshold ($T_{\text{PNO}} = 10^{-9}$) to ensure that the errors introduced by the local approximations are negligible.⁴²

For this comparison, we used basis sets with and without diffuse basis functions since the amount of ESO varies a lot with the inclusion of diffuse functions.^{22,45} For excited states, however, diffuse basis functions are often needed to achieve reasonable accuracy. Therefore, it is desirable to be able to use these basis sets also in QM/MM calculations. Thiel and co-workers⁴⁶ showed that in vacuum, CC2 excitation energies that are calculated with the triple- ζ basis TZVP deviate already for valence states by about 0.2 eV from more accurate results in the aug-cc-pVTZ basis. Already for a qualitatively correct description of Rydberg excitations and mixed-valence/Rydberg states, diffuse functions are unavoidable. Even though the Rydberg character is in solution partially quenched, the corresponding states remain diffuse in character. In ref. 47, some of us have shown that the standard PE scheme is not sufficient for the simulation of the electronically excited states of cholesterol in cyclohexane solution. The second-lowest excited state, S_2 highlights the problem. Table 1 summarizes some results from full QM and QM/MM calculations at the ADC(2) level for this state in a cyclohexane cluster. In the full-QM PNO-ADC(2) calculation with the aug-cc-pVTZ (aTZ) basis, the vertical excitation energy of cholesterol within the cyclohexane cluster is 6.40 eV after subtracting the dispersion contribution δE_{Disp} which in this case

would cause a red-shift of 0.138 eV. A similar PNO-ADC(2) calculation in the cc-pVTZ (TZ) basis, i.e., leaving out the diffuse functions, gives as S_2 a different excited state with a much higher excitation energy of 7.54 eV that deviates from the aTZ result by 1.1 eV. The PE-ADC(2) QM/MM calculation in the aTZ basis set is, however, strongly affected by ESO which leads to the localization of the virtual natural transition orbital (NTO) on a subunit of the MM system and an unphysical small excitation energy of 4.5 eV within the post-SCF linear response reaction field scheme.²¹ In contrast to the standard PE embedding, PE(ECP)-ADC(2) with ECPs on all MM sites can reproduce in the aTZ basis the NoD-PNO-ADC(2) result very accurately with a remaining small deviation of only 0.01 eV.

Table 1: The full-QM NoD-PNO-ADC(2) ($T_{\text{PNO}} = 9$) excitation energy ΔE to the second-lowest singlet electronic excited state S_2 , calculated with mixed aTZ/DZ and TZ/DZ basis sets. The full-QM PNO-ADC(2) S_2 excitation energies are quoted in the parentheses. PE- and PE(ECP)-ADC(2)/aTZ excitation energies ΔE are calculated within the post-SCF linear response (LR) formalism. NTOs are visualized for isosurface values of ± 0.0275 a.u.. Carbon atoms of solvent molecules are shown in green.

Level	S_2	ΔE	occ	vir
full-QM/aTZ	$n\pi \rightarrow \text{O-Ryd.}$	6.40 (6.26)		
full-QM/TZ	$\pi \rightarrow \text{C-Ryd.}$	7.54 (7.08)		
PE/aTZ	$\pi \rightarrow \text{ESO}$	4.53		
PE(ECP)/aTZ	$n\pi \rightarrow \text{O-Ryd.}$	6.41		

To evaluate the transferability and accuracy of the constructed all-electron ECPs for the embedding, we chose for (PE-)ADC(2) a benchmark set that comprises 9 chromophores and in total 15 electronic excitations, 10 local ($\pi\pi^*$, $n\pi^*$) and 5 non-local transitions with charge-transfer CT or Rydberg character (see the NTOs given in the SI). For (PE-)TDHF

we included in addition 15 further excited states (cmp. SI) so that in total of 30 states are considered at this level. The chromophores are embedded (see Fig. 1) in various chemically different environments like a large cluster of water molecules, organic solvents (cyclohexane and acetonitrile), biomolecular (protein and DNA) matrices and a large molecular container, (curcubituril-[7], CB[7]). These examples have been chosen because of their sensitivity to ESO reported in previous excited-state QM/MM studies.^{22,47,48} In all QM/MM calculation, we placed only the chromophore in the QM region, and all other molecules are treated at the MM level. In the PE(ECP) calculations, ECPs are added to all atoms of the MM environment. The full-QM TDHF calculations are done using either the cc-pVDZ (DZ) or the aug-cc-pVDZ (aDZ) basis set for all atoms. For the PE-TDHF excitation energies, the embedding parameters for the atom-centered multipoles and polarizabilities were determined with the LoProp method using an ANO-recontracted version⁴⁹ of the corresponding basis that is used in the full-QM TDHF calculation. These calculations are done with the Dalton program. The full-QM PNO-ADC(2) calculations, PE-ADC(2), and PE(ECP)-ADC(2) calculations are done with the `pnoccsd` and `ricc2` modules of the Turbomole program package. In the full-QM PNO-ADC(2) calculations, mixed basis sets are used with a larger (TZ or aTZ) basis for the chromophore and a smaller DZ basis set for the remaining molecules. Consistently, in the PE-ADC(2) calculations, we used the TZ and aTZ basis for the QM system. For the embedding, we used atomic multipoles obtained from an intrinsic atomic orbital analysis⁵⁰ at the B3LYP/DZ level and isotropic polarizabilities from the parameter set for the D3 dispersion correction.^{51,52} Furthermore, in the PE- and PE(ECP)-ADC(2) calculations, the linear response LR coupling between the correlated wavefunction and the polarizable environment is included at the post-SCF linear response level.^{21,22}

Looking first at the statistical measures for the TDHF benchmark set (see Table 2), we see that for a compact basis set, the overall performance between the PE and PE(ECP) embeddings is comparable, with similar MSD (0.068 and 0.082 eV, respectively) and likewise for the MAD. The largest difference is seen in the maximum deviation, which is significantly

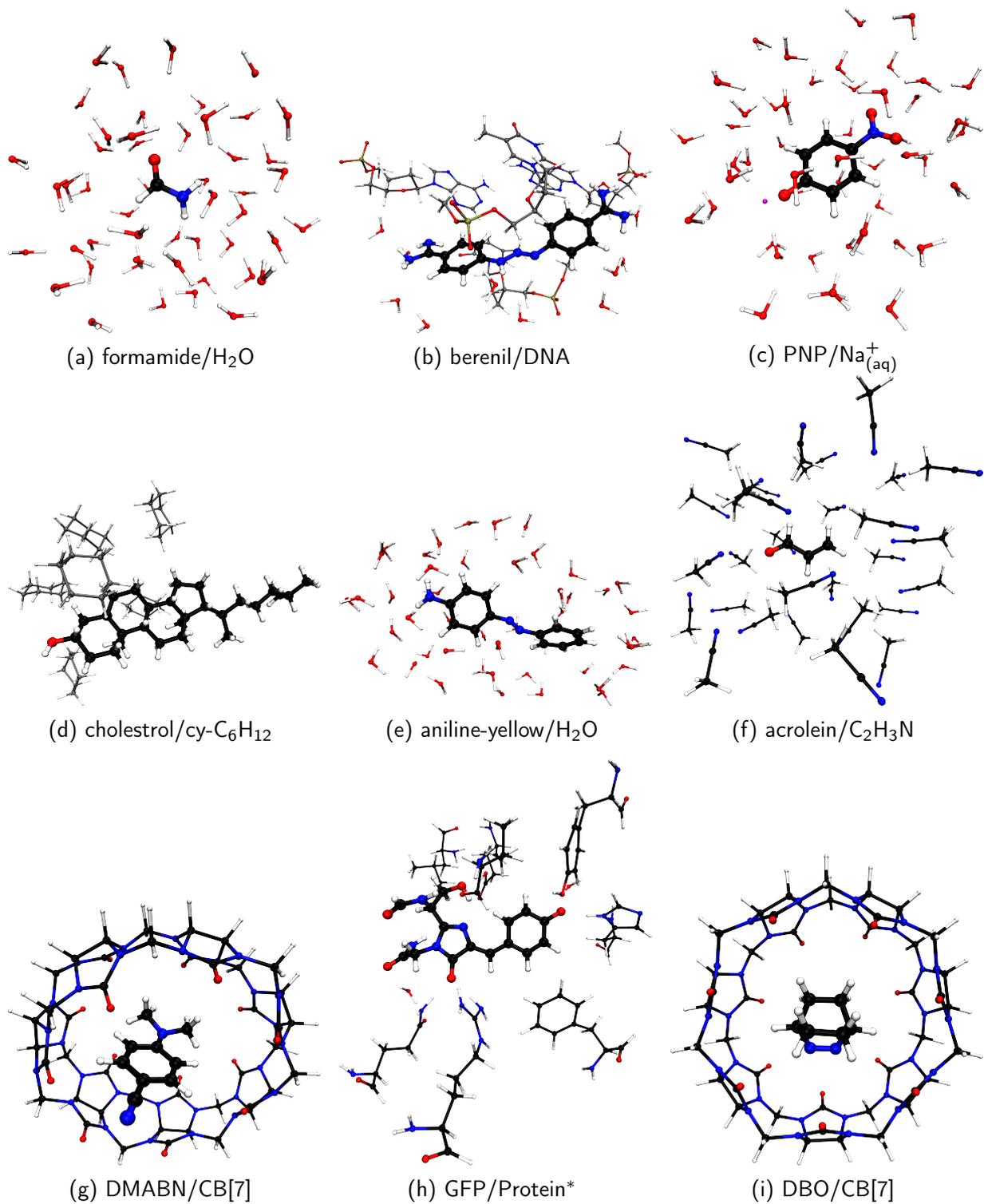


Figure 1: Systems that are considered in this study. The QM and MM subsystems for the QM/MM calculations are distinguished by a "/". (a,f) taken from a classical MD, see SI, (b) taken from a classical MD, see SI, (c) taken from ref. 45, (d) taken from ref. 47, (e) taken from ref. 48, (g,i) taken from ref. 22, (h) taken from ref. 7

Table 2: Statistical measures of excitation energies (eV) for the QM/MM PE and PE(ECP) embedding models combined with TDHF and ADC(2) in the test set. The MSD, MAD, MAX are the mean signed, mean absolute, and maximum deviations, respectively. The TDHF results are reported w.r.t. full-QM TDHF. PE- and PE(ECP)-ADC(2) deviations are compared with the reference PNO-ADC(2) (columns 2–4) and dispersion-corrected NoD-PNO-ADC(2) (columns 5–7) calculations.

Method/Basis Set	$\Delta E_{\text{full-QM}} - \Delta E_{\text{QM/MM}}$			$\Delta E_{\text{full-QM}} - \delta E_{\text{disp}} - \Delta E_{\text{QM/MM}}$		
	MSD	MAD	MAX	MSD	MAD	MAX
PE-TDHF/DZ	0.07	0.07	0.34	–	–	–
PE(ECP)-TDHF/DZ	0.08	0.09	0.77	–	–	–
PE-TDHF/aDZ	–0.72	0.73	–3.83	–	–	–
PE(ECP)-TDHF/aDZ	0.08	0.08	0.40	–	–	–
PE-ADC(2)/TZ	0.13	0.15	0.48	0.02	0.11	0.34
PE(ECP)-ADC(2)/TZ	0.13	0.14	0.54	0.02	0.10	0.39
PE-ADC(2)/aTZ	–0.95	1.02	–2.91	–1.06	1.07	–3.07
PE(ECP)-ADC(2)/aTZ	0.13	0.14	0.54	0.02	0.07	0.37

higher (0.772 eV) when using the PE(ECP) than when using conventional PE (0.335 eV). This deviation is caused by the second excited state in cholesterol; excluding this, the maximum deviation drops down to 0.357 eV.

If diffuse basis functions are included the two embeddings perform very differently. Whereas the MSD and MAD remain similar to the DZ case for the PE(ECP), the conventional PE approach has much larger errors, with the MAD increased from 0.07 to 0.73 eV. The errors are systematically dominated by highly red-shifted states, as indicated by the large, negative MSD. The maximum deviation of –3.8 eV occurs with the third excited state in the DBO@CB[7] system. Similar large errors are present in the DMABN and cholesterol systems.

The full set of excitation energies is depicted in Figure 2. It clearly demonstrates that the inclusion of ECPs significantly improves the quality of the embedding description when using diffuse functions. Some of the excitation energies that should be above 6 eV collapse with PE to unphysically low values if diffuse basis functions allow the electrons to spill out into the MM system. This problem disappears if the ECPs are added. Some states are still

accurately described by the original PE scheme even if diffuse functions are included, as evidenced by the clustering of green points near the $y = x$ line. In such cases, including the ECPs has no negative impact on the quality of the description.

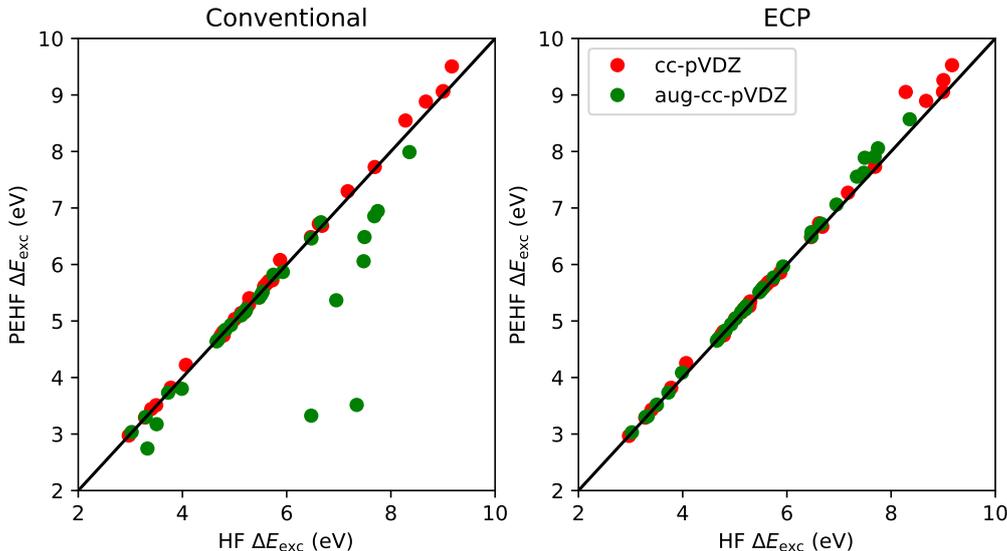


Figure 2: Supermolecular TDHF excitation energies are compared to excitation energies computed with a conventional PE-TDHF (left) or a PE(ECP)-TDHF (right) calculation.

Table 2 also shows the statistical measures for the deviations of PE- and PE(ECP)-ADC(2) excitation energies from the reference full-QM PNO-ADC(2) calculations in the test set. For the TZ basis set, the deviations between the full-QM reference and both the PE and the PE(ECP) results are reasonably small with mean absolute deviations (MAD) of 0.15 eV and 0.14 eV, respectively. After excluding the dispersion contribution (NoD-PNO-ADC(2) results), the MADs reduce to 0.11 eV and 0.10 eV. However, with the aTZ basis, the PE-ADC(2) results have an order of magnitude larger MAD of ≈ 1.0 eV. The large negative mean deviation (MSD) has the same magnitude, indicating again that large redshifts are responsible for most of the errors in the excitation energies. This dramatic failure is again due to ESO which we observe not only for non-local excitations (see Table 1), but also for local $\pi\pi^*$ excitations (NTOs are available in the SI).

In contrast to this, PE(ECP)-ADC(2) performs equally well for basis sets with and

without diffuse functions. In combination with the aTZ basis set, the MSD and MAD of PE(ECP)-ADC(2) calculations are 0.13 eV and 0.14 eV, respectively. Not a single example shows the huge red-shifts from ESO. On the contrary, the MSD between the PE(ECP) and the uncorrected PNO-ADC(2) results is positive, indicating that a significant part of the remaining error is probably due to the missing London dispersion contribution which in most cases lowers the excitation energy. This conclusion is also corroborated by the comparison between the PE(ECP)-ADC(2)/aTZ and the corrected NoD-PNO-ADC(2)/aTZ results for which the MSD is only 0.02 eV, i.e., there is no systematic red- or blue-shift. The MAD and maximum deviation (MAX) of the PE(ECP)-ADC(2)/aTZ from the NoD-PNO-ADC(2) results are, respectively, 0.07 eV and 0.37 eV. The maximum deviation is found for the lowest excitation in DMABN@CB[7]. A test calculation revealed that if the PE(ECP)-ADC(2) calculation is done with multipole moments and polarizabilities from a B3-LYP/DZ Loprop calculation the error decreases for this excitation to 0.04 eV. A systematic study of the performance of different parameterizations of the electrostatic embedding will be subject of forthcoming work.

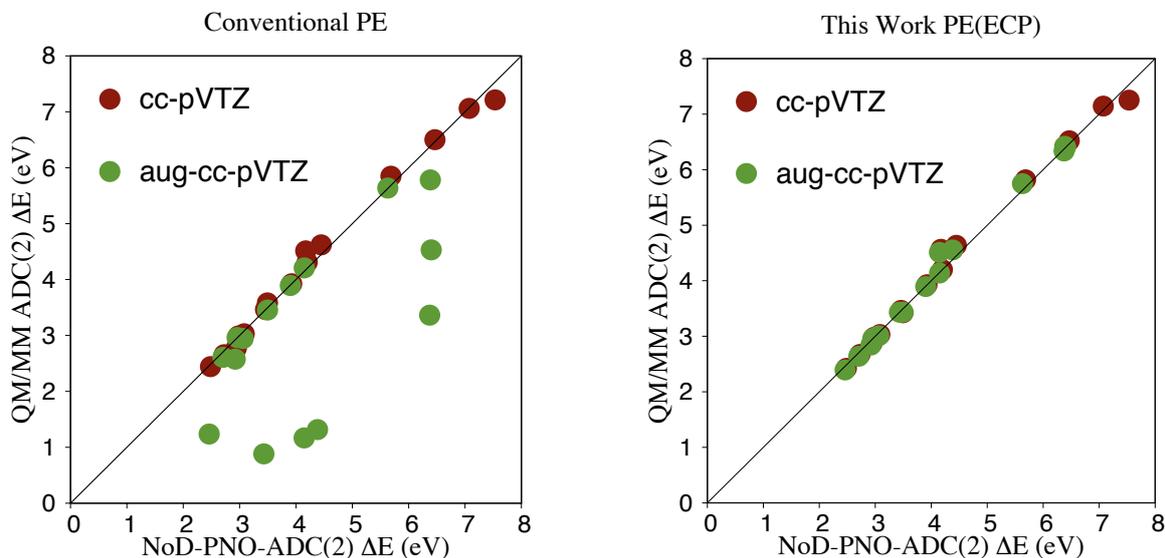


Figure 3: Comparison between (left) PE- and (right) PE(ECP)-ADC(2) QM/MM excitation energies with TZ and aTZ basis sets. The central diagonal line indicates the perfect match.

Figure 3 visualizes the distribution of the results from the QM/MM calculations versus the full QM reference values for the standard TZ and the augmented aTZ basis sets. The correlation factor R^2 between the full-QM and QM/MM results at the ADC(2) levels are with the TZ basis 0.99 with and without ECPs at the MM sites. When including diffuse basis functions, only the PE(ECP) scheme provides physically meaningful results with a correlation coefficient of $R^2 = 0.99$. The results for the PE scheme without ECPs show essentially no correlation ($R^2 = -0.17$) with the full-QM reference values. In the PE calculations without ECPs, the lack of Pauli repulsion causes ESO at least for half of the cases that were considered in this study.

A comparison between aTZ and TZ results in Figure 3 shows that the full QM PNO-ADC(2)/TZ excitation energies are slightly blue-shifted (on average 0.15 eV) relative to those for the aTZ basis. Typically, the higher the excitation energies are, the more important are the diffuse functions to obtain accurate results and the correct physical character of the excited state. Therefore, it is also for QM/MM calculations for systems in a condensed phase important to have the possibility to use large basis sets with diffuse functions without that results are affected by ESO.

In conclusion, our benchmark study on the excitation energies of several chromophores in different realistic molecular environments reveals that ESO is an important and common pitfall in QM/MM calculations on electronic excitations with basis sets that contain diffuse functions even if the QM system has no bonded interactions into the MM region. The PE(ECP) embedding with the above described simple ECPs provides a solution to this problem that is transferable and easy to implement and can be combined with most QM/MM force field without the need to adapt the parametrization of the electrostatic potential. It is also compatible with the ECPs of the pseudobond approach^{38,53–55} and the multi-centered valence-electron ECPs that have been introduced by Slavíček and Martinez³⁹ to treat covalent bonds at the QM/MM boundary. The PE(ECP) embedding gives a consistently robust performance, even when using diffuse basis functions. At the TDHF level it provides results

with an MAD of less than 0.1 eV from full QM calculations. After excluding the dispersion contribution from the full QM PNO-ADC(2) reference values, also the MAD of the PE(ECP)-ADC(2)/aTZ calculations are for the current test set as small as 0.07 eV, which shows that PE(ECP) parameterization is well-suited for this kind of QM/MM calculations. At this point, the most important remaining error in the current PE(ECP) scheme is for vertical excitation energies the missing effect of the dispersion interaction with the environment on the electronic wavefunction of the QM system. Further tuning of the ECPs, in particular, to fit them for specific MM systems seems at this point, at least for vertical excitation energies, not to be necessary. This makes polarizable QM/MM calculations with ECPs a readily accessible, reliable, and cost-effective solution to the electron spill-out problem in excited-state QM/MM calculations.

Acknowledgment

This work has been supported by the Deutsche Forschungsgemeinschaft within the Cluster of Excellence RESOLV (EXC 2033) (AMK) and the Priority Program SPP 1807 (HA 2588/10-1) (CH).

References

- (1) Warshel, A.; Levitt, M. Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme. J. Mol. Biol. **1976**, 103, 227–249.
- (2) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. Angew. Chem. Int. Ed. **2009**, 48, 1198–1229.
- (3) Brunk, E.; Rothlisberger, U. Mixed Quantum Mechanical/Molecular Mechanical Molec-

- ular Dynamics Simulations of Biological Systems in Ground and Electronically Excited States. Chem. Rev. **2015**, 115, 6217–6263.
- (4) List, N. H.; Olsen, J. M. H.; Kongsted, J. Excited States in Large Molecular Systems Through Polarizable Embedding. Phys. Chem. Chem. Phys. **2016**, 18, 20234–20250.
- (5) Morzan, U. N.; de Armiño, D. J. A.; Foglia, N. O.; Ramírez, F.; Lebrero, M. C. G.; Scherlis, D. A.; Estrin, D. A. Spectroscopy in Complex Environments from QM–MM Simulations. Chem. Rev. **2018**, 118, 4071–4113.
- (6) Olsen, J. M. H.; List, N. H.; Kristensen, K.; Kongsted, J. Accuracy of Protein Embedding Potentials: An Analysis in Terms of Electrostatic Potentials. J. Chem. Theory Comput. **2015**, 11, 1832–1842.
- (7) Nâbo, L. J.; Olsen, J. M. H.; Martínez, T. J.; Kongsted, J. The Quality of the Embedding Potential Is Decisive for Minimal Quantum Region Size in Embedding Calculations: The Case of the Green Fluorescent Protein. J. Chem. Theory Comput. **2017**, 13, 6230–6236.
- (8) Thompson, M. A. QM/MMpol: a consistent model for solute/solvent polarization. application to the aqueous solvation and spectroscopy of formaldehyde, acetaldehyde, and acetone. J. Phys. Chem. **1996**, 100, 14492–14507.
- (9) Kongsted, J.; Osted, A.; Mikkelsen, K. V.; Christiansen, O. The QM/MM approach for wavefunctions, energies and response functions within self-consistent field and coupled cluster theories. Mol. Phys. **2002**, 100, 1813–1828.
- (10) Olsen, J. M.; Aidas, K.; Kongsted, J. Excited States in Solution through Polarizable Embedding. J. Chem. Theory Comput. **2010**, 6, 3721–3734.
- (11) Olsen, J. M. H.; Kongsted, J. Adv. Quant. Chem.; Elsevier, 2011; Vol. 61; pp 107–143.

- (12) Caprasecca, S.; Jurinovich, S.; Viani, L.; Curutchet, C.; Mennucci, B. Geometry optimization in polarizable QM/MM models: the induced dipole formulation. J. Chem. Theory Comput. **2014**, 10, 1588–1598.
- (13) Menger, M. F. S. J.; Caprasecca, S.; Mennucci, B. Excited-state gradients in polarizable QM/MM models: an induced dipole formulation. J. Chem. Theory Comput. **2017**, 13, 3778–3786.
- (14) Lipparini, F.; Barone, V. Polarizable Force Fields and Polarizable Continuum Model: A Fluctuating Charges/Pcm Approach. 1. Theory and Implementation. J. Chem. Theory Comput. **2011**, 7, 3711–3724.
- (15) Mortier, W. J.; Genechten, K. V.; Gasteiger, J. Electronegativity Equalization: Application and Parametrization. J. Am. Chem. Soc. **1985**, 107, 829–835.
- (16) Chelli, R.; Procacci, P. A Transferable Polarizable Electrostatic Force Field for Molecular Mechanics Based on the Chemical Potential Equalization Principle. J. Chem. Phys. **2002**, 117, 9175–9189.
- (17) Anisimov, V. M.; Lamoureux, G.; Vorobyov, I. V.; Huang, N.; Roux, B.; MacKerell, A. D. Determination of Electrostatic Parameters for a Polarizable Force Field Based on the Classical Drude Oscillator. J. Chem. Theory Comput. **2005**, 1, 153–168.
- (18) Lopes, P. E.; Huang, J.; Shim, J.; Luo, Y.; Li, H.; Roux, B.; MacKerell Jr, A. D. Polarizable Force Field for Peptides and Proteins Based on the Classical Drude Oscillator. J. Chem. Theory Comput. **2013**, 9, 5430–5449.
- (19) Boulanger, E.; Thiel, W. Toward QM/MM Simulation of Enzymatic Reactions with the Drude Oscillator Polarizable Force Field. J. Chem. Theory Comput. **2014**, 10, 1795–1809.

- (20) Sneskov, K.; Schwabe, T.; Kongsted, J.; Christiansen, O. The Polarizable Embedding Coupled Cluster Method. J. Chem. Phys. **2011**, 134, 104108.
- (21) 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.
- (22) Khah, A. M.; Khani, S. K.; Hättig, C. Excited State Gradients for the QM/MM Polarizable Embedded Second-Order Algebraic Diagrammatic Construction for the Polarization Propagator PE-ADC(2). J. Chem. Theory Comput. **2018**, 14, 4640–4650.
- (23) Scheurer, M.; Herbst, M. F.; Reinholdt, P.; Olsen, J. M. H.; Dreuw, A.; Kongsted, J. QM/MM Methods for Biomolecular Systems. J. Chem. Theory Comput. **2018**, 14, 4870–4883.
- (24) Hedegård, E. D.; List, N. H.; Jensen, H. J. A.; Kongsted, J. The Multi-Configuration Self-Consistent Field Method within A Polarizable Embedded Framework. J. Chem. Phys. **2013**, 139, 044101.
- (25) Li, Q.; Mennucci, B.; Robb, M. A.; Blancafort, L.; Curutchet, C. Polarizable QM/MM Multiconfiguration Self-Consistent Field Approach with State-Specific Corrections: Environment Effects on Cytosine Absorption Spectrum. J. Chem. Theory Comput. **2015**, 11, 1674–1682.
- (26) Eriksen, J. J.; Sauer, S. P. A.; Mikkelsen, K. V.; Jensen, H. J. A.; Kongsted, J. On the Importance of Excited State Dynamic Response Electron Correlation in Polarizable Embedding Methods. J. Comp. Chem. **2012**, 33, 2012–2022.
- (27) Hedegård, E. D.; Reiher, M. Polarizable Embedding Density Matrix Renormalization Group. J. Chem. Theory Comput. **2016**, 12, 4242–4253.

- (28) Laio, A.; VandeVondele, J.; Rothlisberger, U. A Hamiltonian Electrostatic Coupling Scheme for Hybrid Car-Parrinello Molecular Dynamics Simulations. J. Chem. Phys. **2002**, 116, 6941–6947.
- (29) Giovannini, T.; Lafiosca, P.; Cappelli, C. A General Route to Include Pauli Repulsion and Quantum Dispersion Effects in QM/MM Approaches. J. Chem. Theory Comput. **2017**, 13, 4854–4870.
- (30) Laino, T.; Mohammed, F.; Laio, A.; Parrinello, M. An Efficient Real Space Multigrid QM/MM Electrostatic Coupling. J. Chem. Theory Comput. **2005**, 1, 1176–1184.
- (31) Gökcan, H.; Kratz, E.; Darden, T. A.; Piquemal, J.-P.; Cisnero, G. A. QM/MM Simulations with the Gaussian Electrostatic Model: A Density-based Polarizable Potential. J. Phys. Chem. Lett. **2018**, 9, 3062–3067.
- (32) Krauss, M.; Stevens, W. J. Effective Potentials in Molecular Quantum Chemistry. Ann. Rev. Phys. Chem. **1984**, 35, 357–385.
- (33) Hellmann, H. A New Approximation Method in the Problem of Many Electrons. J. Chem. Phys. **1935**, 3, 61.
- (34) Gombás, P. Über die metallische Bindung. Z. Phys. **1935**, 94, 473–488.
- (35) Dolg, M.; Cao, X. Relativistic Pseudopotentials: Their Development and Scope of Applications. Chem Rev. **2012**, 112, 403–480.
- (36) Luaña, V.; Pueyo, L. Simulation of ionic crystals: The ab initio perturbed-ion method and application to alkali hydrides and halides. Phys. Rev. B **2012**, 41, 3800–3814.
- (37) von Lilienfeld, O. A.; Tavernelli, I.; Rothlisberger, U. Variational Optimization of Effective Atom Centered Potentials for Molecular Properties. J. Chem. Phys. **2005**, 122, 014113.

- (38) Zhang, Y.; Lee, T.-S.; Yang, W. A pseudobond approach to combining quantum mechanical and molecular mechanical methods. J. Chem. Phys. **1999**, 110, 46–54.
- (39) Slavíček, P.; Martínez, T. J. Multicentered Valence Electron Effective Potentials: A Solution to The Link Atom Problem for Ground And Excited Electronic States. J. Chem. Phys. **2006**, 124, 084107.
- (40) DiLabio, G. A.; Hurely, M. M.; Christiansen, P. A. Simple One-Electron Quantum Capping Potentials for Use in Hybrid QM/MM Studies of Biological Molecules. J. Chem. Phys. **2002**, 116, 9578.
- (41) Igel-Mann, G.; Stoll, H.; Preuss, H. Pseudopotentials for main group elements (IIIa through VIIa). Mol. Phys. **1988**, 65, 1321–1328.
- (42) Helmich, B.; Hättig, C. A Pair Natural Orbital Implementation of The Coupled Cluster Model CC2 for Excitation Energies. J. Chem. Phys. **2013**, 139, 084114.
- (43) Helmich, B.; Hättig, C. A Pair Natural Orbital Based Implementation of ADC(2)-x: Perspectives and Challenges for Response Methods for Singly And Doubly Excited States in Large Molecules. Comp. Theor. Chem. **2014**, 1040–1041, 35–44.
- (44) Development version of TURBOMOLE, 2019, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989–2007, TURBOMOLE GmbH, since 2007; available from <http://www.turbomole.com>.
- (45) Reinholdt, P.; Kongsted, J.; Olsen, J. M. H. Polarizable Density Embedding: A Solution to the Electron Spill-Out Problem in Multiscale Modelling. J. Phys. Chem. Lett. **2017**, 8, 5949–5958.
- (46) Silver-Junior, M. R.; Sauer, S. P.; Schreiber, M.; Thiel, W. Basis set effects on coupled cluster benchmarks of electronically excited states: CC3, CCSDR(3) and CC2. Mol. Phys. **2010**, 108, 453–465.

- (47) Nåbo, L. J.; Olsen, J. M. H.; List, N. H.; Solanko, L. M.; Wüstner, D.; Kongsted, J. Embedding beyond electrostatics—The role of wave function confinement. J. Chem. Phys. **2016**, 145, 104102.
- (48) Reinholdt, P.; Nørby, M. S.; Kongsted, J. Modeling of Magnetic Circular Dichroism and UV/Vis Absorption Spectra Using Fluctuating Charges or Polarizable Embedding within a Resonant-Convergent Response Theory Formalism. J. Chem. Theory Comput. **2018**, 14, 6391–6404.
- (49) Schwabe, T.; Olsen, J. M. H.; Sneskov, K.; Kongsted, J.; Christiansen, O. Solvation Effects on Electronic Transitions: Exploring the Performance of Advanced Solvent Potentials in Polarizable Embedding Calculations. J. Chem. Theory Comput. **2011**, 7, 2209–2217.
- (50) Knizia, G. Intrinsic Atomic Orbitals: An Unbiased Bridge between Quantum Theory and Chemical Concepts. J. Chem. Theory Comput. **2013**, 9, 4834–4843.
- (51) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. J. Chem. Phys. **2010**, 132, 154104.
- (52) Schröder, H.; Schwabe, T. Efficient Determination of Accurate Atomic Polarizabilities for Polarizable Embedding Calculations. J. Comp. Chem. **2016**, 37, 2052–2059.
- (53) Zhang, Y. Improved pseudobonds for combined ab initio quantum mechanical/molecular mechanical methods. J. Chem. Phys. **2005**, 122, 024114.
- (54) Parks, J. M.; Hu, H.; Cohen, A. J.; Yang, W. A pseudobond parametrization for improved electrostatics in quantum mechanical/molecular mechanical simulations of enzymes. J. Chem. Phys. **2008**, 129, 154106.

- (55) Chaudret, R.; Parks, J. M.; Yang, W. Pseudobond parameters for QM/MM studies involving nucleosides, nucleotides, and their analogs. J. Chem. Phys. **2013**, 138, 045102.

9 Berenil in DNA-Minor-Groove Confinement

In previous chapters 5 and 6, the vacuum excited-state relaxation mechanism of berenil has been clarified in details using the wavefunction-based method ADC(2). It is understood that in vacuum, the S_1 relaxation mechanism of berenil has a bicycle-pedal character, which is a volume-conserving mechanism. However, the weak response to changes of the viscosity and having a volume-conserving relaxation mechanism does not mean that the molecule is insensitive to the environment. The bicycle-pedal type photoisomerization mechanism has also been observed for photo-active yellow protein, which has been controlled by mutating the contacting protein residues. [137] Moreover, the polarity of the environment, which has a non-isotropic effect in case of biomolecular embedding and DNA confinement, may have an influence on excited-state relaxation mechanisms even for the volume-conserving cases. [134, 136]

In this chapter, we aim to study the S_1 relaxation mechanism of berenil in DNA-minor-groove and solvent molecules confinement using the polarizable QM/MM ADC(2) method that we have already developed and implemented for this project, cmp. chapter 7. At this level, the polarization of the environment, not only on the ground state but also with respect to the transition densities of the QM subsystem, has been taken into account. The implementation of excited state energy and gradients allows us to calculate the IRC pathways of berenil in the state of binding to the minor groove of DNA.

In the current work, the initial configuration of berenil, DNA and solvent water molecules

are taken from a classical molecular dynamics simulation. Furthermore, the deexcitation mechanism of berenil in DNA-minor-groove has been studied with three different embedding models to distinguish between the role of DNA-minor-groove as a confined medium and also water molecules of the first solvation shell of berenil which are interacting with berenil via hydrogen-bonds.

9.1 Computational Details

9.1.1 Molecular Dynamics Simulation

The initial structure of berenil in complexation to an AT-rich DNA sequence has been obtained from the X-ray structure that was reported with the PDB code 2gvr (berenil-D(CGCGAATTCGCG)2). All the crystallographic water molecules, co-solvents and counterions were removed and the berenil molecule was saturated by hydrogen atoms to form the well-known dication protonation state. There are two possibilities for the hydrogen saturation of the central linear triazene functional group of berenil. Since the two tautomeric forms are mirror images, only one of them is considered in this study. A protonation state of DNA was obtained with AmberTools. [144]. The DNA-minor-groove bounded structure of berenil, berenil@DNA, has then been solvated in a cubic water box with at least 20 nm distance between the DNA and the box edges. The system was neutralized with the Na⁺ cation, and then, Na⁺ and Cl⁻ ions were added to reproduce the experimental salt concentration, cmp. chapter 5. In the molecular dynamics calculations, the ParmBSC1 parameters [145] of the AMBER force field [146] were used for DNA and the GAFF param-

eters [147] with MRESP point charges [148] (fitted to the B3-LYP/def2-TZVP [149, 150] electrostatic potential) were used for berenil. In this simulation TIP3P water models were used. The following pre-equilibration processes have been carried out: First, a steepest descent minimization for 5000 steps to remove unphysical close contacts, second, an NVT molecular dynamics run for 200 ps with the time steps of 2 fs at ($T = 300K$) while the heavy atoms of DNA and berenil were restrained to their XYZ coordinates, third, an NPT simulation ($T = 300K$ and $P = 1$ bar) with keeping the heavy atoms of DNA under the same restrain as the previous step, fourth, an NPT simulation with removing all the geometrical restrains for 200 ps with 2 fs time step, at $T = 300K$ and $P = 1.0$ bar. At this stage the box size was changed to the averaged value of the latter NPT simulation to obtain the density of bulk water in the NVT simulation. A 20 ns long NVT production run to generate the final trajectory with the Velocity-rescaling thermostat and the time steps of 2 fs were carried out. In the previous step, the shake algorithm was applied to constrain all bonds involving hydrogen atoms. Then, a representative configuration of the trajectory has been used for generating the embedding potential for the QM/MM calculations.

9.1.2 Calculation of Polarizable Potential

The MM subsystem of the QM/MM calculation is partitioned into three layers. First, a polarizable layer includes the DNA and all water and ions within 5 Å of berenil. Second, a multipole layer comprises the water molecules and ions within 15 Å of berenil. Third a point charge layer consists of water molecules and ions within 30 Å of berenil. A graphical representation of the polarizable layer and the total system is given in Fig. 3. The polarizabilities

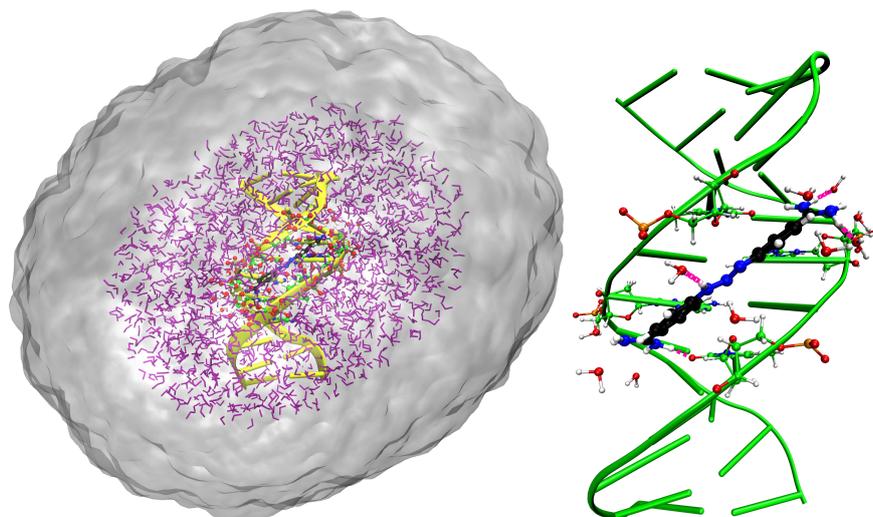


Figure 3: The total system (left) consists of the QM region (berenil) and three MM layers: **(i)** The polarizable embedding layer including DNA (yellow) and all water and ions within a 5 Å of berenil. **(ii)** The multipole embedding layer including water and ions within 15 Å of berenil (magenta). **(iii)** The point charge embedding layer including water molecules and ions within 30 Å of berenil (transparent sphere). The polarizable layer of the MM subsystem (right) with atom colors and carbon atoms of the DNA in green.

and multipole moments of DNA have been obtained via the MFCC [151,152] fragmentation procedure using the pyFraMe package [153], which calls Dalton program [154] for the Lo-prop [155] distributed multipoles and polarizabilities calculations. The polarizabilities and multipoles were calculated at the DFT/B3LYP [149] level with the ANO-recontracted version of aug-cc-pVDZ basis set. [85, 156] For the point charge embedding layer the solvent embedding potential has been used. [157]

9.1.3 QM/MM Calculations

All the QM/MM calculations have been performed for a configuration extracted from the molecular dynamics trajectory. The ground- and excited-state geometry optimizations were carried out at the polarizable PE(ECP)-MP2 and PE(ECP)-ADC(2) levels, respectively (chapter 7), with the SVP basis set [158] and the corresponding auxiliary basis set [159] for the RI- approximation [27]. The vdW parameters for the QM/MM calculations that describe the vdW interactions between the QM and the MM subsystems were adjusted to those of the molecular dynamics calculation. The atoms of the polarizable MM layer were supplemented by the simple ECPs that have been parametrized for the QM/MM calculations to avoid electron spill out, cmp. chapter 8. In both ground- and excited-state geometry optimizations only the QM subsystem (berenil) has been relaxed and all the other atoms were fixed. In the intrinsic reaction coordinate (IRC) calculations, the step size of 1 fs has been chosen. [160] All the QM/MM calculations were done with a local version of the `ricc2` module of the TURBOMOLE program package. [35]

9.1.4 Models of berenil@DNA complex

To study the ultrafast excited-state relaxation processes of the berenil@DNA complex, we focused on three models for the embedding environment of berenil. The first embedding model (EM1) serves as an example of the extreme confinement situation. In this model, all the surrounding molecular environment of berenil are fixed in the space and, thus, the berenil molecule is caged in a fully confined environment. By studying the EM1,

we aimed to examine whether the volume-conserving relaxation motions of berenil can be suppressed by DNA-minor-groove confinement and the surrounding water molecules, or the proposed relaxation processes (cmp. chapter 6) [161] are ultimately independent from the changes in the molecular environment. The effects of DNA embedding alone as a confinement (embedding model 2, EM2) into the excited-state relaxation mechanism have been studied with the EM2 by removing all the water molecules and ions from the environment. The third model (embedding model 3, EM3) scrutinizes on the influence of the first solvation shell relaxation on the excited-state relaxation. In addition to berenil, three water molecules that have a hydrogen-bond with berenil were included in the flexible QM region while the rest of the system including DNA, other water molecules and ions were kept frozen.

9.2 Results and Discussions

The PE-MP2/SVP optimized ground-state geometry of berenil in the EM1 and EM2 are given in Fig. 4. In comparison to the most stable conformer of berenil in vacuum, the *E-azo-s-trans* (ET) conformer (cmp. chapter 6), the relaxed geometry of berenil within both type of confinements do not change substantially and retain its almost planar structure. In the EM1, berenil has 6 hydrogen-bonds with the neighboring molecules and in particular there is a hydrogen-bond between the hydrogen of the triazene bridge and one of the water molecules of the environment. In table 1, the PE-ADC(2)/aug-cc-pVDZ (aDZ) vertical excitation energies, oscillator strengths and characters of excitation for the two lowest

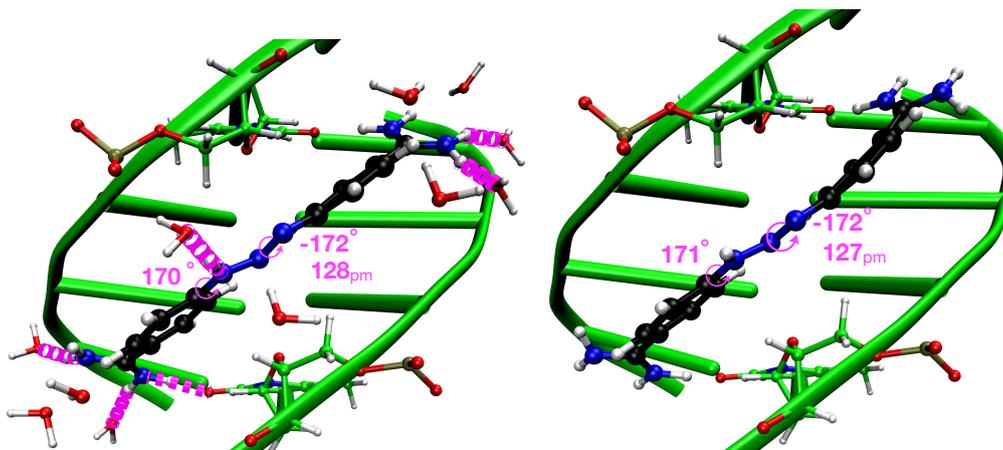
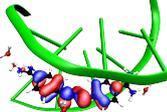
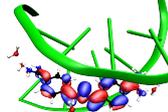
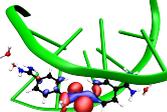
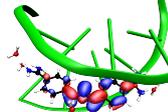
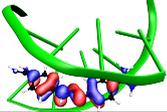
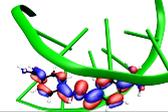
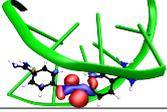
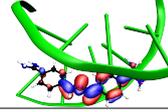


Figure 4: The PE-MP2/SVP ground-state relaxed geometries of berenil in the (left) EM1 and (right) EM2.

singlet excited states of berenil are compared within the EM1 and EM2 models. [162] The characters of excitation are the same for both models. The lowest electronic excitation, S_1 , has a $\pi\pi^*$ character while the second electronic excitation, S_2 , is a dark $n\pi^*$ transition. The excitation energy for the bright $\pi\pi^*$ transition of the EM1 is at 3.25 eV, which is in an excellent agreement with the 3.24 eV maximum of the experimental berenil@DNA steady-state absorption spectrum, cmp. chapter 5. By comparing the calculated $\pi\pi^*$ excitation energies of the EM1 and EM2, one can identify a 0.16 eV solvatochromic red-shift, due to the interactions with the surrounding water and ions. Based on the agreement between the EM1 and the experiment, and also the similar excitation characters, and the relatively small differences between the EM1 and EM2, here we argue that both models are capturing the main features of the $\pi\pi^*$ excitation and thus we can use these models to study the role of confinements to the relaxation processes.

The *in vacuo* deexcitation mechanism on the S_1 potential energy surface of berenil has

Table 1: Vertical excitation energies (ΔE) and oscillator strengths (f) of EM1 and EM2 for S_1 and S_2 at the PE-ADC(2)/aDZ level with PE-MP2/SVP optimized geometry. Excitation characters (Char.), first occupied (OCC) and virtual (VIR) pair of NTOS (isosurface ± 0.05) are given.

Emb. Model	Char.	ΔE (eV)	f	OCC	VIR	Contrib.
1	$\pi\pi^*$	3.25	1.00			96.8 %
1	$n\pi^*$	3.83	0.08			98.2 %
2	$\pi\pi^*$	3.41	1.02			96.1 %
2	$n\pi^*$	3.92	0.09			98.1 %

been studied in chapter 5 by means of intrinsic reaction coordinate (IRC) analysis. The proposed two-step deexcitation mechanism, first elongation of the N=N bond followed by a volume-conserving bicycle-pedal motion of the triazine group, is in agreement with the time-resolved fluorescence up-conversion measurements of the excited-state deexcitation process in solution and also in the complexation to AT-rich DNA, cmp. chapter 5. In Fig. 5, the IRC of the S_1 relaxation process of berenil@DNA is mapped into the linearly independent intramolecular coordinates of the triazine functional group of berenil. The analysis of the IRC path of berenil@DNA unveils that in the early stage of the deexcitation process, the δ_1 and δ_4 bonds are contracted by 4 and 5 pm, respectively, and the δ_3 bond is stretched by ≈ 3 pm. In the next stage of the relaxation process, bond torsions of the triazine evolve with a mechanism that has predominantly a bicycle-pedal type character.

In vacuum (cmp. chapters 5 and 6), due to the approximate planarity of the ET conformer, the IRC calculation that is started from the Franck-Condon point of the S_1 surface, FC- S_1 , leads to a transition state on the S_1 potential energy surface. However, for berenil@DNA the IRC calculations continue to a S_1 minimum without meeting a transition state.

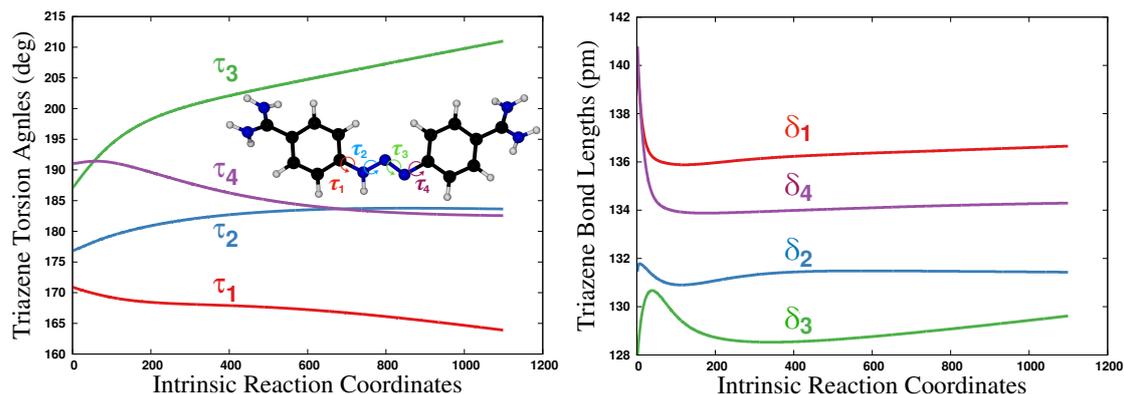


Figure 5: (Left) The definition and evaluation of the triazene group torsions along the IRC pathway from the FC- S_1 geometry of berenil@DNA within the EM1 calculated at PE-ADC(2)/SVP level of theory. (Right) Evaluation of triazene bond lengths along the S_1 IRC of berenil@DNA.

The fully optimized S_1 excited-state structures of berenil@DNA in the two confinement models, EM1 and EM2, are compared in Fig. 6. Within the EM2, berenil is relaxed to a S_1 minimum with a half-way twisted N=N torsion bond ($\tau_3 = -92^\circ$) and a rotated τ_1 angle that facilitates a berenil-DNA hydrogen-bond formation, cmp. Fig. 6(right). The S_1 minimum geometry of berenil in the EM2 (the model without water molecules) is similar to the structure of the S_1 minimum in vacuum, cmp. chapter 6. Hence, it can be concluded that the berenil S_1 relaxation process is insensitive to the DNA-confinement. In contrast, the S_1 optimized geometry of berenil inside the EM1 is considerably different from the minimum geometry of the S_1 state in vacuum, or in a polarizable continuum: the τ_3 angle

is only rotated by 60° and the hydrogen-bond structure of berenil is similar to the ground state.

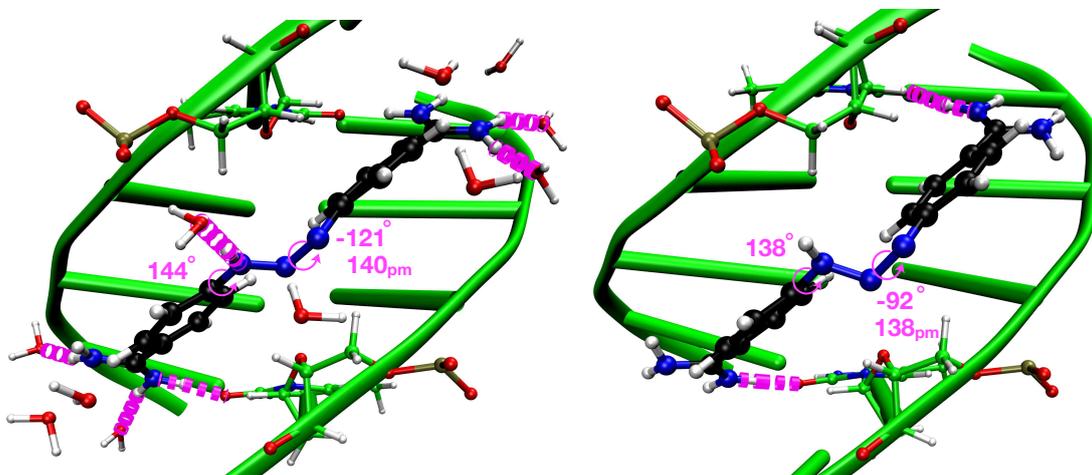


Figure 6: PE-ADC(2)/SVP relaxed geometries of berenil in the lowest singlet electronically excited state S_1 within two environment model, EM1 (left) and EM2 (right), cmp. caption of Fig. 4.

In Fig. 7, the ground- and excited-state optimized structure of berenil are compared by overlaying the structures for the (left) EM1 and (right) EM2. In the same figure, looking at the left overlaid structures and compare it to the right ones make it clear that the relaxation process in both cases are similar to the vacuum relaxation path, which are spatially localized on the triazene group. In other words, the S_1 relaxation of bereni@DNA is a volume-conserving process. However, the relaxation process in the presence of the water molecules network (the EM1) seems to be incomplete. In consequence of the incomplete intramolecular motion under the extremely confined condition (water molecules are fixed in the EM1), the energy gap between the S_0 and S_1 at the S_1 equilibrium geometry is with $25.7\text{ kcal}\cdot\text{mol}^{-1}$ ($\approx 9000\text{ cm}^{-1}$) larger than in vacuum and aqueous solution ($\approx 2800\text{ cm}^{-1}$); but that is still in the range that allows for a non-radiative thermal transition to S_0 .

The intramolecular relaxation process can be completed by the reorganization of the water

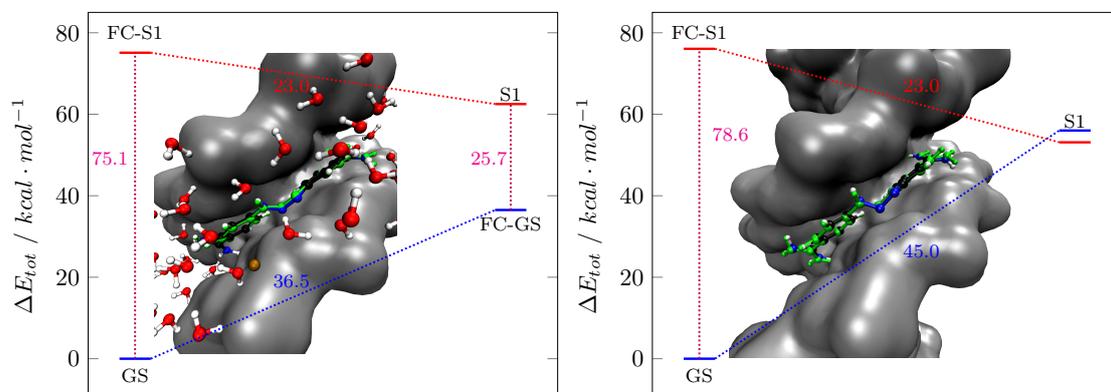


Figure 7: Ground- and excited-state relaxation paths of berenil@DNA in the (left) EM1 and (right) EM2 at PE-ADC(2)/aDZ//PE-ADC(2)/SVP level of theory. Color codes: (solid line) stationary and Franck-Condon points. (dotted line) geometrical changes. (blue) electronic ground-state PES, (red) lowest electronic excited state, (purple) vertical energy gaps. The geometry of the S_1 stationary point (green) is superimposed to the geometry of S_1 Franck-Condon point FC- S_1 (atoms colors).

molecules in the first solvation shell. In the EM3, Fig. 8, three water molecules that are close to berenil were included in the flexible QM region. The latter degree of flexibility allows the azo-bond N=N torsion angle to change to -107° and reduces the energy gap between S_0 and S_1 at the S_1 equilibrium geometry to $7.0 \text{ kcal}\cdot\text{mol}^{-1}$ ($\approx 2400 \text{ cm}^{-1}$) according to the supermolecular PE(ECP)-ADC(2)/aDZ/SV calculation, cmp. Fig. 7. Consequently, the excited state relaxation of berenil@DNA is not totally independent from the solvation shell relaxation, although it is volume-conserving and, therefore, the non-radiative deexcitation mechanism of berenil@DNA can be completed also in the presence of the first solvation shell water molecules. Furthermore, it is known that the dynamics of biological and minor-groove waters are slower than the bulk water; [163–165] hence, as it is observed in the intensity of

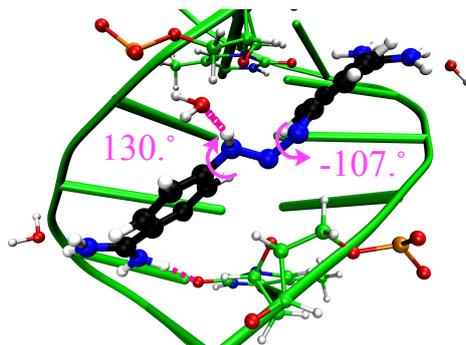


Figure 8: The S_1 optimized structure of a supermolecular complex of berenil and three water molecules within the EM1 calculated at the QM/MM PE(ECP)-ADC(2)/SVP/SV level of theory. The SVP and SV basis sets are used in the QM/MM calculations for berenil and water molecules, respectively. For brevity, the MM water molecules and ions are not shown.

the steady-state fluorescence spectrum, cmp. chapter 5, one may expect a longer excited state life-time and cooling for berenil@DNA compared to berenil in the bulk water.

9.3 Conclusion and Outlook on the Berenil@DNA Project

In this study, the excited state relaxation mechanism of berenil in the DNA-confinement has been investigated using polarizable QM/MM PE-ADC(2) calculations. A volume-conserving bicycle-pedal type relaxation mechanism is observed in both EM1 and EM2. However, for the EM1 the volume-conserving bicycle-pedal motion of berenil couples through the hydrogen-bonds to the water molecules in the first solvation shell. Thus, in this extremely confined model with fixed DNA and surrounding water molecules, the relaxation motion of berenil is moderately hindered by the hydrogen-bounded water molecules of the first solvation shell. From the EM1, we learned that even in the case of volume-

conserving motions, the excited state relaxation process and the relative energy levels of the ground- and excited-state potential energy surfaces are entangled with the local solvent structure via the specific intermolecular interactions. The EM2 clarifies that the volume-conserving relaxation process can be completed in the DNA-minor-groove confinement *i.e.*, the excited-state intramolecular relaxation motion of berenil proceeds to a region where S_1 and S_0 become energetically close. This study can be further advanced by sampling more structures from the molecular dynamics trajectory and the calculation of supermolecular IRC pathways. Furthermore, the bulk solvent relaxation can be include via a polarizable continuum model. In addition, the ground state photo-isomerization pathways to ET and ZT conformers can be studied by both EM1 and EM2.

10 Summary and Conclusion

During my PhD project, I have used computational chemistry methods to investigate the atomistic details of the two-phase excited-state relaxation process of berenil. For this, the *ab initio* second-order method ADC(2) has been employed for the calculation of intrinsic reaction coordinates of berenil. According to the analysis that has been presented in chapter 5, the first phase of the excited-state relaxation starts from the Franck-Condon point of the lowest electronic excited state S_1 . This relaxation mechanism is initially dominated by the *azo*-bond elongation, which leads to a transition state on the S_1 potential energy surface. In the second phase of the relaxation process, the twisting of the triazene moiety via a volume-conserving bicycle-pedal type motion leads to a conical intersection between the ground and the S_1 excited state. This twist can happen clock or anti-clock wise. Comprehensive studies of berenil ground- and excited-state photoisomerization pathways disclosed that from the Franck-Condon region of the ground state, where the intersection between the S_1 and the ground state potential energy surfaces located, the berenil molecule can be photoisomerized non-radiatively to two different *E-azo* and *Z-azo* ground-state conformers via bicycle-pedal and the Hula-twist mechanisms, respectively. The proposed ground- and excited-state relaxation mechanisms of the aromatic triazene berenil have been confirmed by comparing experimental and the theoretical results for ground- and excited-state time-resolved transient absorption data where the computational results were obtained with *ab-initio* calculations at the ADC(2) level. The latter findings together with

the experimental works of our collaborators explain the ultra-fast relaxation of berenil in viscous solvents and in complexation to bio-molecular host molecules. To come to the latter conclusions the results of theoretical modeling and the interpretations of experimental data have been analyzed together, which provides an example of complex problems that can be solved by means of the second-order quantum chemistry method ADC(2) for medium-sized single-reference molecules. Our study on the excited-state relaxation pathways of berenil did not remain limited to *in vacuo* quantum chemistry simulations. With the aid of the polarizable QM/MM ADC(2) method that we developed and implemented in TURBOMOLE to advance this project, the role of the DNA-minor-groove confinement and specific solvent interactions on the volume-conserving excited-state relaxation pathways of berenil have been studied. It could be shown that the non-radiative, volume-conserving bicycle-pedal type deexcitation mechanism of berenil@DNA from the Franck-Condon point of the lowest electronic excited-state S_1 minimum is influenced by the water molecules of the first solvation shell via hydrogen-bonding interactions and, thus, the relaxation mechanism towards the conical intersection of S_1 and the ground-state is coupled to the relaxation of water molecules in the first solvation shell. This is an example of how secondary interactions can modify a volume-conserving relaxation mechanism.

To the best of my knowledge, our post-SCF implementation of PE-ADC(2) excited-state energies and gradients is the first implementation of polarizable QM/MM excited-state gradients for a correlated wavefunction method. In this model, the important polarization-correlation couplings are approximated based on a perturbation expansion of the polarization operator, which in combination with the post-SCF reaction field scheme provide

a versatile approach for implementation of excited-state gradients and properties with a continuum or atomistic polarizable embedding, cmp. chapter 7. The accuracy of the PE-ADC(2) method in reproducing the supermolecular ADC(2) excited-state optimized geometries and fluorescent emission energies have been examined for two supramolecular complexes of DMABN@CB[7] and DBO@CB[7]. From the calculations on the locally-excited S_1 relaxed geometry of the DMABN@CB[7] complex, it was demonstrated that the QM/MM ADC(2) excited-state geometry optimization is a cost-effective alternative for the much more costly super-molecular full QM calculations. The PE-ADC(2) method has been implemented in the `ricc2` module of TURBOMOLE program package, based on an efficient implementation of PERI-CC2 excitation energies and response properties, which in the near future will be available for users of the TURBOMOLE program. It is expected that the efficient implementation of PE-ADC(2) method can be useful to advance the simulations of excited-state dynamics of small- and medium-sized organic molecules in the complex molecular environment. Furthermore, the PE-ADC(2) calculations on the DBO@CB[7] complex with basis sets that are augmented with diffuse functions elucidated the significance of electron spill-out that occurs due to having no effect of the Pauli repulsion by the environment on the electronic wavefunction in the QM/MM calculations. In the PE-ADC(2) excitation energy calculations of DBO@CB[7] with the aug-cc-pVDZ basis set, the electron spill-out issue manifests itself with large unphysical red shifts of excitation energy and localization of virtual NTOs and electron density in the MM system. The latter observation has led us to develop the polarizable embedding with effective core potentials PE(ECP) that can avoid electron spill-out issue in the QM/MM excited-state

calculations with basis sets that include diffuse basis functions.

This line of research has been continued by parametrizing a set of atom-specific and transferable ECPs that are tailored for QM/MM calculations to mimic the Pauli repulsion of all electrons at each MM atom. Benchmarking the PE(ECP)-ADC(2) and the PE-ADC(2) models against the full-QM PNO-ADC(2) supermolecular calculations for various type of chromophores in different molecular environments, *e.g.*, water, organic solvents, proteins and molecular containers, indicates that the PE(ECP) model is a robust solution to the electron spill-out problem. The result was confirmed by benchmarking the PE(ECP)-TDHF and PE-TDHF methods against full-QM TDHF calculations for the same systems using the same parametrized ECPs in collaboration with the theoretical chemistry group of Professor Jacob Kongsted. Moreover, my idea to select the PNO-ADC(2) method as the full-QM reference gave us the opportunity to study the effects of Pauli repulsion and dispersion separately; and hence, make a point about the role of dispersion interactions for excitation energies. At the second-order MP2 and ADC(2) levels of theory, the dispersion effects are included as part of the electron correlation, in contrast to the Hartree-Fock level that has no dispersion contribution. Also, it is known from the perturbation analysis of intermolecular interactions that the dispersion interaction is the next important type of intermolecular interactions after the electrostatics and the polarization ones. [86] For the first time, decomposition of the full-QM PNO-ADC(2) excitation energies to isolate the share of dispersion interaction and comparison of the correlated excitation energies with the PE(ECP)-ADC(2) excitation energies demonstrates that at this level of QM/MM modeling, the most important missing contribution is the dispersion interaction on the

electronic wavefunction. Thus, we can suggest that the next step towards having a more accurate QM/MM model is the inclusion of the dispersion effect.

In conclusion, the implementation of PE-ADC(2) excited-state gradients and the more robust PE(ECP) framework to have a description of Pauli repulsion in the QM/MM calculations pave the way to expand the application of second-order methods to the calculation of excited-state relaxation mechanisms in complex molecular environments. The exemplary work on relaxation pathways of berenil in the solvent and DNA confinements provides an insight about the influence of solvent relaxation into the volume-conserving deexcitation mechanism of berenil.

References

- [1] U. N. Morzan, D. J. A. do Armiño, N. O. Foglia, F. Ramírez, M. C. G. Lebrero, D. A. Scherlis and D. A. Estrin, *Chem. Rev.* **118**, 4071 (2018).
- [2] F. Segatta, L. Cupellini, M. Garavelli and B. Mennucci, *Chem. Rev.* **119**, 9361 (2019).
- [3] C. Curutchet and B. Mennucci, *Chem. Rev.* **117**, 294 (2017).
- [4] K. Hüll, J. Morstein and D. Trauner, *Chem. Rev.* **118**, 10710 (2018).
- [5] W. A. Velema, W. Szymanski and B. L. Feringa, *J. Am. Chem. Soc.* **136**, 2178 (2014).
- [6] A. S. Klymchenko, *Acc. Chem. Res.* **50**, 366 (2017).
- [7] G. Cabré, A. Garrido-Charles, M. Moreno, M. Bosch, M. P. de-la Riva, M. Krieg, M. Gascón-Moya, N. Camerero, R. Gelabert, J. M. Lluch, F. Busqué, J. Hernando, P. Gorostiza, and R. Alibés, *Nature Communications.* **10**, 907 (2019).
- [8] M. Dong, A. Babalhavaeji, S. Samanta, A. A. Beharry and G. A. Woolley, *Acc. Chem. Res.* **48**, 2662 (2015).
- [9] G. Haberhauer, *Chem. Eur. J.* **23**, 9288 (2017).
- [10] P. Reinholdt, S. Wind, D. Wüstner and J. Kongsted, *J. Phys. Chem. A.* **8**, 7313 (2019).

- [11] T. Schultz, J. Quenneville, B. Levine, A. Toniolo, T. J. Martínez, S. Lochbrunner, M. Schmitt, J. P. Shaffer, M. Z. Zgierski, and A. Stolow, *J. Am. Chem. Soc.* **125**, 8098 (2003).
- [12] H. M. D. Bandara and D. C. Burdette, *Chem. Soc. Rev.* **41**, 1809 (2012).
- [13] M. Quick, A. L. Dobryakov, M. Gerecke, C. Richter, F. Berndt, I. N. Ioffe, A. A. Granovsky, R. Mahrwald, N. P. Ernsting, and S. A. Kovalenko, *J. Phys. Chem. B.* **118**, 8756 (2014).
- [14] M. Böckmann, N. L. Doltsinis and D. Marx, *J. Chem. Phys.* **137**, 22A505 (2012).
- [15] M. Böckmann, S. Braun, N. L. Doltsinis and D. Marx, *J. Chem. Phys.* **139**, 084108 (2013).
- [16] L. Laprell, K. Hüll, P. Stawski, C. Schön, S. Michalakis, M. Biel, M. P. Sumser and D. Trauner, *ACS Chem. Neurosci.* **7**, 15 (2016).
- [17] T. W. Gampbell and B. F. Day, *Chem. Rev.* **48**, 299 (1951).
- [18] B. A. Newton, Berenil: A trypanocide with selective activity against extranuclear DNA, in *Antibiotics III - Mechanism of Action of Antimicrobial and Antitumor Agents* (J. W. Corcoran, F. E. Hahn, J. F. Snell, and K. L. Arora, eds.), pp. 34–47. Springer Heidelberg Berlin, 1975.
- [19] C. Dugave and L. Demange, *Chem. Rev.* **103**, 2475 (2003).
- [20] S. Patil and A. Bugarin, *Eur. J. Org. Chem.* **2016**, 860 (2015).

- [21] M. Barra and N. Chen, *J. Org. Chem.* **65**, 5739 (2000).
- [22] P. Zhao, Z. Zhang, P. J. Wang and D. S. Liu, *Physica B.* **404**, 3462 (2009).
- [23] J. Baro, D. Dudek, K. Luther and J. Troe, *Ber. Bunsenges. Phys. Chem.* **87**, 1155 (1983).
- [24] D. S. Pilch, M. A. Kirolos, X. Liu, E. Plum and K. J. Breslauer, *Biochemistry.* **34**, 9962 (1995).
- [25] A. M. Khah. A computational study of the first electronic excited state of berenil in solutions. Master's thesis, Department of Chemistry and Biochemistry, Ruhr University of Bochum, Universität Straße 150, Bochum, July 2016.
- [26] A. Szabo and N. S. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory (Dover Books on Chemistry)*. Dover Publications, 1996.
- [27] C. Hättig and F. Weigend, *J. Chem. Phys.* **113**, 5154 (2000).
- [28] J. Schirmer, *Phys. Rev. A.* **26**, 2395 (1981).
- [29] A. Warshel and M. Levitt, *J. Mol. Biol.* **103**, 227 (1976).
- [30] H. M. Senn and W. Thiel, *Angew. Chem. Int. Ed.* **48**, 1198 (2009).
- [31] E. Brunk and U. Rothlisberger, *Chem. Rev.* **115**, 6217 (2015).
- [32] N. H. List, J. M. H. Olsen and J. Kongsted, *Phys. Chem. Chem. Phys.* **18**, 20234 (2016).

- [33] J. M. H. Olsen and J. Kongsted, *Adv. Quant. Chem.* **61**, 107 (2011).
- [34] T. Schwabe, K. Sneskov, J. M. H. Olsen, J. Kongsted, O. Christiansen and C. Hättig, *J. Chem. Theory Comput.* **8**, 3274 (2012).
- [35] TURBOMOLE V7.3 2018 and V7.4 2019, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989–2007, TURBOMOLE GmbH, since 2007; available from <http://www.turbomole.com>.
- [36] M. Dolg and X. Cao, *Chem. Rev.* **112**, 403 (2012).
- [37] C. Møller and M. S. Plesset, *Phys. Rev.* **46**, 618 (1934).
- [38] T. Helgaker, P. Jørgensen and J. Olsen. *Molecular Electronic-structure Theory*. John Wiley & Sons, Ltd, 2000.
- [39] O. Christiansen, P. Jørgensen and C. Hättig, *Int. J. Quantum Chem.* **68**, 1 (1998).
- [40] C. Hättig, *NIC Series.* **31**, 245 (2006).
- [41] M. Head-Gordon, M. Oumi and D. Maurice, *Mol. Phys.* **96**, 593 (1999).
- [42] O. Christiansen, H. Koch and P. Jørgensen, *Chem. Phys. Lett.* **243**, 409 (1995).
- [43] O. Christiansen, H. Koch, P. Jørgensen and T. Helgaker, *Chem. Phys. Lett.* **263**, 530 (1996).
- [44] A. Köhn and C. Hättig, *J. Chem. Phys.* **119**, 5021 (2003).
- [45] C. Hättig and A. Köhn, *J. Chem. Phys.* **117**, 6939 (2002).

- [46] M. Head-Gordon, R. J. Rico, M. Oumi and T. J. Lee, Chem. Phys. Lett. **219**, 21 (1994).
- [47] J. F. Stanton, J. Gauss, N. Ishikawa and M. Head-Gordon, J. Chem. Phys. **103**, 4160 (1995).
- [48] A. Dreuw and M. Wormit, WIREs Comput. Mol. Sci. **5**, 82 (2015).
- [49] J. Schirmer, Chem. Phys. **329**, 1 (2006).
- [50] C. Hättig, Adv. Quant. Chem. **50**, 37 (2005).
- [51] D. Hršak, A. M. Khah, O. Christiansen and C. Hättig, J. Chem. Theory Comput. **11**, 3669 (2015).
- [52] A. Köhn and A. Tajti, J. Chem. Phys. **127**, 044105 (2007).
- [53] D. Kats, D. Usvyat and M. Schütz, Phys. Rev. A. **83** (2011).
- [54] E. F. Kjørstad and H. Koch, J. Phys. Chem. Lett. **8**, 4801 (2017).
- [55] E. F. Kjørstad and H. Koch, J. Chem. Theory Comput. **15**, 5386 (2019).
- [56] K. Sneskov. *The Polarizable Embedding Coupled Cluster Method Exciting News from Beyond the Vacuum*. PhD thesis, Department of Chemistry Aarhus University, Interdisciplinary Nanoscience Center iNANO, Aarhus University, July 2012.
- [57] S. K. Khani, A. M. Khah and C. Hättig, Phys. Chem. Chem. Phys. **20**, 1634 (2018).

- [58] P. G. Szalay, T. Müller, G. Gidofalvi, H. Lischka and R. Shepard, *Chem. Rev.* **112**, 108 (2012).
- [59] H. Lischka, D. Nachtigollavá, A. J. A. Aquino, P. Szalay, F. Plasser, F. B. C. Machado and M. Barbatti, *Chem. Rev.* **118**, 7293 (2018).
- [60] M. R. Silver-Junior, S. P. Sauer, M. Schreiber and W. Thiel, *Mol. Phys.* **108**, 453 (2010).
- [61] M. Schreiber, M. R. Silva-Junior, S. P. A. Sauer and W. Thiel, *J. Chem. Phys.* **128**, 134110 (2008).
- [62] R. Izsák, *WIREs Comput. Mol. Sci.* **e1445**, 1 (2019).
- [63] H. H. Falden, K. R. Flaster-Hansen, K. L. Bak, S. Rettrup and S. P. A. Sauer, *J. Phys. Chem. A.* **113**, 11995 (2009).
- [64] P.-F. Loos, A. Scemama, A. Blondel, Y. Garniron, M. Caffarel and D. Jacquemin, *J. Chem. Theory Comput.* **14**, 4360 (2018).
- [65] C. Hättig, *J. Chem. Phys.* **118**, 7751 (2003).
- [66] D. Jacquemin, I. Duchemin and X. Blase, *J. Chem. Theory Comput.* **11**, 5340 (2015).
- [67] P. Hohenberg and W. Kohn, *Phys. Rev.* **136**, 864 (1964).
- [68] W. Kohn and L. J. Sham, *Phys. Rev.* **140**, 1133 (1965).
- [69] T. Ziegler, *Chem. Rev.* **91**, 651 (1991).

- [70] A. J. Cohen, P. Mori-Sánchez and W. Yang, *Chem. Rev.* **112**, 289 (2012).
- [71] N. Mardirossian and M. Head-Gordon, *Mol. Phys.* **19**, 2315 (2017).
- [72] N. O. C. Winter, N. K. Graf, S. Leutwyler and C. Hättig, *Phys. Chem. Chem. Phys.* **15**, 6623 (2013).
- [73] R. Send, M. Kühn and F. Furche, *J. Chem. Theory Comput.* **7**, 2376 (2011).
- [74] R. Send, V. R. I. Kaila and D. Sundholm, *J. Chem. Theory Comput.* **7**, 2473 (2011).
- [75] A. Dreuw, J. L. Weisman and M. Head-Gordon, *J. Chem. Phys.* **119**, 2943 (2003).
- [76] A. Dreuw and M. Head-Gordon, *J. Am. Chem. Soc.* **126**, 4007 (2004).
- [77] R. Crespo-Otero and M. Barbatti, *Chem. Rev.* **118**, 7026 (2018).
- [78] M. J. Field, P. A. Bash and M. Karplus, *J. Comp. Chem.* **11**, 700 (1990).
- [79] B. Lunkenheimer and A. Köhn, *J. Chem. Theory Comput.* **9**, 977 (2013).
- [80] N. H. List, P. Norman, J. Kongsted and H. J. A. Jensen, *J. Chem. Phys.* **146**, 234101 (2017).
- [81] A. Stone. *The Theory of Intermolecular Forces*. Oxford, Oxford, UK, 2013.
- [82] L. W. Chung, W. M. Sameera, R. Ramozzi, A. J. Page, M. Hatanaka, G. P. Petrov, T. V. Harris, X. Li, Z. Ke, F. Liu, H.-B. Li, L. Ding, and K. Morukuma, *Chem. Rev.* **115**, 5678 (2015).

- [83] J. Kongsted, A. Osted, K. V. Mikkelsen and O. Christiansen, *Mol. Phys.* **100**, 1813 (2002).
- [84] T. Schwabe, J. M. H. Olsen, K. Sneskov, J. Kongsted and O. Christiansen, *J. Chem. Theory Comput.* **7**, 2209 (2011).
- [85] P. Söderhjelm, J. W. Krogh, G. Karlström, U. Ryde and R. Lindh, *Journal of Computational Chemistry.* **28**, 1083 (2007).
- [86] A. D. Buckingham. *Permanent and Induced Molecular Moments and Long Range Inter-molecular Forces.* Wiley, New York, 1967.
- [87] Z. Jing, C. Liu, S. Y. Cheng, R. Qi, B. D. Walker, J.-P. Piquemal and P. Ren, *Ann. Rev. Biophys.* **48**, 371 (2019).
- [88] P. Reinholdt, M. S. Nørby and J. Kongsted, *J. Chem. Theory Comput.* **14**, 6391 (2018).
- [89] J. Huang, A. C. Simmonett, F. C. Pickard, A. D. MacKerell and B. R. Brooks, *J. Chem. Phys.* **147**, 161702 (2017).
- [90] T. Giovannini, R. R. Riso, M. Ambrosetti, A. Puglisi and C. Cappelli, *J. Chem. Theory Comput.* **15**, 2233 (2019).
- [91] S. Caprasecca, S. Jurinovich, L. Viani, C. Curutchet and B. Mennucci, *J. Chem. Theory Comput.* **10**, 1588 (2014).

- [92] M. F. S. J. Menger, S. Caprasecca and B. Mennucci, *J. Chem. Theory Comput.* **13**, 3778 (2017).
- [93] M. A. Thompson, *J. Phys. Chem.* **100**, 14492 (1996).
- [94] K. Sneskov, T. Schwabe, J. Kongsted and O. Christiansen, *J. Chem. Phys.* **134**, 104108 (2011).
- [95] L. J. Nàbo, J. M. H. Olsen, T. J. Martínez and J. Kongsted, *J. Chem. Theory Comput.* **13**, 6230 (2017).
- [96] M. Caricato, *J. Chem. Phys.* **135**, 074113 (2011).
- [97] M. Caricato, *J. Chem. Theory Comput.* **8**, 5081 (2012).
- [98] J. Kongsted, A. Osted, K. V. Mikkelsen and O. Christiansen, *Mol. Phys.* **100**, 1813 (2002).
- [99] J. M. Mewes, Z. Q. You, M. Wormit, T. Kriesche, J. M. Herbert and A. Dreuw, *J. Phys. Chem. A.* **119**, 5446 (2015).
- [100] B. Lunkenheimer and A. Köhn, *J. Chem. Theory Comput.* **9**, 977 (2013).
- [101] N. H. List, M. T. P. Beerepooti, J. M. H. Olsen, B. Gao, K. Ruud, H. J. A. Jensen and J. Kongsted, *J. Chem. Phys.* **142**, 034119 (2015).
- [102] J. M. H. Olsen and J. Kongsted, *Adv. Quant. Chem.* **61**, 107 (2011).

- [103] K. Sneskov, T. Schwabe, O. Christiansen and J. Kongsted, *Phys. Chem. Chem. Phys.* **13**, 18551 (2011).
- [104] J.-M. Mewes, Z.-Q. You, M. Wormit, T. Kriesche, J. M. Herbert and A. Dreuw, *J. Phys. Chem. A* **119**, 5446 (2015).
- [105] J.-M. Mewes, J. M. Herbert and A. Dreuw, *Phys. Chem. Chem. Phys.* **19**, 1644 (2017).
- [106] M. Scheurer, M. F. Herbst, P. Reinholdt, J. M. H. Olsen, A. Dreuw and J. Kongsted, *J. Chem. Theory Comput.* **14**, 4870 (2018).
- [107] A. Klamt and G. Schüürmann, *J. Chem. Soc., Perkin Trans. 2* **5**, 799 (1993).
- [108] S. K. Khani, R. Faber, F. Santoro, C. Hättig and S. Coriani, *J. Chem. Theory Comput.* **15**, 1242 (2018).
- [109] J. Ángyán, *Chem. Phys. Lett.* **241**, 51 (1995).
- [110] T. Kumpulainen, B. Lang, A. Rosspeintner and E. Vauthey, *Chem. Rev.* **117**, 10826 (2016).
- [111] J.-C. Mialocq and T. Gustavsson, Investigation of femtosecond chemical reactivity by means of fluorescence up-conversion, in *New Trends in Fluorescence Spectroscopy*, pp. 61–80. Springer Berlin Heidelberg, 2001.
- [112] J. Knorr, P. Sokkar, S. Schott, P. Costa, W. Thiel, W. Sander, E. Sanchez-Garcia and P. Nuernberger, *Nature Communications*. **7** (2016).

- [113] A. J. Orr-Ewing, J. Chem. Phys. **140**, 090901 (2014).
- [114] M. Park, C. H. Kim and T. Joo, J. Phys. Chem. A. **117**, 370 (2013).
- [115] P. K. Verma, F. Koch, A. Steinbacher, P. Nuernberger and T. Brixner, Journal of the American Chemical Society. **136**, 14981 (2014).
- [116] A. Douhal, Chem. Rev. **104**, 1955 (2004).
- [117] D. Marx and J. Hutter. *Ab Initio Molecular Dynamics*. Cambridge University Press, 2009.
- [118] Y. Sugita and Y. Okamoto, Chemical Physics Letters. **314**, 141 (1999).
- [119] A. Laio and M. Parrinello, Proc. Natl. Acad. Sci. **99**, 12562 (2002).
- [120] A. Barducci, M. Bonomi and M. Parrinello, WIREs Comput. Mol. Sci. **1**, 826 (2011).
- [121] C. D. Fu, L. F. L. Oliveira and J. Pfaendtner, J. Chem. Theory Comput. **13**, 968 (2017).
- [122] T. Taketsugu and M. S. Gordon, J. Chem. Phys. **103**, 10042 (1995).
- [123] T. Tsutsumi, Y. Harabuchi, Y. Ono, S. Maeda and T. Taketsugu, Phys. Chem. Chem. Phys. **20**, 1364 (2018).
- [124] K. Fukui, J. Chem. Phys. **74**, 4161 (1970).
- [125] K. Fukui, Acc. Chem. Res. **14**, 363 (1981).

- [126] S. Maeda, Y. Harabuchi, Y. Ono, T. Taketsugu and K. Morokuma, *Int. J. Quantum Chem.* **115**, 258 (2014).
- [127] A. Tachibana and K. Fukui, *Theor. Chim. Acta.* **49**, 321 (1978).
- [128] K. Ishida, K. Morokuma and A. Komornicki, *J. Chem. Phys.* **66**, 2153 (1977).
- [129] H. P. Hratchian and H. B. Schlegel, *J. Phys. Chem. A.* **106**, 165 (2002).
- [130] H. P. Hratchian and E. Kraka, *J. Chem. Theory Comput.* **9**, 1481 (2013).
- [131] A. N. Sheppard and O. Acevedo, *J. Am. Chem. Soc.* **131**, 2530 (2009).
- [132] M. Hirsch, W. Quapp and D. Heidrich, *Phys. Chem. Chem. Phys.* **1**, 5291 (1999).
- [133] B. L. Feringa, *Acc. Chem. Res.* **34**, 504 (2001).
- [134] D. H. Waldeck, *Chem. Rev.* **91**, 415 (1991).
- [135] F. Renth, J. Bahrenburg and F. Temps, Ultrafast photoswitching dynamics of azobenzenes with intra- and intermolecular constraints, in *Photon-Working Switches*, pp. 237–259. Springer Japan, 2017.
- [136] J. Conyard, I. A. Heisler, Y. Chan, P. C. B. Page, S. R. Meech and L. Blancafort, *Chem. Sci.* **9**, 1803 (2018).
- [137] Y. O. Jung, J. H. Lee, J. Kim, M. Schmidt, K. Moffat, V. Šrajter and H. Ihee, *Nature Chemistry.* **5**, 212 (2013).

- [138] K. Pande, C. D. M. Hutchison, G. Groenhof, A. Aquila, J. S. Robinson, J. Tenboer, S. Basu, S. Boutet, D. P. DePonte, M. Liang, T. A. White, N. A. Zatsepin, O. Yefanov, D. Morozov, D. Oberthuer, C. Gati, G. Subramanian, D. James, Y. Zhao, J. Koralek, J. Brayshaw, C. Kupitz, C. Conrad, S. Roy-Chowdhury, J. D. Coe, M. Metz, P. L. Xavier, T. D. Grant, J. E. Koglin, G. Ketawala, R. Fromme, V. rajer, R. Henning, J. C. H. Spence, A. Ourmazd, P. Schwander, U. Weierstall, M. Frank, P. Fromme, A. Barty, H. N. Chapman, K. Moffat, J. J. van Thor, and M. Schmidt, *Science*. **352**, 725 (2016).
- [139] A. Warshel, *Nature*. **260**, 679 (1976).
- [140] I. Schapiro, O. Weingart and V. Buss, *J. Am. Chem. Soc.* **131**, 16 (2009).
- [141] M. Barbatti, M. Ruckebauer, J. J. Szymczak, B. Sellner, M. Vazdar, I. Antol, M. Eckert-Maksić and H. Lischka, Model systems for dynamics of π -conjugated biomolecules in excited states, in *Handbook of Computational Chemistry*, pp. 1697–1739. Springer International Publishing, 2017.
- [142] R. S. H. Liu, *Acc. Chem. Res.* **34**, 555 (2001).
- [143] J. J. Szymczak, M. Barbatti and H. Lischka, *J. Chem. Theory Comput.* **4**, 1189 (2008).
- [144] R. Salomon-Ferrer, D. A. Case and R. C. Walker, *WIREs Comput. Mol. Sci.* **3**, 198 (2013).

- [145] I. Ivani, P. D. Dans, A. Noy, A. Pérez, I. Faustino, A. Hospital, J. Walther, P. Andrio, R. Goñi, A. Balaceanu, G. Portella, F. Battistini, J. L. Gelpí, C. González, M. Vendruscolo, C. A. Laughton, S. A. Harris, D. A. Case, and M. Orozco, *Nature Methods*. **13**, 55 (2015).
- [146] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, *J. Am. Chem. Soc.* **117**, 5179 (1995).
- [147] J. Wang, W. Wang, P. A. Kollman and D. A. Case, *J. Mol. Graph. Model.* **25**, 247 (2006).
- [148] C. I. Bayly, P. Cieplak, W. D. Cornell and P. A. Kollman, *J. Phys. Chem.* **97**, 10269 (1993).
- [149] A. D. Becke, *Phys. Rev. A.* **38**, 3098 (1988).
- [150] F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.* **7**, 3297 (2005).
- [151] X. He and J. Z. H. Zhang, *J. Chem. Phys.* **122**, 031103 (2005).
- [152] P. Söderhjelm, F. Aquilante and U. Ryde, *J. Phys. Chem. B.* **113**, 11085 (2009).
- [153] J. M. H. Olsen. *Pyframe: Python tools for fragment-based multiscale embedding*, 2018.
- [154] K. Aidas, C. Angeli, K. L. Bak, V. Bakken, R. Bast, L. Boman, O. Christiansen, R. Cimiraglia, S. Coriani, P. Dahle, E. K. Dalskov, U. Ekström, T. Enevoldsen,

- J. J. Eriksen, P. Ettenhuber, B. Fernández, L. Ferrighi, H. Fliegl, L. Frediani, K. Hald, A. Halkier, C. Hättig, H. Heiberg, T. Helgaker, A. C. Hennum, H. Hettema, E. Hjertenæs, S. Høst, I.-M. Høyvik, M. F. Iozzi, B. Jansik, H. J. A. Jensen, D. Jonsson, P. Jørgensen, J. Kauczor, S. Kirpekar, T. Kjærgaard, W. Klopper, S. Knecht, R. Kobayashi, H. Koch, J. Kongsted, A. Krapp, K. Kristensen, A. Ligabue, O. B. Lutnæs, J. I. Melo, K. V. Mikkelsen, R. H. Myhre, C. Neiss, C. B. Nielsen, P. Norman, J. Olsen, J. M. H. Olsen, A. Osted, M. J. Packer, F. Pawłowski, T. B. Pedersen, P. F. Provasi, S. Reine, Z. Rinkevicius, T. A. Ruden, K. Ruud, V. V. Rybkin, P. Salek, C. C. M. Samson, A. S. de Merás, T. Saue, S. P. A. Sauer, B. Schimmelpfennig, K. Sneskov, A. H. Steindal, K. O. Sylvester-Hvid, P. R. Taylor, A. M. Teale, E. I. Tellgren, D. P. Tew, A. J. Thorvaldsen, L. Thøgersen, O. Vahtras, M. A. Watson, D. J. D. Wilson, M. Ziolkowski, and H. Ågren, *WIREs Comput. Mol. Sci.* **4**, 269 (2014).
- [155] L. Gagliardi, R. Lindh and G. Karlstöm, *J. Chem. Phys.* **121**, 4494 (2004).
- [156] T. Schwabe, J. M. H. Olsen, K. Sneskov, J. Kongsted and O. Christiansen, *J. Chem. Theory Comput.* **7**, 2209 (2011).
- [157] E. Boulanger and W. Thiel, *J. Chem. Theory Comput.* **10**, 1795 (2014).
- [158] F. Weigend, A. Köhn and C. Hättig, *J. Chem. Phys.* **116**, 3175 (2001).
- [159] F. Weigend, M. Häser, H. Patzelt and R. Ahlrichs, *Chem. Phys. Lett.* **294**, 143 (2005).

- [160] A. Hellweg, *J. Comp. Chem.* **34**, 1835 (2013).
- [161] L. Grimmelsmann, V. Schuabb, B. Tekin, R. Winter and P. Nuernberger, *Phys. Chem. Chem. Phys.* **20**, 18169 (2018).
- [162] T. H. Dunning Jr., *J. Chem. Phys.* **90**, 1007 (1989).
- [163] S. K. Pal, J. Peon and A. H. Zewail, *Proc. Natl. Acad. Sci.* **99**, 1763 (2002).
- [164] S. K. Pal, L. Zhao and A. H. Zewail, *Proc. Natl. Acad. Sci.* **100**, 8113 (2003).
- [165] S. K. Pal, L. Zhao, T. Xia and A. H. Zewail, *Proc. Natl. Acad. Sci.* **100**, 13746 (2003).

Appendices

Supplementary Information:

Calculations on Excited States with Simple Pseudopotentials

Alireza Marefat Khah,^{*,†} Peter Reinholdt,^{*,‡} Jógvan Magnus Huagaard Olsen,[¶]

Jacob Kongsted,^{*,‡} and Christof Hättig^{*,†}

[†]*Quantum Chemistry Group, Ruhr University of Bochum, Germany*

[‡]*Department of Physics, Chemistry and Pharmacy, University of Southern Denmark, Odense, Denmark*

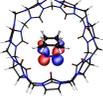
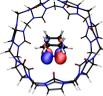
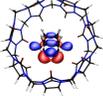
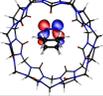
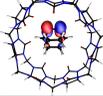
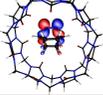
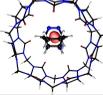
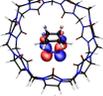
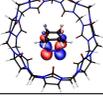
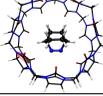
[¶]*Hylleraas Centre for Quantum Molecular Sciences, Department of Chemistry, UiT The Arctic University of Norway, N-9037 Tromsø, Norway*

E-mail: alireza.marefatkhah@rub.de; reinholdt@sdu.dk; kongsted@sdu.dk;

christof.haettig@rub.de

Natural Transition Orbitals

DBO@CB[7] : PNO-ADC(2)/aug-cc-pVTZ/cc-pVDZ, dispersion is excluded based on the EDA analysis. PE(ECP)-ADC(2)/aug-cc-pVTZ/IAO-D3(eCP-all). PE-ADC(2)/aug-cc-pVTZ/IAO-D3. Natural transition orbitals are visualized with 0.025 isosurface value. Excitations with electron spilling-out issue are labeled with ESO.

Level	Char.	VEE	OCC	VIR	Contrib.
PNO-noDisp	$S_1^{n \rightarrow \pi^*}$	3.431			99.8 %
PNO-noDisp	$S_2^{\sigma \rightarrow \pi^*}$	6.104			99.6 %
PE(ECP)	$S_1^{n \rightarrow \pi^*}$	3.433			99.8 %
PE(ECP)	$S_2^{n \rightarrow \sigma^*}$	5.801			99.3 %
PE	$S_1^{n \rightarrow \text{ESO}}$	0.883			99.9 %
PE	$S_2^{n \rightarrow \text{ESO}}$	1.091			99.9 %

BRN@DNA : PNO-ADC(2)/aug-cc-pVTZ/cc-pVDZ. dispersion is excluded based on the EDA analysis.
 PE(ECP)-ADC(2)/aug-cc-pVTZ/IAO-D3(ecp-all). PE-ADC(2)/aug-cc-pVTZ/IAO-D3. Natural transition
 orbitals are visualized with 0.025 isosurface value. Excitations with electron spilling-out issue are labeled
 with ESO.

Level	Char.	VEE	OCC	VIR	Contrib.
PNO-noDisp	$S_1^{\pi \rightarrow \pi^*}$	2.705			94.3 %
PNO-noDisp	$S_2^{\pi \rightarrow \pi^*}$	3.062			95.0 %
PE(ECP)	$S_1^{\pi \rightarrow \pi^*}$	2.640			93.8 %
PE(ECP)	$S_2^{\pi \rightarrow \pi^*}$	3.001			94.5 %
PE	$S_1^{\pi \rightarrow \pi^*}$	2.608			93.9 %
PE	$S_2^{\pi \rightarrow \text{ESO}}$	2.946			92.8 %

PNP : PNO-ADC(2)/aug-cc-pVTZ/cc-pVDZ. dispersion is excluded based on the EDA analysis.
 PE(ECP)-ADC(2)/aug-cc-pVTZ/IAO-D3(ecp-all). PE-ADC(2)/aug-cc-pVTZ/IAO-D3. Natural transition
 orbitals are visualized with 0.025 isosurface value. Excitations with electron spilling-out issue are labeled
 with ESO. For Na cation the Loprop HF/cc-pVDZ polarizabilities has been used.

Level	Char.	VEE	OCC	VIR	Contrib.
PNO-noDisp	$S_1^{\pi \rightarrow \pi^*}$	2.923			97.3 %
PNO-noDisp	$S_2^{\pi \rightarrow \pi^*}$	3.497			99.5 %
PE(ECP)	$S_1^{\pi \rightarrow \pi^*}$	2.841			97.7 %
PE(ECP)	$S_2^{\pi \rightarrow \pi^*}$	3.434			99.1 %
PE	$S_1^{\pi \rightarrow \pi^*}$	2.576			97.6 %
PE	$S_2^{\pi \rightarrow \pi^*}$	3.454			96.8 %

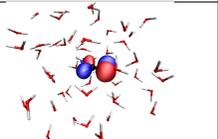
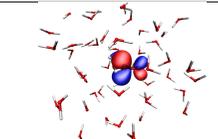
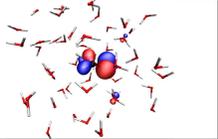
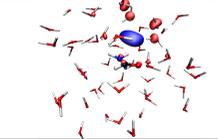
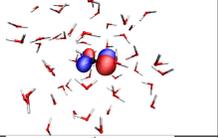
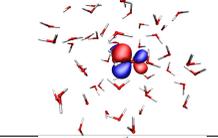
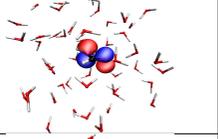
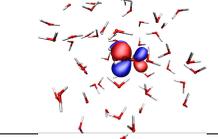
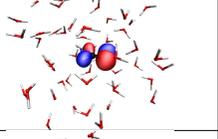
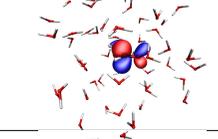
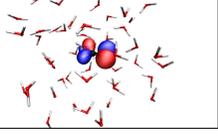
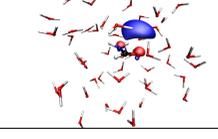
Cholesterol: PNO-ADC(2)/aug-cc-pVTZ/cc-pVDZ. dispersion is excluded based on the EDA analysis. PE(ECP)-ADC(2)/aug-cc-pVTZ/IAO-D3(esp-all). PE-ADC(2)/aug-cc-pVTZ/IAO-D3. Natural transition orbitals are visualized with 0.025 isosurface value. Excitations with electron spilling-out issue are labeled with ESO.

Level	Char.	VEE	OCC	VIR	Contrib.
PNO-noDisp	$S_1^{\pi \rightarrow \text{Ryd.}}$	6.373			99.3 %
PNO-noDisp	$S_2^{\pi \rightarrow \text{Ryd.}}$	6.401			91.4 %
PE(ECP)	$S_1^{\pi \rightarrow \text{Ryd.}}$	6.335			81.2 %
PE(ECP)	$S_2^{\pi \rightarrow \text{Ryd.}}$	6.407			56.8 %
PE	$S_1^{\pi \rightarrow \text{ESO}}$	3.363			99.9 %
PE	$S_2^{\pi \rightarrow \text{ESO}}$	4.530			99.9 %

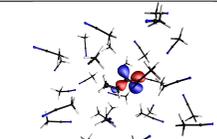
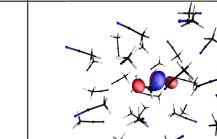
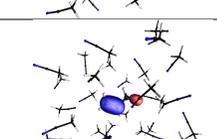
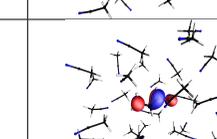
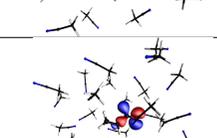
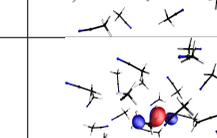
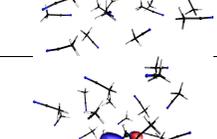
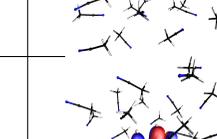
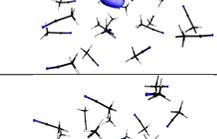
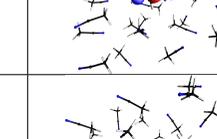
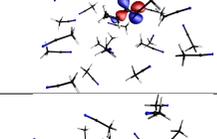
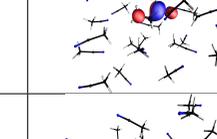
ANILINE YELLOW : PNO-ADC(2)/aug-cc-pVTZ/cc-pVDZ. dispersion is excluded based on the EDA analysis. PE(ECP)-ADC(2)/aug-cc-pVTZ/IAO-D3(esp-all). PE-ADC(2)/aug-cc-pVTZ/IAO-D3. Natural transition orbitals are visualized with 0.025 isosurface value. Excitations with electron spilling-out issue are labeled with ESO.

Level	Char.	VEE	OCC	VIR	Contrib.
PNO-noDisp	$S_1^{n \rightarrow \pi^*}$	2.956			99.7 %
PNO-noDisp	$S_2^{\pi \rightarrow \pi^*}$	4.147			97.1 %
PE(ECP)	$S_1^{n \rightarrow \pi^*}$	2.960			99.7 %
PE(ECP)	$S_2^{\pi \rightarrow \pi^*}$	4.138			96.5 %
PE	$S_1^{n \rightarrow \pi^*}$	2.964			99.7 %
PE	$S_2^{\pi \rightarrow \pi^*}$	4.209			94.5 %

Formamide : PNO-ADC(2)/aug-cc-pVTZ/cc-pVDZ. dispersion is excluded based on the EDA analysis.
 PE(ECP)-ADC(2)/aug-cc-pVTZ/IAO-D3(eCP-all). PE-ADC(2)/aug-cc-pVTZ/IAO-D3. Natural transition
 orbitals are visualized with 0.025 isosurface value. Excitations with electron spilling-out issue are labeled
 with ESO.

Level	Char.	VEE	OCC	VIR	Contrib.
PNO-noDisp	$S_1^{n \rightarrow \pi^*}$	5.630			99.9 %
PNO-noDisp	$S_2^{n \rightarrow R_{yd}}$	7.487			89.5 %
PE(ECP)	$S_1^{n \rightarrow \pi^*}$	5.749			99.9 %
PE(ECP)	$S_2^{\pi \rightarrow \pi^*}$	7.187			97.3 %
PE	$S_1^{n \rightarrow \pi^*}$	5.637			99.7 %
PE	$S_2^{n \rightarrow ESO}$	6.412			83.3 %

Acrolein : PNO-ADC(2)/aug-cc-pVTZ/cc-pVDZ. dispersion is excluded based on the EDA analysis.
 PE(ECP)-ADC(2)/aug-cc-pVTZ/IAO-D3(eCP-all). PE-ADC(2)/aug-cc-pVTZ/IAO-D3. Natural transition
 orbitals are visualized with 0.025 isosurface value. Excitations with electron spilling-out issue are labeled
 with ESO.

Level	Char.	VEE	OCC	VIR	Contrib.
PNO-noDisp	$S_1^{n \rightarrow \pi^*}$	3.900			99.8 %
PNO-noDisp	$S_2^{\pi \rightarrow \pi^*}$	6.385			94.3 %
PE(ECP)	$S_1^{n \rightarrow \pi^*}$	3.896			99.7 %
PE(ECP)	$S_2^{\pi \rightarrow \pi^*}$	6.421			96.1 %
PE	$S_1^{n \rightarrow \pi^*}$	3.894			99.7 %
PE	$S_2^{n \rightarrow ESO}$	5.782			99.8 %

GFP: PNO-ADC(2)/aug-cc-pVTZ/cc-pVDZ. dispersion is excluded based on the EDA analysis.
 PE(ECP)-ADC(2)/aug-cc-pVTZ/IAO-D3(ecp-all). PE-ADC(2)/aug-cc-pVTZ/IAO-D3. Natural transition orbitals are visualized with 0.025 isosurface value. Excitations with electron spilling-out issue are labeled with ESO. Out-of core (the embedded chromophore) excitations are labeled with Off-core.

Level	Char.	VEE	OCC	VIR	Contrib.
PNO-noDisp	$S_1^{\pi \rightarrow \pi^*}$	2.459			97.6 %
PNO-noDisp	$S_2^{\text{Off-Chor.}}$	3.727			99.9 %
PE(ECP)	$S_1^{\pi \rightarrow \pi^*}$	2.393			97.8 %
PE(ECP)	$S_2^{n \rightarrow \pi^*}$	3.042			99.8 %
PE	$S_1^{n \rightarrow \text{ESO}}$	1.237			99.8 %
PE	$S_2^{\pi \rightarrow \pi^*}$	2.428			97.0 %

SYSTEM : PNO-ADC(2)/aug-cc-pVTZ/cc-pVDZ. dispersion is excluded based on the EDA analysis.
 PE(ECP)-ADC(2)/aug-cc-pVTZ/IAO-D3(ecp-all). PE-ADC(2)/aug-cc-pVTZ/IAO-D3. Natural transition orbitals are visualized with 0.025 isosurface value. Excitations with electron spilling-out issue are labeled with ESO.

Level	Char.	VEE	OCC	VIR	Contrib.
PNO-noDisp	S_1^{LE}	4.146			93.1 %
PNO-noDisp	S_2^{ICT}	4.383			97.0 %
PE(ECP)	S_1^{LE}	4.512			83.7 %
PE(ECP)	S_2^{ICT}	4.555			94.4 %
PE	S_1^{ESO}	1.167			99.7 %
PE	S_2^{ESO}	1.316			99.7 %

PE and PE(ECP) Computational Costs

Table S1: PE- and PE(ECP)-ADC(2)/aTZ/DZ CPU-times for excitation energy calculations of two lowest singlet electronic excited states. The numbers are in minutes. The calculations for small or large chromophores have been done on different machines with different numbers of nodes and different CPUs; however, always with the same type of machine for a particular system.

System	$\mathcal{T}(\text{PE})$	$\mathcal{T}(\text{PE}(\text{ECP}))$	Ratio.
DBO	751	851	1.13
BRN	12582	10087	0.80
PNP	348	312	0.90
CHL	146309	171437	1.17
ANL	2168	2019	0.93
FRM	10	9	0.90
ACR	29	22	0.76
GFP	15292	17816	1.16
DMABN	1353	1974	1.46

Parametrized ECPs of He, Ne and Ar

All the numbers and values are given in TURBOMOLE input format.

Table S2: He ecp-all. All-electron ECP were fitted to reproduce with zero electron the orbital energies of the 2s2p3s3p shells in the d-aug-cc-pV5Z basis within 6 digits (and for other basis sets with a less digits)

ncore=	2	lmax=	2
d			
	0.0000	2	1.0000
s-d			
	2.4200	2	0.5098
p-d			
	-0.43590	2	0.49165

Table S3: Ne ecp-all. All-electron ECP were fitted to reproduce with zero electron the orbital energies of the 3s3p3d4s4p4d shells in the d-aug-cc-pwCV5Z basis within 6 digits (and for other basis sets with a less digits)

ncore=	10	lmax=	3
d			
	0.0000	2	1.0000
s-d			
	54.5100	2	2.0475
p-d			
	1.46500	2	0.44815
d-f			
	-0.8380	2	0.49205

Table S4: Ar ecp-all. All-electron ECP were fitted to reproduce with zero electron the orbital energies of the 4s4p4d5s5p5d shells in the aug-cc-pV(6+d)Z basis.

ncore=	18	lmax=	3
f			
	0.0000	2	1.0000
s-f			
	275.00	2	1.6410
p-f			
	1.9000	2	0.2733
d-f			
	-3.4000	2	0.44000