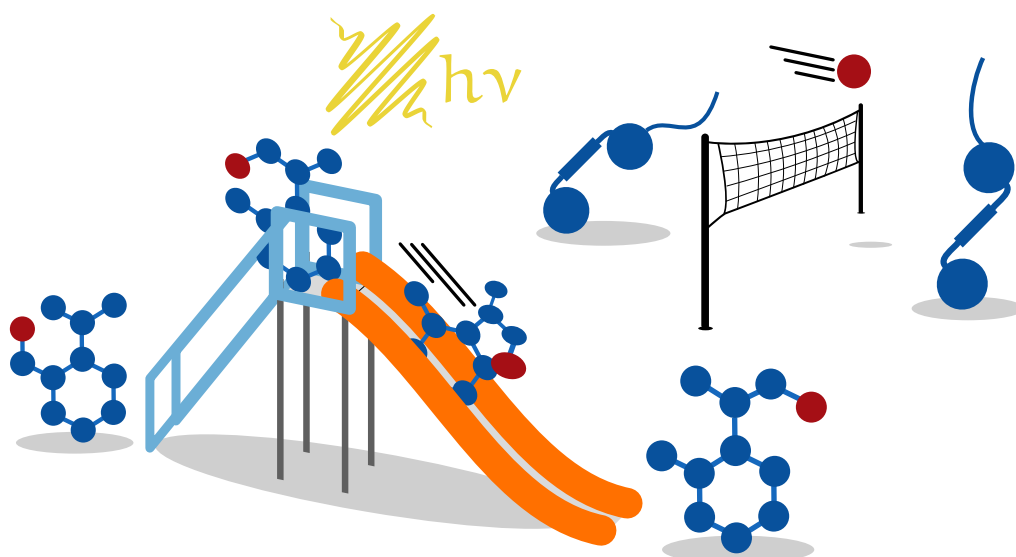


Full-Dimensional Photodynamics Simulations

From Photoisomerizations to Excited-State Proton Transfer Reactions



Dissertation

in partial fulfillment of the requirements
for the degree

Dr. rer. nat.

of the Faculty of Mathematics and Natural Sciences
at Kiel University

submitted by

Tim Raeker

Kiel, April 2018

Full-Dimensional Photodynamics Simulations
From Photoisomerizations to Excited-State Proton Transfer Reactions

Dissertation, submitted by Tim Raeker

Kiel – April 12, 2018

First Referee: Prof. Dr. Bernd Hartke
Second Referee: Prof. Dr. Friedrich Temps

Date of Oral Examination: June 14, 2018
Approved for Publication: June 14, 2018

Abstract

Photochemical reactions are of great interest in both applied and fundamental research. Photoactive molecules are studied for their electronic properties and are also being designed for use in specific applications or even molecular machines.

Two photochemical reactions are investigated in this thesis: *trans*⇌*cis* photoisomerization (“switching”) and excited-state intramolecular proton transfer (ESIPT).

The first reaction type is represented by the newest generation of bridged-azobenzene derivatives, indandiazocine (ID) and diindandiazocine (DID). Both ID and DID were designed to achieve chiral *trans*→*cis* isomerization as an upgrade over their parent system diazocine. This predicted feature is validated by calculations of the actual photodynamics of these systems. In a second project, azobenzene and the three aforementioned bridged azobenzenes are used as motors in an artificial cilium that served as a prototype for a molecular particle transport machine. With ferulic acid, azobenzene-containing carbohydrate macrocycles and axitinib, three more systems – as part of cooperative projects – are also investigated for their photoinduced switching properties.

The first system that was found to perform ESIPT was salicylic acid (SAC) in 1956. However, full-dimensional calculations of SAC ESIPT dynamics are presented for the first time here. A deactivation pathway through a conical intersection from the first electronically excited state is found to play a major role in the ESIPT process of SAC. Several SAC derivatives are studied for their dynamical properties as well. The excited-state deactivation *via trans*⇌*cis* isomerization of the so-called “crane” unit was found to be a common motif for these ESIPT systems.

With 7-hydroxy-4-methylquinoline-8-carbaldehyde (HMQCA) a more complex ESIPT system is also investigated. HMQCA utilizes the crane-switching to allow for a PT to a different site of the molecule than the initial proton donor. Two additional systems are designed and investigated with the aim to improve the excited-state properties of HMQCA.

The computational model of choice for calculating the photodynamics is semiempirical quantum mechanics (SEQM) coupled to floating-occupation configuration interaction. SEQM relies on parameters that are fitted to reproduce a set of molecular properties. However, the studied ESIPT systems – or excited-state properties in general – are not part of such a set. Hence, deficiencies were encountered when comparing the obtained results to the results of *ab initio* QM methods. Consequently, a newly implemented particle swarm algorithm is tested as means of optimizing SEQM parameters. First results of this endeavor will be presented.

Kurzzusammenfassung

Photochemische Reaktionen sind sowohl in der Anwendungs- als auch Grundlagenforschung von großem Interesse. Photoaktive Moleküle werden auf ihre elektronischen Eigenschaften hin untersucht und werden auch für bestimmte Anwendungen oder sogar für den Einsatz in molekularen Maschinen entworfen.

In dieser Dissertation werden zwei photochemische Reaktionen untersucht: *trans*⇌*cis* Photoisomerisierung (“Schalten”) und der intramolekulare Protonentransfer im elektronisch angeregten Zustand (*excited-state intramolecular proton transfer*, ESIPT).

Der erste Reaktionstyp wird anhand der neuesten Generation von verbrückten Azobenzol-Derivaten, Indandiazocin (ID) und Diindandiazocin (DID), vorgestellt. Sowohl ID als auch DID wurden entworfen, um eine chirale *trans*→*cis*-Schaltung als Verbesserung gegenüber dem Stammsystem Diazocin zu erreichen. Diese vorhergesagte Eigenschaft kann durch Berechnung der Photodynamik dieser Systeme bestätigt werden. In einem zweiten Projekt finden Azobenzol und die drei vorherigen verbrückten Azobenzole Verwendung als Motoren in einer künstlichen Cilie, die als Prototyp für eine molekulare Partikeltransport-Maschine dient. Als Teil von Kooperationsprojekten werden auch Ferulasäure, Azobenzol-enthaltende Kohlenhydratmakrocyclen und Axitinib auf ihre Photoschaltungs-Eigenschaften hin untersucht.

1956 wurde mit Salicylsäure (*salicylic acid*, SAc) das erste Molekül gefunden, das einen ESIPT ausführen kann. Trotzdem werden erst in dieser Dissertation zum ersten Mal volldimensionale Rechnungen der ESIPT-Dynamik dieses Moleküls vorgestellt. Es wurde herausgefunden, dass ein Deaktivierungspfad durch eine konische Durchschneidung aus dem ersten elektronisch angeregten Zustand eine große Rolle im ESIPT-Prozess von SAc spielt. Einige weitere SAc-Derivate werden auf deren dynamische Eigenschaften hin untersucht. Die Deaktivierung des angeregten Zustands über *trans*⇌*cis* Isomerisierung der sog. “Kran”-Einheit konnte als häufiges Motiv ausgemacht werden.

Mit 7-Hydroxy-4-methylchinolin-8-carbaldehyd (HMQCA) wird auch ein komplexeres ESIPT-System untersucht. Dieses nutzt die Kran-Schaltung, um einen PT zu einer anderen Stelle, die nicht die anfängliche Protonendonator-Stelle ist, im Molekül zu ermöglichen. Zwei weitere System wurden entworfen und mit dem Ziel untersucht die Eigenschaften von HMQCA im angeregten Zustand zu verbessern.

Das rechnerische Modell der Wahl zur Photodynamik-Berechnung ist semiempirische Quantenmechanik (SEQM) verbunden mit der *floating-occupation-configuration-interaction*-Methode. SEQM ist auf Parameter angewiesen, die daran angepasst sind, einen Satz von bestimmten molekularen Eigenschaften zu reproduzieren. Allerdings sind die untersuchten ESIPT-Systeme, bzw. Eigenschaften in angeregten Zuständen im Allgemeinen, nicht Teil einer solchen Sammlung. Daher wurden einige Mängel ausgemacht, wenn man die erhaltenen Ergebnisse mit denen von *ab-initio*-QM-Methoden vergleicht. Folglich wird ein Partikelschwarm-Algorithmus als Mittel zur Parameteroptimierung von SEQM-Methoden vorgestellt. Die ersten Ergebnisse dieses Unterfangens werden vorgestellt.



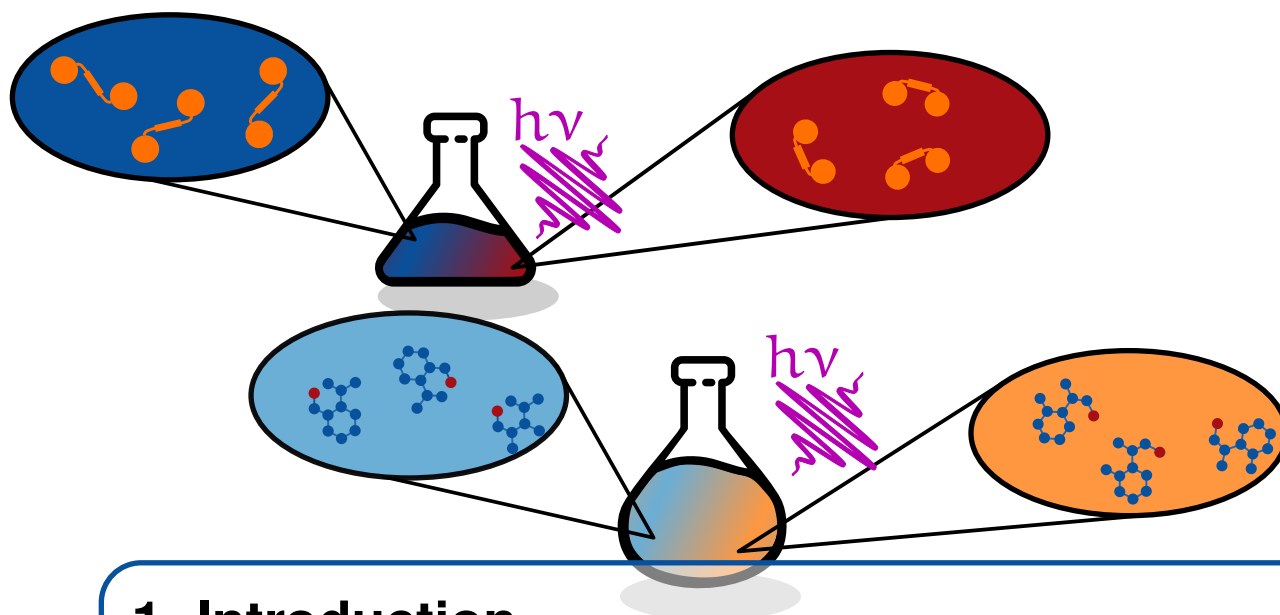
Contents

1	Introduction	1
1.1	Photochemistry	1
1.1.1	Photoisomerizations	2
1.1.2	Excited-State Intramolecular Proton Transfer	4
1.1.3	Photochemistry on the Computer	7
2	Theory	9
2.1	Hartree-Fock Theory	10
2.2	Semiempirical Hamiltonians	12
2.3	Correlation Methods	13
2.4	Density Functional Theory	18
2.5	Force Fields	20
2.6	Molecular Dynamics	22
2.6.1	Non-Adiabatic Dynamics	23
2.7	Techniques	26
2.7.1	Optimization Processes	26
2.7.2	Preparing Photodynamics	28
2.7.3	Running Photodynamics	28
2.7.4	Calculating Quantum Yields and Isomer Populations	29

3	Photoisomerizations	31
3.1	Indandiazocines: Unidirectional Molecular Switches	32
3.1.1	Scope of the project	32
3.1.2	Publication Data	33
3.2	Simulating Particle Transport	34
3.2.1	Scope of the project	34
3.2.2	Publication Data	34
3.3	Ultrafast Dynamics of Trans- and Cis-Ferulic Acid	35
3.3.1	Scope of the project	35
3.3.2	Publication Data	35
3.3.3	Additional Information	36
3.4	Azobenzene in Carbohydrate Macrocycles	37
3.5	Static Properties of a Kinase Inhibitor	44
3.6	Summary	46
4	Excited-State Proton Transfer	49
4.1	Salicylic Acid	50
4.1.1	Scope of the project	50
4.1.2	Publication Data	51
4.1.3	Additional Information	52
4.2	Further Benchmarking	55
4.3	Structural Analogues of Salicylic Acid	57
4.4	Fused-Ring Systems	61
4.4.1	Quinoline Backbone	61
4.4.2	Crane-Substitution Effects	68
4.4.3	Phenanthridine Backbone	71
4.5	Summary	74
5	Parameter Optimization	77
5.1	Semiempirical Parameters	78
5.2	Algorithm	81
5.2.1	Application	83
5.3	Test Cases	84
5.4	Summary and Outlook	90
6	Supervised Projects	93
6.1	Björn Jansen: Photodynamics of Unidirectional Cilia	93

6.2	Dominik Behrens: Particle Transport Simulations of Artificial Cilia	95
6.3	Christopher Witt: Photodynamics of a Molecular Pincer	95
6.4	Jennifer Müller: Photodynamics of Salicylic Acid Derivatives	100
6.4.1	Additional Information	101
6.5	Lasse Jannsen: Photodynamics of a Blue Fluorescence Chromophore	105
7	Outlook	113
A	Appendix	I
A.1	PM6* Parameters	I
A.2	Carbohydrate Macrocycles	II
A.2.1	Energies	II
A.2.2	Spectra	II
A.3	Salicylic Acid Derivatives	IV
A.3.1	DFT Scans	IV
A.4	Scan of the Crane Rotation of MMIMQO	V
A.5	PSO Routines	V
A.5.1	getNeighborIndices	V
A.5.2	checkParticleDistance	VI
A.5.3	performLocalSearch	VI
A.5.4	checkForPlateau	VI
A.5.5	shrinkSurface	VII
A.6	Optimized Parameters Sets	VII
A.6.1	Minimal Reference Data Set	VII
A.6.2	Larger CI Setup	VIII
A.6.3	Larger Reference Data Set	VIII
A.6.4	Including Scan Data	IX
A.7	BFP Chromophore Scans	X
A.8	Orbitals	XI
A.8.1	HMQCA	XI
A.8.2	MMIMQO	XII
A.8.3	MIMPO	XIII

*This page is intentionally left blank; but this text is printed here,
therefore this is not a blank page.*



1. Introduction

The core work in chemistry is about reactions of atoms and molecules. The aim of those reactions is to synthesize new molecules with interesting and conceivably useful properties.

However, chemical reactions are not restricted to synthesis, i.e., bringing two or more molecules together. Reactions can also happen within a single molecule, resulting in geometric changes that can either be rearrangements of atoms or just conformational isomerization¹.

Both aforementioned single-molecule reactions have in common that they can be a reaction of a molecule to an external stimulus that triggers movement² of atoms. When this stimulus is a photon, we refer to this movement as photodynamics.

1.1 Photochemistry

In photodynamics, a molecule gets excited to an energetically higher electronic state by energy transfer from a photon. In this excited state the system experiences a force, ideally favoring a change of its structure but practically often hindered by an energy barrier. Studying *if*, *how* and *how fast* such a change happens and where the reaction ends after de-excitation are the integral parts of photochemistry.

Experimentally, this is ideally achieved by looking at the positions of the atoms directly over the course of time. But photochemical reactions happen in the range of femto- and picoseconds,

¹ Strictly speaking, chemical reactions require breaking and forming of bonds between atoms which is not the case in conformational isomerization.

² In addition to the movement that is in any case present due to the temperature being greater than 0 K and also to the internal – within the reaction vessel – stimuli of Brownian motion.

hence coining the term “ultrafast reactions”. Unfortunately, current reaction microscopes still lack either temporal or spatial resolution to track photodynamics.

Such reactions are therefore usually tracked by their change in electronic properties, because of the high temporal resolution that is achieved by certain spectroscopical methods (but lacking spatial resolution)³.

Alternatively, they are accompanied or driven by simulations on the computer, where spatial resolution does not matter, because each atom can be treated individually. The time scales of such reactions are advantageous as well, because the dynamics of molecules are calculated at discrete time steps, thus fewer time steps mean less calculations.

In the following, two classes of photochemical reactions will be presented and how they can be studied on the computer.

1.1.1 Photoisomerizations

A typical non-bond-breaking reaction induced by photoexcitation is *trans*⇌*cis* isomerization – in the following defined as “switching” – at a double bond between two atoms.

Regarding this reaction in general, probably all photoisomerizations around double bonds happen in the same fashion: After electronic excitation from a (bonding) orbital⁴ to a non-bonding π^* orbital of the double bond in question, the bond order is lowered and thus allows for rotation.

This rotation is driven by the gradient in the electronically excited state ($S_{i>0}$). To return to the electronic ground state (S_0), double-bond isomerizations usually involve conical intersections (CoIns), which are places where S_1 and S_0 come in touch⁵. In photoisomerizations these CoIns are typically found at a rotational angle of about 90°. [2] At this stage – and assuming that the system even reaches a CoIn –, the reaction is at a branching point, because when the system has traveled through the CoIn, there are ideally equal chances that the system follows the S_0 gradient either back to the initial structure or keeps the direction of momentum towards the photoproduct. Any deviations from this ratio comes from a non-symmetrical branching on S_0 , favoring one pathway over the other. If the system does not reach a CoIn, it may also return to S_0 *via* fluorescence. A graphical summary of these processes is depicted in Fig. 1.1.

One of the most prominent, non-naturally-occurring⁶ molecule that performs photoisomerization is azobenzene (AB). Even though its photodynamical properties have been known for a long time,^[3] AB and its derivatives are still part of modern applied^[4–11] and computational chemistry.^[12–19] Interesting properties of AB are its fast photoisomerization rate, high photostability and ease of chemical use to tune its properties or integrating it into larger structures.

³ It must be mentioned, though, that there are experimental setups that, e.g., utilize femtosecond-resolved X-ray spectroscopy. For example, in Ref. 1 the photosystem I protein was investigated using this technique with a temporal resolution of down to 10 fs. However, the spatial resolution was 8.5 Å, which is sufficient for a system of this size (~200 nm to 2 μm) but not for typical photoactive molecules like the ones being presented in this thesis.

⁴ This must not necessarily be a π orbital of the corresponding double bond but can also be an n orbital. Examples for these types will be studied in this thesis.

⁵ Of course CoIns can also exist between excited states.

⁶ Otherwise it would probably be the retinal molecule that allows our eyes to see.

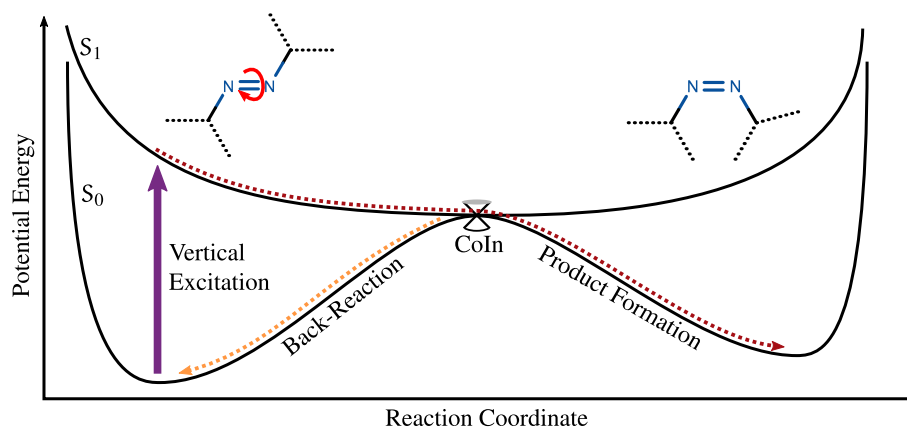


Figure 1.1.: Qualitative scheme of a photoisomerization on the example of an azo group dihedral rotation: After vertical electronic excitation from S_0 to S_1 the system follows the negative gradient (dashed line) towards the new minimum back to S_0 through a conical intersection (CoIn). A back-reaction to the reactant-side may also happen at the CoIn.

Of course AB is not without deficiencies: Because of the rather simple reactive core – the $-N=N-$ azo bridge – the reaction path is not well defined.^[20,21] Although one may be favored over the other, the photoisomerization can happen *via* two pathways: Rotation^[22,23] or inversion^[20,24] (Fig. 1.2, left). Even though both perform different conformational changes, they do eventually end up at the same geometry.

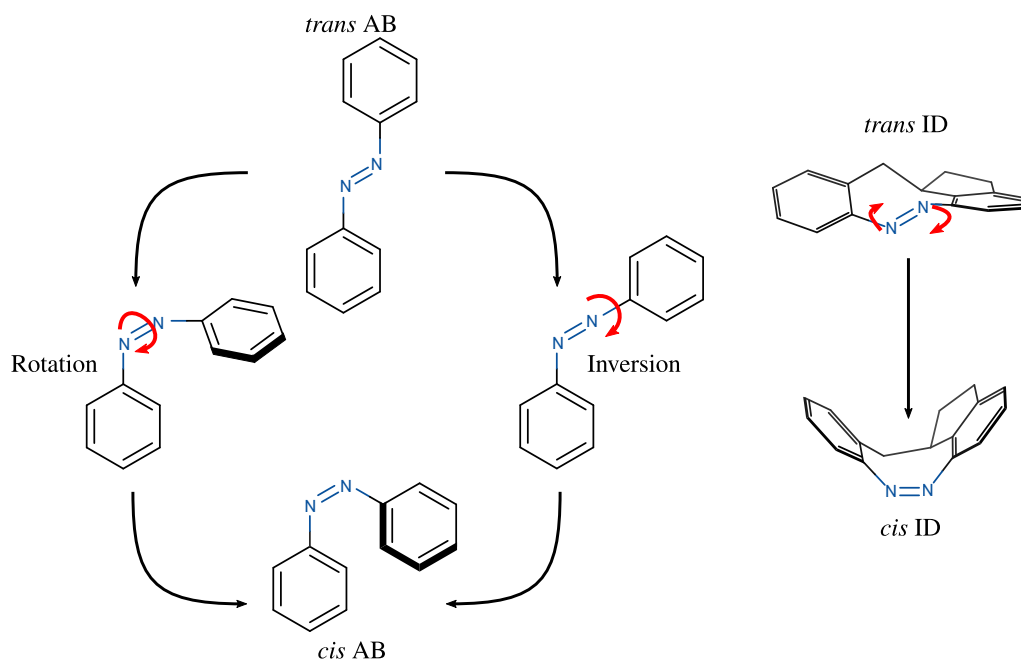


Figure 1.2.: Reaction scheme from *trans* to *cis* azobenzene (AB) following the two controversial pathways of either rotation or inversion (left). In comparison, the uni-directional *trans*→*cis* reaction of indandiazocine (ID) (right).

To tackle this problem, a bridged AB derivative was conceived, synthesized^[25] and analyzed^[26–28] at Kiel University within the Collaborative Research Center 677 “Function by Switching”. The additional bridge confines the isomerization so that there is only one accessible pathway, which in this case is neither inversion nor rotation but something in between, typically described as “hula twist” (Fig. 1.2, right).^[26,28–31] This enhancement also led to improved reaction rates, since the ethylenic bridge introduces strain to the molecular frame, and also made the *cis* (*Z*) form more stable than *trans* (*E*).

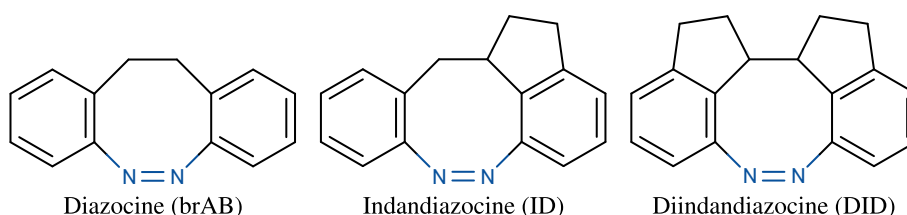


Figure 1.3.: Sketches of the three bridged-azobenzenes: Diazocine, indandiazocine, diindandiazocine.

The most recent generation of AB derivatives are indandiazocines (ID) which introduce a second or third ethylenic bridge, respectively, and thus a chiral center to the system^[32] (Fig. 1.3). Besides even more improved dynamical properties, the $E \rightarrow Z$ reaction pathway is additionally confined, which results in more control over the reaction.

Especially the latter point is important when such systems are to be used in larger molecular frameworks, e.g., molecular machines, which are to exhibit certain properties for specified applications. For example, molecular machines that are to transport particles, i.e., other molecules, should do this in a directed – predictable – way. This is only achieved when the “motor” of the machine features directed motion. The resulting transport would otherwise be random which does not serve any benefit over the inherent Brownian motion of a particle in solution. Being restricted to only one of two possible $E \rightarrow Z$ directions, ID and DID are superior to AB and brAB in this regard.

Generally, well-studied systems are crucial in designing systems with specific properties.

1.1.2 Excited-State Intramolecular Proton Transfer

One of many examples for bond-breaking photon-triggered reactions would be proton transfer (PT). When it comes to bond-breaking reactions, the transfer of a proton⁷ – as the smallest atom that can be dislocated – is the fastest known in chemistry.

PT are also one of the most often occurring reaction types in nature because they are very fast and exhibit high photostability. A prominent intermolecular PT system is DNA. There, a proton between two adjacent nucleobases can be transferred as means of deactivation, i.e., to return to S_0 , after electronic excitation with UV light.^[34] In the following only the excited-state intramolecular PT (ESIPT) will be presented, but most of the general reaction properties hold true for an intermolecular PT as well.^[35,36]

⁷ It is also possible to transfer hydrogen atoms. However, although there are suggested ways of checking which of the two is transferred (cf., e.g., Ref. 33), no distinction between the two will be made for the remainder of the thesis.

1.1 Photochemistry

To achieve the capability to perform an ESIPT, a molecule has to fulfill the following two conditions: i) there must be a proton donor and a proton acceptor site, which both are typically – but not exclusively – oxygen or nitrogen atoms, and they must be geometrically close, ideally forming a hydrogen bond; ii) upon photo-induced excitation to an electronically excited state the donor-hydrogen bond must get weakened and the acceptor-hydrogen bond strengthened, until eventually the proton is fully transferred to the acceptor, forming a new chemical bond (Fig. 1.4).

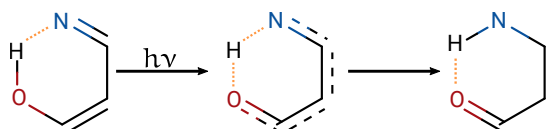


Figure 1.4.: Qualitative reaction scheme of an ESIPT with a six-membered ring intermediate structure. The dotted lines indicate hydrogen bonds; the dashed lines indicate the intermediate electron densities.

The critical part of the reaction is the initial strengthening of the hydrogen bond in the electronically excited state and if this results in overcoming the dissociation barrier – if there is one present at all – of the donor-hydrogen bond. Experimentally, this process is displayed at first in a red-shift of the corresponding frequency of the donor-hydrogen stretch vibration after photoexcitation, before it eventually disappears entirely.

Because protons are easily influenced by the polarity and proticity of the environment, e.g., solvent or tissue – when it comes to biological environments –, ESIPT systems are useful as probes by detecting changes in their optical properties. For example, in rigid media like polymers, ESIPT reactions exhibit high fluorescence quantum yields (0.1 to 0.6),^[37] i.e., how many fluorescence photons are obtained per excitation photon. In liquid phase, though, low fluorescence quantum yields are found,^[38–42] indicating that there are accessible non-radiative deactivation pathways for ESIPT systems through a CoIn, induced by larger geometric changes that are hindered when embedded in an environment. For all systems studied in this work, these larger geometric changes are rotations of a certain part of the molecule – which is basically again a photoisomerization (Fig. 1.5). This subsequent part after the ESIPT allows for more sophisticated setups of the molecule, e.g., when an additional proton acceptor site is introduced. In such a case, the proton could be transported over a distance that is farther than that of only one chemical bond. The proposed reaction scheme of 7-hydroxy-4-methylquinoline-8-carbaldehyde (HMQCA) serves as an example for these coupled processes (Fig. 1.6).^[43] In the context of molecular machines, such a system can be compared to a simple crane, transporting a proton from one end of the molecule to the other.

While ESIPT systems alone are well studied, ESIPT-crane systems are relatively new, promoted by the theoretical work of Sobolewski and co-workers^[43–46] and experimental studies.^[47] Recent experiments on two ESIPT-crane systems employed time-resolved infrared spectroscopy, in which the bond-breaking and creation can be easily observed.^[48,49] This technique is as close as it comes to a reaction microscope; although single molecules can still not be resolved directly, the ensemble-averaged change of the binding pattern can. Depending on the change of the electronic properties after the ESIPT, time-resolved absorption spectroscopy may also be suited for this task.^[50]

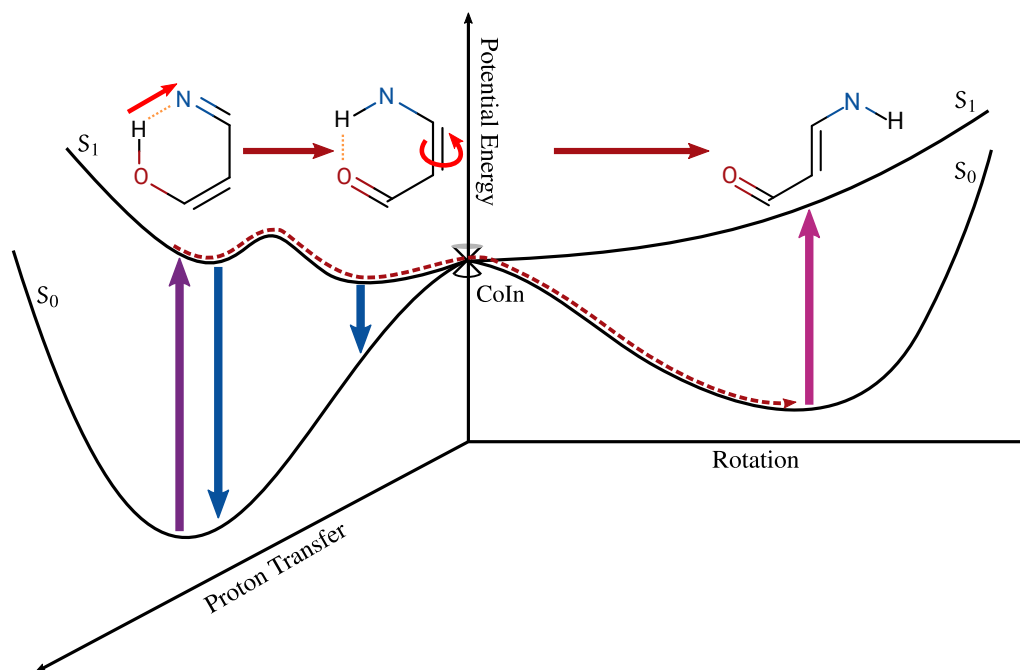


Figure 1.5: Qualitative scheme of an ESIPT and subsequent rotation. After vertical electronic excitation (VE, upward arrows) from S_0 to S_1 the system first follows the negative gradient (dashed line) towards the minimum of the PT structure over an energy barrier. Unless this structure is confined by another barrier, it may travel through a conical intersection (CoIn) *via* rotation of a “proton crane” group; otherwise, if either the system after VE or after PT are stable minima, they may return to S_0 by fluorescence (downward arrows). Photoexcitation from the rotated conformation may reverse the reaction process to recover the initial structure.

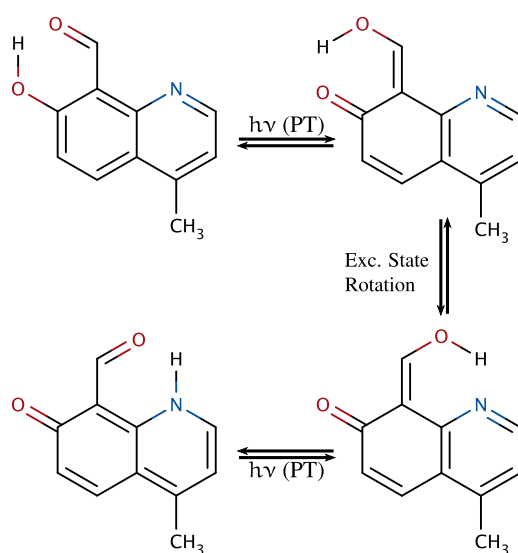


Figure 1.6: Reaction scheme of 7-hydroxy-4-methylquinoline-8-carbaldehyde (HMQCA): After the ESIPT, the protonated aldehyde group is able to rotate making it possible to transfer a proton to a second acceptor site (quinoline nitrogen).

1.1.3 Photochemistry on the Computer

As already mentioned, (photo)chemistry can also be studied on the computer. For this task, there are nowadays a variety of methods and techniques for a chemist to choose from.

The most basic technique is to calculate the energy of reactants and products in order to determine relative energies. These sole numbers allow for a prediction for which “side” of the reaction is the more stable one. However, this neglects the fact that reactions typically need to overcome energy barriers. Hence, with the height of the barrier, there are at least three energies needed for information on *if* a reaction can happen as proposed.

The *how* is a much more difficult question to answer. Because this also requires information of everything that happens in between the three points.

There are two basic approaches for this, depending on *what* is happening in the reaction. If the reaction is rather simple, i.e., it is obvious what the main reaction coordinate is, then changing this coordinate and gathering the energy along the way (“scanning”) is the typical approach. For a more complicated reaction which involves movements/rearrangements of many atoms at once, one usually starts at the top of the energy barrier, i.e., the transition state, and follows the steepest way down to either the reactant or the product side⁸. The latter technique yields the minimum energy path (MEP) or intrinsic reaction coordinate (IRC) path. Scanning, although also yielding a path, must be not necessarily yield an MEP because of the added constraint to the followed coordinate. The three presented techniques – energy calculations, MEP/IRC, scans – are established standard tools in computational chemistry when it comes to reactions⁹.

However, this still neglects that “real-life” molecules are constantly moving and thus all experimentally obtained properties are a result of this movement. Calculating the molecular dynamics (MD) is the fourth and most important tool in studying reactions on the computer. Nevertheless, it also the most rarely used, because reactions take time. And the longer this time is, the longer the computation gets as well, because the structural changes are calculated for small, discrete time steps¹⁰.

When it comes to photochemistry, it is advantageous that such reactions are typically much faster. But they also include electronically excited states that need to be investigated in addition to the ground state. The latter part is what makes MD simulations of photodynamics complicated and (computationally) demanding.

Here, full-dimensional MD, i.e., allowing the movement of all atoms in the molecule, of the two aforementioned types of photochemical reactions, *trans*⇌*cis* isomerizations and ESIPT, will be presented. For all but two of the ESIPT systems, this is the first time this has been done.

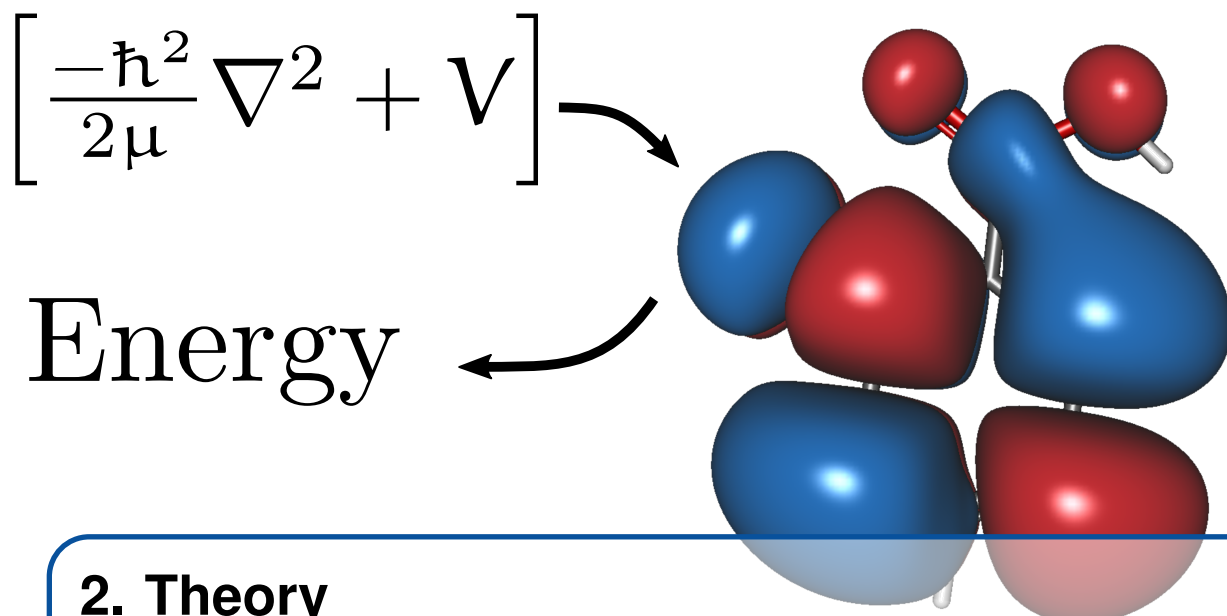
Besides the techniques, methods need to be chosen that can treat the electronic ground state as well as any excited states; but ideally with the constraint of relatively low computational demand.

⁸ For more detailed information the author refers the reader to textbooks on computational chemistry, e.g., Ref. 51.

⁹ Geometry optimizations are probably more common than MEP/IRC or scans, yet they also only yield an energy, albeit for an optimized geometry.

¹⁰ For example, all MD calculations in this thesis were performed with a time step of 0.1 fs.

In the following chapter, the theory behind the methods used throughout the thesis will be introduced, to an extent that highlights the advantages and disadvantages for each of it. After that, the results for the photoisomerization systems will be presented in Chapter 3, followed by the results of ESIPT(-crane) switches in Chapter 4. Means of optimizing parameters for the employed semiempirical quantum mechanical method will be discussed in Chapter 5. Supervised projects during the author's studies are summarized in Chapter 6. Finally, an outlook of future challenges and upcoming advances in method development and hardware utilization will be given in Chapter 7.



2. Theory

The basis of theoretical and computational chemistry is Schrödinger's quantum mechanics (QM). The central equation of this theory is the Schrödinger equation,

$$\hat{H}(\mathbf{r}, \mathbf{R})\Psi(\mathbf{r}, \mathbf{R}) = E\Psi(\mathbf{r}, \mathbf{R}), \quad (2.1)$$

here written in its time-independent form, with $\hat{H}(\mathbf{r}, \mathbf{R})$ being the Hamilton operator, $\Psi(\mathbf{r}, \mathbf{R})$ the total wave function (eigenfunction) of a system and E the energy of the system (eigenvalue), and the positions of the electrons (\mathbf{r}) and nuclei (\mathbf{R}).

The non-relativistic Hamilton operator is a sum of the kinetic (\hat{T}) and potential ($\hat{V} = V(\mathbf{r}, \mathbf{R})$) energy operators of the system:

$$\hat{H}(\mathbf{r}, \mathbf{R}) = \hat{T} + V(\mathbf{r}, \mathbf{R}). \quad (2.2)$$

Those operators can be further expanded into the kinetic energy operators of the nuclei (\hat{T}_{nu}) and electrons (\hat{T}_{el}), and the potential energy operators of the nuclear-nuclear ($V_{nu}(\mathbf{R})$) and electron-electron repulsion ($V_{el}(\mathbf{r})$) and nuclear-electron attraction ($V_{nu,el}(\mathbf{r}, \mathbf{R})$) (Eq. 2.3).

$$\hat{H}(\mathbf{r}, \mathbf{R}) = \hat{T}_{nu} + \hat{T}_{el} + V_{nu}(\mathbf{R}) + V_{el}(\mathbf{r}) + V_{nu,el}(\mathbf{r}, \mathbf{R}) \quad (2.3)$$

Calculating the energy of a system with the total Hamiltonian can be achieved analytically only for some very simple model systems. Approximations and/or numerical calculations are required for any typical chemical system with many electrons.

The first approximation – commonly known as the Born-Oppenheimer approximation^[52] – to tackle the solution of the Schrödinger equation is to assume electron movement at fixed nuclei

positions, which results in $\hat{T}_{nu} = 0$ and $V_{nu}(\mathbf{R})$ being constant. This is equivalent to re-formulating the wave function $\Psi(\mathbf{r}, \mathbf{R})$ as a product of an electronic wave function ($\psi(\mathbf{r})$) and a nuclear wave function ($\Xi(\mathbf{R})$).^[53] Hence, only the electronic part of the Hamiltonian may be considered at first, which needs only an electronic eigenfunction ($\psi(\mathbf{r}, \mathbf{R})$) and yields the electronic energy $E_{el}(\mathbf{R})$ of the system (Eq. 2.4).

$$\hat{H}_{el}(\mathbf{r}, \mathbf{R})\psi(\mathbf{r}) = E_{el}(\mathbf{R})\psi(\mathbf{r}) \quad (2.4)$$

$E_{el}(\mathbf{R})$ is the energy that is usually called and drawn as the potential energy surface¹ (PES) on which the nuclei move.

In the following, means of calculating $E_{el}(\mathbf{R})$ and concepts of further approximations will be presented. For in-depth descriptions, the author is referring to textbooks on theoretical and computational chemistry^[51, 54, 55] and the respective publications for each method.

2.1 Hartree-Fock Theory

The total wave function $\Psi(\mathbf{r}, \mathbf{R})$ and even the N -electron wavefunction $\psi(\mathbf{r})$ are the great unknowns in quantum chemistry. In a bottom-up approach, $\psi(\mathbf{r})$ may be expressed in N one-electron wave functions $\phi_i(\mathbf{r})$, or molecular orbitals (MOs). Hartree's *ansatz* for this expression is a simple product,^[56, 57]

$$\psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = \phi_1(\mathbf{r}_1)\phi_2(\mathbf{r}_2)\cdots\phi_N(\mathbf{r}_N). \quad (2.5)$$

This expression would allow an exact solution of the SE if it were not for the electron-electron repulsion term $V_{el}(\mathbf{r})$ in the Hamiltonian.^[58] To solve this, $V_{el}(\mathbf{r})$ is replaced by an effective potential $V_{el}^{eff}(\mathbf{r})$,^[56, 57] in which one electron interacts with the field of all other electrons. A direct solution is now no longer possible, because $V_{el}^{eff}(\mathbf{r})$ depends on the solution of the electronic Schrödinger equation within this field. Applying this potential energy operator results in a new $\psi(\mathbf{r})$ which has an updated V_{el}^{eff} . This procedure needs to be iterated several times until an effective electron field is achieved that does not change $\psi(\mathbf{r})$ as an approximation to the eigenfunction of the Schrödinger equation. The procedure is therefore called the self-consistent field (SCF) method, which has been the foundation of most quantum chemical methods for decades.

The Hartree product *ansatz* also has flaws regarding the antisymmetry property of wave functions, which was fixed by Slater by using a determinant instead of a simple product,^[59, 60]

$$\Phi(\mathbf{r}) = \psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(\mathbf{r}_1) & \phi_2(\mathbf{r}_1) & \cdots & \phi_N(\mathbf{r}_1) \\ \phi_1(\mathbf{r}_2) & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \phi_1(\mathbf{r}_N) & \cdots & \cdots & \phi_N(\mathbf{r}_N) \end{vmatrix}. \quad (2.6)$$

¹ Actually, the PES also contains the nuclear repulsion energy $V_{nu}(\mathbf{R})$. However, the calculation of $V_{nu}(\mathbf{R})$ is a simple Coulomb repulsion between the positively charged nuclei and also method-independent.

2.1 Hartree-Fock Theory

The electronic Hamilton operator² is the sum of a one-electron operator \hat{h}_i ,

$$\hat{h}_i = \hat{T}_{i,el} - \sum_{\alpha}^{nuclei} \frac{Z_{\alpha}}{|\mathbf{R}_{\alpha} - \mathbf{r}_{i\alpha}|}, \quad (2.7)$$

and a two-electron operator \hat{g}_{ij} ,

$$\hat{g}_{ij} = \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}. \quad (2.8)$$

Fock has shown that when doing the calculus with this Hamilton operator on a many-electron wave function, the two-electron operator is split into two terms: one describing the Coulomb interaction as expected, hence called Coulomb term (\hat{J}), and one that arises from taking care of the permutation, hence called exchange term (\hat{K}).^[61]

Combining the approaches of Hartree and Fock leads to the Hartree-Fock (HF) method, in which the electronic energy of an MO (ε_i) is obtained by applying the Fock operator \hat{f}_i ,

$$\hat{f}_i = \hat{h}_i + \sum_i^N (\hat{J}_i - \hat{K}_i), \quad (2.9)$$

to each of the MOs,

$$\hat{f}_i \phi_i(\mathbf{r}) = \varepsilon_i \phi_i(\mathbf{r}). \quad (2.10)$$

As stated before, each independent calculation of ε_i also needs information of all other MOs $\phi_{j \neq i}$ because of the two-electron operators in \hat{f}_i , which is why the MOs need to be calculated iteratively using the SCF procedure. The total electronic HF energy (E_{el}^{HF}) – using the Bra-Ket notation – is given as

$$E_{el}^{HF} = \sum_i^N \langle \phi_i | \hat{h}_i | \phi_i \rangle + \frac{1}{2} \sum_{ij}^N (\langle \phi_j | \hat{J}_i | \phi_j \rangle - \langle \phi_j | \hat{K}_i | \phi_j \rangle). \quad (2.11)$$

This all neglects the problem that even the MOs are unknown for non-model systems, which is why $\phi(\mathbf{r})$ needs to be broken down even further. The next smaller step after a molecule is an atom, so an MO can therefore be expressed in the so called linear combination of atomic orbitals (LCAO) $\chi(\mathbf{r})$ (Eq. 2.12).

$$\phi_i(\mathbf{r}) = \sum_{\nu}^{\Pi} C_{i\nu} \chi_{\nu}(\mathbf{r}), \quad (2.12)$$

with C_{ν} being the so-called MO coefficients. It is noted that Eq. 2.12 is only exact when $\Pi = \infty$. When Π is finite, this becomes an approximation and introduces the so-called basis set error. The $\chi_{\nu}(\mathbf{r})$ are typically not “true” (hydrogen-like) AOs³, therefore these functions are simply called basis functions (BFs). Selecting BFs is the crucial point for any quantum chemical calculation, because the better the BFs are in describing the MOs, whether it being optimized in shape or just in their number used, the better the many-electron wavefunction and the smaller the basis set error.

² For the remainder of this thesis all mathematical expressions are given in atomic units, hence many physical constants become unity. The spin of electrons will be also ignored in the following. For proper inclusion of spin, the author refers the reader to the aforementioned textbooks on QM.

³ The term AO is only used for historical reasons (cf. Refs. 62, 63).

Inserting the LCAO *ansatz* into Eq. 2.10 leads to

$$\hat{f}_i \sum_{\nu}^{\Pi} C_{i\nu} \chi_{\nu} = \epsilon_i \sum_{\nu}^{\Pi} C_{i\nu} \chi_{\nu}. \quad (2.13)$$

and the total HF energy changes to $E_{el}^{LCAO-HF}$

$$\begin{aligned} E_{el}^{LCAO-HF} &= \sum_i^N \sum_{\alpha\beta}^{\Pi} C_{i\alpha} C_{i\beta} \langle \chi_{\alpha} | \hat{h} | \chi_{\beta} \rangle \\ &+ \frac{1}{2} \sum_{ij}^N \sum_{\alpha\beta\gamma\delta}^{\Pi} C_{i\alpha} C_{i\beta} C_{j\gamma} C_{j\delta} (\langle \chi_{\alpha} \chi_{\gamma} | \hat{g} | \chi_{\beta} \chi_{\delta} \rangle - \langle \chi_{\alpha} \chi_{\gamma} | \hat{g} | \chi_{\delta} \chi_{\beta} \rangle) \end{aligned} \quad (2.14)$$

Applying the scheme of Roothaan^[64] simplifies Eq. 2.14 to a matrix-eigenvector problem

$$\mathbf{FC} = \mathbf{SC}\epsilon \quad (2.15)$$

with \mathbf{F} the Fock matrix,

$$F_{\alpha\beta} = \langle \chi_{\alpha} | \hat{f} | \chi_{\beta} \rangle = \langle \chi_{\alpha} | \hat{h} | \chi_{\beta} \rangle + \sum_i^{N/2} \sum_{\gamma\delta}^{\Pi} C_{j\gamma} C_{j\delta} (\langle \chi_{\alpha} \chi_{\gamma} | \hat{g} | \chi_{\beta} \chi_{\delta} \rangle - \langle \chi_{\alpha} \chi_{\gamma} | \hat{g} | \chi_{\delta} \chi_{\beta} \rangle), \quad (2.16)$$

and \mathbf{S} the overlap matrix of the BFs,

$$S_{\alpha\beta} = \langle \chi_{\alpha} | \chi_{\beta} \rangle, \quad (2.17)$$

and \mathbf{C} the vector containing the MO coefficients.

2.2 Semiempirical Hamiltonians

The computationally expensive part of the HF method are the two-electron integrals of \hat{g}_{ij} , which scale to the fourth power with the number of basis functions used (Π^4). To treat bigger systems with more atoms one has to come up with more approximations.

In chemistry, usually only the valence electrons take part in a reaction and the core electrons are mostly unaffected by any change in electron density induced by atom rearrangement. Hence, only taking the valence electrons into account reduces the amount of electrons that must be taken care of in the Fock equations. To compensate the neglected core electrons the nuclear charge is decreased.

In a next step the number of BFs per nucleus is pre-defined. For most compounds in (organic) chemistry it is sufficient to use a minimal basis set of s- and p-type orbitals, resulting in one BF for every hydrogen atom and four for every second and third row element. To increase the accessible compound space, d-type orbitals may also be considered.

At last, the product of BFs of the same electron but on different atoms are set to zero, which is called the zero differential overlap approximation. Because of this, many one-electron and

2.3 Correlation Methods

two-electron integrals can be neglected, thus greatly reducing the computational demand of the method. Further, the remaining integrals are substituted by analytical functions and parameters which are much simpler to evaluate. For example, the two-electron integral between two electrons in s-type orbitals is given as

$$\langle ss|s's' \rangle = \sqrt{R_{AB}^2 + \frac{1}{4} \left(\frac{1}{G_{ss,A}} + \frac{1}{G_{ss,B}} \right)^2}, \quad (2.18)$$

with R_{AB} being the distance between nuclei A and B , and G_{ss} their parameter for the s-type orbital repulsion.^[65,66]

When applying these approximations to the HF method, we enter the class of the semiempirical QM (SEQM) methods. All approximations lead to errors that can be compensated by optimizing the parameters within the newly added functions. Further parameters include, e.g., the remaining one-electron integrals or the shape of the basis functions for each atom – for a complete list of the parameters and functions for one type of semiempirical Hamiltonian see Sec. 5.1.

In this work, the Austin Model 1 (AM1) Hamiltonian and parametrization^[67] is used, as well as two reparametrizations: One as a general update to the parameters – called Recife Model 1 (RM1)^[68] –, and one that was made specifically for azobenzenes.^[69] Examples for a reparametrization and additional changes to the underlying Hamiltonian are the Parametric Models of generation x (PM x , $x = 3, 5, 6, 7$). For example, the newest model, PM7,^[70] utilizes optimized parameters as well as Grimme’s dispersion correction (see below).

Because of the parametrization, semiempirical methods are no longer “*ab initio*” like HF, where all information is calculated without any presumptions of the system and solely with natural constants. Semiempirical methods may therefore be unsuited for a given system if it was not taken into consideration during the initial parametrization process. But if a system was part of such a training set, semiempirical methods can outperform HF and possibly even higher-level methods (see below). The computational demand of semiempirical methods scales with Π^3 . However, given that only 4 BFs per second or third row element are used, the effect of increasing the size of the system is less drastic when compared to typical basis sets for QM methods which employ 14 (def2-SVP), 29 (def2-QZVP) or even more BFs for the carbon atom.^[71]

2.3 Correlation Methods

In the aforementioned HF and SEQM methods only one Slater determinant (SD) (cf. Eq. 2.6) is used as the total electronic wavefunction $\psi(\mathbf{r})$. One SD may be seen as the mathematical representation of one electronic configuration of the system. The HF SD will be referenced in the following as $\Phi_0(\mathbf{r})$. Ideally, one electronic configuration would correspond to one *electronic state*⁴ (S_i) and each electronic state has its unique PES.

⁴ Practically, it is a linear combination of multiple configurations.

Of course there are more than only one electronic state and also interactions between them in a molecule. When there is a large⁵ energy difference between two electronic states, this interaction may be neglected, which is the essential part of the second Born-Oppenheimer approximation. In such a case $\Phi_0(\mathbf{r})$ may be a good approximation to $\psi(\mathbf{r})$, yielding an energy E_0 that may be up to 99% of the exact energy. The missing 1% in energy may be accounted for by the so called *electron correlation energy*. This energy comes from the instantaneous interaction of electrons that is neglected in HF by not using the explicit two-electron operator. There are two options to recover this missing energy: i) To consider more electronic configurations, i.e., more SD, or ii) to fix the error of not using the exact two-electron operator \hat{g}_{ij} that led to an approximated electron-electron interaction.

Starting off with the first option is pretty straightforward: To get more electronic configurations, electrons need to be moved from occupied orbitals to unoccupied – or virtual – orbitals in $\Phi_0(\mathbf{r})$. If all combinatorial possibilities are explored we end up in the full configuration interaction (FCI) method, in which all possible SDs are considered. Depending on the number of electrons moved, the new SDs may be grouped as “singly excited” SDs ($\Phi_i^a(\mathbf{r})$) in which one electron is moved from MO i to MO a , or “doubly excited” ($\Phi_{ij}^{ab}(\mathbf{r})$) and so on. The FCI wave function $\psi_{FCI}(\mathbf{r})$ may then be written as a linear combination of these new SDs:

$$\psi_{FCI}(\mathbf{r}) = C_0^{CI} \Phi_0(\mathbf{r}) + \sum_i C_i^{CI} \Phi_i^a(\mathbf{r}) + \sum_{ij} C_{ij}^{CI} \Phi_{ij}^{ab}(\mathbf{r}) + \dots \quad (2.19)$$

Eq. 2.19 is another basis set expansion, in which the BFs are now the SDs. In order to get $\psi_{FCI}(\mathbf{r})$ the expansion coefficients C^{CI} , which are called CI coefficients, need to be calculated.

As it was the case before, a basis set expansion is only exact if an infinite amount of basis functions is used. For FCI that would mean that there needs to be an infinite amount of electronic configurations, which is only the case if there is an infinite amount of combinatorial possibilities for moving electrons to orbitals. Because the number of electrons is definitely not infinite in a system, only the number of orbitals may be so. But since the number of basis functions dictate the number of orbitals, it is impossible to perform an exact FCI calculation. In practice, FCI is possible because basis sets use a finite amount of BFs but this method is still too computationally expensive for typical molecules⁶.

To make CI feasible the number of determinants can be restricted based on the excitation level. In that case, a CI method only utilizing $\Phi_0(\mathbf{r})$ and all derived singly excited determinants may be called CIS; if also all doubly excited determinants are used CISD, and so on. The computational demand of CI methods scales at least with Π^5 for CIS and increases drastically with the excitation level, e.g., Π^{10} for CISDTQ.

⁵ Note that “large” may be not a quantifiable term.

⁶ For example, an FCI calculation for the N_2 molecule with 10 electrons in 34 MOs already leads to $9.68 \cdot 10^9$ Slater determinants.^[72]

2.3 Correlation Methods

A different *ansatz* than adding more SDs to the system is to fix the error (Δ_{corr}) of using the HF Hamiltonian (\hat{H}_{HF}) – which is the sum of all Fock operators – instead of the exact Hamiltonian (\hat{H}_{exact}) (Eq. 2.20).

$$\Delta_{corr} = \hat{H}_{exact} - \hat{H}_{HF} \quad (2.20)$$

This can be done by applying perturbation theory under the assumption that Δ_{corr} is relatively small. Starting from the *ansatz* that the wave function may be expanded as

$$\psi(\mathbf{r}) = \lambda^{(0)}\psi^{(0)}(\mathbf{r}) + \lambda^{(1)}\psi^{(1)}(\mathbf{r}) + \dots, \quad (2.21)$$

the energy becomes

$$E = \lambda^{(0)}E^{(0)} + \lambda^{(1)}E^{(1)} + \dots. \quad (2.22)$$

By inserting Eq. 2.21 into the Schrödinger equation, we eventually⁷ end up with general expressions to calculate the correction energy $E^{(n)}$.

The first-order correction using this *ansatz* would yield the same energy as HF. Starting from the second-order correction ($E^{(2)}$) we gain an improved energy that includes some of the missing electron correlation energy:

$$E^{(2)} = \sum_{i>j}^{N/2} \sum_{a>b}^{\Pi-N/2} \frac{\langle \Phi_0 | \hat{H}^{(0)} | \Phi_{ij}^{ab} \rangle \langle \Phi_{ij}^{ab} | \hat{H}^{(0)} | \Phi_0 \rangle}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b}, \quad (2.23)$$

with $\hat{H}^{(0)}$ being the unperturbed, zeroth-order Hamiltonian, i.e., \hat{H}_{HF} , Φ_0 the SD obtained from HF, Φ_{ij}^{ab} the double-excited determinants and ϵ_X the HF energy of the corresponding MO ϕ_X – so basically all the information that is available from an HF calculation. The part that comes on top of the HF calculation are the integrals of the numerator in Eq. 2.23, which are the computationally expensive part of the calculation, scaling with the fifth power of the number of basis functions used (Π^5). The application of perturbation theory to QM was done by Møller and Plesset,^[73] thus the method is called MPn, with n being the order of correction employed. For example, Eq. 2.23 gives the MP2 energy E^{MP2} .

Because both classes of methods presented for recovering correlation energy depend on only $\Phi_0(\mathbf{r})$, they are called single-reference methods. They only work well when $\Phi_0(\mathbf{r})$ is already a sufficient approximation to $\psi(\mathbf{r})$.

CI or MPn methods do not optimize the MOs within $\Phi_0(\mathbf{r})$. But when the system evolves on its PES, it may eventually get to points where there is strong coupling to a higher electronic state. To describe any other electronic state, it would need its own set of optimized MOs and thus new SDs $\Phi_i(\mathbf{r})$. However, there is also a special class of methods that does exactly that, which is called multi-configuration SCF (MCSCF). And in the same fashion as the HF wave function was improved by CI or MPn methods, the MCSCF $\Phi_i(\mathbf{r})$ – with each having its own optimized set of MOs – may also be corrected by either setting up additional excited determinants from multiple

⁷ The details of how to get the corrected energies are not of importance to understand the computational demand.

SDs as reference, or by calculating the energy corrections using perturbation theory. The former method is called multi-reference CI (MRCI) and the latter complete active space perturbation theory of order n (CASPT n), when applying the n th order of correction to a subset of “active” orbitals. These orbitals and their electrons are usually the ones being easily excited and not staying passive during the reaction, in contrast to core electrons and orbitals. If all combinations of electrons and orbitals within this selection are considered, it is called complete active space (CAS). Using this approach would also allow for truncated MRCI expansions, i.e., CAS-CIS or CAS-CISD. The CAS approach can also be applied to MCSCF, which is called CASSCF.

Multi-reference methods are highly accurate means of solving the Schrödinger equation for multiple electronic states, but therefore also even more computationally demanding than their single-reference counterparts. In addition, multi-reference methods require careful setup by choosing an optimal active space, and this choice is not easy to automate.

With the so called floating-occupation CI (FOCI) approach, a multi-reference wave function can be approximated without actually performing a real multi-reference *ansatz*. Usually, orbitals can be occupied by zero, one or two electrons. In the FOCI approach, the electrons within a given active space – which consists of both fully-occupied and non-occupied orbitals at first – is redistributed over all orbitals following a Gauss-like curve based on the energies of the MOs^[74] (Fig. 2.1). The width of the curve can be set by a parameter.

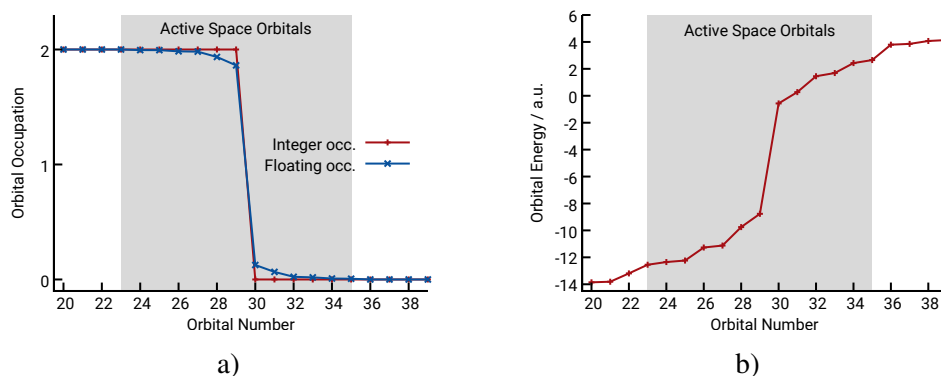


Figure 2.1.: Example of the FOCI approach to an active space consisting of 12 electrons in 12 orbitals ([12,12]) and an orbital-energy width parameter of 0.1. a) “Standard” integer occupation in comparison to the floating-occupation orbital-occupation numbers. The orbitals outside the active space are unaffected by the FOCI approach. b) The energies for the corresponding orbitals.

This approach allows taking some of the former non-occupied orbitals into consideration for the wave function in the electronic ground state without the need to set up any excited SDs, therefore making FOCI not nearly as computationally demanding as a full-fledged MRCI calculation. Coupling this approach to semiempirical methods allows for low-cost, multi-reference-like calculations, which can perform on the same level as CASSCF.^[74] A downside of this approach would be that the intention of getting the best single SD representation of the electronic ground state – by means of optimizing the MO coefficients with the SCF – is no longer satisfied; in exchange for

2.3 Correlation Methods

MO coefficients that are somewhat optimized for excited states, thus improving the excited state determinants which makes FOCI superior to conventional (truncated) CI expansions.

By calculating the energy for each SD the energy difference between the two electronic states for the exact same nuclear coordinates can be obtained, which is called vertical excitation energy (VEE, $\Delta E_{S_i \rightarrow S_j}$).

In the same fashion as semiempirical Hamiltonians are an approximation to HF, there are methods that use approximative methods for calculating excitation energies to lower the computational demand. One example for this is the second-order coupled cluster (CC2) method. Coupled cluster (CC) is another method that generates excited SDs from a reference wave function and follows the same naming scheme as CI methods, e.g., CCSD for CC with single and double excitations. The only difference is how these are generated: If the excitation levels are written as a cluster operator \hat{T}_{CC} , the CI wave function ψ_{CI} may be written as

$$\psi_{CI} = \hat{T}_{CC} \Phi_0, \quad (2.24)$$

the CC wave function Ψ_{CC} , in comparison, as

$$\Psi_{CC} = e^{\hat{T}_{CC}} \Phi_0, \quad (2.25)$$

both with

$$\hat{T}_C = \hat{T}_1 + \hat{T}_2 + \dots \quad (2.26)$$

Instead of calculating CI coefficients, CC calculates cluster amplitudes t ,

$$\hat{T}_1 = \sum_{ia} t_a^i \hat{\tau}_i^a, \quad \hat{T}_2 = \sum_{ijab} t_{ab}^{ij} \hat{\tau}_{ij}^{ab}, \quad (2.27)$$

with $\hat{\tau}$ being the excitation operators.

In this picture the MP2 energy – for the sake of comparison (cf. Eq. 2.23) – may be written as

$$E^{MP2} = \langle \Phi_0 | (\hat{H}_{exact} - \hat{H}_{HF}) \hat{T}_2 | \Phi_0 \rangle = \sum_{ijab} t_{ab}^{ij} \langle \Phi_0 | \hat{H}_{exact} - \hat{H}_{HF} | \Phi_{ij}^{ab} \rangle, \quad (2.28)$$

with

$$t_{ab}^{ij} = \frac{\langle \Phi_{ij}^{ab} | \hat{H}_{exact} - \hat{H}_{HF} | \Phi_0 \rangle}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b}. \quad (2.29)$$

Using response theory – see below – on Eq. 2.28 gives us the aforementioned CC2 energy

$$E^{CC2} = \left\langle \Phi_0 \left| (\hat{H}_{exact} - \hat{H}_{HF}) \left(\hat{T}_2 + \frac{1}{2} \hat{T}_1 \hat{T}_1 \right) \right| \Phi_0 \right\rangle. \quad (2.30)$$

The difference between the MP2 energy (Eq. 2.23) and the CC2 energy (Eq. 2.30) can be found in the operator, where CC2 has an additional product of single-excitation cluster operators that need to be evaluated. This addition allows calculation of single-type excitations for VEE at an increased

computational demand of Π^6 . But in contrast to multi-reference methods, CC2 does not require any additional input on the amount or type of orbitals to be considered as active space, therefore making it rather popular.

Response theory – in short, for details see Ref. 75 – is a means to get excited-state information from a time-dependent perturbation of the ground-state wavefunction. The “response” of this perturbation is usually truncated at first order, leading to linear response (LR) methods. A spectrum is obtained from a Fourier transformation of the response. The poles of the spectrum lie at the excitation energies of the system. Response theory allows for relatively easy calculation of excited states and their properties, in comparison to full-fledged multi-reference calculations, but with good accuracy only for low-lying electronic states and – in the case of LR methods – only for single-type excitations.

All of the aforementioned correlation methods suffer from a rather large computational demand, mainly because of the two-electron integrals that need to be calculated. To optimize this part of a calculation, one may employ the so called resolution of identity (RI) *ansatz*.

In short, the product of two BFs $\chi_\alpha\chi_\beta$ (cf. Eq. 2.16) may be approximated by an expansion in a series of auxiliary BFs χ_{aux} , reducing the former four-center integrals to three- and two-center ones,^[76] which reduces the scaling by at least one order. There are of course modern RI approximations, e.g., RI chain of spheres exchange for the Coulomb integrals (RIJCOSX),^[77] which utilize some more mathematical tricks to further reduce the scaling. The error introduced by this approximation is negligible^[76,78,79] (depending on the size of the auxiliary basis set, some meV), so that RI may be used in any calculation without additional benchmarking.

2.4 Density Functional Theory

All the previous methods have in common that they calculate the electronic energy using an unknown electronic wave function.

Although there are many ways to approximate ψ , it still has an unpleasant disadvantage: Every electron in the system increases the dimensionality of ψ by three dimensions – four if spin is included –, as

$$\psi = \psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots, \mathbf{r}_N). \quad (2.31)$$

To reduce the complexity of the problem, Hohenberg and Kohn have shown^[80] that using the density $n(\mathbf{r})$ of an N -electron system (lhs. of Eq. 2.32) allows the exact calculation of the energy of the electronic ground state as well as all excited states, while at the same time reducing the complexity down to a three-dimensional problem, regardless of the amount of electrons in the system. Hence, this method is density functional theory (DFT).

$$n(\mathbf{r}) = N \int |\psi|^2 d\tau. \quad (2.32)$$

2.4 Density Functional Theory

However, at the present day there is still no functional based on $n(\mathbf{r})$ that yields accurate results, because even though Hohenberg and Kohn proved the existence of a density functional, no one knows how to obtain it. The two problematic parts are the functionals for the kinetic energy and exchange-correlation (XC), which both need to be approximated.

Kohn and Sham reintroduced orbitals ($\psi^{KS}(\mathbf{r})$) to fix the kinetic energy,^[81] at the cost of losing the feature of a solely three-dimensional problem.

$$n(\mathbf{r}) = \sum_i |\psi_i^{KS}(\mathbf{r})|^2 \quad (2.33)$$

The summation in Eq. 2.33 also assumes non-interacting orbitals. The energy of each KS orbital is then calculated as

$$\left[-\frac{1}{2}\nabla^2 + \hat{V}_{nu,el}(\mathbf{r}, \mathbf{R}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\tau + E_{XC}(\mathbf{r}) \right] \psi^{KS}(\mathbf{r}) = \epsilon \psi^{KS}(\mathbf{r}), \quad (2.34)$$

in which an SCF-type of calculation is reintroduced, because $n(\mathbf{r})$ is needed to compute the orbitals which are needed to compute the density. Because this approach is more accurate than orbital-free DFT, Kohn-Sham DFT (KS-DFT) is nowadays usually what is meant when talking about DFT.

Modern KS-DFT is all about finding a good formulation of the exchange-correlation (XC) functional (E_{XC}), to add the missing many-electron interaction energy E_{XC} . There are DFT-only approaches to add E_{XC} by taking only $n(\mathbf{r})$ into account (local density approximation, LDA) or additionally the gradient of the density (generalized gradient approximation, GGA). However, one popular approach is to add the exchange energy from HF to the KS-DFT energy, thus coining the term ‘‘hybrid DFT’’. But this comes at the cost of an HF calculation in addition to LDA and/or GGA DFT calculations.

Considering this, many different hybrid functionals were created, with the most popular being B3LYP^[82] and PBE0.^[83] The difference between all hybrid DFT functionals is their parametrization. For example, the B3LYP XC-energy is given as

$$E_{XC}^{B3LYP} = E_x^{LDA} + a_0(E_x^{HF} - E_x^{LDA}) + a_x(E_x^{GGA} - E_x^{LDA}) + E_c^{LDA} + a_c(E_c^{GGA} - E_c^{LDA}), \quad (2.35)$$

with $a_0 = 0.2$, $a_x = 0.72$ and $a_c = 0.81$, whereas in comparison the PBE0 XC-energy is given as

$$E_{XC}^{PBE0} = \frac{1}{4}E_x^{HF} + \frac{3}{4}E_x^{PBE} + E_c^{PBE}. \quad (2.36)$$

The computational cost of such hybrid KS-DFT methods are mostly determined by the HF part.

DFT as it is, is only made for electronic ground-state calculations, but may be expanded by response theory to yield a ‘‘time-dependent’’ DFT (LR-TDDFT) method or be coupled to the random phase approximation (RPA-DFT),^[84] to also enable the calculation of VEE.

Most DFT functionals exhibit a systematically wrong description of van der Waals interactions. A simple, yet effective improvement was presented by Grimme's D3 correction (including Becke-Johnson (BJ) damping for long-range interaction).^[85,86] It introduces an additive force-field term to the total DFT energy and contains optimized parameters for most elements. Its introduction does not have any considerable impact on computational time, but increases the quality of the obtained energy, therefore it is recommended in any DFT calculation nowadays.

2.5 Force Fields

All the aforementioned methods can be summarized as "means to calculate the electronic energy". The calculation of this energy was presented as difficult and time-consuming, but – ironically – in some parts we already know the results before even starting the calculation – albeit only qualitatively.

Let us take any molecular covalent bond. As chemists we know without any calculation that there must be an energy minimum at some distance R_e . We also know that upon shortening or elongation the energy must rise. And if a bond is stretched too long, there will be dissociation and the energy will stay constant from that point on. Instead of employing the methods mentioned above we can simply fit a model to our empirical finding. In the case of a molecular bond, this could be the Morse potential, which is able to reproduce our guess. This model now basically allows us to effectively skip any calculations of electrons – and thus time and computer resources.

Of course, there are more components to consider in molecules than just covalent bonds, but with the same argumentation, by intuition, we can define more models for angles, dihedrals and non-covalent interactions.

Taking all of this into consideration leads to the so called force field (FF) or molecular mechanics (MM) methods. These methods do not take electrons into consideration but use models to reproduce the expected effects. A usual force field uses five terms to calculate the energy which handle the majority of effects in a molecule (Eq. 2.37).

$$\begin{aligned}
 E_{MM} = & \sum_{bond} \frac{1}{2} k_R (R - R_{eq})^2 \\
 & + \sum_{angle} \frac{1}{2} k_\alpha (\alpha - \alpha_{eq})^2 \\
 & + \sum_{dihed} \frac{1}{2} k_T [1 + \cos(n\omega + \gamma_{eq})] \\
 & + \sum_{i=1}^{N-1} \sum_{j=i+1}^N \left\{ \epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right\}.
 \end{aligned} \tag{2.37}$$

The prevalent non-covalent interactions are of Coulomb and van der Waals type, with the latter being described by a Lennard-Jones (LJ) potential.

2.5 Force Fields

However, there is a catch in all of this: In order to apply such a model to any system, it needs to be fitted. And to do this, there must be parameters that need to be tuned to yield a good agreement. This was already the challenge in the SEQM methods, but in contrast to those, the parameters in force fields are much more tangible: For the bonded interactions (the first three terms in Eq. 2.37) we need force constants (k_X) and equilibrium values (X_{eq}), and LJ parameters (ϵ, σ) and point charges q_i for the non-bonded ones.

Naturally, there are many force fields that have their own parametrization. There are general force fields with parameters for the whole range of elements in the periodic table,^[87] and special force fields that, e.g., only describe water.^[88] The area of application is only dependent on the parametrization. Then of course the terms in the energy expression are also not set in stone. The bond energies in Eq. 2.37 are expressed as harmonic oscillators – which is fine as long as no bond dissociations are of interest – but they can easily be exchanged with Morse potentials, which are more accurate from a chemical point of view. Such “improvements” can be done for every term, but also additional ones can be added, that, e.g., couple bond and angle terms.

Force fields are commonly used in biochemistry where large molecules, like proteins or DNA, and very long simulated times (up to milliseconds and longer), need to be considered. Such calculations are not possible with today’s electronic structure methods – including SEQM.

In the scope of photochemistry, electrons are needed to correctly describe electronic states. But force fields come in handy there, too, e.g, when simulating solvent molecules, for which the electronic structure is not of interest but mainly its sterical effect. Furthermore, QM and MM based methods can also be combined for the treatment of a single molecule if only one (relatively small) part of the molecule is of importance and the remaining part is just a passive bystander. This combination of methods is often abbreviated as QM/MM.

The “art” of QM/MM is to get the separation of QM and MM part as well as the interaction between them right (Eq. 2.38).^[89,90]

$$\hat{H}_{el} = \hat{H}_{QM} + \hat{H}_{QM/MM} + E_{MM} \quad (2.38)$$

The separation gets important, when the two parts are connected by a covalent bond. If the system is split there, the wave function is also effectively split. To make the wave function behave nicely at that point, it needs to converge to zero. Practically, this is done by transforming the QM frontier atoms to hydrogen atoms. This effectively cuts the bond from the QM point-of-view. To make the MM part still see the old bond information, this frontier atom gets the number of electrons and mass it had before. The remaining interaction is usually taken care of by non-bonding terms, like Coulomb or van der Waals potentials. The remaining energy of the MM part is calculated as shown in Eq. 2.37.

2.6 Molecular Dynamics

All methods so far yielded the PES of a (static) system, from which nearly all other molecular properties can be derived. But molecules in real-life are constantly moving, so calculating a property from one single structure of a molecule is not representative. At this point there are two possibilities. From physical chemistry we know that a system is usually found in an energy minimum, so, statistically, observed properties of a molecule are mostly obtained from an energetically stable structure. This means that the calculation of properties from only one structure is reasonable when the system is in a minimum on the PES.

To get there, the structure of the system, described by the positions of the nuclei (\mathbf{R}), needs to be manipulated in such a way that the energy decreases until no change in energy is observed (Eq. 2.39), or simply, until the gradient of the energy with respect to the coordinates becomes zero.

$$\frac{E_{el,1} - E_{el,2}}{\mathbf{R}_1 - \mathbf{R}_2} = \frac{\Delta E_{el}}{\Delta \mathbf{R}} \stackrel{!}{=} 0 \quad (2.39)$$

The second possibility is to actually propagate the dynamics of the system and then (or on-the-fly) calculate the properties at every structural change. For this, the gradient is also needed. This might raise the question, why this is not common practice. First, depending on the size of the system and the method, calculating $E_{el}(\mathbf{R})$ might get expensive in terms of computational time (see above); second, for a statistically meaningful result, many structural changes need to be calculated (see Fig. 2.2), which means more calculations of $E_{el}(\mathbf{R})$, and therefore more computational effort.

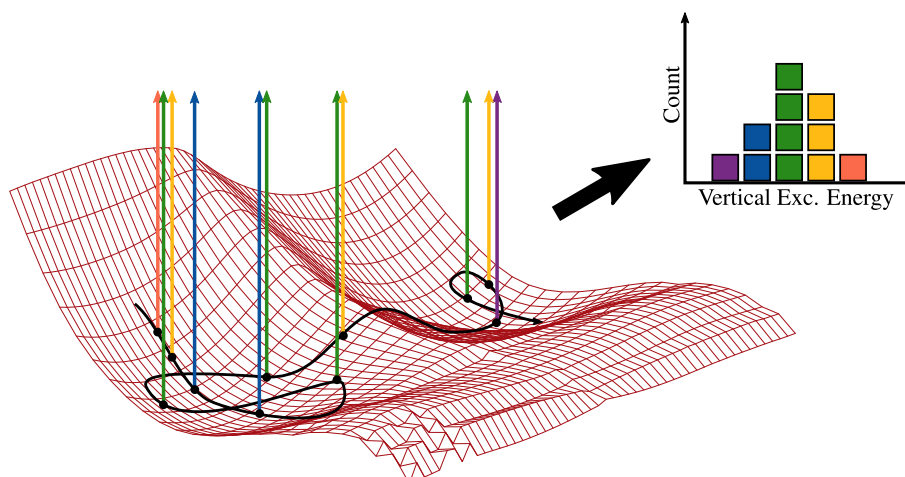


Figure 2.2.: Qualitative depiction of a trajectory (black line) evolving on a PES. At each position, exemplarily shown as black circles, properties of the system may be calculated, e.g., vertical excitation energies (arrows). Binning the excitation energies into discrete ranges would lead to a simulated UV spectrum.

This is why molecular dynamics (MD) are only performed with methods which allow for a relatively fast calculation of $E_{el}(\mathbf{R})$, i.e., force-fields, SEQM or DFT. But recent advances in hardware and method development also allows for MD with CASSCF.

The simplest algorithm for MD is the Verlet algorithm:^[91] At time step i the change of the structure, i.e., the structure at the next time step x_{i+1} , after a time step Δt can be calculated from the current (x_i) and previous (x_{i-1}) structures, the mass m and the current gradient of the energy (F_i) (Eq. 2.40).

$$x_{i+1} = 2 \cdot x_i - x_{i-1} + \frac{(\Delta t)^2 \cdot F_i}{m}. \quad (2.40)$$

The previous structure x_{i-1} can be derived from an initial velocity v , since

$$v = \frac{x_i - x_{i-1}}{\Delta t} \quad \curvearrowright \quad x_{i-1} = x_i - \Delta t \cdot v, \quad (2.41)$$

and F_i needs to be calculated from any QM or MM method (see below, Eq. 2.42). Temperature may be introduced by either tuning the initial velocities according to a Boltzmann distribution and/or scaling the velocities during the propagation down or up, respectively, to a specified kinetic energy.

A superior and more common variant to Eq. 2.40 is the velocity form of the Verlet algorithm.^[92]

In a classical picture, nuclei can be viewed as marbles rolling on a given surface. The shape of the surface is only dictated by the PES, which is the electronic energy of the system, according to the Born-Oppenheimer separation (see above). The dynamics of nuclei moving on only one PES are called *adiabatic dynamics*.

In the adiabatic picture, changing the PES, i.e., changing the electronic state, is also allowed if the transition is not induced by nuclear motion. However, photochemistry may involve CoIns, where there clearly are transitions between states, induced by nuclear motion. This is called *non-adiabatic dynamics*.

2.6.1 Non-Adiabatic Dynamics

Instead of solving the Schrödinger equation for the nuclei, in the semi-classical *ansatz*, the nuclei are only propagated according to Newton's Law of Motion, in which the force \mathbf{F} is given by the negative derivative of the current PES with respect to the nuclear coordinates (Eq. 2.42).

$$\mathbf{F} = - \frac{\partial}{\partial \mathbf{R}} E_i(\mathbf{R}) \quad (2.42)$$

In the case of many electronic states Eq. 2.42 changes to

$$\mathbf{F} = - \left\langle \psi_i(\mathbf{r}) \left| \frac{\partial}{\partial \mathbf{R}} \hat{H}_{el} \right| \psi_j(\mathbf{r}) \right\rangle + (E_j - E_i) \mathbf{d}_{ji}, \quad (2.43)$$

with \mathbf{d}_{ji} the non-adiabatic coupling vectors (NCV),

$$d_{ji} = \left\langle \psi_j \left| \frac{\partial}{\partial \mathbf{R}} \right| \psi_i \right\rangle. \quad (2.44)$$

Eq. 2.43 can be rewritten as

$$\mathbf{F} = -\sum |a_i|^2 \frac{\partial}{\partial \mathbf{R}} E_i + \sum_{ij} a_i^* a_j (E_j - E_i) \mathbf{d}_{ji}, \quad (2.45)$$

which in comparison to Eq. 2.42 shows that the force is dependent on the gradients of all PES at once, scaled by a factor $|a_i|^2$, which is the population of each electronic state. The a_i factors need to be computed on-the-fly by solving the time-dependent Schrödinger equation for the electrons at each time step during the dynamics.^[93]

Letting a system evolve on a PES using \mathbf{F} as described in Eqs. 2.43 and 2.45 would lead to an effective potential, because the system is influenced by all PES, even at places where coupling is unphysical. Practically, nuclei are under the influence of only one gradient, which is the one of the PES they are currently evolving on. Only after a change of the electronic state, the gradient changes.

In the so-called surface hopping method by Tully (TSH),^[94] the state populations $|a_i|^2$ define the probability that the nuclei are evolving on state i . The a_i coefficients are used in the density matrix formalism $\rho_{ij} = a_i^* a_j$. From the off-diagonal elements of this matrix and the NCV, transition probabilities $P_{i \rightarrow j}$ can be calculated. With $P_{i \rightarrow j}$, a change of state – or surface hop – is defined as

$$\sum_{i=1}^{j-1} P_{i \rightarrow j} \leq \mathfrak{R} \leq \sum_{i=1}^j P_{i \rightarrow j}, \quad (2.46)$$

with \mathfrak{R} being a random number between 0 and 1. For more details see Refs. 93 and 94.

The simplicity of TSH also allows to couple it to any method that can calculate multiple adiabatic states.^[95–98]

However, the standard TSH algorithm presents two problems. First, chances of many hops must be suppressed or else this would cause the nuclei to move again on an effective averaged PES. This problem is solved in the fewest switching TSH algorithm.^[99] Second, even though the transition probabilities are calculated on-the-fly – based on the actual density matrix at any given time – hops may occur also again at unphysical points on the PES because of the random nature of the algorithm. An example for this may be hops to an excited state when the system is already back in S_0 , oscillating near the energy minimum. Even though the energy difference between S_0 and S_1 , and therefore the elements of the density matrix, do not favor the transition, it may still happen.

A fix for the latter problem was proposed by Granucci and Persico with their so-called quantum decoherence correction (QDC).^[100] Surface hops at late stages of the propagation are due to missing decoherence in the semi-classical *ansatz*. In QM, there would be electron wave packets present on every electronic state, according to the state-specific population prescribed by the density matrix. A transition may only occur when there is an overlap between two wave packets. Since all wave packets are moving at the same time, a large overlap can only occur for an extended time when the momentum of each has the same direction. In the semi-classical picture these wave packets are approximated by the ensemble of trajectories, which are usually propagated individually without

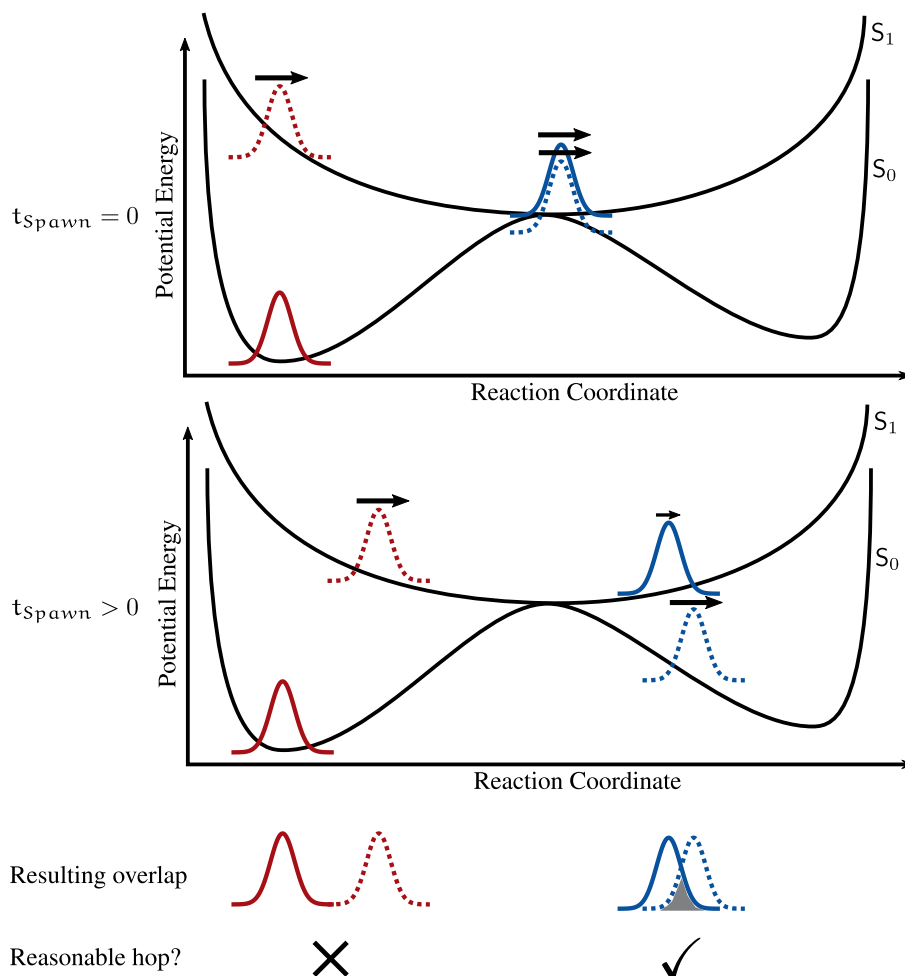


Figure 2.3.: Qualitative scheme of the quantum decoherence correction scheme using rigid Gaussians. Top: Two initial Gaussians (red and blue, solid lines) on their current surface, as well as their auxiliary Gaussians (dashed lines) at the same coordinates. Bottom: Position of the Gaussians after some propagation time and the resulting overlap. Arrows indicate the direction of momentum.

interaction. In the QDC this QM feature is mimicked by rigid Gaussian functions that are spawned at different time intervals during the propagation of a trajectory on every electronic state adjacent to the current. For an arbitrary amount of time steps, these Gaussians are propagated alongside the nuclei, each following its own gradient of the PES they were spawned. Afterwards, the overlap of each Gaussian pair is calculated. If it is above a given threshold and Eq. 2.46 is fulfilled, the hop may happen.

This process is graphically summarized in Fig. 2.3.

The QDC is i) straightforward to add to any currently implemented TSH code and ii) does not add much computational demand, because the gradients are computed anyways and the wave packets do not change their shape – which is, of course, an approximation.

For information on non-adiabatic dynamics besides the TSH approach the author refers to a recent review article.^[101]

2.7 Techniques

The main focus of the results presented in the following chapter will be on semiempirical quantum mechanical (SEQM) calculations utilizing FOCI as means of getting information on electronically excited states. All calculations of this type were performed with a development version of MOPAC2002^[102] that was extended by Granucci and Persico to allow for on-the-fly surface-hopping SEQM dynamics.

Non-SEQM calculations (DFT, MP2 and CC2) were mainly performed with the TURBOMOLE v7.0^[103] program package. These typically served as benchmarks for, e.g., relative energies or vertical excitation energies. Non-default settings for these calculations will be explicitly stated in the text. The three methods were always used with the RI approximation as implemented in the program.

In the following, further practical techniques on how to run and how to analyze the calculations are going to be presented.

2.7.1 Optimization Processes

Three types of optimizations were performed: Geometry optimizations (geoOpts), meta optimizations and parameter optimizations. All three types are fundamentally different as their names suggest, hence they shall be introduced briefly.

A geoOpt optimizes the molecular structure according to the negative gradient – derivative of the energy with respect to all nuclear coordinates – of its PES. The structure is updated iteratively according – but not limited to⁸ – the gradient until it eventually reaches a value of zero. GeoOpts can be performed to either find a minimum or a saddle-point/transition state of the PES, because in both cases the requirement of $\partial E / \partial \mathbf{R} = 0$ is fulfilled. Assuming that structural minima are of interest, a geoOpt may only find local minima that depend on the initial geometry (Fig. 2.4).

For most molecules this is not problematic, because the optimal structures can be guessed by chemical intuition, but for very large systems – see, e.g., the systems presented in Sec. 3.4 – this approach may fail.

Restricted geoOpts are also used as means to calculate one or two dimensions of a PES of the system. This is done by fixing one or two selected degrees of freedom (DOF), e.g., bond lengths, dihedral angles, etc., to a certain value and then performing a geoOpt (“let the system relax on its PES”) in all but these one or two DOF, respectively. From the relaxed structure, the DOF is then changed and another geoOpt is performed. This process is repeated from a given starting value for one or two DOF, respectively, up to a final value and a given increment size – this process will be called “relaxed scanning” in the remainder of this thesis. In contrast, a “fixed scan” does not include the relaxation part.

Meta optimizations are to optimize a computational setup. This technique is applied to the SEQM calculations for some of the systems presented in this thesis.

⁸ For more details on the different gradient-following approaches the author refers to standard textbooks on computational chemistry, e.g., Refs. 51, 54.

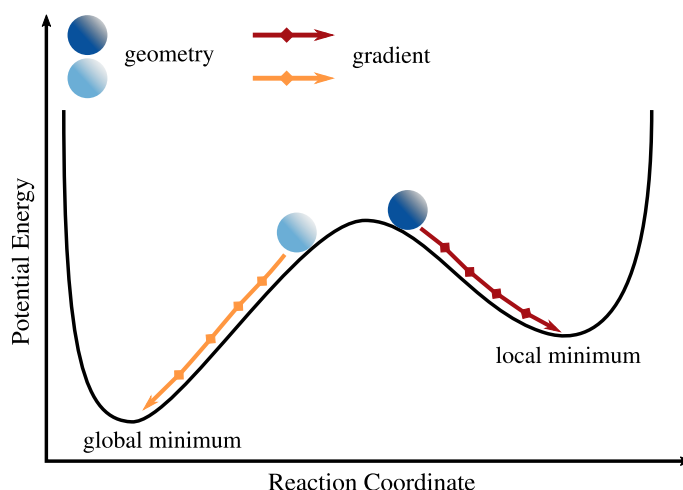


Figure 2.4.: Scheme for showcasing the features of a typical geometry optimization on the example of a one-dimensional PES. The minimum to which the system (balls) is getting optimized depends on the starting point on the PES.

The aim of a meta optimization is to find a suitable setup among many possible combinations that best reproduces certain reference values. The combinatorial space consists of possible semiempirical Hamiltonians implemented in the development version of MOPAC (AM1, PM3, PM6*⁹ and RM1), certain CI excitation levels (CIS, paired-electron CISD (PECI), CISD) and AS sizes (typically starting from 6 electrons in 6 orbitals ([6,6]) up to [12,12]). One may dub this as “computational alchemy”, but it is necessary because there is no real hierarchy among the available Hamiltonians¹⁰ as well as increasing the AS and/or CI excitation level¹¹. The best setup among all possible ones is the one showing the smallest sum of absolute differences (“fitness”) to supplied reference values. Practically, a meta optimization for a simple test case would be performed as following:

- i) Generation of reference data, e.g., vertical excitation energies for a given structure
- ii) Preparation of a general input file for MOPAC
- iii) Selection of the combinatorial space
- iv) Automatic execution of each calculations
- v) For each calculation the VEE are parsed and compared with the reference values
- vi) Calculation of the fitness
- vii) All combinations are sorted according to their fitness values

⁹ The parameters implemented in the development version of MOPAC are not the published ones for the “real” PM6 in Ref. 104. In addition, PM6 should also utilize an extended Hamiltonian which is not the case in the present PM6* (it uses the AM1/PM3 one). Interestingly, it usually performs rather well in a meta optimization and is therefore also sometimes used as starting point for the parameter optimizations. The implemented parameters are shown in the appendix (Sec. A.1).

¹⁰ It could be argued that a general hierarchy comes from the recency of the parametrization – leading to the following hierarchy: AM1 < PM3 < (PM6*?) < RM1. However, known deficiencies for some of the presented Hamiltonians are listed in Ref. 51; but in what way these influence a system that was not part of the parametrization process is unclear, thus a hierarchy based on theoretical quality can not be given.

¹¹ Both should theoretically give a systematic improvement of the wave function, though. But tests by the author could not confirm this for semiempirical FOCI calculations.

Of course, the reference data can consist of more than just VEE, e.g., relative energies of two structures, gradients, electronic state characters, PES profiles, etc. A specific input needs to be prepared for each data point that will be used with a given semiempirical FOCI setup.

Lastly, a parameter optimization optimizes the underlying parameters of a semiempirical Hamiltonian at a given computational setup. Chapter 5 is specifically dedicated to this technique.

2.7.2 Preparing Photodynamics

To start the photodynamics, the following three ingredients are needed:

- i) A collection of structures
- ii) Atom-wise momenta of each structure
- iii) Vertical excitation energies of each structure

Typically, these data are gathered from one single ground-state trajectory. The collection of structures are simply the structures during the dynamics sampled at a given time interval. The momenta for each atom are also a known quantity and are part of the standard output of the program. Both the structures and momenta are needed, because the photodynamics use just one structure from the S_0 trajectory as input to initialize the propagation in an electronically excited state, without actually exciting the system from S_0 . To simulate an instantaneous, vertical electronic excitation the momenta should be retained as well to make it realistic.

The VEEs at each structure on S_0 are needed to select the structures that would correspond to some specified excitation wavelength. For example, if, on average, the $S_0 \rightarrow S_1$ excitation corresponds to 300 nm, all structures are selected for S_1 -dynamics that would exhibit the same vertical excitation energy within a given window (usually ± 20 nm).

2.7.3 Running Photodynamics

Depending on whether the dynamics are propagated in S_0 or initially in an excited state $S_{i>0}$, the employed setup is slightly different.

The S_0 MD includes Brownian motion. This changes the way the force \mathbf{F} is calculated – from Newton’s equation to Langevin’s by adding a friction and random noise term:^[105,106]

$$\mathbf{F} = -\frac{\partial}{\partial \mathbf{R}} E_i(\mathbf{R}) - m\gamma\mathbf{v}(t) + \xi(t), \quad (2.47)$$

with γ being a friction coefficient, m the mass of the nucleus, $\mathbf{v}(t)$ the velocity and $\xi(t)$ a random, white-noise force. γ was set to 13.5 ps^{-1} as proposed by Granucci and Persico for azobenzene dynamics. The velocities are typically scaled to accommodate for a temperature of 298.15 K. Brownian dynamics are needed, because the trajectories use the optimized structure as the initial structure. This, however, would make Eq. 2.42 zero which would also cancel the acceleration term in the Verlet algorithm (cf. Eq. 2.40). The two additional terms in Eq. 2.47 therefore help to initiate the movement of the system. Additionally, the surface-hopping evaluation was deactivated during the propagation.

The excited-state dynamics used Newton's equation of motion to let the system follow the "pure" PES without any external forces. However, a stochastic thermostat was added that kept the kinetic energy at 298.15 K.^[107] This ensures that i) the system has some minimum kinetic energy throughout the propagation and ii) the system does not gain too much energy when following a steep gradient. The latter effect would otherwise result in "hot" S_0 dynamics. This could allow the system to overcome the reaction's energy barrier in S_0 which is not the process that shall be observed.

The QDC of Granucci and Persico is used with an overlap criterion of 0.1 as proposed in Ref. 100. Other adjustable parameters were kept at their default values.

2.7.4 Calculating Quantum Yields and Isomer Populations

After the calculation of a set of trajectories has finished, the quantum yields (QYs) and isomer populations can be evaluated.

A QY defines the amount of trajectories that perform a specific reaction. For ESIPT reactions this would at least be the formation of the proton-acceptor bond. If this distance reaches a value of $1 \pm 0.2 \text{ \AA}$, the ESIPT has probably happened.

But instead of checking each trajectory for each system by actually looking at the geometries at each time step¹², the evaluation was done in an automated fashion. A script checks the actual value at each time step for specified DOF and compares it to the provided criteria for each reaction, e.g., the aforementioned proton-acceptor distance.

Of course this needs to be restricted, because even when the distance criterion is met for only one time step ("snapshot") of the trajectory, it would have counted the trajectory as reactive. To only consider stable forms/products, the DOF criterion needs to be held for a minimum amount of snapshots of the trajectory. The longer, the better. The typical procedure was to start with a small amount of snapshots first, to see how long each form is actually stable during the dynamics. This allows to find a "natural gap" which can be taken as the optimized frame count. For example, when the majority of trajectories stay in the product form for >1000 snapshots and some only for <50 , the latter ones may be excluded from the actual QYs in a second evaluation, because these are not stable. The obtained QYs are thus only an estimate of the actual performance of the ensemble, but relatively easy to get. A drawback of this method would be that the automated check only yields information on *if* the process has happened and not *when* or *for how long*¹³ the system is in the product conformation.

More precise is the direct evaluation of the isomer populations because it allows answering the *when* and *for how long* questions. For the evaluation, one again determines DOF that define each of the expected isomers. All trajectories are then scanned for these DOF and to which of the defined isomer they correspond to at each timestep. Ideally, the DOF are defined in such a way that it is

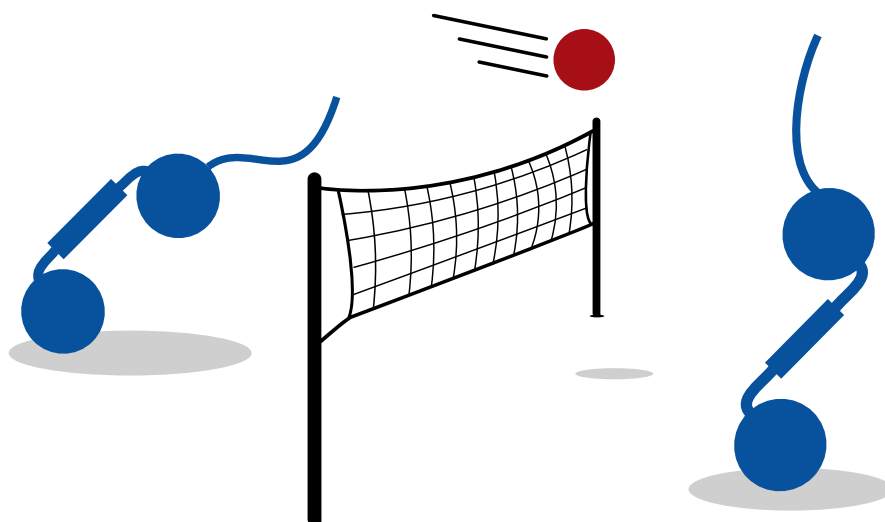
¹² Typically 200 trajectories per reaction were used and each simulated picosecond corresponds to 10000 changes of the structure when applying a 0.1 fs time step.

¹³ Indirectly it does, though. But this is just because how the script was built.

always possible to ascribe a structure to an isomer definition. After the evaluation, these isomer assignments are averaged across all trajectories to give a time-evolution of each isomer during the dynamics. Ideally, the maxima of each isomer evolution would correspond to the QY but only if the snapshot-criterion is neglected¹⁴.

However, the isomer evaluation as implemented for this thesis relies on a prior QY calculation to divide the complete set of trajectories into “reactive” and “unreactive” subsets according to one criterium or an overlap between two or more. This distinction is not strictly necessary, however, it allows to identify differences between the two sets and how each of it behaves over time.

¹⁴ For example, if the isomer evaluation reveals an 80% abundance of one specific isomer, e.g., the product, for the length of 30 snapshots, a QY calculation on the same set of trajectories would not necessarily yield the same value when the snapshot-restriction was set to more than 30 snapshots.



3. Photoisomerizations

In this chapter the results of calculations for systems that are able to perform photoisomerizations/photoswitching are going to be presented.

The first two sections present publications on azobenzene (AB)-based systems. AB is one of the workhorses of the CRC677 and AB-based systems were extensively studied in the author's workgroup by former members,^[26–28] since the discovery of a bridged AB derivative (brAB) in 2009.^[25]

The popularity arose from the fact that for AB a specifically optimized set of semiempirical parameters from Granucci and Persico was available,^[69] which allowed computationally low-cost calculations but with high-quality results. With their expanded version of MOPAC it was also possible to use this setup to perform surface-hopping MD calculations without much restriction in molecular size, when compared to other electronic structure methods. The MD simulations were able to replicate many experimental findings, illustrating the quality of the reparametrization.

Based on this, further studies in this AB-field were performed by the author. The foundation of these works were laid by the QM/MM study on an artificial cilium.^[108] A cilium can be used for light-driven particle transport, so using AB as the photoactive motor unit seemed reasonable. Further synthetic advances from the organic chemistry workgroup of Prof. Herges at Kiel University led to a new generation of brAB derivatives. These new systems, indandiazocine (ID) and diindandiazocine (DID), restrict $E \rightarrow Z$ isomerization to only one direction, which was verified by direct surface-hopping MD calculations that are presented in Sec. 3.1.

These new chromophores were then used in a new cilium project in two Bachelor theses: The first to update the cilia with the two new motor units (Sec. 6.1) and the second to use these new systems for actual particle transport and surface interaction simulations (Sec. 6.2). Both studies lead to the publication shown in Sec. 3.2.

Unrelated to those QM/MM studies, the author took part in three cooperative projects on photoisomerizations of different systems. The first is about the *trans*⇌*cis* switching of ferulic acid which led to the publication presented in Sec. 3.3. The second project is about AB as part of several carbohydrate macrocycles, presented in Sec. 3.4. The third project is about axitinib as a photoswitchable kinase inhibitor (Sec. 3.5).

3.1 Indandiazocines: Unidirectional Molecular Switches

3.1.1 Scope of the project

The idea behind this publication was to validate the experimentalists' claim of the unidirectional *E*→*Z* isomerization of ID and DID.^[32] These two systems can be seen as the next generation of bridged azobenzene derivatives with additional one (ID) or two (DID) ethylene bridges. This additional linking introduces one or two, respectively, chiral carbon atoms to the systems.

Both the claim of the unidirectional switching direction and the effect of the added ethylene bridges could be studied by calculations of the excited state dynamics using [14,13]-FOCI-AM1 as the standard SEQM FOCI setup for AB-based systems. All evaluations and the interpretation was done in comparison to the well-studied parent-system diazocine.^[25–28]

3.1.2 Publication Data

Authors Tim Raeker and Bernd Hartke

Title Indandiazocines: unidirectional molecular switches

Submitted January 26, 2015

Accepted January 29, 2015

Reference ScienceOpen Research, **2015**.

DOI:

10.14293/S2199-1006.1.SOR-CHEM.ARDTLN.v1

Contribution Setup and execution of all quantum chemical calculations (FOCI-AM1 and MP2/CC2 as benchmark), as well as parsing, evaluation and interpretation of the obtained static and dynamical data; major contributions to writing the article

3.2 Simulating Particle Transport

3.2.1 Scope of the project

This project is the follow-up work from the author's first publication on azo-TATA systems.^[108]

In the first generation there were some deficiencies regarding the dynamics of the system – the $E \rightarrow Z$ isomerization could not be observed –, which was back then ascribed to the stability of the AB-motor unit in its E conformation. It was later found out that the QM/MM setup could be changed from having a larger QM part – formed by the azo motor and tail unit – to an azo-only QM part and everything else treated by the force field. This setup was first checked on the first-generation AB-TATA and did indeed lead to an improved performance. Because of the advancements in AB-derivatives in the years after publication of Ref. 108, the newest bridged-azobenzenes were implemented as motor units as well.

These systems can also be utilized as molecular machines to facilitate particle transport. ID-TATA was used as a test case on how to study such a process and what needs to be taken care of during computational setup.

The latter part was taken directly from the results of D. Behrens Bachelor's thesis which is presented in Sec. 6.2. The investigation of the photodynamical properties of each artificial cilium was performed on an improved set of trajectories from B. Jansen, whose Bachelor's thesis is summarized in Sec. 6.1.

3.2.2 Publication Data

Authors Tim Raeker, Björn Jansen, Dominik Behrens and Bernd Hartke

Title Simulations of Optically Switchable Molecular Machines for Particle Transport

Submitted 25.01.2018

Accepted 28.02.2018

Reference *J. Comput. Chem.* **2018**, published online.
DOI: 10.1002/jcc.25212

Contribution Major contributions to writing the publication, FOCI-AM1 calculation of the non-surface-including systems, MP2 and DFT energy calculations for the Pearson correlation coefficients

3.3 Ultrafast Dynamics of Trans- and Cis-Ferulic Acid

3.3.1 Scope of the project

This project was about new insights of the excited state properties of ferulic acid (FA) as one of the many possible organic sunscreen ingredients. Although *trans* FA (tFA) was already subject to experimental^[109,110] and theoretical^[111] studies, the results led to controversial views on what is happening in the electronically excited states. This is because there are many close-lying excited states of $n\pi^*$, $\pi\pi^*$ and $\pi\sigma^*$ character to which different reactions might be assigned. One of these is *trans* \rightleftharpoons *cis* photoisomerization of the propenoic acid group.

Utilizing ultrafast transient absorption and fluorescence spectroscopy, the dynamics of tFA were re-investigated, and for the first time also the dynamics of *cis* FA (cFA) in aqueous solution. To compare with Ref. 109, the mono- and dianionic forms of both isomers (tFA⁻, tFA²⁻, cFA⁻ and cFA²⁻) were included in the present study as well.

For all tFA species and cFA only small geometric changes could be deduced from the changes in electronic properties. Only the obtained data for cFA²⁻ suggested a barrierless excited-state deactivation pathway which probably results in formation of tFA²⁻.

In order to aid in understanding the experimental results, QM calculations on the DFT level of theory were performed for the electronic ground state as well as the first two excited states.

3.3.2 Publication Data

Authors	Shuangqing Wang, Sebastian Schatz, Mayra Christina Stuhldreier, Hendrik Böhnke, Joss Wiese, Carsten Schröder, Tim Raeker, Bernd Hartke, Julia K. Keppler, Karin Schwarz, Falk Renth, Friedrich Temps
Title	Ultrafast dynamics of UV-excited trans- and cis-ferulic acid in aqueous solutions
Submitted	August 4, 2017
Accepted	November 2, 2017
Reference	<i>Phys. Chem. Chem. Phys.</i> 2017 , <i>19</i> , 30683-30694. DOI: 10.1039/C7CP05301K
Contribution	Execution and evaluation of all TDDFT calculations, contributions to writing the electronic supplementary information.

3.3.3 Additional Information

Given the many excited-state pathways of FA, it was also tested if the photodynamics can be calculated for the neutral species utilizing a SEQM setup. To find a good setup, a meta optimization was performed utilizing the ground-state energies and vertical excitation energies of tFA and cFA from Ref. 111.

Table 3.1.: Results of the meta optimization for FA. Reference energies taken from Ref. 111: Vertical excitation energies on the EOM-CCSD level of theory, ΔE on CASPT2 level. State character reference values set to arbitrary values. All energies in eV.

System	Property	Reference values	[10,6]-FOCISD-PM3
tFA	$\Delta E_{S_0 \rightarrow S_1}$	4.19	4.11
	$\Delta E_{S_0 \rightarrow S_2}$	4.80	4.83
	$\Delta E_{S_0 \rightarrow S_3}$	4.90	5.38
	$S_1 \pi\pi^*$ character	0.8	0.69
	$S_2 n\pi^*$ character	0.4	0.08
	$S_3 \pi\pi^*$ character	0.4	0.06
	$\Delta E_{tFA-cFA}$	0.2021	0.2113
cFA	$\Delta E_{S_0 \rightarrow S_1}$	4.20	4.13
	$\Delta E_{S_0 \rightarrow S_2}$	4.80	4.76
	$\Delta E_{S_0 \rightarrow S_3}$	4.82	5.50
	$S_1 \pi\pi^*$ character	0.8	0.69
	$S_2 \pi\pi^*$ character	0.4	0.67
	$S_3 n\pi^*$ character	0.4	0.58

As it turned out, there is no setup – out of all combinations with four Hamiltonians (AM1, PM3, PM6* and RM1), three CI excitation levels (CIS, PECl, CISD) and all possible active space size starting from an [8,8] active space to [12,12] – that can get all state characters right.

The best fitting setup was found to use the PM3 Hamiltonian, a CISD excitation level and a [10,6] active space, i.e., five occupied orbitals and one unoccupied. But even this setup fails at describing the S_2 and S_3 states of tFA with the same characters as the reference. For cFA the state characters were reproduced correctly. It must be noted, though, that the reference values for this purpose are chosen rather arbitrarily, knowing that higher excitations may be described worse than S_1 . But at least for cFA the found CI coefficients for the corresponding determinants are dominant (>0.5). The [10,6]-FOCISD-PM3 setup was used in subsequent MD calculations of the photodynamics of tFA and cFA starting from S_1 and a propagation time of 20 ps. The 200 initial structures of each isomer were taken from a 15 ps ground-state trajectory at 298.15 K.

14% of the trajectories starting from tFA form a stable cFA isomer. Other possible dihedral rotations within the propylene chain, as investigated for several cinnamic acid derivatives in Ref. 112, can be seen as well, but are typically not as stable. The population inversion (PI) time of the trajectory ensemble is ~ 1.8 ps and only 83% of all trajectories reached S_0 at the end of the propagation.

3.4 Azobenzene in Carbohydrate Macrocycles

When starting from cFA, the isomerization QY is increased to 40%, as well as a faster PI time of 0.8 ps and a final S_0 population of 90%.

The parameters for determining the cFA \rightleftharpoons tFA QY were set as reaching a CC=CC dihedral of $0\pm 30^\circ$ for reaching cFA and $180\pm 30^\circ$ for reaching tFA, respectively. The second criterion was that these conformations should be stable for at least 100 fs, but all structures that fulfilled this criterion stayed in their respective product conformation for 1.4 to 1.9 ps.

However, the results must not be seen as final because of the aforementioned deficiencies of the semiempirical setup. Nevertheless, the photoisomerization of neutral FA could be simulated and qualitatively fits the finding from the publication in which cFA was attributed a more effective reaction pathway than tFA.

Further work on this topic should increase the reference data set with the information from the two anionic forms of FA, which were already calculated on the RPA-TDDFT level of theory. The results for the photodynamics of (di)anionic FA could then be compared to neutral FA to see if any qualitative differences that are present in the experiment can be reproduced. In addition, the SEQM calculations could be re-done utilizing a similar implicit solvent method like it was used in the publication.

Considering the computational demand of a SEQM/MM setup (cf. Sec. 3.2), introduction of explicit water molecules to the setup would i) be feasible and ii) also allow for investigations of sterical effects on the dynamics. Especially the latter point played an important role in the photoswitching dynamics of diazocine that led to better agreement with the experiment.^[28]

3.4 Azobenzene in Carbohydrate Macrocycles

This section presents the theoretical part of a publication that is in the process of being written¹.

When azobenzene is part of a macromolecular ring system, the switching capabilities of the chromophore can be used to change the cavity size. This would allow to switch the ability of complexation of an ion or small molecule inside the ring “on” or “off”. This feature could be applied to, e.g., medicinal processes to specifically determine or to manually trigger the release of drugs. Additionally, when the photoswitch is covalently connected to a chiral molecule, a transfer of asymmetry occurs, thus making the azobenzene unit optically active². For example, this feature was described with azobenzenes substituted with a chiral sulfoxide group,^[113] or with azobenzene embedded into chiral frames.^[114–117] The resulting chiral azobenzene is therefore a chiroptical switch, exhibiting two distinct CD patterns corresponding to the *trans* and *cis* isomers, respectively. Importantly, the chirality transfer enables the selection of one *cis* helical conformer (M or P), hence resulting in a directed photoisomerization.

The conjugation of an azobenzene moiety with an enantiopure molecule is an attractive approach as it provides optically pure products and no racemate resolution is needed. In particular, the

¹ Preliminary author list: Tim Raeker, Bernd Hartke, Frank Sönnichsen, Guillaume Despras.

² In contrast to the “inherent” chirality of, e.g., ID and DID.

preparation of chiral and photoresponsive macrocycles opens up a way for the design of smart receptors for chiral molecules and chiroptical materials for various applications, e.g., data storage. In addition, the unidirectional photoisomerization of the azo bond can also be exploited in the conception of molecular machines.

In the junior research group of G. Despras (part of the Lindhorst group) at Kiel University, AB was implemented in two carbohydrate-based macrocycles (Fig. 3.1). Notably, the macrocyclization was performed by pre-organizing the *trans* form of the open precursor – isothiocyanate-armed azobenzene glycosides – *via* photoisomerization of the azobenzene unit to the *cis* form.

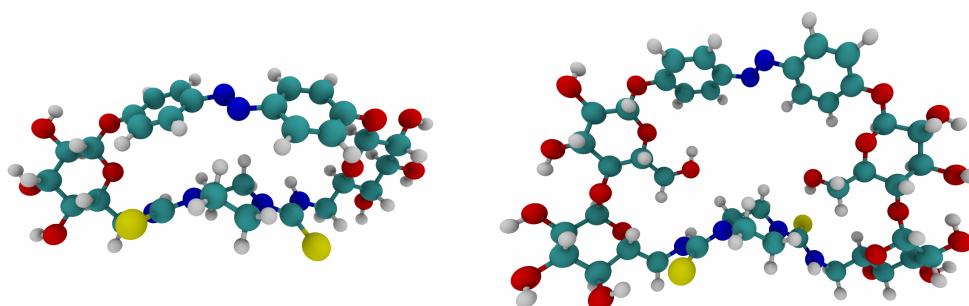


Figure 3.1.: Exemplary structures for the two studied carbohydrate. (Left: Azobenzene-glucose macrocycles; right: Azobenzene-maltose macrocycles.) Both in their *trans* conformation.

When AB is in the *cis* form, there is axial chirality which effectively allows the macrocycles to be distinguished by two different *cis* forms, *cisP* and *cisM*. Experimentally, chiral systems can be studied by circular dichroism (CD) spectroscopy. Such spectra were obtained for both *cis* and *trans* forms of a glucose macrocycle (**G**) and maltose macrocycle (**M**), respectively.^[117] According to these data, the AB unit exhibits a significant CD signal in the *trans* form, due to a twist induced by the chiral frame. Upon photoisomerization, the CD pattern changes remarkably, showing a stronger twist characteristic of a single helical conformer. We assumed, based on literature data, that **cis-G** is in the M-conformation while **cis-M** is in the P-conformation.

In order to exactly determine which of the two possible *cis* isomers is present after irradiation of the *trans* systems, CD spectra were calculated theoretically. The computational setup was as follows: The initial structures were prepared and optimized using a force field. These structures were then further optimized on the DFT level of theory, employing the B3LYP functional, the D3BJ dispersion correction and the def2-SVP basis set. The same functional was already employed in a study on azobenzene-containing cyclopeptides and yielded good results.^[114] From the DFT-optimized structures the vertical excitation energies were calculated using RPA-TDDFT as implemented in TURBOMOLE v7.0.

A second set of structures was also investigated in which the carbohydrate hydroxyl groups were acetylated (**GAc**, **MAc**), as well as a third set of the open precursors to the macrocycles in both possible *cis* forms (**GAcNCS**, **MAcNCS**). Experimental CD spectra are also recorded for

these systems.^[118] The latter set was used to check if there is any preorientation effect present that determines the outcome of the ring-closure reaction.

Because intensities in calculated spectra are not quantitatively reliable the following comparison will only take the qualitative features, i.e., the sign of each band in the CD spectrum, into consideration. Even the position of each band is highly dependent on the given structure which itself is highly dependent on the method and basis set employed. The calculations only yield stick spectra. They are additionally convoluted with a Gaussian function for comparison with the experimental spectra with an arbitrary full-width half-maximum of 30 nm.

Glucose Systems

The spectra of **GAcNCS** (Fig. 3.2c) reveal that the spectrum of **cisM-GAcNCS** fits the experimental one nicely when looking at the signs of the two dominant peaks above 300 nm: A positive peak at ~ 450 nm and a negative at ~ 310 nm which are both recovered in the calculated spectrum. As expected, the corresponding *cisP* structure gives the exact opposite spectrum.

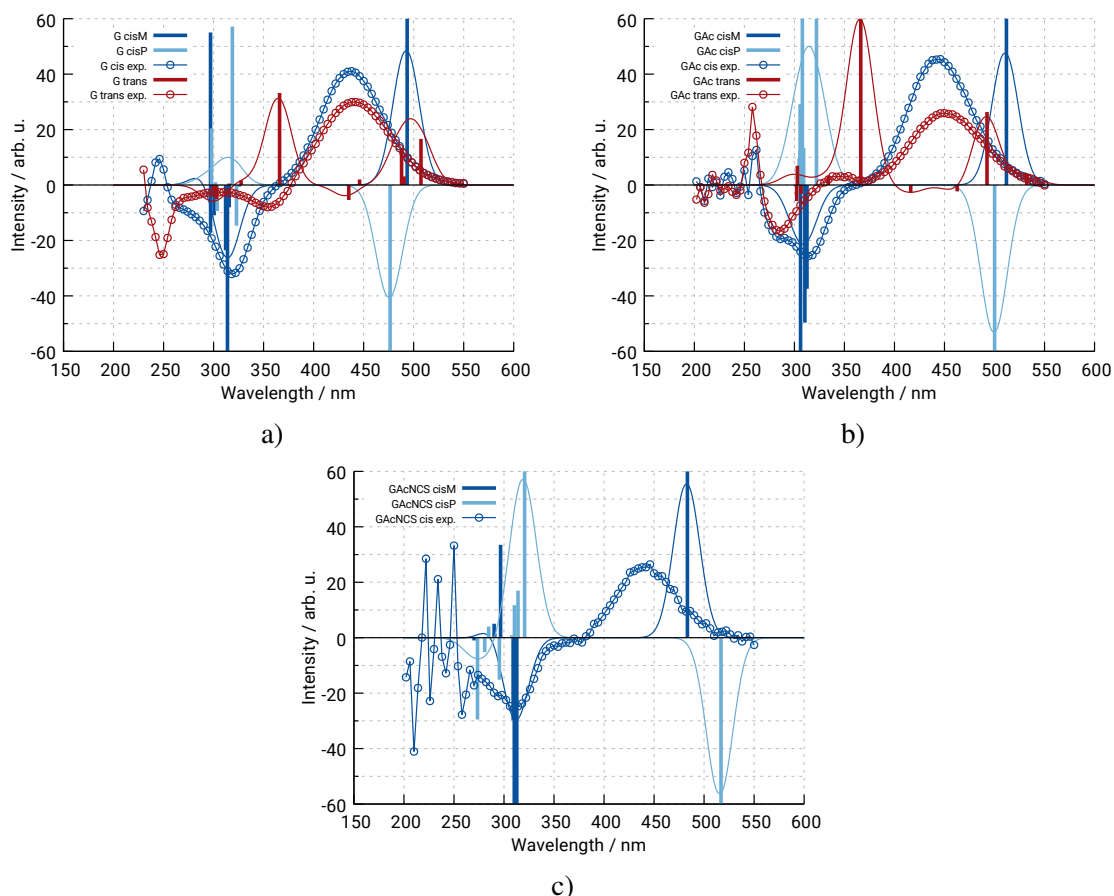


Figure 3.2.: Calculated CD spectra of all three isomers (*trans*, *cisM*, *cisP*) of **G**: a) **G**, b) **GAc**, c) **GAcNCS**. All spectra are compared to obtained experimental spectra.

The same features of these spectra are present for **G** (Fig. 3.2a) and **GAc** (Fig. 3.2b) in both the calculated and experimental spectra, thus confirming that in both cases the measured *cis* forms

feature M helicity. The calculated CD spectra for the *trans* macrocycles also fit the experimental ones rather nicely³.

In all cases it is obvious that the calculated positions of these bands are red-shifted in comparison to the experiment. But this shift is constant for all systems and isomers and can therefore be disregarded – as it is probably the result of both methodical error and basis set error.

When looking at the relative energies for each of the three groups of glucose structures (Table 3.2), in all but one case the *cisM* forms are energetically favored over *cisP*. Only the linear *cisP* form is heavily stabilized over *cisM*. This value can not be correct, because both experimental and calculated CD spectra predict the *cisM* form to be present.

Table 3.2.: Relative energies of all glucose-containing systems. Each relative energy is given with respect to the *trans* structure of the corresponding group (*cisM* for the open system). All values in eV.

System	Isomer	ΔE	System	Isomer	ΔE
G	<i>trans</i>	0	GAc	<i>trans</i>	0
	<i>cisM</i>	0.001		<i>cisM</i>	0.843
	<i>cisP</i>	0.293		<i>cisP</i>	0.919
GAcNCS	<i>cisM</i>	0			
	<i>cisP</i>	-0.401			
	<i>cisProt</i>	0.226			

Why is this not very problematic? The structures came initially from a force field optimization in which the parameters heavily influence the obtained geometry. In addition, these systems are relatively large and therefore have many DOF. Both features were already encountered during the investigation of the azobenzene TATA systems using a SEQM/MM setup (cf. the publication presented in Sec. 3.2). When the MM parts of the systems were included during the reference calculations with DFT or MP2, the resulting energies were worse than when this part was neglected completely.

Without having repeated this procedure for the carbohydrate systems, this discrepancy might be present here as well. The “active core” of the macrocycles is again AB, which gives rise to the signals in the CD spectrum. Since the calculated spectra fit the experimental ones rather well, this core part seems to be well described structurally. The remaining part of each system is much larger in size and as such can influence the energy of the system much more. This is why the wrong relative energy for both *cisM* and *cisP* in the open form is attributed to a non-ideal initial structure, which led to a local minimum for **cisM-GAcNCS** that is higher in energy, probably due to steric reasons. Figs. 3.3a and 3.3b show the two pre-optimized structures: **CisP-GAcNCS** features one rotated carbohydrate ring which results in opposing NCS groups, thus less steric hindrance than **cisM-GAcNCS**, where these two groups are close to each other. This large difference can not

³ Except for the positive signal at 350 nm in Fig. 3.2a. No obtained CD spectrum from manipulated **transG** structures – with the restriction to having a lower energy than the two *cis G* structures – exhibited a negative sign for this signal.

be overcome by a simple geometry optimization towards a local minimum. And even though **cisP-GAcNCS** presents a sterically favored conformation, it would not be useful for a ring-closure reaction in which the two NCS groups need to be close.

In an attempt to fix this, one of the glucose rings in **cisP-GAcNCS** was rotated (**cisProt-GAcNCS**, Fig. 3.3c) so that both NCS groups are right below the azobenzene core. After geometry optimization, **cisProt-GAcNCS** is now energetically less favored than **cisM-GAcNCS**, while at the same time being structurally more similar. Even though **cisProt-GAcNCS** still exhibits the same CD spectra pattern, its energy now also fits the experimental result.

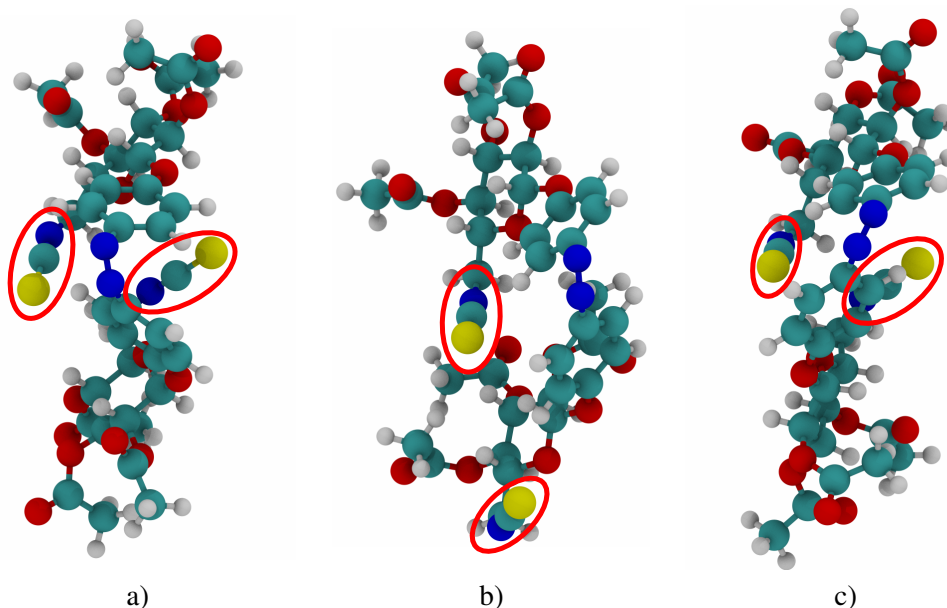


Figure 3.3.: Comparison of the two *cis* structures of **GAcNCS**: a) **cisM-GAcNCS** and b) **cisP-GAcNCS**. The positions of the NCS groups are marked by red circles. **cisM-GAcNCS** features the NCS groups close to each other, whereas in **cisP-GAcNCS** these groups are opposing. This steric difference could explain why *cisP* isomer is more stabilized than *cisM*. c) The *cisP* rotamer (**cisProt-GAcNCS**) which has a less favoring energy but with a non-azo part comparable to **cisM-GAcNCS**.

Maltose Systems

The main features in all three experimental CD spectra for the *cis*-maltose-containing systems are a negative peak at ~ 440 nm and a positive at ~ 310 nm. In all three cases – **M** (Fig. 3.4a), **MAc** (Fig. 3.4b) and **MAcNCS** (Fig. 3.4c) – this pattern is only present in the *cisP* forms when comparing with the calculated spectra.

When starting from the MM optimized structures, more apparent discrepancies between both calculated and theoretical spectra and relative energies were found. In the following, only the results of the improved structures (spectra and energies) are presented. The relative energies of the initial structures are presented in the appendix (Sec. A.2.1) for comparison.

The same features of the spectra of **MAcNCS** are also present for **M** (Fig. 3.4a) and **MAc** (Fig. 3.4b) in both the calculated and experimental spectra, thus confirming that in both cases the

measured *cis* forms feature P helicity. The calculated CD spectra for the *trans* macrocycles also fit the experimental ones rather nicely, including the different sign of the prominent band at ~ 450 nm for **trans-M** and **trans-MAc**.

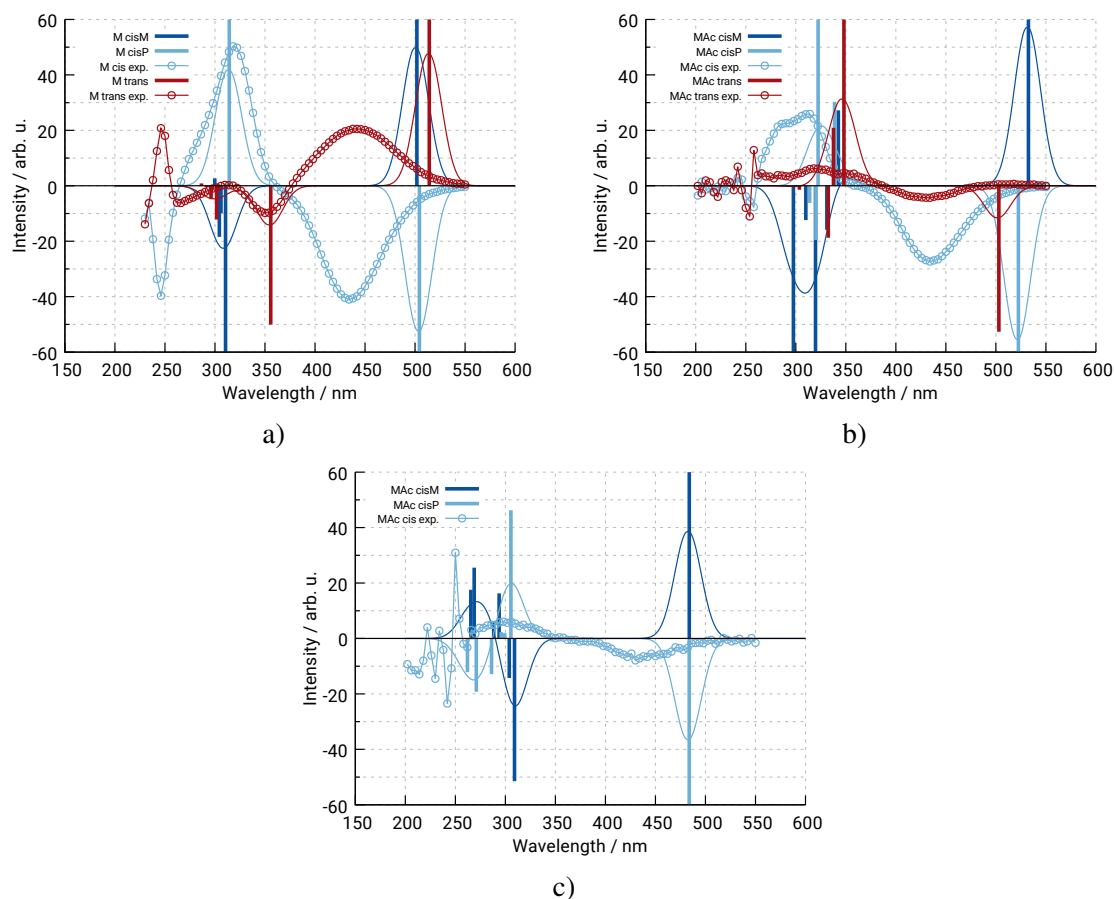


Figure 3.4.: Calculated CD spectra of all three isomers (*trans*, *cisM*, *cisP*) of **M**: a) **M**, b) **MAc**, c) **MAcNCS**. All spectra are compared to obtained experimental spectra.

The relative energies confirm the results from the CD spectra for both **M** and **MAcNCS** in which the *cisP* structures are energetically favored over *cisM* (Table 3.3), albeit the energy difference between **cisM-MAcNCS** and **cisP-MAcNCS** being nearly zero. However, for **MAc** this property could not be reproduced.

In the case of **cisP-GAcNCS** it was rather obvious how the structure needed to be changed in order to make it more similar to **cisM-GAcNCS**. Applying the same procedure to the maltose-containing systems also led to better results. However, these intuition-driven approaches followed the trial-and-error principle, because of the many more rotational DOF in these systems. Therefore the energies must not be seen as the final answer at this point.

There are two ways out of this dilemma: i) a pure computational approach would involve conformational analysis, i.e., rotate every dihedral so that the energy gets minimized – this is of course not possible for a system this size on the DFT level of theory; ii) the more elegant

Table 3.3.: Calculated relative energies of all improved maltose-containing systems. Each relative energy is given with respect to the *trans* structure of the corresponding group (*cisM* for the open system). All values in eV.

System	Isomer	ΔE	System	Isomer	ΔE
M	<i>trans</i>	0	MAc	<i>trans</i>	0
	<i>cisM</i>	0.429		<i>cisM</i>	0.224
	<i>cisP</i>	0.189		<i>cisP</i>	0.283
MAcNCS	<i>cisM</i>	0			
	<i>cisP</i>	-0.095			

solution would involve applied chemistry by getting structural information from obtaining the crystal structure with X-ray spectroscopy.

The first approach requires a more in-depth investigation which is out of scope for this thesis. But given the fact that conformational analysis is a standard tool in computational biochemistry to find optimal conformations for large molecules, there should be enough resources on how to perform such a type of calculation. An alternative to this would be running electronic ground-state MD simulations up to the nanosecond regime in order to sample the thermally accessible conformational space. Such calculations are currently being performed by the author and his cooperation partner.

Only the experimental alternative would give the final solution to which *cis* isomer is present in the experiment as well as what the real structures of both *cis* and *trans* isomers are. This is of course out of scope for a thesis on computational chemistry and hence the first alternative is being pursued in the meantime.

A quick solution for the relative energies was the addition of an implicit solvent through the conductor-like screening model (COSMO) as implemented in TURBOMOLE v7.0. By setting a dielectric constant of 47.2 and a refractive index of 1.479, the properties of dimethyl sulfoxide (DMSO) – as the solvent that was used in the experiment – were mimicked. The structures were allowed to relax on the new PES under the influence of the solvent which actually resulted in the correct energy ordering of **cisM-MAc** and **cisP-MAc**. The relative energies for all other system did not show any changes in their order. The improved energies are presented in Table 3.4. The calculated CD spectra do not show any qualitative differences, thus these are only presented in the appendix (Sec. A.2.2).

Table 3.4.: Calculated relative energies of all azobenzene-carbohydrate systems including an implicit treatment of solvent (DMSO). Each relative energy is given with respect to the *trans* structure of the corresponding group (*cisM* for the open systems). All values in eV.

System	Isomer	ΔE	System	Isomer	ΔE
G	<i>trans</i>	0	GAc	<i>trans</i>	0
	<i>cisM</i>	0.317		<i>cisM</i>	0.709
	<i>cisP</i>	0.476		<i>cisP</i>	0.785
GAcNCS	<i>cisM</i>	0			
	<i>cisP</i>	0.163			
M	<i>trans</i>	0	MAc	<i>trans</i>	0
	<i>cisM</i>	0.397		<i>cisM</i>	0.029
	<i>cisP</i>	0.231		<i>cisP</i>	0.022
MAcNCS	<i>cisM</i>	0			
	<i>cisP</i>	-0.046			

3.5 Static Properties of a Kinase Inhibitor

This section presents the theoretical part of a publication that is in the process of being written⁴.

The third cooperative project originates from the Peifer group at the Institute of Pharmaceutical Chemistry at Kiel University, where they investigate the approved drug axitinib (**Ax**) regarding its applicability as a photoswitchable kinase inhibitor. **Ax** is a stilbene-like chromophore that is able to perform an $E \rightleftharpoons Z$ isomerization upon irradiation with UV light.^[119,120] For the indazole ring of the *Z* isomer an interesting NH-tautomerism has been observed which leads to a *Z*-2H-tautomer that is stabilized by a hydrogen bond. These reactions are summarized in Fig. 3.5.

Interestingly, although **E-Ax** is an approved drug for second-line therapy of renal cell carcinoma, the $E \rightarrow Z$ isomerization is only briefly described in the literature and there are no reports on the bioactivity or toxicity of **Z-Ax**.

The scope of the project is to explore if the $E \rightleftharpoons Z$ isomerization can be used to switch the inhibitory effect of **Ax** “on” and “off” triggered by light.

Experimentally, UV spectra have been recorded for both **E-Ax** and **Z-Ax**. **Z-Ax** has a less intense and broader absorption band and is red-shifted compared to **E-Ax**. The red-shifted absorption band of **Z-Ax** is unusual compared to stilbene or other stilbene-like systems, where the absorption band of the *Z* isomer is often hidden under the more intense band of the *E* isomer which makes a selective excitation of the *Z* isomer difficult. We attribute the red-shift to interactions between the 2H-indazole and the pyridine nitrogen atom in the *Z* configuration. To proof this hypothesis, vertical excitation energies from the optimized structures on the RPA-TDDFT level of theory (B3LYP-D3BJ/def2-TZVP) were calculated.

⁴ Preliminary author list: Dorian Schmidt, Linda Heintze, Theo Rodat, Jantje Weber, Tim Raeker, Rebecca Horbert, Christian Renn and Christian Peifer.

3.5 Static Properties of a Kinase Inhibitor

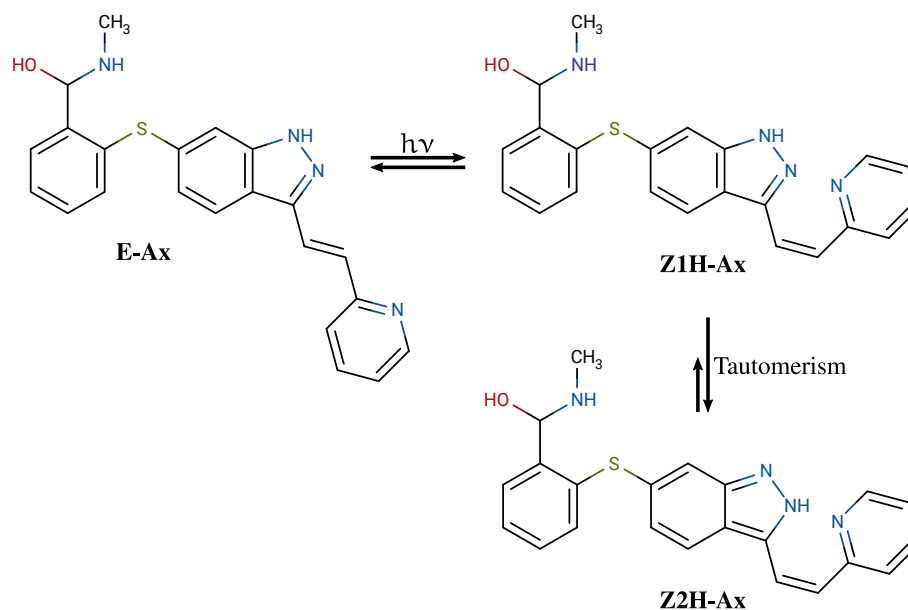


Figure 3.5.: Reaction scheme of the **E-Ax** \rightleftharpoons **Z-Ax** photoisomerization and equilibrium of the Z-1H- and Z-2H-indazole annular tautomers.

The data suggest that only **Z2H-Ax** is responsible for the red-shifted absorption band. According to the relative energies of the optimized structures, **Z2H-Ax** is also favored over **Z1H-Ax** in S_0 . Both results are summarized in Table 3.5. However, NMR data proof that both tautomers are in equilibrium in solution which leads to the assumption that there are important solvent effects that need to be considered.

Table 3.5.: Relative energies and vertical excitation energies for the three studied systems using the B3LYP-D3BJ/def2-TZVP optimized structures. The values in parentheses are the relative energies when including the zero point vibrational energy. All values in eV.

Isomer	Property	Ax	AxMe	AxPhen
<i>E</i>	$\Delta E_{S_0 \rightarrow S_1}$	3.53	3.38	3.47
	$\Delta E_{S_0 \rightarrow S_2}$	3.70	3.60	3.68
<i>Z1H</i>	ΔE_{E-Z1H}	0.26 (0.25)	0.26 (0.26)	0.16 (0.18)
	$\Delta E_{S_0 \rightarrow S_1}$	3.63	3.46	3.54
	$\Delta E_{S_0 \rightarrow S_2}$	3.73	3.62	3.74
<i>Z2H</i>	ΔE_{E-Z2H}	0.10 (0.08)	0.37 (0.38)	0.26 (0.28)
	$\Delta E_{S_0 \rightarrow S_1}$	3.01	2.96	3.25
	$\Delta E_{S_0 \rightarrow S_2}$	3.45	3.48	3.66

Two **Ax** derivatives were also investigated using the same computational setup. For both compounds intramolecular hydrogen bonding is not possible due to the presence of a methyl group instead of the proton (**AxMe**, Fig. 3.6a) or the absence of the nitrogen atom of the pyridine ring

(**AxPhen**, Fig. 3.6b), respectively. In both cases the 1H-indazole or the 1-Me-indazole should be energetically favored over the 2H- or 2-Me-indazole, respectively.

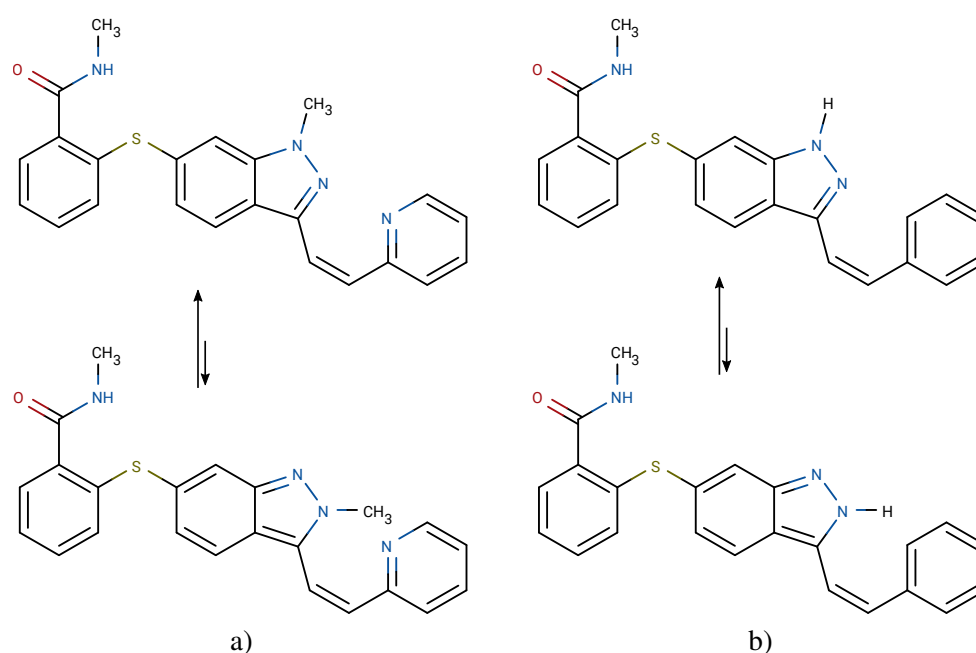


Figure 3.6.: Reaction schemes of the two **Ax** derivatives in their *Z* conformations. a) Substitution of the proton with a methyl group (**AxMe**) favors the 1-Me-indazole over the tautomer due to steric effects. b) Substitution of the pyridine ring with a phenyl ring (**AxPhen**) disables the formation of a hydrogen bond completely.

And indeed, the relative energies (cf. Table 3.5) of the *Z* isomers show the now favored *Z1H*-equivalent structure to be more stable, thus an experimental UV spectra of both systems should not exhibit a red-shift after irradiation of the *E* isomer according to the vertical excitation energies.

As it was the case with the previous project, the main question could be answered by computational chemistry. Until further results are obtained, the presented theoretical results are a valuable addition to the experimental observations.

3.6 Summary

This chapter presented two publications on photoisomerizations of azobenzene derivatives.

The first one presented a theoretical investigation of the two newest members in the bridged-azobenzene family, **ID** and **DID** (Sec. 3.1). The photodynamics of both systems were calculated using the [14,13]-FOCI-AM1 setup that was specifically parametrized to azobenzene. The calculations revealed that these two systems exhibit a restricted *E*→*Z* reaction, which all former systems did not. Hence, these new bridged azobenzenes are a welcome addition to the portfolio of possible photoactive motors for use in molecular machines.

In the second publication azobenzene and its derivatives were used in an artificial cilium (Sec. 3.2). In these systems, the chromophores served as motor units that lead to a large amplitude

3.6 Summary

motion that could be harnessed to achieve particle transport. A SEQM/MM setup was used to study systems consisting of more than 100 atoms. This setup allowed an extensive investigation on how to study molecular machines that are driven by photodynamics and what difficulties may arise when, e.g., dealing with surfaces and with actually transporting particles. This publication was based on the Bachelor's theses of B. Jansen and D. Behrens (Secs. 6.1 and 6.2).

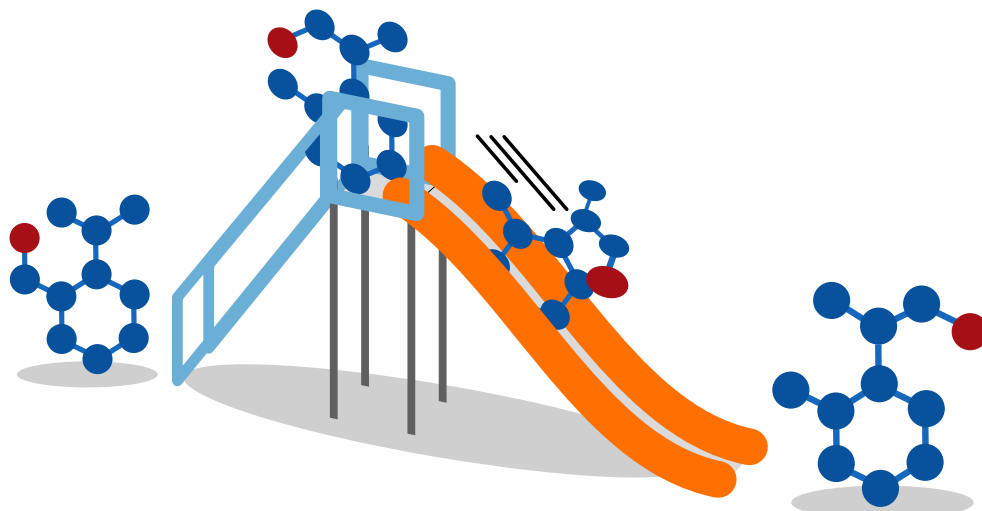
Sec. 3.3 presented a publication on the photoisomerization process of ferulic acid. Here, the experimental findings on the photodynamics of **tFA** and **cFA** were aided with QM calculations on the electronic ground state and their excited states. The data suggest that a pathway for the *trans*⇒*cis* isomerization in the excited states exist.

Sections 3.4 and 3.5 presented preliminary results on cooperations with experimentalists in the organic and pharmaceutical chemistry at Kiel University.

The first project is about azobenzene in carbohydrate macrocycles and how computational chemistry can aid in finding the chiral configuration of the systems in the experiment. For this endeavor, calculated and experimental CD spectra were compared. All optimized geometries yielded CD spectra that qualitatively agree with the respective experimental ones. Thus, it was found out that for the glucose systems the prevalent *cis* isomers is *cisM* and for the maltose systems *cisP*. However, the relative energies of the structures did not fit these findings. Some improvements were made in the non-azobenzene parts of all systems to make the structures within each of the three groups – unprotected macrocycles, acetylated macrocycles and open precursors – more similar. These improved structures did lead to better relative energies in all but one case. The addition of an implicit solvent model helped in this regard. Both the improvement of the structures and the addition of solvent did not change the qualitative agreement of the spectra, thus confirming that only the azobenzene part – as the chromophoric core – is of importance for reproducing qualitative features of the experimental CD spectra.

The latter project presented axitinib as a kinase inhibitor which can perform photochemical *E*⇒*Z* isomerization and features a proton tautomerism at the indazole moiety when in the *Z* form (Sec. 3.5). In order to find out about the energetics of the three isomers involved, RPA-TDDFT calculations were performed. The theoretical data fits the currently available experimental data (UV spectra) qualitatively so that predictions of the stable form after photoisomerization could be made. Suggested modifications from the experimentalists were tested as well, showing the opposite behavior when it comes to the stability of the tautomeric forms.

*This page is intentionally left blank; but this text is printed here,
therefore this is not a blank page.*



4. Excited-State Proton Transfer

This chapter deals with different ESIPT switches, each with different “levels” of complexity of the reaction process. ESIPT-switching is a new project within the CRC677 joining organic, physical and theoretical chemists. The primary aim for the theoretical part was to establish a computational setup that allows to study the ESIPT processes by means of MD, for the long-term goal of predicting optimally designed new ESIPT systems for subsequent synthesis and analysis.

Salicylic acid (SAC) and its derivatives are relatively simple systems, only showing the proton transfer in the excited state followed by relaxation to the ground state *via* a conical intersection. This behavior was already predicted by static calculations of SAC.^[121] Sec. 4.1 presents a publication on the first full-dimensional MD simulation of the ESIPT process of SAC. Furthermore, the SEQM FOCI surface-hopping approach is also compared to the results of OM2/MRCI dynamics of salicylidene methylamine (SMA).

Sec. 4.2 presents a second benchmark with the ESIPT process of 4-(2-hydroxybenzylidene)-1H-imidazol-5(4H)-one (OHBI) for which reference data on the OM2/MRCI level of theory are also available. These benchmarks were performed in order to determine if SEQM FOCI works as reliably for PT reactions as it does for azobenzene photoswitching.

After having established the semiempirical FOCI approach for ESIPT dynamics simulations, four SAC-like systems were also studied with different proton acceptor and donor sites as well as a different backbone and how this influences the PES and dynamics (Sec. 4.3).

The next level of complexity is covered by fused-ring systems, i.e., ESIPT switches with quinoline and phenanthridine backbones (Sec. 4.4). The first system is 7-hydroxy-4-methylquinoline-8-carbaldehyde (HMQCA) which was already studied by Sobolewski for its static ESIPT properties

(Sec. 4.4.1).^[43,45] This is the first time that ESIPT properties of HMQCA are studied with MD calculations.

Two systems were designed based on HMQCA to improve the ESIPT process: In the first system the aldehyde crane was substituted with a methyl imine group (Sec. 4.4.2) and the second features a phenanthridine backbone which substitutes the quinoline backbone (Sec. 4.4.3).

4.1 Salicylic Acid

4.1.1 Scope of the project

One aim of this project was to test the SEQM FOCI setup for ESIPT reactions which was already well established for azobenzene-related calculations.^[27,28,108] Because ESIPT reactions feature breaking and forming of covalent bonds during the reaction, the semiempirical CI setup needed proper benchmarking against higher-level electronic structure methods.

For this, (TD)DFT was the method of choice, since it has proven to be a robust method for ground- and excited-state properties of simple organic molecules with feasible computational demand.

The second aim was to present the first calculations of the ESIPT process of SAc by means of direct surface-hopping MD.

Two deactivation channels were found in the first electronically excited state. One, in which a simple back-transfer of the proton in the electronic ground state was observed, and the second one triggering a rotation of the carboxylic acid group before the proton transfer in the ground state. This second pathway was already proposed by Domcke and Sobolewski by static calculations along the reaction coordinate.^[121]

As benchmark of the employed method, the surface-hopping dynamics of salicylidene methylamine (SMA) were calculated as well. For SMA there are reference values available of both static and dynamical properties.^[122] The reference method was OM2/MRCI, another semiempirical CI method that was developed by Thiel and co-workers.^[123,124] A [12,12]-FOCIS+pD-PM6* setup was able to reproduce their findings with satisfying accuracy.

4.1.2 Publication Data

Authors Tim Raeker and Bernd Hartke

Title Full-Dimensional Excited State Intramolecular Proton Transfer Dynamics of Salicylic Acid

Submitted April 6, 2017

Accepted July 18, 2017

Reference *J. Phys. Chem. A* **2017**, 121, 5967-5977.
DOI: 10.1021/acs.jpca.7b03261

Contribution Setup and execution of all quantum chemical calculations (FOCI-RM1/PM6* and DFT as benchmark), as well as parsing, evaluation and interpretation of the obtained static and dynamical data; major contributions to writing the article

4.1.3 Additional Information

The hand-selected Slater determinants for the CI part of the calculation in the publication later turned out to be missing six determinants. These determinants accounted for all double excitations from the highest occupied MO (HOMO) of the system. The same setup, consisting of all single and paired-double CI (PECI) excitations can also be accomplished by a keyword in the input file of the program.

The impact of these missing determinants shall be evaluated by recalculating the PES of **SAC** and comparing them to the old ones.

The one-dimensional (1D) scan of the O_2-H_1 ¹ distance reveals a slight, but constant, shift of the energy for the ground state PES (Fig. 4.1a). The same effect is seen in S_1 (Fig. 4.1b), but in contrast, the PT form is now energetically favored over the structure at the Franck-Condon region. But in both the original and PECI-based cases, the rotation of the COOH group was initiated at the end of the scan, which is why the old setup did not even allow the fully proton transferred form to be reached with a planar **SAC**. The energy barriers are about the same height, but when using the PECI setup it is shifted to a slightly shorter O_2-H_1 distance (1.2 Å).

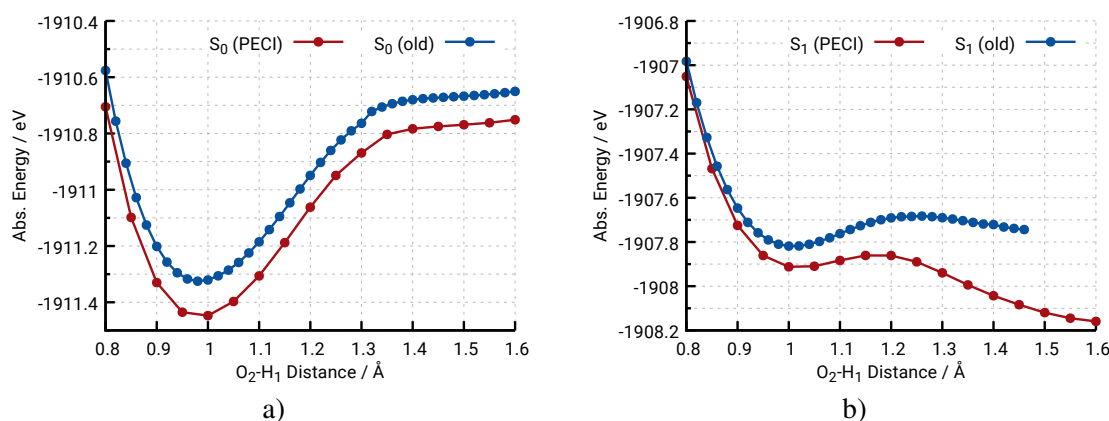


Figure 4.1.: Comparison of the 1D PES scans of the O_2-H_1 distance of **SAC** using the setup from the publication (old) and the new keyword-based approach (PECI): a) Following the S_0 gradient, b) following the S_1 gradient.

The two-dimensional (2D) PES scan of S_0 of the O_2-H_1 and $O_2 \cdots O_6$ distances (Fig. 4.2a) does not show significant changes in comparison to the original scan (Fig. 2a in the publication). The 2D S_1 PES, however, is shaped qualitatively differently after the energy barrier at an O_2-H_1 distance of about 1.2 Å (Fig. 4.2b) due to CO(H)OH rotation at elongated O_2-H_1 distances, which lowers the energy considerably. The rotation did not happen in the original approach during the 2D scan.

Ultimately, although both 1D and 2D scans show qualitative differences when using the keyword-based approach, the only noticeable effect that leads to this discrepancy is that **SAC** now seems to be more prone to initiate the rotation of the COOH group in S_1 .

¹ The numbering is the same as in the publication.

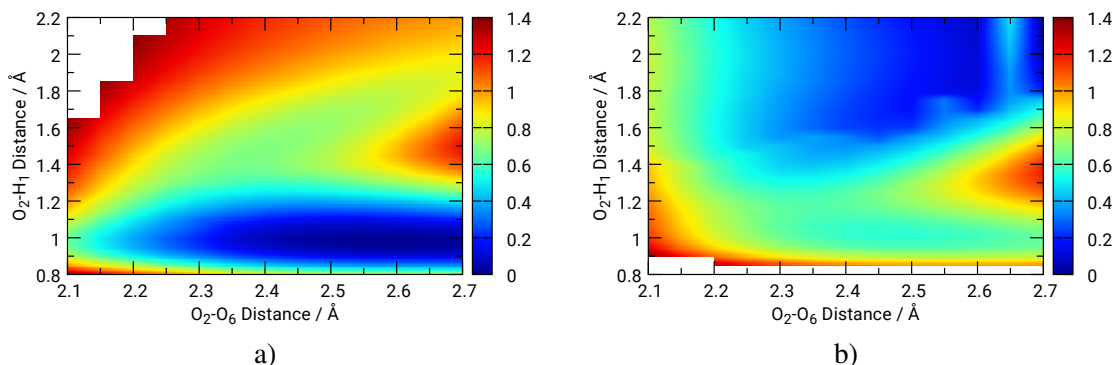


Figure 4.2.: Comparison of the 2D PES scans of the O_2-H_1 and $O_2 \cdots O_6$ distances of **SAC** using the new keyword-based approach (PECI): a) following the S_0 gradient, b) following the S_1 gradient. The energy is restricted to 1.4 eV for a direct comparison with Figs. 2a and 2b from the publication shown in Sec. 4.1.2. All following 2D PES plots use the same color code for the energy range from 0 eV to 1.4 eV.

To investigate to what extent this influences the results from the photodynamics, they were also recalculated using the same initial structures as in the publication, with the only change being the Peci keyword instead of the hand-selected determinants.

Averaging again over 200 trajectories yields the following results: i) the S_0 state still does not get fully populated within the 45 ps of propagation time (93%); ii) when applying the same conditions for determining the QYs as shown in Table 4 in the publication, 55% of the ensemble now performs the ESIPT (4% more), 34% transfer the wrong proton (H_8) back to O_2 (6% more), and 78% of all trajectories now exhibit the $C_3C_4C_5O_6$ rotation (5% more). The QYs are summarized in Table 4.1.

Table 4.1.: Calculated QYs for specific events of the ESIPT process of **SAC** for the computational setup of the publication in Sec. 4.1.2 and the Peci-keyword approach.

Rule	Event	Description	QY (old)	QY (PECI)
R1	ESIPT	$H_1 \cdots O_6 = 1.0 \pm 0.2 \text{ \AA}$ for $\geq 50 \text{ fs}$	0.51 ± 0.04	0.55 ± 0.04
R2	CoIn	$C_3C_4C_5O_6 = 90 \pm 30^\circ$ for $\geq 50 \text{ fs}$	0.72 ± 0.03	0.77 ± 0.03
R3	CCCO rot.	$C_3C_4C_5O_6 = 180 \pm 30^\circ$ for $\geq 300 \text{ fs}$	0.73 ± 0.03	0.78 ± 0.03
R4	“Wrong isomer”	$H_8 \cdots O_2 = 1.0 \pm 0.2 \text{ \AA}$ for $\geq 300 \text{ fs}$	0.28 ± 0.03	0.34 ± 0.03
R5	R1 \rightarrow R2		0.51 ± 0.04	0.55 ± 0.04
R6	R1 \rightarrow R3		0.34 ± 0.03	0.36 ± 0.03
R7	R1 \rightarrow R2 \rightarrow R4		0.28 ± 0.03	0.34 ± 0.03

A second improvement to the dynamics of **SAC** was achieved by restricting the starting structures to be planar – as it was done for **SAMe** (cf. Sec. 6.4.1). The evaluation of this third ensemble proved that this was the main improvement for the dynamics. Not only is S_0 repopulated in about 13 ps, but nearly all trajectories (97%) perform the ESIPT according to rule R1 in Table 4.1. The trajectories fulfilling rule R7 only increase to 45%, though.

In summary, the introduction of the Peci keyword introduced a different shape of the S_1 PES after the ESIPT which also resulted in slightly improved dynamical properties in terms of larger QY. However, the effect is within the statistical error margin, which can be probably attributed to the fact that the ESIPT barrier is 0.05 eV smaller. A more careful selection of the starting structures has a more noticeable effect.

The photodynamics of **SMA** from the Supplementary Information of the publication were also improved using the Peci approach as well as using the RM1 parametrization.

The PES of the ESIPT process will be investigated in the same way as before. Starting with the 1D scan of the O_2-H_1 distance (Fig. 4.3) reveals that, in contrast to both **SAC** and **SAMe**, i) the PT leads to a stable minimum – after a relatively large energy barrier of 0.3 eV – on the S_0 surface without triggering a rotation of the methyl imine (CNMe) group; ii) the ESIPT happens without any energy barrier in S_1 , triggering the rotation of the CNMe group after 1.6 Å.

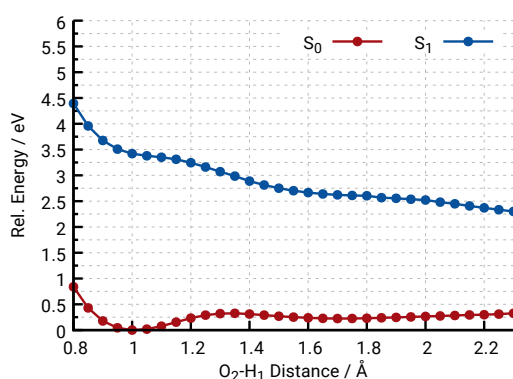


Figure 4.3.: One-dimensional PES scans of the O_2-H_1 distance of **SMA** using the keyword-based approach with Peci excitation level. The energies are given relative to the minimum on S_0 .

Both features are also seen on the 2D PES, where both the additional S_0 minimum and the non-stable Franck-Condon region on the S_1 surface are recognizable (Fig. 4.4).

Averaging over 200 surface-hopping trajectories, each propagated for 45 ps, reveals that **SMA** is the fastest and most efficient **SAC** derivative². The ensemble presents a complete deactivation to S_0 in 510 fs (PI at 230 fs), a 99% ESIPT QY and 80% of the trajectories perform the rotation of the CNMe group. The same values were already obtained using the old setup, with the exception of the ESIPT QY that increased by 30%.

Because of the superior photodynamical properties, **SMA** was also used as a test system for investigating the back-reaction from the rotated, proton-transferred form on S_0 back to the initial conformation – as studied in Ref. 122.

A set of 200 surface-hopping trajectories with the initial structures seeded from a ground-state trajectory of the rotated conformation, reveals that this reaction direction is highly efficient as well:

² Even compared to the structural analogues presented in the following section.

4.2 Further Benchmarking

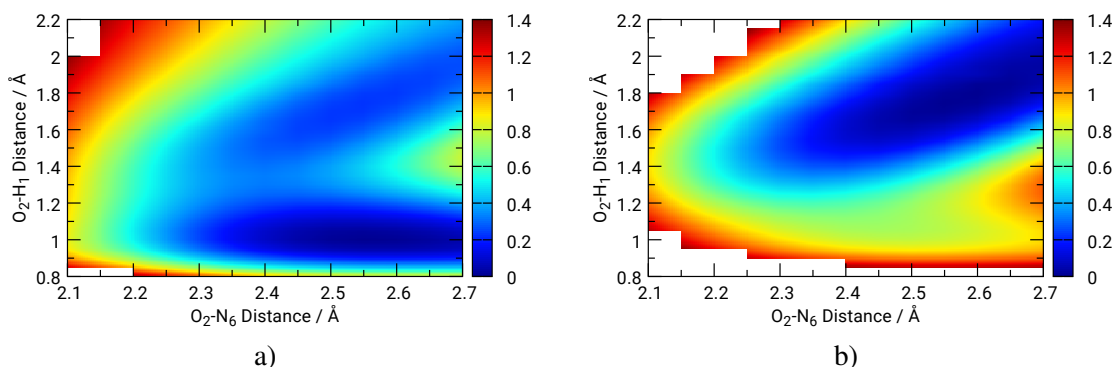


Figure 4.4.: Comparison of the 2D PES scans of the phenolic O_2-H_1 distance and O_2-N_6 distance of **SMA** using the new keyword-based approach (PECI): a) following the S_0 gradient, b) following the S_1 gradient. The energies are restricted to be less or equal to 1.4 eV for a better visual comparison to all other shown 2D PES.

S_0 is repopulated within 535 fs (PI at 150 fs) and 64% of all trajectories rotated back to the initial position before returning the proton (H_1) to the former proton-donor site (O_2). The faster PI time of the reverse ESIPT process is in good agreement with the results from the OM2/MRCI dynamics of Spörkel *et al.* with 96.8 fs, which was also faster than the “forward” reaction at 142 fs.^[122]

To study the effect of different proton donor and acceptor sites on the ESIPT reaction process, more derivatives of **SAc** were investigated in J. Müller’s Bachelor’s thesis, which is summarized in Sec. 6.4.

4.2 Further Benchmarking

Because **SMA** is structurally similar to **SAc**, a second benchmarking system was chosen that exhibits a different structural motif: A seven-membered ring, shaped by the hydrogen-bond between H_1 and N_7 .

4-(2-hydroxybenzylidene)-1H-imidazol-5(4H)-one (**OHBI**, Fig. 4.5) was studied by Cui *et al.*^[125] with the semiempirical OM2/MRCI method of Thiel and coworkers.^[123,124]

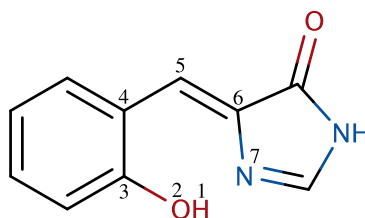


Figure 4.5.: Sketch of 4-(2-hydroxybenzylidene)-1H-imidazol-5(4H)-one (**OHBI**).

OHBI is one of many (truncated) model systems for the chromophoric core within the green fluorescent protein (GFP). The established deactivation pathway for these systems after electronic excitation to S_1 ($\pi\pi^*$) involves a CoIn between S_0 and S_1 . The CoIn was found at a twisted

$C_3C_4C_5C_6$ dihedral, but in contrast to, e.g., azobenzene, not at an angle of 90° but less, therefore favoring the *cis* conformation after return to S_0 (cf. Ref. 125 and references therein). Ref. 125 was chosen as the benchmark because it is the only study on the photodynamics of GFP chromophores that also utilizes a semiempirical Hamiltonian.

The computational setup for the following calculations was chosen to be [12,12]-FOPECI-RM1 as in the previous section. The seeding trajectory for the surface-hopping trajectories was propagated for 15 ps on S_0 . From this, 200 trajectories were initialized on S_1 and propagated for 2 ps as in Ref. 125. There will be no detailed evaluation of the ESIPT process of OHBI in the following, but only three aspects that will be compared between the present and reference set of trajectories.

The first quality criterion is the S_1/S_0 population inversion (PI) time which was found to be 330 fs in Ref. 125. Experimentally found S_1 decay times³ for 4-(2-hydroxybenzylidene)-1,2-dimethyl-1H-imidazol-5(4H)-one⁴ (OHBDI) were found to be 270 fs and 230 fs in acetonitrile and dichloromethane solution, respectively. The [12,12]-FOPECI-RM1 dynamics reveal a PI time of 195 fs – which corresponds to a decay time of 238 fs – which fits the experimental values for OHBDI nicely and is also in the same order as the OM2/MRCI result. The difference between the two theoretical values most likely comes from the different underlying Hamiltonians. Additionally, statistical errors when averaging over a set of trajectories play an important role.

As a sidenote, according to the state populations of the OM2/MRCI dynamics (cf. Fig. 8 in Ref. 125), the ensemble never fully returns to S_0 , with $\sim 15\%$ of the trajectories still being in S_1 after 2 ps. The [12,12]-FOPECI-RM1 trajectories are completely de-excited after 550 fs.

The second quality criterion is the *cis*→*trans* QY of the central carbon bond ($C_3C_4C_5C_6$) which was experimentally found to be less than 5% for OHBDI.^[126] In the OM2/MRCI dynamics no *cis*→*trans* isomerization was observed and the [12,12]-FOPECI-RM1 trajectory ensemble shows a QY of $7\pm 2\%$ ⁵. Considering that not all of those trajectories show a stable *trans* form, this second criterion is also fulfilled by the SEQM FOCI setup.

The third and last criterion is the $C_4C_5C_6N_7$ dihedral rotation which both trajectory ensembles perform. No QY was given for the OM2/MRCI ensemble – Ref. 125 only states that it happens –, but $16\pm 3\%$ of the [12,12]-FOPECI-RM1 perform this rotation. A qualitative difference between the two sets is that in the former case the rotation only happens after the ESIPT, i.e., when H_1 had returned to O_2 . Here, the same rotation happens also while H_1 is still attached to N_7 . Given the lack of information in Ref. 125 if this rotation is also a possible deactivation channel for **OHBI** it is unclear how the importance of the third criterion shall be rated.

However, overall both semiempirical CI approaches seem to be in good agreement. This supports the use of SEQM FOCI for ESIPT reactions.

³ The connection between PI times and decay times can be expressed as $[\text{PI time}]/\ln(2)=\tau$ according to unimolecular rate theory.^[125]

⁴ Although this is a different system, the electronic influence of the two additional methyl groups on the dynamics can be probably neglected.

⁵ The deviation was calculated as $\sqrt{QY(1-QY)/N}$ with N being the number of trajectories.

4.3 Structural Analogues of Salicylic Acid

All structural analogues/derivatives of salicylic acid so far had an oxygen atom as hydrogen donor. Also, the proton stayed at the respective crane group of the system, after it had rotated.

In the following, four more systems shall be investigated, with three having nitrogen atoms as hydrogen donor and acceptor, and one that has a second proton accepting site for the rotated conformation⁶ (Fig. 4.6). All systems utilize the methyl imine (CNMe) crane of **SMA** which performed the most efficient ESIPT process so far.

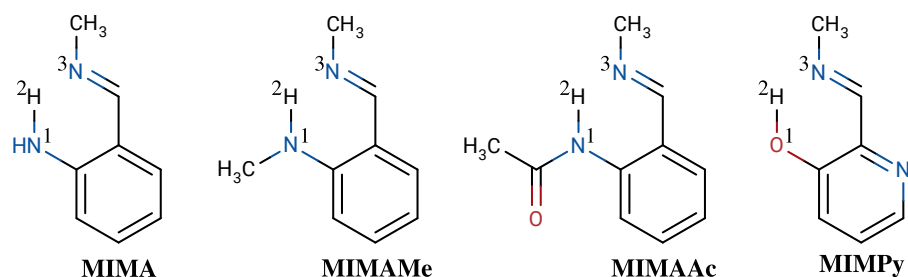


Figure 4.6.: Schemes of the four new structural analogues of SAc studied for their photo-dynamical properties. Three are based on 2-[(1E)-(methylimino)methyl]aniline (**MIMA**) with methyl (**MIMAMe**) and acetamide (**MIMAAc**) substitutions, respectively; one uses a pyridine backbone instead of a phenyl one: 2-[(1E)-(methylimino)methyl]pyridin-3-ol (**MIMPy**).

To decide which of the three N-donor-N-acceptor systems might be most suited as an ESIPT switch, they will first be compared by their change in N_1-H_2/O_1-H_2 bond strength indicated by the respective change in vibrational frequency in the excited state.

Table 4.2.: Calculated vibrational frequencies on the [12,12]-FOPECI-RM1 level of the N_1-H_2/O_1-H_2 stretch band in S_0 , S_1 and S_2 . All values in cm^{-1} .

	MIMA	MIMAMe	MIMAAc	MIMPy
S_0	3031	2796	2698	2639
S_1	2419	n.a.	1257	n.a.
S_2	3021	2784	2638	3063

From the vibrational frequencies summarized in Table 4.2 it can be deduced that any ESIPT would happen exclusively in S_1 because there is no shift in S_2 . **MIMAMe** and **MIMPy** performed the PT during the geometry optimization in S_1 , meaning that there is no barrier in the excited state hindering the reaction. **MIMA** and **MIMAAc** both show red-shifted vibrational frequencies, meaning that the S_1 minimum is broader – thus, favoring the elongation of the N_1-H_2/O_1-H_2 bond length – but still leads to an energy barrier.

⁶ This system may also be seen as a derivative of 3-hydroxy picolinic acid (3HPA) which is also currently being studied experimentally by means of ultrafast transient vibrational spectroscopy.^[49] There is also static computational data available.^[44]

Additionally, the 1D N_1 -H₂/O₁-H₂ distance scans in S_0 and S_1 were calculated which lead to the same findings (Fig. 4.7).

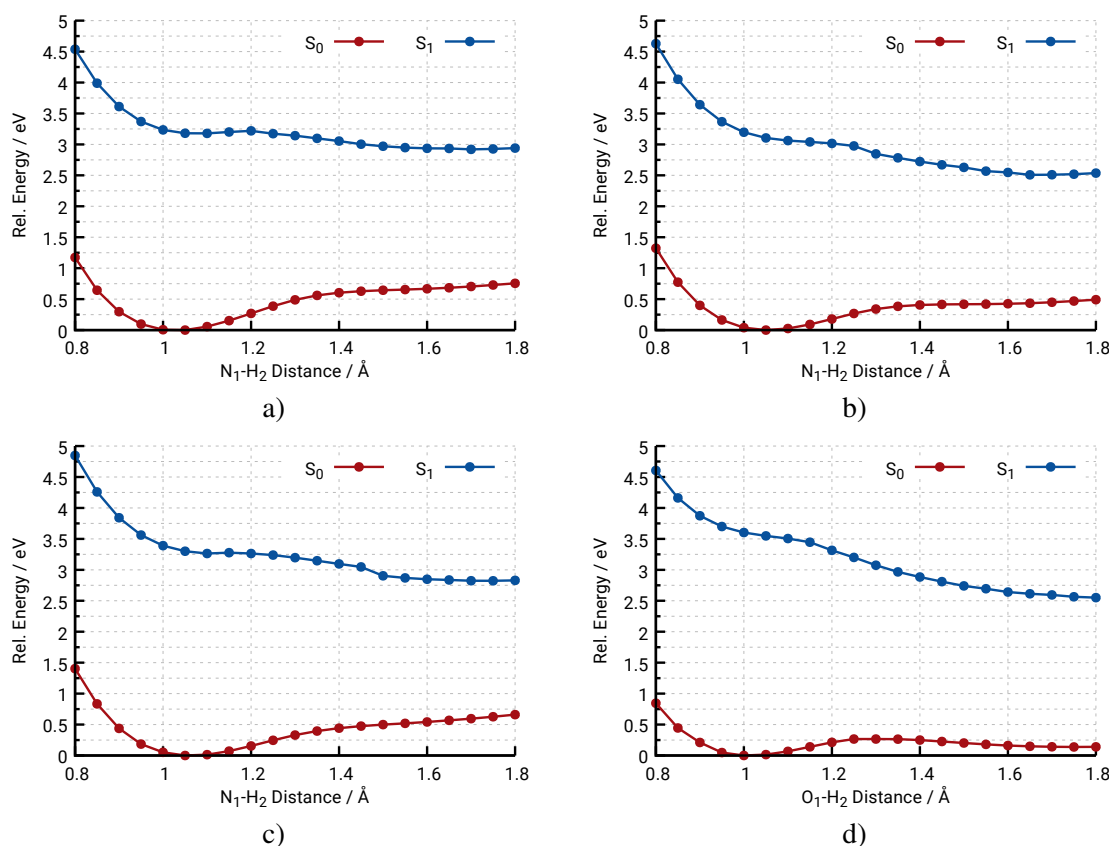


Figure 4.7.: Calculated 1D PES for the four studied structural analogues of SAC: a) **MIMA**, b) **MIMAMe**, c) **MIMAAc** and d) **MIMPy**. The energies are given relative to the minimum on S_0 .

For **MIMA** and **MIMAAc** there is a small energy barrier which makes the calculation of the N_1 -H₂/O₁-H₂ vibrations possible at all. As expected, the ES IPT processes of **MIMAMe** and **MIMPy** are not hindered by a barrier. Interestingly, all three N-donor-N-acceptor system do not exhibit a minimum on S_0 for the proton-transferred structure, in contrast to the two O-donor-N-acceptor systems **SMA** and **MIMPy**.

Subsequently, geometry optimizations were performed on the PBE0-D3BJ/def2-TZVPP level on S_1 to benchmark the PES; and the S_1 scans shown in Fig. 4.7 were recalculated using PBE0-D3BJ/def2-SVP⁷.

For **MIMA** the geometry optimization failed because the rotation of the CNMe crane started during the optimization. As it is the case for all previous ES IPT systems, this results in the system to reach a CoIn which DFT is not able to describe, therefore failing at the optimization. This also tells us that there is no barrier for the ES IPT in S_1 , which [12,12]-FOPECI-RM1 predicted, though.

⁷ A smaller basis set was used for these calculations because i) a large amount of structures needed to be screened and ii) only the relative energies of the scan are of interest, thus systematic errors within a set of structures can be neglected.

4.3 Structural Analogues of Salicylic Acid

However, the RM1 barrier is also very small so that the barrier may also come from the restriction of the DOF during the scan. The DFT N_1 -H₂/O₁-H₂ scans of the three remaining systems also reveal a barrierless ESIPT pathway in S_1 but without also initiating CNMe rotation. These scans are presented in the appendix (Sec. A.3.1). These findings confirm the good agreement of the SEQM FOCI methods with calculations on the DFT level of theory.

The calculation of the S_0 seeding trajectories for all N-donor-N-acceptor systems showed a rotation of the CNMe crane independent of the ESIPT. Therefore, there were less than 200 suitable starting structures for the photodynamics of **MIMA** (59) and **MIMAAc** (197) that could be extracted from the S_0 MD before the rotation had started. Therefore, the differences between the starting structures are relatively small, hence most of the trajectories should perform similarly. However, given that the hydrogen bond between H₂ and N₃ is the important motif for the ESIPT process, a qualitative correct behavior of the trajectories should be discernible⁸. There is no other major structural change that could have happened in S_0 that would not have eliminated the possibility of an ESIPT – hence even if there were 200 starting structures there would be no major structural difference between them anyways.

The results of the dynamics are summarized in Table 4.3.

Table 4.3.: Results of the photodynamics of **MIMA**, **MIMAMe** and **MIMAAc**, this includes QYs of the ESIPT process, as well as the times for 50% (PI) and 100% S_0 population. The same results for **SMA** are given in comparison.

	MIMA	MIMAMe	MIMAAc	SMA
ESIPT QY	0.97±0.02	0.65±0.03	0.95±0.02	0.99±0.01
PI / fs	320	175	290	230
return to S_0 / fs	570	630	635	510

In terms of overall efficiency, **MIMA** and **MIMAAc** show virtually the same QYs of the ESIPT process compared to **SMA** and also perform very similar in terms of de-excitation to S_0 . The reason why **MIMAMe** seem to stick out in terms of QY was found to be that the ESIPT form is less stable because the system also has a much smaller half-life of S_1 population than the other systems. The criterion for determining the QY for all system was set to a H₂-N₃ distance of 1 ± 0.02 Å that should be stable for 100 fs. When lowering the duration to 50 fs, the QY for **MIMAMe** increases to 0.91 ± 0.02 . For comparison, the QYs for **MIMA** and **MIMAAc** only increase by 1% and 2%, respectively. The additional ~30% of trajectories for **MIMAMe** do not perform the rotation of the CNMe group and thus seem to favor the return to S_0 without the CoIn along the pathway of the crane-rotation. This was already observed in the deactivation dynamics of **SAC**.

Nevertheless, based on these results, all three systems seem to be excellent candidates for ultrafast ESIPT switches.

⁸ As was already seen for the ESIPT dynamics of **SAC**, including the rotated structures as starting structures for the excited-state dynamics only has negative impact on the general (average) performance of the ensemble.

The dynamics of **MIMPy** also showed a high ESIPT QY of 0.98 ± 0.01 , however, the expected second PT to the nitrogen acceptor N_4 after the crane-rotation did not happen. The complete de-excitation to S_0 also was considerably longer than for the other systems with 3.2 ps – however, the PI time was 310 fs.

A scan of the $N_3-H_2 \cdots N_4$ DOF revealed a large energy barrier on both S_0 and S_1 of ~ 1 eV, explaining the behavior of the trajectories (Fig. 4.8a). Recalculating the scan on the TDDFT level of theory confirms this finding (Fig. 4.8b).

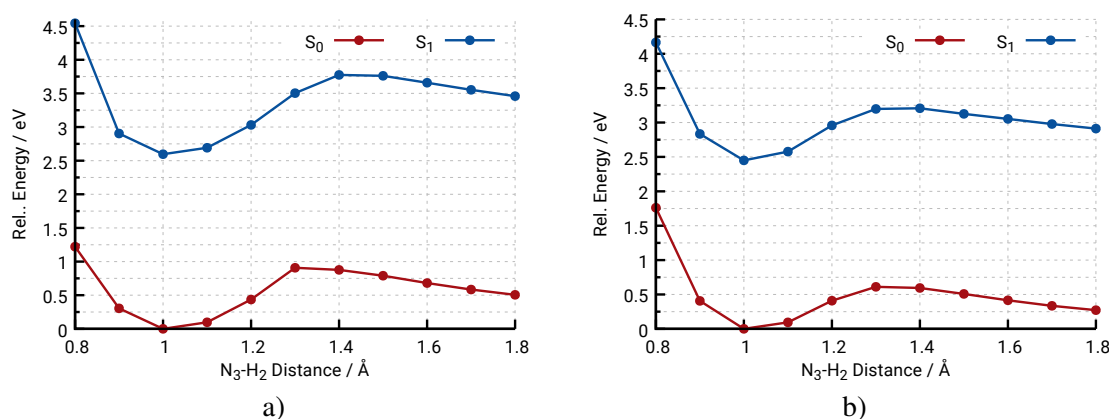


Figure 4.8.: Calculated 1D PES for the $N_3-H_2 \cdots N_4$ DOF of **MIMPy**. The shown PES for S_0 and S_1 were obtained by following the gradient on that respective state. a) Using [12,12]-FOPECI-RM1, b) using PBE0-D3BJ/def2-SVP. The energies are given relative to the minimum on S_0 .

But this is a correct chemical behavior, because the nitrogen atom in the aromatic ring is less basic than the imine group, thus favoring the N_3-H_2 bond over H_2-N_4 . Additionally, the $N_3-H_2 \cdots N_4$ “bridge” is part of a five-membered ring. For the PT to happen the crane must bend considerably which is much more pronounced than in the six-membered ring formed by the initial $O_1-H_2 \cdots N_3$ bridge, which also contributes to the energy barrier (Fig. 4.9).

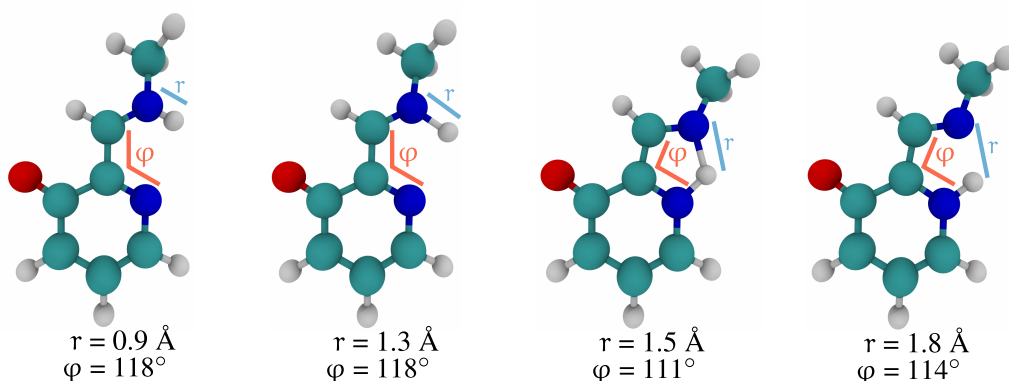


Figure 4.9.: Structures and PT coordinates during the scan of **MIMPy** shown in Fig. 4.8.

To include the capability of **MIMPy** to perform this second PT substituent effects should be investigated as it was done for **SMA**^[127] or **HMQCA**.^[45] Given that [12,12]-FOPECI-RM1

seems to perform on a same level of accuracy as TDDFT for these types of systems, a relatively low-cost optimization process to tune the properties of **MIMPy** or ESIPT-systems in general could be attempted. Such a tuning-process has already been done in the author's work group.^[128]

However, in contrast to all other black-box QM methods, only a SEQM FOCI setup also allows for investigations on the dynamics of the designed systems.

4.4 Fused-Ring Systems

4.4.1 Quinoline Backbone

Increasing the level of complexity of the reaction process, **HMQCA** facilitates the rotation after the ESIPT as means of proton transport to another acceptor site⁹. Previous studies looked at the shape of the PES with static calculations^[43] and at the effects of substitution.^[45] The results of these works suggested the following mechanism (Fig. 4.10): Starting with an excitation to S_2 ($\pi\pi^*$), the electronic density rearranges to allow the proton transfer to happen; relaxation to S_1 ($n\pi^*$) takes place and the proton is transferred to the aldehyde oxygen; this triggers the rotation of the "crane" as it was the case for **SAc**. However, for **HMQCA** the 180° rotation is now the desired pathway, because upon return to S_0 via a CoIn the proton is transferred to the nitrogen atom (N_4) in the backbone of the system.

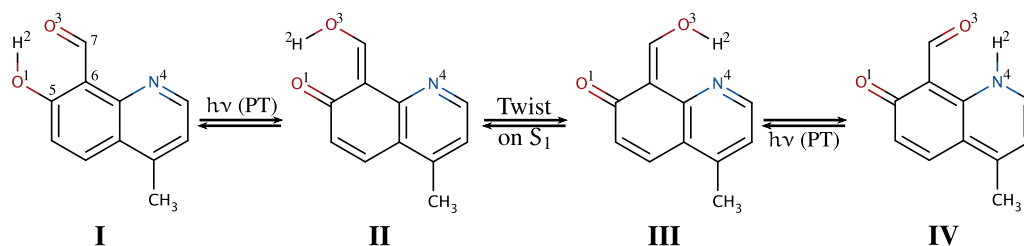


Figure 4.10.: Reaction scheme for the ESIPT-switching process of **HMQCA**. For all following systems I, II, III and IV will correspond to the same stages of the reaction as shown here.

For **HMQCA** a meta optimization was attempted before the calculations because reference data on the CC2 level of theory^[43] suggested that the ESIPT is started in S_2 rather than S_1 , because of the characters of the excited states: S_1 is an $n\pi^*$ state and S_2 $\pi\pi^*$. A first naive approach using the available data from Ref. 43 revealed that the PM6* Hamiltonian, a near-minimal active space consisting of six electrons in five orbitals and a CISD excitation level ([6,5]-FOCISD-PM6*) is able to reproduce the reference energies of **I-HMQCA** and **IV-HMQCA** (Table 4.4). Unfortunately, this is not the [12,12]-FOPECI-RM1 setup from the previous systems.

First, 1D scans of both I→II PT (Fig. 4.11a) and III→IV PT (Fig. 4.11b) were performed to check the expected energy barriers. The ESIPT (I→II), although being initialized in S_2 , seems to happen after internal conversion to S_1 because the calculated energy barrier is much smaller in that

⁹ The difference to **MIMPy** would be that **HMQCA** features a six-membered-ring motif for the second PT which seems to be favorable.

Table 4.4.: Calculated static properties of the I and IV form of **HMQCA** using [6,5]-FOCISD-PM6* compared to the results of CC2/aug-cc-pVDZ calculations from Ref. 43. All values in eV.

System	Property	[6,5]-FOCISD-PM6*	CC2
I	$\Delta E_{S_0 \rightarrow S_1}$	3.85	3.71
	$\Delta E_{S_0 \rightarrow S_2}$	3.94	3.87
	$\Delta E_{S_0 \rightarrow S_3}$	4.51	4.53
	ΔE_{I-IV}	0.26	0.42
IV	$\Delta E_{S_0 \rightarrow S_1}$	2.92	2.96
	$\Delta E_{S_0 \rightarrow S_2}$	3.42	3.03
	$\Delta E_{S_0 \rightarrow S_3}$	4.00	3.70

state (~ 0.1 eV in S_1 vs. ~ 0.4 eV in S_2). The II-form at 1.7 \AA is also more stable in the excited state than the structure in the Franck-Condon region. After rotation of the CO(H)H group, **HMQCA** follows a steep gradient towards a CoIn to S_0 . This would also allow the system to overcome the small energy barrier of ~ 0.1 eV. However, the minimum of **IV-HMQCA** is only 0.02 eV lower in energy than the one of **III-HMQCA**. Both features lead to the assumption that both forms of **HMQCA** can coexist in S_0 . Interestingly, in both S_1 and S_2 the gradient does not seem to favor the backwards reaction from **IV-HMQCA** to **III-HMQCA**.

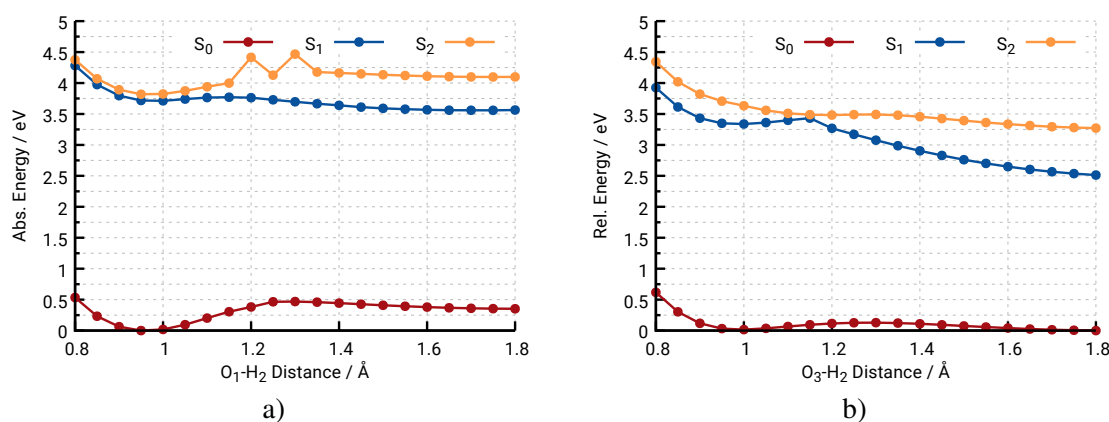


Figure 4.11.: One-dimensional PES scans of the I→II (a) and III→IV (b) PTs of **HMQCA** using [6,5]-FOCISD-PM6*. The scan for each state followed its respective gradient. The spikes in a) at 1.2 and 1.3 \AA correspond to optimizations that led to minima on the PES that do not fit to the minima found by the other steps of the scan. The energies are given relative to the minimum on S_0 .

When the photodynamics are calculated – 200 trajectories and 25 ps propagation time – using the [6,5]-FOCISD-PM6* setup, the expected reaction is correctly reproduced. The trajectories are initialized in S_2 , because this is the allowed transition. Half of the trajectories reach S_1 in 62 fs showcasing the very efficient deactivation pathway of **HMQCA** in the excited states (Fig. 4.12a). S_0 gets populated after 400 fs and after 7.1 ps half the population has returned to the ground state.

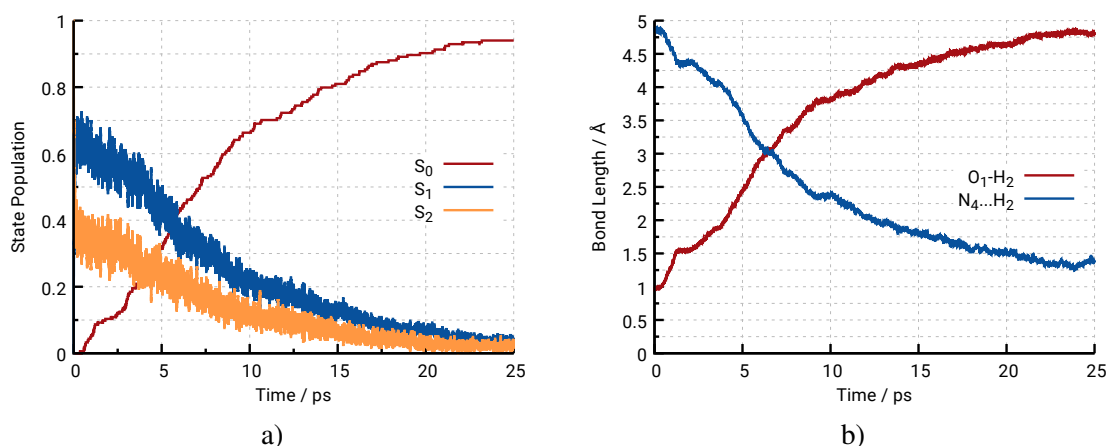


Figure 4.12.: Ensemble-averaged results of the surface-hopping MD simulations of **HMQCA** using [6,5]-FOCISD-PM6*. a) State populations. b) Time-evolution of the O₁-H₂ and H₂⋯N₄ distances.

At this point, however, 16% of the trajectories are still in S_2 . The calculated QY is 0.59 ± 0.04 for reaching a H₂⋯N₄ distance of 1 ± 0.2 Å, i.e., completing the whole reaction from **I-HMQCA** → **IV-HMQCA**.

The evolution of the important DOF of this reaction, i.e., the O₁-H₂ and H₂⋯N₄ distances, support this result (Fig. 4.12b)

Because these values are averaged they only allow for a qualitative conclusion. Therefore the four individual species – I, II, III, IV of **HMQCA** – are tracked separately during the propagation in each state. Table 4.5 lists the specifications that were used to define each isomer.

Table 4.5.: DOF specifications that define each of the four **HMQCA** isomers: I, II, III, IV.

Isomer	Structural Identifiers
I	H ₂ -O ₁ = 1.0 ± 0.15 Å
II	H ₂ ⋯O ₃ = 1.0 ± 0.15 Å; $-30^\circ \leq \text{C}_5\text{C}_6\text{C}_7\text{O}_3 \leq 90^\circ$
III	H ₂ ⋯O ₃ = 1.0 ± 0.15 Å; $90^\circ \leq \text{C}_5\text{C}_6\text{C}_7\text{O}_3 \leq 210^\circ$
IV	H ₂ ⋯N ₄ = 1.0 ± 0.15 Å

The MD propagation starts in S_2 with all systems being in the I-form of **HMQCA** (Fig. 4.13c). After 64 fs, half of the I isomers reached S_1 . The first **II-HMQCA** is formed after 248 fs in S_1 , however, at no point during the dynamics is there a noticeable accumulation of **II-HMQCA**. The first **III-HMQCA** is formed after 407 fs and **IV-HMQCA** after 580 fs, both in S_0 . No **I-HMQCA** reached S_0 which means that all systems have at least performed the first PT. The accumulation of **III-HMQCA** in S_0 confirms the assumption that both **III-HMQCA** and **IV-HMQCA** are stable minima in the electronic ground state.

Thus, the ES IPT-switching process of **HMQCA** seems to be a very efficient process which involves two internal conversions: The first happens close to the Franck-Condon region in S_2 that

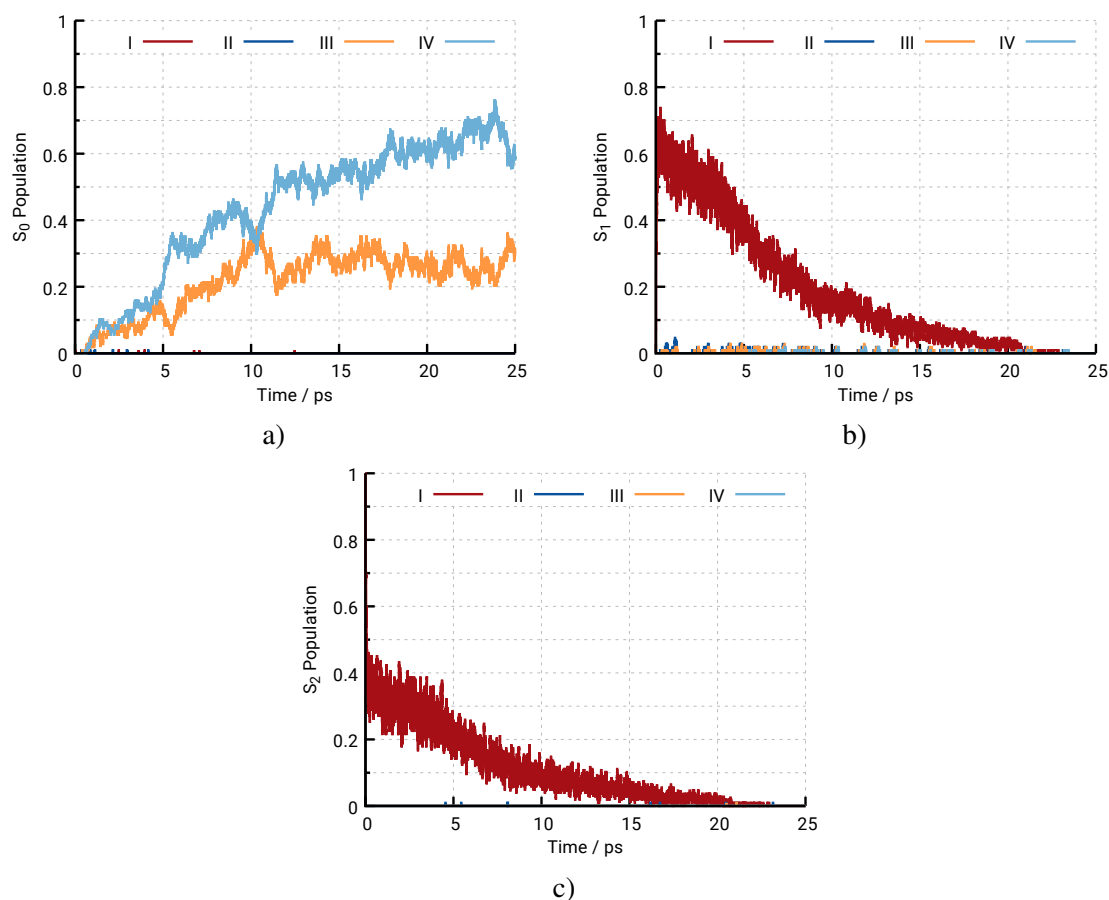


Figure 4.13.: Time-evolution of the four **HMQCA** isomers during the surface-hopping MD calculation starting in S_2 . The populations are shown separately for S_0 (a), S_1 (b) and S_2 (c).

requires only small geometric changes, and the second is induced by the rotation of the aldehyde group leading to a CoIn. This is shown exemplarily for one selected trajectory in Fig. 4.14.

This trajectory represents a complete ES IPT-switching process in about 3.2 ps. In the first 2 ps the system hops frequently between S_2 and S_1 , which is not surprising given the small energy difference of less than 0.1 eV. Eventually, the proton gets transferred. This immediately triggers the rotation of the aldehyde group that leads to the formation of **III-HMQCA** after 2.4 ps. The S_1/S_0 CoIn is reached after 2.2 ps during the rotation. **IV-HMQCA** is formed after another 600 fs *via* a ground-state PT. In the remaining propagation – not shown in Fig. 4.14 – the proton is transferred six more times to the aldehyde group. This indicates that the energy barrier between **IV-HMQCA** and **III-HMQCA** can be easily overcome. It is noted that this particular run did not use a thermostat, thus after isomerization the system has a large amount of kinetic energy that leads to an easier ground-state PT than it probably should.

To confirm this effect, the trajectories were recalculated with an added thermostat. With the new set of trajectories the $H_2 \cdots N_4$ QY turned out to be within the error margin of the non-thermostat run with 0.53 ± 0.04 . The overall behavior of the trajectories is also qualitatively the same in

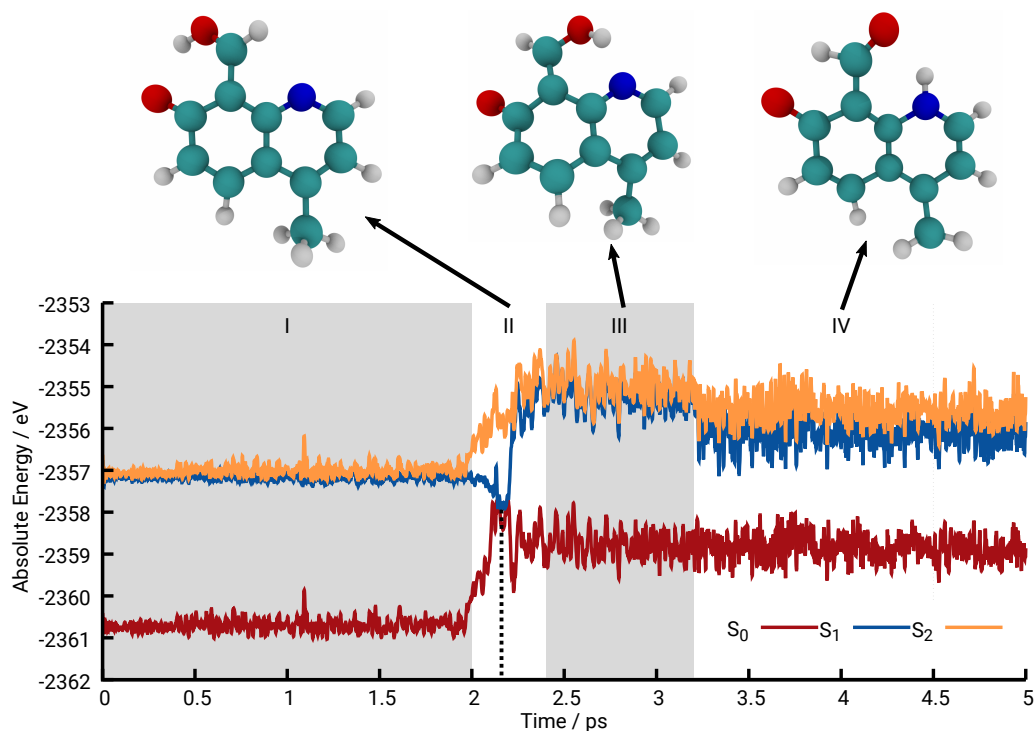


Figure 4.14.: State energies and structures for one example trajectory of **HMQCA**. The ESIP for this particular trajectory happens after 2 ps, which triggers the rotation of the aldehyde group and leads to the CoIn shortly after 2.1 ps. At 2.4 ps the system reaches the III-form which is stable for ~ 600 fs until it eventually reaches the desired photoproduct form (IV).

comparison. The only two important changes are that the electronic de-excitation now happens faster (Fig. 4.15a) and the IV-form is now more stable – as expected – indicated by the larger population of **IV-HMQCA** in S_0 at the end of the propagation (Fig. 4.15b).

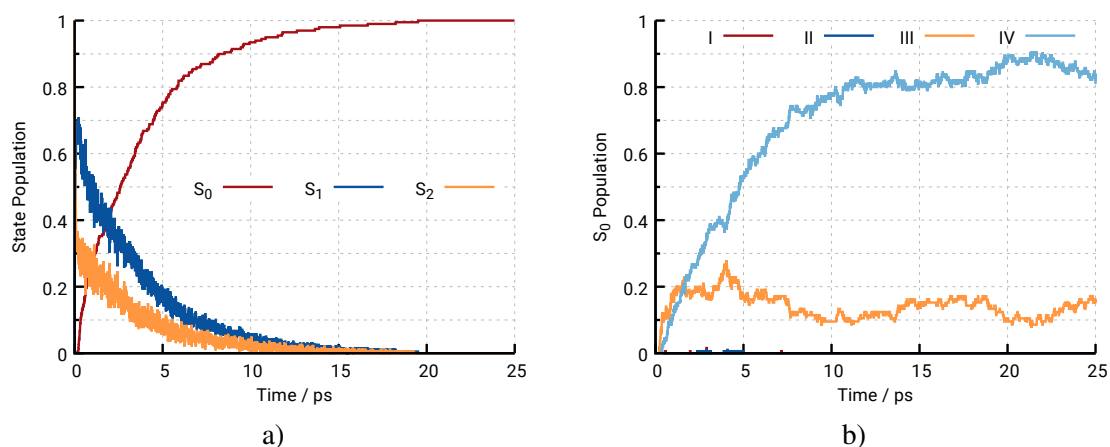


Figure 4.15.: Results for the recalculated **HMQCA** dynamics using [6,5]-FOCISD-PM6* with the added thermostat. a) Ensemble-averaged state populations; b) ensemble-averaged isomer population in S_0 .

The back-reaction starting from **IV-HMQCA** in S_2 was also investigated. In accordance with the III→IV scan (cf. Fig 4.11b) the energy barrier hinders an effective IV→III ESIPT. Only 22% of the 200 trajectories reached the III-form during the propagation. The rotation of the aldehyde group was not triggered either, resulting in only a 45% population of S_0 after 25 ps. However, after 2.5 ps the trajectory ensemble completely de-populated S_2 , confirming that there is an accessible CoIn between S_1 and S_2 along the ESIPT reaction coordinate.

At the beginning of this section it was said that the meta optimization approach was naive. This is because it only took the energies of two ground-state forms of **HMQCA** into account. **HMQCA** is the first ESIPT system studied in this thesis so far that did not start its excited state dynamics in S_1 but in S_2 , because of the forbidden $n \rightarrow \pi^*$ excitation. However, in the [6,5]-FOCISD-PM6* setup the first two excited states are not dominated by the according $n\pi^*$ or $\pi\pi^*$ SDs, but as a mixture of both, which is not backed up by own DFT (PBE0-D3BJ/def2-TZVPP) and reference *ab initio* calculations.^[43] Both calculations predict a dominant $n\pi^*$ character for S_1 and $\pi\pi^*$ character for S_2 . Getting the state characters right is important so that the electronic wavefunction is correct in the respective state to give meaningful results from the dynamics when compared to experimental results.

In order to find a potentially more suited SEQM FOCI setup, important points on the PES were benchmarked on the DFT level (PBE0-D3BJ/def2-TZVPP). This includes geometry optimizations of the four isomers shown in the proposed mechanism (Fig. 4.10), their excitation energies to S_1 and S_2 and their corresponding excited-state characters. These results were then used in the meta optimization scheme (Table 4.6).

The best fitting setup was found to be [10,12]-FOCIS-AM1. However, none of all tested combinations could reproduce all state characters correctly. The results for **II-HMQCA** and **IV-HMQCA** in S_2 are omitted from the table, because the geometry optimizations did not converge¹⁰. The obtained fitness for this setup was 63.31. This value alone can not be used as a quality criterion. However, it can be compared to the value for [6,5]-FOCISD-PM6* which is 25014.39. This high value is due to missing assignments of five $n\pi^*$ characters¹¹ which contributed a value of 5·(5·999) to the fitness¹². When adjusting the fitness value by removing these five entries, the fitness reduces to 39.39. In comparison, when neglecting the same five $n\pi^*$ assignments, the [10,12]-FOCIS-AM1 fitness becomes 47.23. This means that the remaining reference values are better recovered by [6,5]-FOCISD-PM6* than by the newly found setup.

Before the dynamics were run, scans of both I→II PT and III→IV PT are performed on the SEQM level of theory and compared against DFT reference calculations (PBE0-D3BJ/def2-SVP). Comparing both 1D scans obtained with the [10,12]-FOCIS-AM1 (Figs. 4.16a and 4.16b) and the ones using the [6,5]-FOCISD-PM6* setup reveals that all calculated energy barriers are now

¹⁰ This is because the geometry optimization probably led to a CoIn which is deduced from degenerate S_1 and S_2 energies at the point of failure. This is same behavior as encountered for the SAc analogues in Sec. 4.3.

¹¹ The entries are: The S_1 character for **I-HMQCA** optimized on S_0 , the S_1 character for **I-HMQCA** optimized on S_2 and the S_1 characters for **III-HMQCA** optimized on S_0 , S_1 and S_2 .

¹² Values that could not be parsed from a calculation are automatically set to 999. Because the state character received a weighting factor 5 and there are five missing entries, thus the factor of 25.

Table 4.6.: Results of the DFT (PBE0-D3BJ/def2-TZVPP) benchmarks and meta optimization for important points and properties of the PES of **HMQCA**. The best fitting SEQM setup was found to be [10,12]-FOCIS-AM1. All energies of the optimized structures (“ S_X opt.”) are given relative to the S_0 energy of **I-HMQCA**. All energies in eV. The excited-state characters S_X ($n\pi^*/\pi\pi^*$) are given as the absolute value of the CI coefficient for the corresponding SD. The corresponding orbital representations are given in the appendix (Sec. A.8.1).

Isomer	Property	TDDFT	SEQM	Isomer	Property	TDDFT	SEQM
I	$\Delta E_{S_0 \rightarrow S_1}$	3.59	3.42	III	S_0 opt.	0.35	0.41
I	$\Delta E_{S_0 \rightarrow S_2}$	3.89	3.53	III	$\Delta E_{S_0 \rightarrow S_1}$	3.17	3.45
I	S_1 ($n\pi^*$)	0.89	0.41	III	$\Delta E_{S_0 \rightarrow S_2}$	3.56	3.53
I	S_2 ($\pi\pi^*$)	0.93	0.52	III	S_1 ($n\pi^*$)	0.95	0.53
I	S_1 opt.	2.52	3.31	III	S_2 ($\pi\pi^*$)	0.98	0.62
I ^a	$\Delta E_{S_0 \rightarrow S_1}$	2.49	2.81	III	S_1 opt.	3.24	3.59
I ^a	$\Delta E_{S_0 \rightarrow S_2}$	3.79	3.42	III ^a	$\Delta E_{S_0 \rightarrow S_1}$	2.56	2.92
I ^a	S_1 ($n\pi^*$)	0.96	0.51	III ^a	$\Delta E_{S_0 \rightarrow S_2}$	3.10	3.26
I ^a	S_2 ($\pi\pi^*$)	0.93	0.39	III ^a	S_1 ($n\pi^*$)	0.96	0.53
I ^b	S_2 opt.	3.68	3.88	III ^a	S_2 ($\pi\pi^*$)	0.98	0.65
I ^b	$\Delta E_{S_0 \rightarrow S_1}$	3.52	3.25	III ^b	S_2 opt.	3.36	3.81
I ^b	$\Delta E_{S_0 \rightarrow S_2}$	3.57	3.49	III ^b	$\Delta E_{S_0 \rightarrow S_1}$	2.75	3.13
I ^b	S_1 ($n\pi^*$)	0.91	0.00	III ^b	$\Delta E_{S_0 \rightarrow S_2}$	3.05	3.19
I ^b	S_2 ($\pi\pi^*$)	0.95	0.00	III ^b	S_1 ($n\pi^*$)	0.97	0.03
II ^c	S_0 opt.	–	–	III ^b	S_2 ($\pi\pi^*$)	0.98	0.00
II	$\Delta E_{S_0 \rightarrow S_1}$	–	–	IV	S_0 opt.	0.29	0.42
II	$\Delta E_{S_0 \rightarrow S_2}$	–	–	IV	$\Delta E_{S_0 \rightarrow S_1}$	3.03	2.85
II	S_1	–	–	IV	$\Delta E_{S_0 \rightarrow S_2}$	3.14	3.22
II	S_2	–	–	IV	S_1 ($n\pi^*$)	0.91	0.09
II	S_1 opt.	3.64	3.42	IV	S_2 ($\pi\pi^*$)	0.95	0.04
II ^a	$\Delta E_{S_0 \rightarrow S_1}$	3.03	3.11	IV	S_1 opt.	2.98	2.98
II ^a	$\Delta E_{S_0 \rightarrow S_2}$	3.33	3.30	IV ^a	$\Delta E_{S_0 \rightarrow S_1}$	2.09	2.30
II ^a	S_1 ($n\pi^*$)	0.97	0.54	IV ^a	$\Delta E_{S_0 \rightarrow S_2}$	2.51	3.00
II ^a	S_2 ($\pi\pi^*$)	0.97	0.67	IV ^a	S_1 ($n\pi^*$)	0.96	0.17
II	S_2 opt.	3.72		IV ^a	S_2 ($\pi\pi^*$)	0.95	0.00
II ^b	$\Delta E_{S_0 \rightarrow S_1}$	3.27		IV	S_2 opt.	3.10	
II ^b	$\Delta E_{S_0 \rightarrow S_2}$	3.27		IV ^b	$\Delta E_{S_0 \rightarrow S_1}$	2.48	
II ^b	S_1 ($n\pi^*$)	0.69		IV ^b	$\Delta E_{S_0 \rightarrow S_2}$	2.48	
II ^b	S_2 ($\pi\pi^*$)	0.71		IV ^b	S_1 ($n\pi^*$)	0.95	
				IV ^b	S_2 ($\pi\pi^*$)	0.95	

a: Structure was optimized on S_1 , *b*: Structure was optimized on S_2 , *c*: Not a stable minimum on S_0 .

much higher: The I→II energy barrier increased to >1 eV and the barriers in all three states for the III→IV reaction are now >0.5 eV. In contrast, the III→IV S_0 scan with the [6,5]-FOCISD-PM6* setup (Fig. 4.11b) fits the DFT scan nicely, with both showing **IV-HMQCA** as more stable than **III-HMQCA**.

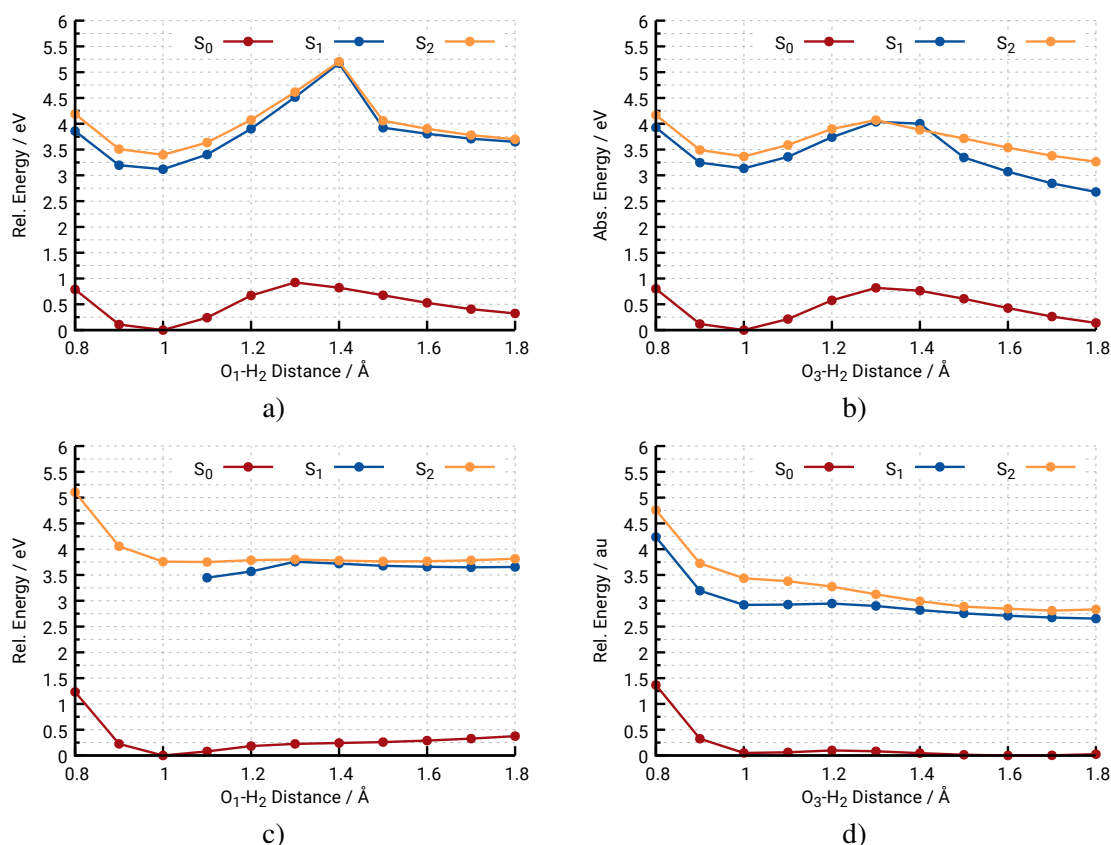


Figure 4.16.: One-dimensional PES scans of the I→II and III→IV PTs of **HMQCA** using [10,12]-FOCIS-AM1 (a,b) and PBE0-D3BJ/def2-SVP (c,d), respectively. The scan for each state followed its respective gradient. The energies are given relative to the minimum on S_0 .

As expected from the scan results, the evaluation of the dynamics does not present any ESIPT and, thus, also no switching of the aldehyde crane.

At this point the first problematic deficiencies of SEQM were found that could not be fixed by changing the CI setup or switching to any of the available parametrizations (AM1, RM1, PM3 and PM6*). Interestingly, azobenzene shows the same excited state pattern, i.e., a $n\pi^*$ S_1 and a $\pi\pi^*$ S_2 state, but a calculation using the standard, non-reparametrized AM1 Hamiltonian can reproduce this correctly. So in principle it should be possible to treat systems with this order of states using SEQM FOCL.

4.4.2 Crane-Substitution Effects

After having found a superior proton acceptor in **SMA** in comparison to its parent system **SAC**, the methyl imine (CNMe) group was used on the **HMQCA** backbone in hopes to enhance the dynamic properties of the system. The properties of this newly designed system – 4-methyl-8-[(1E)-(methylimino)methyl]quinolin-7-ol (**MMIMQO**) – are investigated in the following (Fig. 4.17).

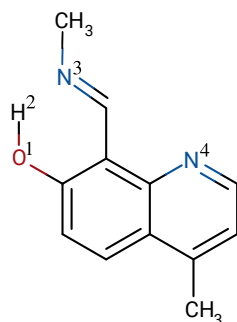


Figure 4.17.: Sketch of **MMIMQO**. An improved ESIPT switch based on the quinoline backbone of **HMQCA** utilizing the methyl imine crane of **SMA**.

The semiempirical setup was again determined *via* a meta optimization approach using DFT reference data (PBE0/def2-TZVPP) of the optimized geometries and their vertical excitation energies as reference. These values together with the values obtained using the SEQM FOCI setup are presented in Table 4.7.

The best fitting setup was found to be [12,12]-FOCIS-PM6*. However, this setup is not able to reproduce the state characters of all predicted $n\pi^*$ states, showcases deficiencies for the VEE and also gets the relative energy for the S_0 structures of **II-MMIMQO** and **III-MMIMQO** wrong. In the cases where even the $\pi\pi^*$ characters were not reproduced, the corresponding determinant just did not correspond to an excitation from the highest occupied MO (HOMO)- π orbital to the lowest unoccupied MO (LUMO)- π^* orbital but to different π and/or π^* orbitals. It is noted, though, that the SEQM energies are for the non-relaxed TDDFT optimized structures, i.e., only the one point of the PES was calculated which must not correspond to a minimum. Therefore scans of the ESIPT and ground-state PT are performed anyways. The obtained structures are then further benchmarked against relaxed scans using TDDFT (PBE0-D3BJ/def-SVP).

The scans for the I \rightarrow II and III \rightarrow IV PTs clearly show the advantageous ESIPT properties of the CNMe crane. There is no barrier in S_1 for the I \rightarrow II reaction (Fig. 4.18a). However, **II-MMIMQO** is still more stable than **I-MMIMQO** in S_0 even when the structure is relaxed on the PM6* PES. This leads to the formation of **II-MMIMQO** during the seeding dynamics in S_0 which defeats the purpose of the ESIPT. Nevertheless, the rotation of the CNMe crane can not happen in S_0 as easily¹³. The DFT scans (Fig. 4.18c) also exhibit a small energy barrier for both PTs in S_2 but a less steep gradient in S_1 .

The second downside of this system is that the IV-form is not a stable minimum on the ground state PES, but the III-form, effectively disabling the PT in S_0 (Fig. 4.18c). This feature is also present in the DFT scans (Fig. 4.18d). However, in hindsight, this could be the same chemically correct phenomenon that was already observed in **MIMPy**.

¹³ Actually, the scan of the crane rotation in S_0 from 0° to 180° does not reveal any energy barrier but only a steady increase in energy. This is in stark contrast to the barrierless rotation in S_1 . The scans are presented in the appendix (Sec. A.4).

Table 4.7.: Results of the TDDFT (PBE0-D3BJ/def2-TZVPP) benchmarks and meta optimization for important points and properties of the PES of **MMIMQO**. The best fitting SEQM setup was found to be [12,12]-FOCIS-PM6*. All energies of the optimized structures (“ S_X opt.”) are given relative to the S_0 energy of **I-MMIMQO**. All energies in eV. The excited-state characters S_X ($n\pi^*/\pi\pi^*$) are given as the absolute value of the CI coefficient for the corresponding SD. The corresponding orbital representations are given in the appendix (Sec. A.8.2).

Isomer	Property	TDDFT	SEQM	Isomer	Property	TDDFT	SEQM
I	$\Delta E_{S_0 \rightarrow S_1}$	3.84	3.35	III	S_0 opt.	0.16	-0.17
I	$\Delta E_{S_0 \rightarrow S_2}$	4.13	3.43	III	$\Delta E_{S_0 \rightarrow S_1}$	3.27	3.22
I	S_1 ($\pi\pi^*$)	0.93	0.61	III	$\Delta E_{S_0 \rightarrow S_2}$	3.32	3.57
I	S_2 ($n\pi^*$)	0.96	0.00	III	S_1 ($n\pi^*$)	0.96	0.00
I	S_1 opt.	3.64	3.56	III	S_2 ($\pi\pi^*$)	0.97	0.13
I ^a	$\Delta E_{S_0 \rightarrow S_1}$	3.05	3.03	III	S_1 opt.	3.11	3.10
I ^a	$\Delta E_{S_0 \rightarrow S_2}$	3.60	3.14	III ^a	$\Delta E_{S_0 \rightarrow S_1}$	2.55	2.92
I ^a	S_1 ($n\pi^*$)	0.97	0.59	III ^a	$\Delta E_{S_0 \rightarrow S_2}$	2.98	3.21
I ^a	S_2 ($\pi\pi^*$)	0.95	0.58	III ^a	S_1 ($n\pi^*$)	0.57	0.08
II	S_0 opt.	0.08	-0.07	III ^a	S_2 ($\pi\pi^*$)	0.58	0.07
II	$\Delta E_{S_0 \rightarrow S_1}$	3.58	3.15	IV	S_0 opt.	0.42	0.22
II	$\Delta E_{S_0 \rightarrow S_2}$	3.66	3.56	IV	$\Delta E_{S_0 \rightarrow S_1}$	3.01	2.68
II	S_1 ($\pi\pi^*$)	0.97	0.00	IV	$\Delta E_{S_0 \rightarrow S_2}$	3.24	3.06
II	S_2 ($n\pi^*$)	0.95	0.00	IV	S_1 ($\pi\pi^*$)	0.95	0.68
II	S_1 opt.	3.28	2.89	IV	S_2 ($n\pi^*$)	0.94	0.00
II ^a	$\Delta E_{S_0 \rightarrow S_1}$	2.86	2.81	IV	S_1 opt.	2.93	2.51
II ^a	$\Delta E_{S_0 \rightarrow S_2}$	3.34	3.29	IV ^a	$\Delta E_{S_0 \rightarrow S_1}$	2.17	1.94
II ^a	S_1 ($\pi\pi^*$)	0.58	0.63	IV ^a	$\Delta E_{S_0 \rightarrow S_2}$	2.68	2.83
II ^a	S_2 ($n\pi^*$)	0.53	0.45	IV ^a	S_1 ($\pi\pi^*$)	0.97	0.68
				IV ^a	S_2 ($n\pi^*$)	0.96	0.00

a: Structure optimized on S_1

With the [10,12]-FOCIS-AM1 setup, 200 surface-hopping trajectories were initialized in S_1 and propagated for 2.5 ps given the ultrafast deactivation that was already observed in **SMA**. From the trajectories it can be deduced that **MMIMQO** really does show a faster ESIPT reaction process, including the rotation of the crane indicated by the decrease of the $H_1 \cdots N_4$ distance (Fig. 4.19a).

The ultrafast de-excitation of **MMIMQO** is represented by the PI time of ~ 300 fs and complete return to S_0 after 800 fs (Fig. 4.19b). The first 200 fs also present accumulation of S_2 population. This can be explained by the calculated 1D PES of the I \rightarrow II PT in Fig. 4.18a: At the Franck-Condon region S_1 and S_2 are very close so that hops to the higher state are possible. However, at that point the systems did not gain enough kinetic energy due to the gradient in S_1 that would help them in overcoming the energy barrier in S_2 . Therefore, they are reflected and can return to S_1 where the ESIPT can happen.

Although **MMIMQO** might not be a better ESIPT-crane system because of the two mentioned deficiencies – **II-MMIMQO** being stable in S_0 and **IV-MMIMQO** not –, one can conclude that

4.4 Fused-Ring Systems

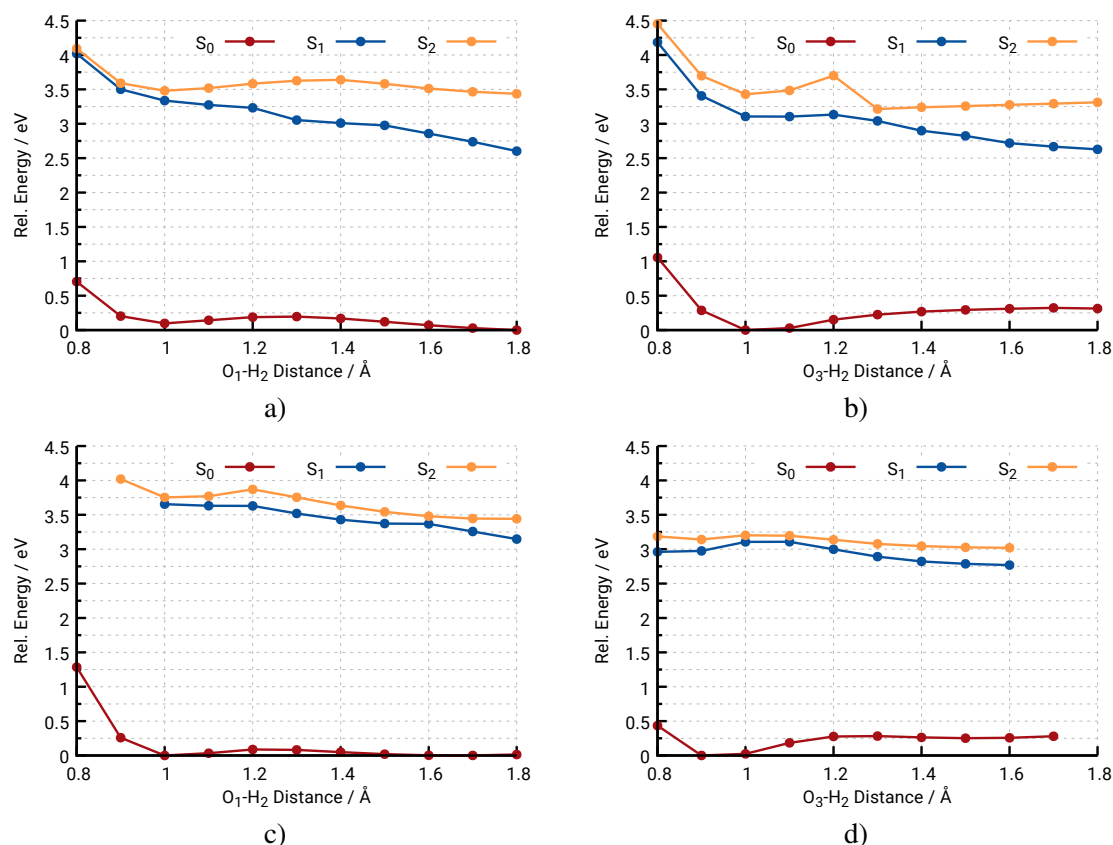


Figure 4.18.: One-dimensional PES scans of the I→II and III→IV PTs of **MMIMQO** using [10,12]-FOCIS-AM1 (a,b) and PBE0-D3BJ/def2-SVP (c,d), respectively. The scan for each state followed its respective gradient. The energies are given relative to the S_0 minimum.

the O-donor-N-acceptor-pattern seems to be more effective in the first steps of the reaction – due to the gradient in the excited states – than the former O-donor-O-acceptor-pattern of the parent systems **HMQCA**, **Sac** or **SAMe**.

The same feature was also found for 7-hydroxy(8-oxazine-2-one)quinoline in a relaxed scan on the CC2/cc-pVDZ level of theory.^[45] There, the ESIPT is also predicted to happen on a barrierless pathway exclusively in S_1 which is of $\pi\pi^*$ character (cf. Fig. 2 in Ref. 45). This system differs from **MMIMQO** energetically only in a stable IV-form.

Therefore it can be concluded that SEQM FOCI recovers the same qualitative behavior as higher-level methods predict and that the found deficiencies should only be ascribed to the design of the system.

4.4.3 Phenanthridine Backbone

In an attempt to improve **MMIMQO**, the backbone was extended by using phenanthridine (Fig. 4.20). This new system – 4-[(1E)-(methylimino)methyl]phenanthridine-3-ol (**MIMPO**) – has a more electron-rich backbone than quinoline. As such, it could potentially lead to a more

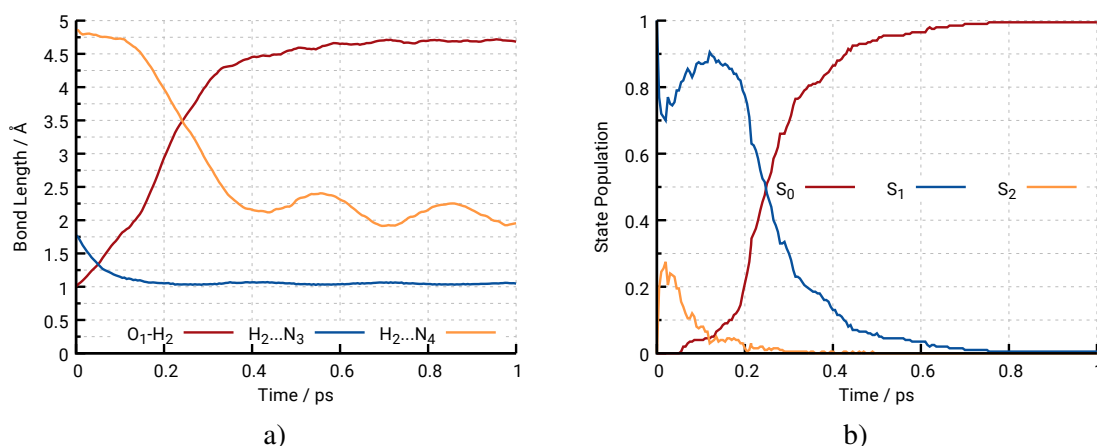


Figure 4.19.: Ensemble-averaged properties of the photodynamics of **MMIMQO**: a) Time-evolution of the O₁-H₂, H₂...N₃ and H₂...N₄ distances and b) state populations.

stable IV-form which was predicted for the system with an aldehyde crane¹⁴. If this substitution also helps for a system with a CNMe crane will be investigated in the following.

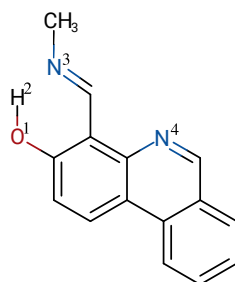


Figure 4.20.: Sketch of **MIMPO**. This system utilizes a phenanthridine backbone which extends the π system of the backbone, thus lowering the basicity of the backbone-nitrogen atom.

The SEQM FOCI setup was again determined by a meta optimization approach against DFT (PBE0-D3BJ/def2-TZVPP) reference data of geometry optimizations. The best fitting one turned out to be [12,10]-FOCISD-PM6*. Both reference and obtained values of the training set are presented in Table 4.8.

In the case of **MIMPO** the SEQM approach seems to yield a better agreement with the TDDFT reference data in terms of energies. As a qualitative difference to **MMIMQO** the excited-state dynamics probably only involve S₁, because the state is always dominantly described by a $\pi\pi^*$ determinant¹⁵. This is the reason why additional S₂ calculations were omitted for **MIMPO**.

The scans of the I→II and III→IV PTs were performed in the SEQM level and benchmarked against TDDFT (PBE0-D3BJ/def2-SVP).

¹⁴ This specific system was not investigated in this thesis because of the deficiencies with the order of states of **HMQCA**.

¹⁵ This is why the [12,12]-FOPECI-RM1 setup also performs well for this system. However, it is not the best by only a small margin: The fitness value for [12,12]-FOPECI-RM1 is 12.78 and for [12,10]-FOCISD-PM6* 12.13.

Table 4.8.: Results of the TDDFT (PBE0-D3BJ/def2-TZVPP) benchmarks and meta optimization for important points and properties of the PES of **MIMPO**. The best fitting SEQM setup was found to be [12,10]-FOCISD-PM6*. All energies of the optimized structures (“ S_X opt.”) are given relative to the S_0 energy of **I-MIMPO**. All energies in eV. The excited-state characters S_X ($n\pi^*/\pi\pi^*$) are given as the absolute value of the CI coefficient for the corresponding SD. The corresponding orbital representations are given in the appendix (Sec. A.8.3).

Isomer	Property	TDDFT	SEQM	Isomer	Property	TDDFT	SEQM
				III	S_0 opt.	0.15	-0.00
I	$\Delta E_{S_0 \rightarrow S_1}$	3.60	3.61	III	$\Delta E_{S_0 \rightarrow S_1}$	2.96	3.29
I	$\Delta E_{S_0 \rightarrow S_2}$	4.07	3.91	III	$\Delta E_{S_0 \rightarrow S_2}$	3.21	3.58
I	S_1 ($\pi\pi^*$)	0.90	0.57	III	S_1 ($\pi\pi^*$)	0.97	0.65
I	S_2 ($n\pi^*$)	0.82	0.04	III	S_2 ($n\pi^*$)	0.85	0.06
I	S_1 opt.	3.44	3.47	III	S_1 opt.	2.92	3.19
I ^a	$\Delta E_{S_0 \rightarrow S_1}$	3.27	3.29	III ^a	$\Delta E_{S_0 \rightarrow S_1}$	2.54	2.97
I ^a	$\Delta E_{S_0 \rightarrow S_2}$	3.86	3.68	III ^a	$\Delta E_{S_0 \rightarrow S_2}$	2.90	3.39
I ^a	S_1 ($\pi\pi^*$)	0.93	0.59	III ^a	S_1 ($\pi\pi^*$)	0.97	0.65
I ^a	S_2 ($\pi'\pi^*$)	0.67	0.51	III ^a	S_2 ($n\pi^*$)	0.87	0.06
II	S_0 opt.	0.02	0.04	IV	S_0 opt.	0.61	0.58
II	$\Delta E_{S_0 \rightarrow S_1}$	3.26	3.22	IV	$\Delta E_{S_0 \rightarrow S_1}$	2.27	2.26
II	$\Delta E_{S_0 \rightarrow S_2}$	3.57	3.74	IV	$\Delta E_{S_0 \rightarrow S_2}$	2.89	3.07
II	S_1 ($\pi\pi^*$)	0.96	0.63	IV	S_1 ($\pi\pi^*$)	0.98	0.64
II	S_2 ($n\pi^*$)	0.97	0.06	IV	S_2 ($n\pi^*$)	0.93	0.03
II	S_1 opt.	3.10	2.97	IV	S_1 opt.	2.47	2.49
II ^a	$\Delta E_{S_0 \rightarrow S_1}$	2.77	2.77	IV ^a	$\Delta E_{S_0 \rightarrow S_1}$	1.60	1.61
II ^a	$\Delta E_{S_0 \rightarrow S_2}$	3.27	3.52	IV ^a	$\Delta E_{S_0 \rightarrow S_2}$	2.34	2.87
II ^a	S_1 ($\pi\pi^*$)	0.88	0.60	IV ^a	S_1 ($\pi\pi^*$)	0.98	0.65
II ^a	S_2 ($n\pi^*$)	0.53	0.14	IV ^a	S_2 ($n\pi^*$)	0.97	0.02

^a: Structure optimized on S_1

The large backbone of the systems leads to a larger energy barrier in S_0 when it comes to the I→II PT in comparison to the one of **MMIMQO**. However, this barrier is still not high enough, because **II-MIMPO** is again already formed during the S_0 MD calculations and **IV-MIMPO** is still not a stable minimum on S_0 . Hence, there were no further attempts in investigating the ES IPT-switching properties of **MIMPO**.

Nevertheless, the obtained DFT reference data can still serve a valuable purpose for parameter optimization attempts.

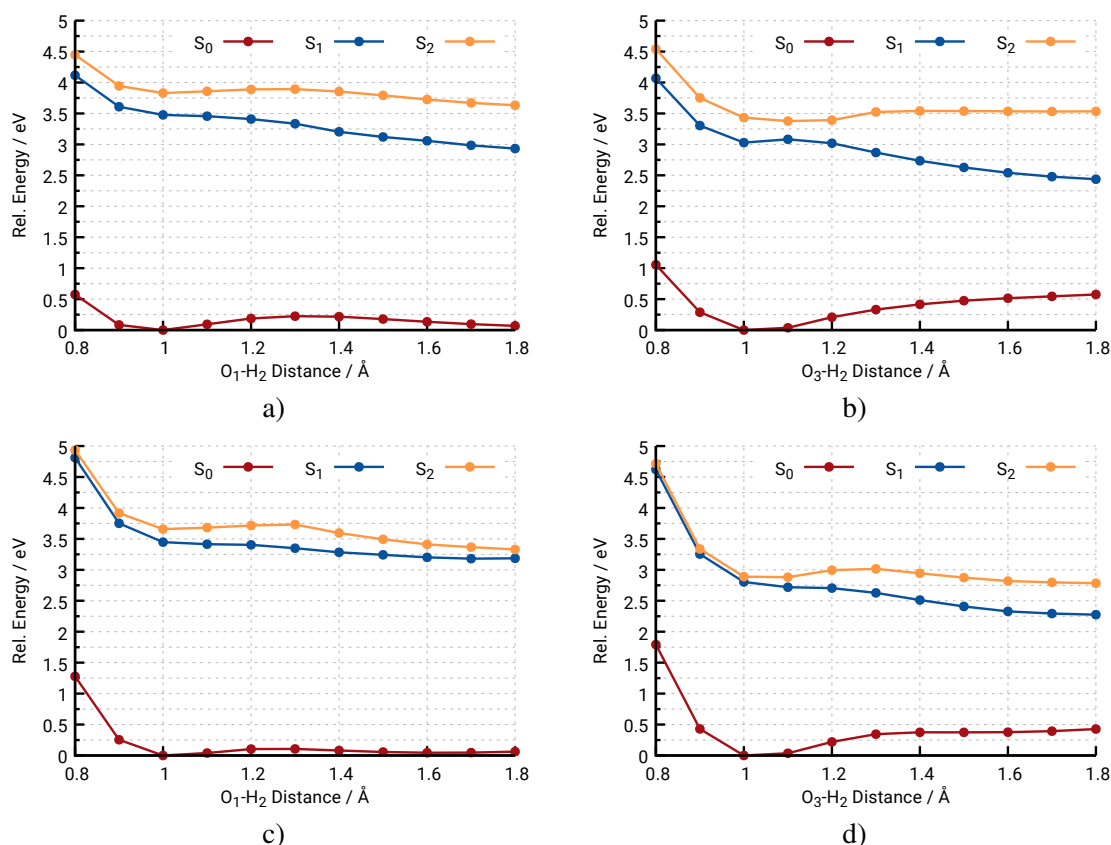


Figure 4.21.: One-dimensional PES scans of the I→II and III→IV PTs of **MIMPO** using the [12,10]-FOCISD-PM6* (a,b) setup and PBE0-D3BJ/def2-SVP (c,d) setup, respectively. The scan for each state followed its respective gradient.

4.5 Summary

In this chapter the photodynamical properties of several ESIPT switches were presented. While some of them had already been studied by static calculations, none except for the two benchmark systems were studied with MD calculations so far.

The SEQM FOCI setup was tested against OM2/MRCI MD calculations of the ESIPT and switching processes of **SMA** and **OHBI**. This benchmark was necessary because using SEQM FOCI for ESIPT systems is not as established as using it for *trans*⇌*cis* switching of azobenzene. A [12,12]-FOPECI-RM1 setup was able to reproduce the reference dynamics with overall satisfying accuracy. In fact, this setup also proved to be reliable for all systems in which S₁ is of ππ* character.

Starting with **SAc** as the historically first known ESIPT system, a proposed radiationless pathway could be reproduced by semiempirical MD calculations for the first time which led to the publication presented in Sec. 4.1. Some deficiencies regarding the CI setup of the calculation were later found. However, recalculations of the same trajectories presented in the publication with correct paired-double CI excitations only exhibited differences within the error margin of the original setup. A more careful selection of the starting structures for the excited-state dynamics were found to play a more important role.

4.5 Summary

Using the knowledge regarding the computational setup, several structural analogues of **SAC** were investigated as well (Sec. 4.3). The systems were obtained by substitution of the proton donor and acceptor sites consisting of different oxygen- and nitrogen-containing groups. Out of the seven¹⁶ studied systems, **SMA** proved to be the superior one in terms of overall efficiency of the reaction.

HMQCA was chosen as a system that performs a more complex ESIPT-switching process because of the larger backbone which also included a second proton accepting site. This system has also been investigated by means of static calculations so that reference data was available. The expected reaction of **HMQCA** was successfully reproduced with excited-state MD calculations. However, the investigation also revealed a first deficiency of a SEQM FOCI method: The characters of the electronically excited states of **HMQCA** were calculated as $n\pi^*$ for S_1 and $\pi\pi^*$ for S_2 using *ab initio* and DFT methods, which could not be resolved as clearly with SEQM. In fact, none of the available SEQM parameters and Hamiltonians in MOPAC were able to correctly reproduce this feature for this system.

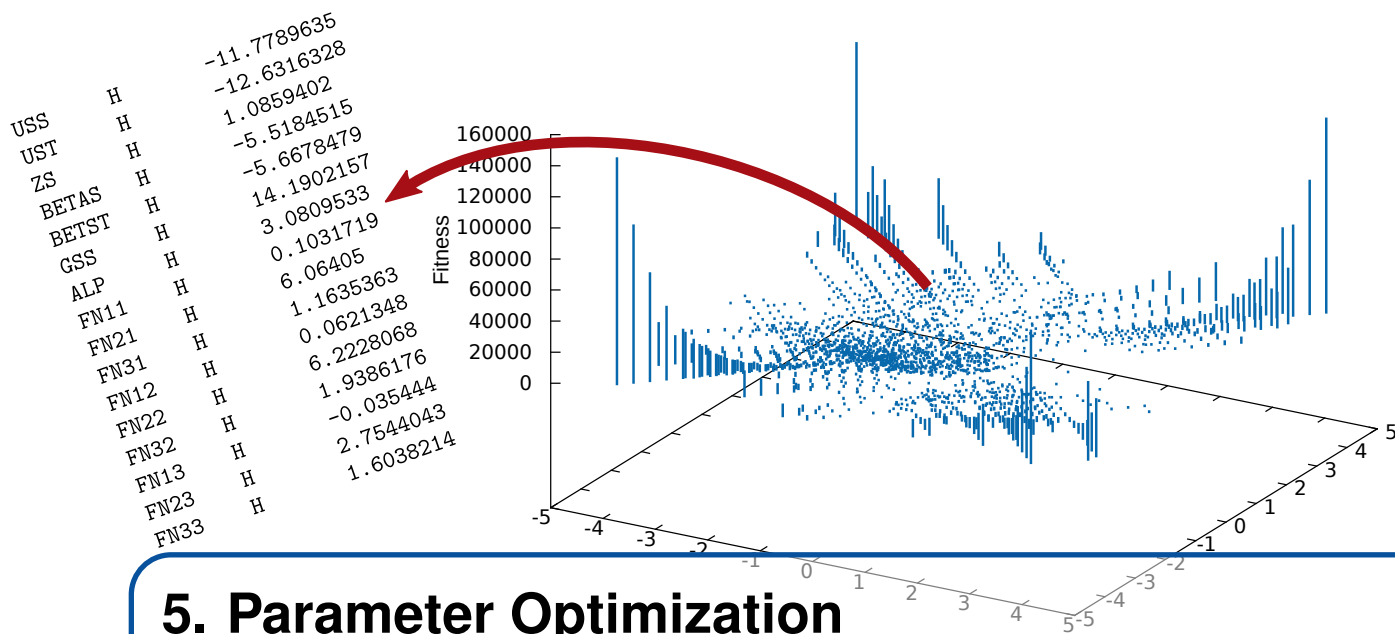
As a subsequent design study, the aldehyde-crane of **HMQCA** was substituted with the methyl imine-crane of **SMA** which led to faster dynamics. In this **MMIMQO** system, however, the second PT to the nitrogen acceptor in the backbone could not be observed in MD calculations and 1D scans of the PES. The aim of substituting the proton-crane with a more effective one was a success, though.

In an attempt to fix the non-stable IV-form a phenanthridine backbone was introduced which could have helped in stabilizing the proton at the backbone-nitrogen. Yet, 1D scans on both the SEQM and DFT levels did not confirm this. Also, the investigations revealed that the proton-transferred form of both **MMIMQO** and **MIMPO** are already stable minima in S_0 which defeats the purpose of ESIPT switches. Nevertheless, even though the two newly designed systems do not serve as effective ESIPT switches, the obtained reference data is still valuable for parameter optimizations of semiempirical Hamiltonians.

In summary, this chapter revealed both successful and somewhat deficient simulations of ESIPT processes for various kinds of systems. For systems in which the dynamics primarily happens in S_1 that is also of $\pi\pi^*$ character, the [12,12]-FOPECI-RM1 setup proved to be the tool of choice. When the dynamics also involve $n\pi^*$ states, though, the current parameters are struggling. This is the reason why the following chapter will present ways of optimizing a set of semiempirical parameters with the focus on ESIPT switches.

¹⁶ Including the systems from the Bachelor's thesis of Jennifer Müller (Sec. 6.4).

*This page is intentionally left blank; but this text is printed here,
therefore this is not a blank page.*



5. Parameter Optimization

Using semiempirical methods requires careful benchmarking of the employed setup, because of three reasons:

- i) “Semiempirical method” should be called “semiempirical parametrization” instead, because the underlying Hamiltonian stays more or less the same for methods like AM1 or PMx. The parametrization is performed for some set of reference data. As long as the studied molecule is part of this reference data, any values obtained should be trustworthy. If this Hamiltonian and its parametrization is used on “exotic” systems, any value needs to be validated.
- ii) All non-specific semiempirical parametrizations were performed for electronic ground state properties. So even if the studied system is part of the reference data set, any excited state information again needs to be validated.
- iii) Even if excited state information is included in the reference data, parametrization is only done for one specific setup of, e.g., CI space and excitation level. Changing these, again, requires validation of the results.

The general computational setup that was used for the azobenzene-containing studies in this work is a prime example of a specific reparametrization of a semiempirical Hamiltonian. Using high-quality reference data on the multireference level of theory, the parameters of the AM1 Hamiltonian were optimized to reproduce ground-state as well as excited-state energies and geometries of azobenzene.^[69] To cope with the excited states, a FOCI active space of 14 electrons in 13 orbitals – but with 94 hand-selected determinants – was employed for the reparametrization. Because these new azobenzene parameters were already thoroughly benchmarked against higher levels of theory, one can be rather safe to assume good results on other azobenzene-based systems.

But this is the sole example of a specific reparametrization to a photoswitch¹, so that one is usually stuck with general semiempirical Hamiltonians, which always require tinkering and validation.

As it was stated in the previous chapter, there are certain features of the reaction paths, e.g., $n\pi^*$ and $\pi\pi^*$ PES interactions, that are not reproducible from reference calculations with the Hamiltonians implemented in the MOPAC development version. Thus, the success of the azobenzene-specific parametrization led to the endeavor for an ESIPT-specific parametrization of the AM1-type Hamiltonian.

In the following section all the “dials”, i.e. the semiempirical parameters and their meanings, will be described.

5.1 Semiempirical Parameters

Both AM1 and RM1 parametrizations use the same Hamiltonian, which shall be presented briefly.

The total energy is defined as sum of electronic and nuclear energy,

$$E_{tot} = E_{el} + \sum_{A<B} E_{AB}^{nuc}. \quad (5.1)$$

The electronic energy $E_{el}^{AM1/RM1}$ was taken from the MNDO Hamiltonian^[65] and is defined as

$$E_{el}^{AM1/RM1} = \frac{1}{2} \sum_{\mu} \sum_{\nu} P_{\mu\nu} (H_{\mu\nu} + F_{\mu\nu}), \quad (5.2)$$

with $P_{\mu\nu}$ being elements of the density matrix,^[130] $H_{\mu\nu}$ the one-electron part and $F_{\mu\nu}$ the two-electron part of the Fock matrix. In the typical notation of semiempirical Hamiltonians, μ and ν describe the AOs of atom A, and λ and σ the AOs centered on atom B. The Fock matrix elements are defined as:

$$F_{\mu\mu} = U_{\mu\mu} + \sum_B V_{\mu\mu,B} + \sum_{\nu}^A P_{\nu\nu} [(\mu\mu, \nu\nu) - \frac{1}{2}(\mu\nu, \mu\nu)] + \sum_B \sum_{\lambda,\sigma} P_{\lambda\sigma} (\mu\mu, \lambda\sigma), \quad (5.3)$$

$$F_{\mu\nu} = \sum_B V_{\mu\nu,B} + \frac{1}{2} P_{\mu\nu} [3(\mu\nu, \mu\nu) - (\mu\mu, \nu\nu)] + \sum_B \sum_{\lambda,\sigma} P_{\lambda\sigma} (\mu\nu, \lambda\sigma), \quad (5.4)$$

with $U_{\mu\mu}$ being the one-center one-electron energies, $(\mu\mu, \nu\nu)$ the one-center two-electron integrals (Coulomb integrals, $g_{\mu\nu}$), $(\mu\nu, \mu\nu)$ the exchange integrals ($h_{\mu\nu}$) and $V_{\mu\nu,B}$ the two-center one-electron attractions between the electrons on atom A to the core of atom B.

$$F_{\mu\lambda} = \beta_{\mu\lambda} - \frac{1}{2} \sum_{\nu}^A \sum_{\sigma}^B P_{\nu\sigma} (\mu\nu, \lambda\sigma), \quad (5.5)$$

¹ But not for a reparametrization to improve excited-state information. For example, the parameters for carbon in the azobenzene parameter set were taken from a specific reparametrization to benzene and its excited state properties.^[129]

5.1 Semiempirical Parameters

with $\beta_{\mu\lambda}$ being the two-center one-electron core-resonance integrals and $(\mu\nu, \lambda\sigma)$ the two-center two-electron repulsion integrals. A detailed description of these integrals is given in Ref. 66.

The core energy of the AM1 Hamiltonian^[67] is defined as

$$E_{AB}^{core} = Z_A Z_B \gamma_{ss} + [1 + F(A) + F(B)], \quad (5.6)$$

with Z_A being the core charge of atom A, $\gamma_{ss} = (s^A s^A, s^B s^B)$ the valence shell charge distribution and $F(A)$ the core-repulsion function (CRF) of atom A. The CRF for atom A is defined as

$$F(A) = e^{-\alpha_A R_{AB}} + \sum_i K_{A_i} \exp[L_{A_i} (R_{AB} - M_{A_i})^2], \quad (5.7)$$

with α being a scaling factor, R_{AB} the distance between cores A and B and $K_{A_i}, L_{A_i}, M_{A_i}$ the scaling factor, exponent and center, respectively, of Gaussian i , $1 \leq i \leq 4$.

All adjustable parameters for these equations are gathered in Table 5.1.

Table 5.1.: List of all parameters used in the AM1/RM1 Hamiltonian and their description.

Parameter	Equiv.	Unit	Description
U_{SS}	USS/UST	eV	monocenter and one-electron integrals of s and p orbitals
U_{PP}	UPP/UPT	eV	
β_S	BETAS/BETST	eV	bicenter and one-electron resonance integrals of s and p orbitals
β_P	BETAP/BETPT	eV	
g_{SS}	GSS	eV	two-electron monocenter repulsion integral between $s-s$, $s-p$, $p-p$ and $p-p'$ orbitals
g_{SP}	GSP	eV	
g_{PP}	GPP	eV	
g_{P2}	GP2	eV	
h_{SP}	HSP	eV	two-electron monocenter exchange integral between $s-p$ orbitals
ζ_S	ZS	bohr ⁻¹	s and p Slater atomic orbital exponent
ζ_P	ZP	bohr ⁻¹	
α	ALP	Å ⁻¹	core-core repulsion factor
K_1	FN11		Gaussian multiplier of core-core repulsion
K_2	FN12		
K_3	FN13		
K_4	FN14		
L_1	FN21	Å ⁻²	Gaussian exponent of core-core repulsion
L_2	FN22	Å ⁻²	
L_3	FN23	Å ⁻²	
L_4	FN24	Å ⁻²	
M_1	FN31	Å	Gaussian center of core-core repulsion
M_2	FN32	Å	
M_3	FN33	Å	
M_4	FN34	Å	

5.2 Algorithm

Due to the number of parameters and their interdependency, optimization and tuning by hand is not possible. For this, algorithms for such multi-dimensional problems need to be employed. In contrast to, e.g., geometry optimizations, the gradients of the parameter space are not known, therefore only numerical optimizations² can be used. Such methods do not need any prior information about the parameter space and can be used for basically any optimization task. One example would be the simplex/simulated annealing (SIMPASA) algorithm,^[131] which was used for the azobenzene-specific parameter optimization.^[69]

For this work a different type of algorithm was used, namely particle swarm optimization (PSO).^[132]

In a PSO, as the name suggest, a swarm of individual particles is propagated on an n -dimensional surface to find its global minimum. The height of the surface is given by a function, which evaluates the “fitness” of a particle at its coordinates on the surface. Since no information about the surface is available at the start of the propagation, no gradient can be followed towards a minimum. Therefore, the driving-force of the particle is communication and memorization. A particle is able to remember its own position at lowest height – best fitness – and can exchange this information with particles in the vicinity – “neighborhood”³. With this information – and some random perturbation – the particle will move along a vector that is a sum of the vectors towards its own and the best known fitness of the neighborhood.

To put this into an equation, the basic form of a PSO can be given as

$$\begin{aligned}\vec{v}_i + 1 &= \vec{v}_i + \vec{U}(0, P_1)(\vec{p}_i - \vec{x}_i) + \vec{U}(0, P_2)(\vec{p}_g - \vec{x}_i), \\ \vec{x}_{i+1} &= \vec{x}_i + \vec{v}_{i+1},\end{aligned}\tag{5.8}$$

with \vec{v}_i and \vec{x}_i being the so-called velocity and position of a particle at iteration step i , $\vec{U}(0, P_1)$ a generator for a random number between 0 and P_1 , \vec{p}_i and \vec{p}_g the positions of the best fitness of the particle and neighborhood, respectively. The initial values, \vec{x}_0 and \vec{v}_0 , are generated randomly within restrictions for each dimension. There are some variations of Eq. 5.8 which introduce “friction” by adding a scaling factor to the velocity (parameter $\chi_2(i)$ in Eq. 5.9) or for having a fully-informed particle swarm (FIPS), in which all particles can communicate with each other.^[132]

For this work a highly adjustable PSO algorithm was programmed. Its underlying equation of motion is presented in Eq. 5.9, with the same basic ingredients of Eq. 5.8 but added iteration-

² It is noted that there actually are methods to derive analytic gradients for such cases. However, given that the optimization of semiempirical parameters was not the central aim of this thesis and the implementation and adaptation to this particular problem is not trivial, it was not being pursued further.

³ The selection which particles belong to the neighborhood of a particle is usually only done once before the propagation starts. Sec. A.5.1 in the appendix explains this part of the PSO in more detail.

progress dependencies for some of the parameters.

$$\begin{aligned}
 \vec{v}_{i+1} &= \chi_2(i)\vec{v}_i \\
 &\quad + \vec{U}(0, \chi_2(i)P_1)(\vec{p}_i - \vec{x}_i)\chi_3(i) \\
 &\quad + \vec{U}(0, \chi_4(i)P_2)(\vec{p}_g - \vec{x}_i)\chi_5(i), \\
 \vec{x}_{i+1} &= \vec{x}_i + \chi_1(i)\vec{v}_{i+1}.
 \end{aligned}
 \tag{5.9}$$

All χ parameters – besides being constant – can either be increased or decreased in a linear or exponential fashion with the iteration number of the optimization. With these added parameters it can be decided whether the particles gain more efficiency in converging (decreasing \vec{v}) or are able to explore the parameter space (increasing \vec{v})⁴. Also, the importance – scaling – of the best fitness position \vec{p}_i of a particle and the global/neighborhood’s best fitness \vec{p}_g can be made iteration-dependent.

Because PSO is a non-deterministic global optimization method the chance of finding the “best solution” in a given problem space is never 100% within a finite number of iterations. Rigorously increasing the iteration number can help but is usually not an option i) because of the computational cost per fitness evaluation (see below) and ii) because there is no guarantee that every iteration reaches points in the problem space that were never accessed before. Thus, covering the complete problem space is not feasible with non-deterministic methods. The only resort is to start many optimization runs with varying initial conditions and to see if they show a tendency to converge to the same solution. In contrast, deterministic optimization algorithms will always find the best possible solution, but are only feasible for a small number of dimensions due to an exponential scaling in computational effort.

A general outline of the algorithm of this program is given below, with all toggle-able routines being newly implemented routines to the core algorithm. A description of these routines is presented in the appendix (Sec. A.5).

At this point program developers would usually explain their reasons for choosing a specific program/scripting language. The PSO as presented here was written in Perl, mainly because of the reason it is the author’s language of choice when it comes to automation on the computer. However, any other programming/scripting language may be suited for the PSO. This is because the computational demand of the algorithm itself is relatively low. The core of the PSO algorithm just performs some vector additions, so there is no need to employ an object-oriented or close-to-hardware-level language.

The computational demand – regarding both time and computer resources – only comes from the fitness evaluation (see below), which is not driven by the algorithm.

⁴ The dependence on the iteration number is the core feature of the simulated annealing approach. Thus, this PSO implementation is on par with SIMPSA algorithm mentioned above.

```

initializeParticles;
for nIterations do
  if toggled then
    | getNeighborIndices;
  end
  updateVelocity;
  if toggled then
    | checkParticleDistance;
  end
  calculateFitness;
  checkForNewMinima;
  if toggled then
    | performLocalSearch;
  end
  if toggled then
    | checkForPlateau;
  end
  if toggled then
    | shrinkSurface;
  end
  if currentFitness < finalFitness then
    | return finalBestParticleCoordinates
  end
end
return currentBestParticleCoordinates

```

Algorithm 1: Particle Swarm Optimization as implemented by the author.

5.2.1 Application

To optimize semiempirical parameters with the PSO algorithm one needs to construct a fitness function that translates the coordinates of N_{part} particles on the parameter surface to a fitness value. Each of N_{param} parameters that are to be optimized, corresponds to one dimension of the surface. The position of particle in each dimension is used here as a scaling factor for the corresponding parameter⁵.

The fitness function takes the coordinates of one particle at a time. From the coordinates, the new parameters are created by scaling the reference parameters. With the new parameter set a series of j calculations, e.g., energy calculations, geometry optimizations, DOF scans or dynamics, are performed as prepared by the user. The result (q_j^{new}) of each calculation is compared to a supplied reference value (q_j^{ref}), e.g., relative energies, vertical excitation energies, gradients, bond lengths, etc., taken from other calculations. The weighted sum of absolute differences (fitness function, Eq. 5.10) finally yields the fitness value F at the coordinates of a particle, which needs to

⁵ For example, if each dimension is given a specific range, e.g., [0.9;1.1] and a particle is at position $x_d = 0.955$ in dimension d , the corresponding parameter with its initial value $P_{original,d}$ will be scaled by a factor corresponding to the position of the particle to give the new parameter $P_{new,d} = x_d \cdot P_{original,d}$. Therefore, the size/range of each dimension directly translates to the maximum allowed change for each parameter during the optimization.

be returned to the PSO algorithm.

$$F = \sum_j (|q_j^{new} - q_j^{ref}|) \cdot w_j \quad (5.10)$$

The number and complexity of reference calculations is what determines the computational effort of the parameter optimization, because each complete evaluation has to be performed per particle per iteration. This complexity is slightly reduced by having introduced parallelization of the algorithm, so that in the best-case scenario all particles can be evaluated at the same time.

To avoid unnecessary calculations, the “immediate fallback” feature – as presented in Ref. 133 – was implemented for the fitness evaluation. Whenever the fitness sum gets higher than the original fitness obtained with the unchanged parameters, further calculations for the current particle are immediately stopped.

Besides optimization of SEQM parameters, the PSO algorithm – in general, not the newly implemented one by the author – has also been recently used to obtain new force-fields for water,^[134] to fit analytical expressions to the PES of an iodine-water system^[135] or for global structure optimizations.^[136,137]

5.3 Test Cases

To test out the PSO, Lasse Jansen performed first calculations in an advanced practical course supervised by the author, for a system for which computational data on the CASPT2 level of theory was used as reference values.^[138] The results of this investigation are presented in Sec. 6.5.

Note that the study on parameter optimizations happened alongside the ESIPT studies, thus the starting points concerning the FOCI SEQM setup and/or the amount of reference data might not be the same as presented in Chapter 4. However, the purpose of this section is to provide preliminary results on this project showing that it is indeed possible to optimize semiempirical parameters with PSO to certain systems. In addition, because the parameters are getting optimized, utilization of more resource-friendly FOCI parameters – excitation level and active space size – can get compensated.

In the following, five attempts at optimizing the semiempirical parameters – using the newly implemented PSO algorithm – to reference data for **HMQCA** will be presented.

Small CI Space

For the first reparametrization process of **HMQCA** a small reference data set was used and the employed FOCI SEQM setup was chosen to be [8,8]-FOCISD-RM1. The settings for the PSO were as follows:

- 91 parameters were included in the optimization, i.e., all parameters for H, C, N and O
- Each parameter was allowed to change by up to $\pm 10\%$ of its original value

5.3 Test Cases

- Eight particles on eight central processing unit (CPU) threads were propagated for 1000 iterations⁶
- Two of the newly added subroutines were used: plateauBuster (checked every 25 iterations) and localSearch (10 searches in a 0.002 unit radius, every 50 iterations). Both were started after 500 iterations.
- the velocity was restricted to [-0.005;0.005]
- P_2 was set to 2, and $\chi_7(i)$ linearly increases P_2 to 4 after 500 iterations every 50 iterations

The reference data was obtained from CC2/aug-cc-pVTZ calculations⁷.

Table 5.2 lists the original values as well as the optimized values for vertical excitation energies, relative energies and state characters, as well as the resulting fitness values. The RM1opt1 parameters are given in the appendix (Sec. A.6.1).

Table 5.2.: Results of the parameter optimization for **HMQCA**. Reference energies are taken from CC2/aug-cc-pVTZ calculations. The fitness value corresponds to the sum of absolute difference from the reference values. The state characters were treated with a weighting factor of 5. All energies of the optimized structures (“ S_X opt.”) are given relative to the S_0 energy of **I-HMQCA**. All energies in eV. The excited-state characters S_X ($n\pi^*/\pi\pi^*$) are given as the absolute value of the CI coefficient for the corresponding SD.

Isomer	Property	Ref.	[8,8]-FOCISD-RM1	[8,8]-FOCISD-RM1opt1
I	$\Delta E_{S_0 \rightarrow S_1}$	3.69	3.5381	3.2677
I	$\Delta E_{S_0 \rightarrow S_2}$	3.80	3.7615	3.7123
I	S_1 ($n\pi^*$)	0.58	0.0186	0.4816
I	S_2 ($\pi\pi^*$)	0.63	0.1550	0.5705
III	S_0 opt.	0.38	0.1942	0.3880
III	$\Delta E_{S_0 \rightarrow S_1}$	3.13	3.6063	3.0959
III	$\Delta E_{S_0 \rightarrow S_2}$	3.57	3.7343	3.5069
III	S_1 ($n\pi^*$)	0.68	0.0237	0.5325
III	S_2 ($\pi\pi^*$)	0.68	0.1726	0.6608
IV	S_0 opt.	0.31	0.4834	0.2325
IV	$\Delta E_{S_0 \rightarrow S_1}$	2.94	2.6952	2.6981
IV	$\Delta E_{S_0 \rightarrow S_2}$	3.05	3.1952	2.6989
IV	S_1 ($n\pi^*$)	0.75	n.a.	0.2390
IV	S_2 ($\pi\pi^*$)	0.77	0.1128	0.6464
Fitness			$\sim 12.117^a$	~ 6.082

a: The missing $n\pi^*$ value for IV accounted for 5.999 to the fitness, this amount was removed for better comparison

The overall fitness could be decreased by a factor of 2. However, while some entries could be optimized, e.g., all state characters or relative energies, some got worse, e.g., the VEE of **I-HMQCA** or the ones for IV which are now degenerate. Nevertheless, the overall result is encouraging.

⁶ This is a very small amount of iterations. But this is just the first test case.

⁷ These are calculations performed by the author, not the ones from Ref. 43 which used a smaller basis set.

Larger CI Space

In the previous chapter the [12,12]-FOPECI-RM1 setup performed remarkably well for all systems that featured a S_1 with $\pi\pi^*$ character but not for **HMQCA** and the two derived systems. In a second attempt to optimize the parameters for **HMQCA** the same SEQM FOCI setup was used. The same settings for the PSO algorithm were used as in the first attempt with the exception of the iteration number which was increased to 5000. The results are gathered in Table 5.3.

Originally, the [12,12]-FOPECI-RM1 setup yielded a worse fitness value than the former [8,8]-FOCISD-RM1. After the parameters were optimized, though, the fitness value could be lowered even further.

Table 5.3.: Results of the parameter optimization for **HMQCA**. Reference energies are taken from CC2/aug-cc-pVTZ calculations. The fitness value corresponds to the sum of absolute difference from the reference values. The state characters were treated with a weighting factor of 5. All energies of the optimized structures (“ S_X opt.”) are given relative to the S_0 energy of **I-HMQCA**. All energies in eV. The excited-state characters S_X ($n\pi^*/\pi\pi^*$) are given as the absolute value of the CI coefficient for the corresponding SD.

Isomer	Property	Ref.	[12,12]-FOPECI-RM1	[12,12]-FOPECI-RM1opt2
I	$\Delta E_{S_0 \rightarrow S_1}$	3.69	3.4518	3.2056
I	$\Delta E_{S_0 \rightarrow S_2}$	3.80	3.4814	3.5197
I	S_1 ($n\pi^*$)	0.58	0.0462	0.4378
I	S_2 ($\pi\pi^*$)	0.63	0.0002	0.5942
III	S_0 opt.	0.38	0.1755	0.2564
III	$\Delta E_{S_0 \rightarrow S_1}$	3.13	3.4991	3.3174
III	$\Delta E_{S_0 \rightarrow S_2}$	3.57	3.6601	3.5259
III	S_1 ($n\pi^*$)	0.68	0.0036	0.5276
III	S_2 ($\pi\pi^*$)	0.68	0.1730	0.6044
IV	S_0 opt.	0.31	0.2198	0.3653
IV	$\Delta E_{S_0 \rightarrow S_1}$	2.94	2.9056	3.0018
IV	$\Delta E_{S_0 \rightarrow S_2}$	3.05	3.2054	3.0291
IV	S_1 ($n\pi^*$)	0.75	n.a.	0.6278
IV	S_2 ($\pi\pi^*$)	0.77	0.0263	0.6422
Fitness			$\sim 13.204^a$	~ 4.542

a: The missing $n\pi^*$ value for IV accounted for 5.999 to the fitness, this amount was removed for better comparison

To see if this new set of parameters also works well outside of the supplied reference points, 1D scans of the I \rightarrow II and III \rightarrow IV PTs were performed (Fig. 5.1).

The scans reveal that, although the parameters got better in reproducing the reference values, they got worse everywhere else on the PES – compare Fig. 5.1 with Fig. 5.2. For example, the RM1opt2 parameters predict **II-HMQCA** at an O_1 -H₂ distance of 1.6 Å to be more stable than **II-HMQCA** at 1 Å in S_0 , or the drop in energy during the III \rightarrow PT which was because the quinoline ring dissociated (Fig. 5.1b).

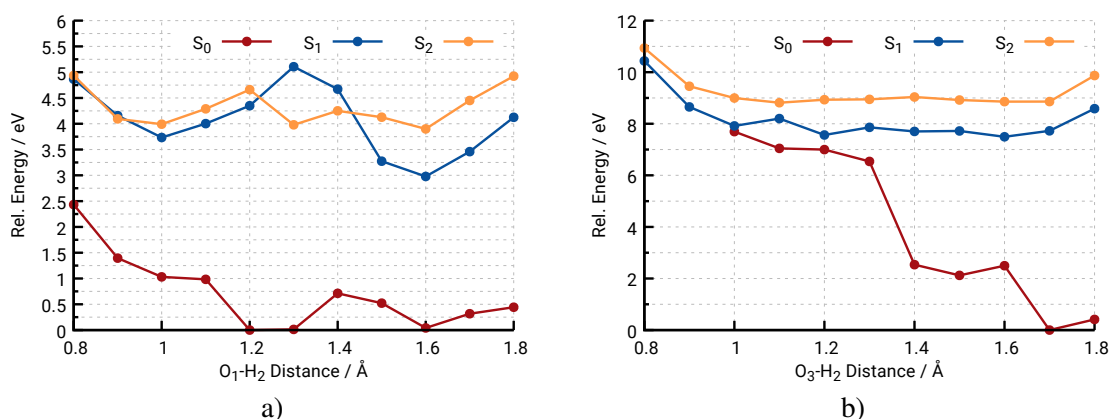


Figure 5.1.: One-dimensional PES scans of the I→II (a) and III→IV (b) PT of **HMQCA** using the RM1opt2 parameters. The energy range in b) was set to the same as in a) for easier comparison – this cuts the scan result for S_0 . However, after 1.3 Å the structure dissociates. Note the different energy range for b). The energies given relative to the S_0 minimum.

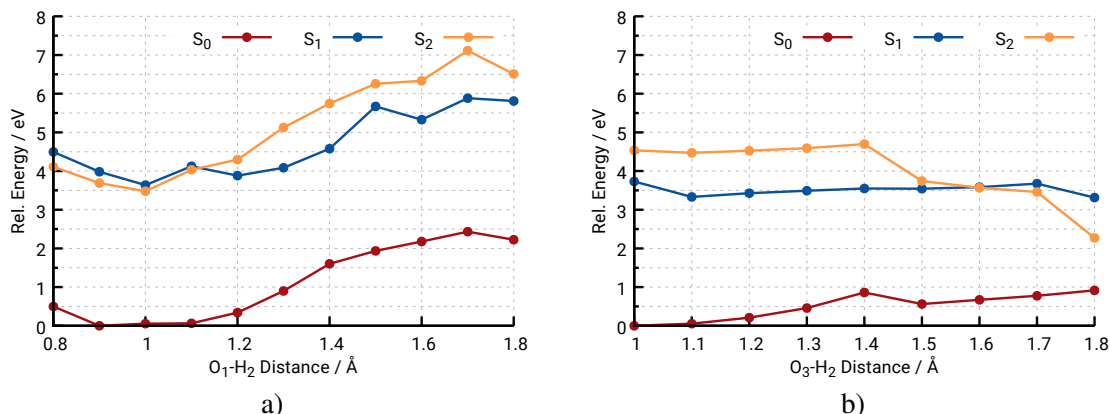


Figure 5.2.: One-dimensional PES scans of the I→II (a) and III→IV (b) PT of **HMQCA** using the RM1 parameters. The energies given relative to the S_0 minimum.

Omitting Carbon Parameters

The quinoline dissociation can probably be appointed to faulty parameters for the carbon atoms. Because these already worked well for all other ESIPT system studied with [12,12]-FOPECI-RM1, all carbon parameters were left out in a third attempt to obtain optimized RM1 parameters with this the reference data set from the previous to runs.

However, while the fitness could be lowered, subsequent scans of the I→II and III→IV PTs revealed that the PES got even worse than in this third attempt. This manifested itself again in dissociation of the quinoline backbone. Thus, no results from this scan will be shown here and the carbon atoms will no longer be left out of the optimization process.

Larger Reference Data Set

The fourth attempt at optimizing the RM1 parameters to **HMQCA** utilized the reference data set that was presented in Sec. 4.4.1 (Table 4.6). The results are listed in Table 5.4.

Table 5.4.: Results of the fourth parameter optimization of **HMQCA**. The SEQM method was [12,12]-FOPECI-RM1opt4. All energies of the optimized structures (“ S_X opt.”) are given relative to the S_0 energy of **I-HMQCA**. All energies in eV. The excited-state characters S_X ($n\pi^*/\pi\pi^*$) are given as the absolute value of the CI coefficient for the corresponding SD.

Isomer	Property	TDDFT	SEQM	Isomer	Property	TDDFT	SEQM
				III	S_0 opt.	0.35	0.03
I	$S_0 \rightarrow S_1$	3.59	2.99	III	$S_0 \rightarrow S_1$	3.17	3.05
I	$S_0 \rightarrow S_2$	3.89	3.67	III	$S_0 \rightarrow S_2$	3.56	3.59
I	S_1 ($n\pi^*$)	0.89	0.40	III	S_1 ($n\pi^*$)	0.95	0.55
I	S_2 ($\pi\pi^*$)	0.93	0.59	III	S_2 ($\pi\pi^*$)	0.98	0.59
I	S_1 opt.	2.52	3.19	III	S_1 opt.	3.24	2.62
I ^a	$S_0 \rightarrow S_1$	2.49	2.49	III ^a	$S_0 \rightarrow S_1$	2.56	2.24
I ^a	$S_0 \rightarrow S_2$	3.79	3.51	III ^a	$S_0 \rightarrow S_2$	3.10	3.19
I ^a	S_1 ($n\pi^*$)	0.96	0.51	III ^a	S_1 ($n\pi^*$)	0.96	0.59
I ^a	S_2 ($\pi\pi^*$)	0.93	0.57	III ^a	S_2 ($\pi\pi^*$)	0.98	0.60
I ^b	S_2 opt.	3.68	3.75	III ^b	S_2 opt.	3.36	3.43
I ^b	$S_0 \rightarrow S_1$	3.52	3.07	III ^b	$S_0 \rightarrow S_1$	2.75	2.65
I ^b	$S_0 \rightarrow S_2$	3.57	3.44	III ^b	$S_0 \rightarrow S_2$	3.05	3.14
I ^b	S_1 ($n\pi^*$)	0.91	0.43	III ^b	S_1 ($n\pi^*$)	0.97	0.59
I ^b	S_2 ($\pi\pi^*$)	0.95	0.61	III ^b	S_2 ($\pi\pi^*$)	0.98	0.62
II ^c	S_0 opt.	–	–	IV	S_0 opt.	0.29	0.13
II	$S_0 \rightarrow S_1$	–	–	IV	$S_0 \rightarrow S_1$	3.03	3.03
II	$S_0 \rightarrow S_2$	–	–	IV	$S_0 \rightarrow S_2$	3.14	3.04
II	S_1	–	–	IV	S_1 ($n\pi^*$)	0.91	0.54
II	S_2	–	–	IV	S_2 ($\pi\pi^*$)	0.95	0.65
II	S_1 opt.	3.64	3.01	IV	S_1 opt.	2.98	2.98
II ^a	$S_0 \rightarrow S_1$	3.03	2.54	IV ^a	$S_0 \rightarrow S_1$	2.09	2.15
II ^a	$S_0 \rightarrow S_2$	3.33	3.39	IV ^a	$S_0 \rightarrow S_2$	2.51	2.35
II ^a	S_1 ($n\pi^*$)	0.97	0.58	IV ^a	S_1 ($n\pi^*$)	0.96	0.60
II ^a	S_2 ($\pi\pi^*$)	0.97	0.62	IV ^a	S_2 ($\pi\pi^*$)	0.95	0.64

a: Structure was optimized on S_1 , *b*: Structure was optimized on S_2 , *c*: Not a stable minimum on S_0 .

The resulting fitness for this parameter set was 39.96. For comparison, the initial [12,12]-FOPECI-RM1 fitness before the optimization was 84.98 and the [10,12]-FOCIS-AM1 fitness 63.31. Overall, the parameter optimization was successful. The RM1opt4 parameters are given in the appendix (Sec. A.6.3). To test the performance of the optimized parameters, the two benchmark PT scans were recalculated.

When comparing the new scans (Fig. 5.3) with the scans from the second optimization attempt (Fig. 5.1), it is noticeable that the new ones feature a smoother PES. The I→II scan (Fig. 5.3a) now

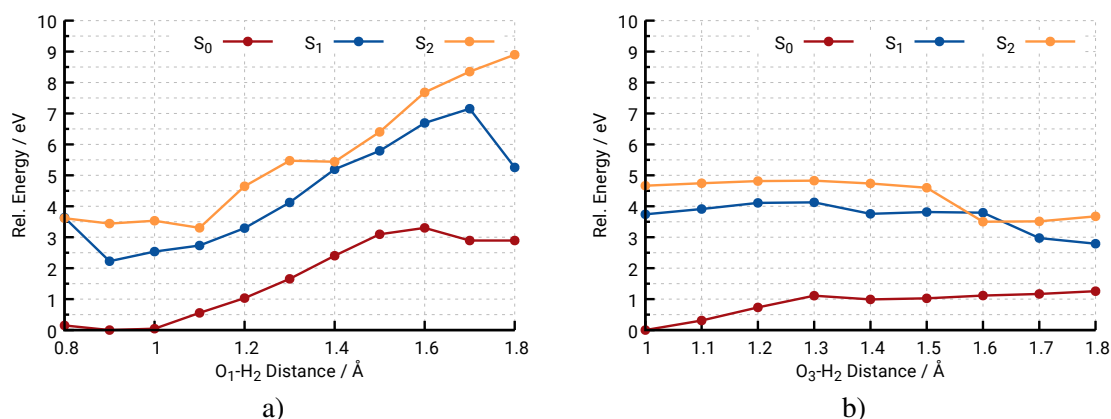


Figure 5.3.: One-dimensional PES scans of the I→II (a) and III→IV (b) PT of **HMQCA** using the RM1opt4 parameters. The energies given relative to the S₀ minimum.

also shows a rise in energy when reaching **II-HMQCA**. However, both excited states do not favor an ESIPT. Interestingly, the new parameters now favor **III-HMQCA** over **IV-HMQCA** (Fig. 5.3b).

Even though the optimization still did not yield a better PES in terms of reproduction of the DFT results, it is clear that a larger reference data set is needed in order to do so.

Including Scan Data

In the fifth and last attempt the reference data set was extended by the energies from the 1D PES scans on the TDDFT level (cf. Figs. 4.16c and 4.16d). The obtained parameters are presented in the appendix (Sec. A.6.4).

Table 5.5 lists the results for the best out of 20 optimizations. Even though this attempt led to a decrease of the fitness from an initial value of 10066.896⁸ to 50.454, the resulting PES still does not feature a shape that is as smooth as the original PES. The 1D PES scan for the I→II PT (Fig. 5.4a) now reveals a (possibly) unstable **II-HMQCA** structure, but this PT also leads to a dissociation of the quinoline ring in S₁. This leads to the assumption that the parameters for the carbon atom are rather sensitive towards slight changes in their values. Given that omitting these parameters during the parameter optimization was found not to be an option, special care should be taken in future optimization attempts. Interestingly, the III→IV PT (Fig. 5.4b) did not result in such a behavior. However, even though the reference data of the DFT PT scans were part of the optimization process, the qualitative shape of the resulting PES – for S₀, S₁ and S₂ – is comparable to the previous parameter optimization attempt including only the optimized minima on the PES (cf. Fig. 5.3b).

Because none of the five optimization attempts yielded a smoothly-shaped PES of the three involved electronic states, no subsequent calculations on the ESIPT-switching dynamics of **HMQCA** with any of the new parameter sets were attempted.

⁸ This large value is due to two missing $n\pi^*$ descriptions accounting for 5.999 each.

Table 5.5.: Results of the fifth parameter optimization of **HMQCA**. The SEQM method was [12,12]-FOPECI-RM1opt5. All energies of the optimized structures (“ S_X opt.”) are given relative to the S_0 energy of **I-HMQCA**. All energies in eV. The excited-state characters S_X ($n\pi^*/\pi\pi^*$) are given as the absolute value of the CI coefficient for the corresponding SD.

Isomer	Property	TDDFT	SEQM	Isomer	Property	TDDFT	SEQM
I	$S_0 \rightarrow S_1$	3.59	3.31	III	S_0 opt.	0.35	0.23
I	$S_0 \rightarrow S_2$	3.89	3.67	III	$S_0 \rightarrow S_1$	3.17	3.30
I	S_1 ($n\pi^*$)	0.89	0.43	III	$S_0 \rightarrow S_2$	3.56	3.66
I	S_2 ($\pi\pi^*$)	0.93	0.56	III	S_1 ($n\pi^*$)	0.95	0.54
I	S_1 opt.	2.52	3.44	III	S_2 ($\pi\pi^*$)	0.98	0.53
I ^a	$S_0 \rightarrow S_1$	2.49	2.52	III	S_1 opt.	3.24	3.03
I ^a	$S_0 \rightarrow S_2$	3.79	3.48	III ^a	$S_0 \rightarrow S_1$	2.56	2.41
I ^a	S_1 ($n\pi^*$)	0.96	0.59	III ^a	$S_0 \rightarrow S_2$	3.10	3.28
I ^a	S_2 ($\pi\pi^*$)	0.93	0.54	III ^a	S_1 ($n\pi^*$)	0.96	0.59
I ^b	S_2 opt.	3.68	3.82	III ^a	S_2 ($\pi\pi^*$)	0.98	0.58
I ^b	$S_0 \rightarrow S_1$	3.52	3.35	III ^b	S_2 opt.	3.36	3.77
I ^b	$S_0 \rightarrow S_2$	3.57	3.40	III ^b	$S_0 \rightarrow S_1$	2.75	2.86
I ^b	S_1 ($n\pi^*$)	0.91	0.46	III ^b	$S_0 \rightarrow S_2$	3.05	3.21
I ^b	S_2 ($\pi\pi^*$)	0.95	0.59	III ^b	S_1 ($n\pi^*$)	0.97	0.59
II ^c	S_0 opt.	–	–	III ^b	S_2 ($\pi\pi^*$)	0.98	0.59
II	$S_0 \rightarrow S_1$	–	–	IV	S_0 opt.	0.29	0.34
II	$S_0 \rightarrow S_2$	–	–	IV	$S_0 \rightarrow S_1$	3.03	3.13
II	S_1	–	–	IV	$S_0 \rightarrow S_2$	3.14	3.13
II	S_2	–	–	IV	S_1 ($n\pi^*$)	0.91	0.50
II	S_1 opt.	3.64	3.42	IV	S_2 ($\pi\pi^*$)	0.95	0.62
II ^a	$S_0 \rightarrow S_1$	3.03	2.76	IV	S_1 opt.	2.98	3.29
II ^a	$S_0 \rightarrow S_2$	3.33	3.42	IV ^a	$S_0 \rightarrow S_1$	2.09	2.16
II ^a	S_1 ($n\pi^*$)	0.97	0.57	IV ^a	$S_0 \rightarrow S_2$	2.51	2.42
II ^a	S_2 ($\pi\pi^*$)	0.97	0.60	IV ^a	S_1 ($n\pi^*$)	0.96	0.59
				IV ^a	S_2 ($\pi\pi^*$)	0.95	0.62

a: Structure was optimized on S_1 , *b*: Structure was optimized on S_2 , *c*: Not a stable minimum on S_0 .

5.4 Summary and Outlook

This first try of optimizing semiempirical parameters to ESIPT systems must not be seen as an in-depth investigation but more of a proof-of-principle process.

A versatile PSO algorithm was programmed for this purpose. In contrast to available PSO algorithms in the Perl scripting language,^[139,140] this new PSO takes advantage of CPU threading and several new subroutines that are potentially useful in exploring a given parameter surface or speeding up convergence.

The PSO-driven parameter optimization was successfully tested on **HMQCA**. The success was the overall reduction of the fitness sum, mainly because the new parameters recovered the $n\pi^*$ character for the S_1 states. Unfortunately, that did not lead to a better description of the PES of

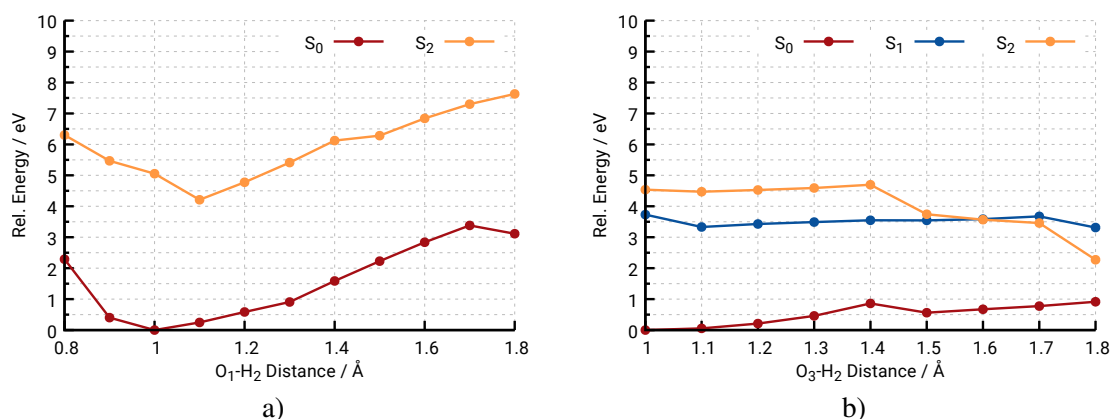


Figure 5.4.: One-dimensional PES scans of the I→II (a) and III→IV (b) PT of **HMQCA** using the RM1opt5 parameters. The S₁ scan in (a) led to a dissociation of the quinoline ring and is therefore not shown. The energies given relative to the S₀ minimum.

HMQCA which was tested on two scans corresponding to the I→II and III→IV PTs (cf. Fig. 4.10). However, increasing the amount of reference data for the parametrization process did lead to slightly smoother PES-descriptions.

The question at this point is how much information must be recovered by the parameters. Is structural information enough or only energies? Any dynamical reference data like quantum yields? In the five presented attempts, the structures obtained by CC2 and DFT reference calculations were not allowed to relax on the SEQM PES. On the one hand, this was done because the reference properties correspond to a specific structure, on the other hand the structures correspond to minima on the CC2 and DFT PES, respectively, thus allowing the structures to also relax to minima of the RM1 PES can help in obtaining better results.

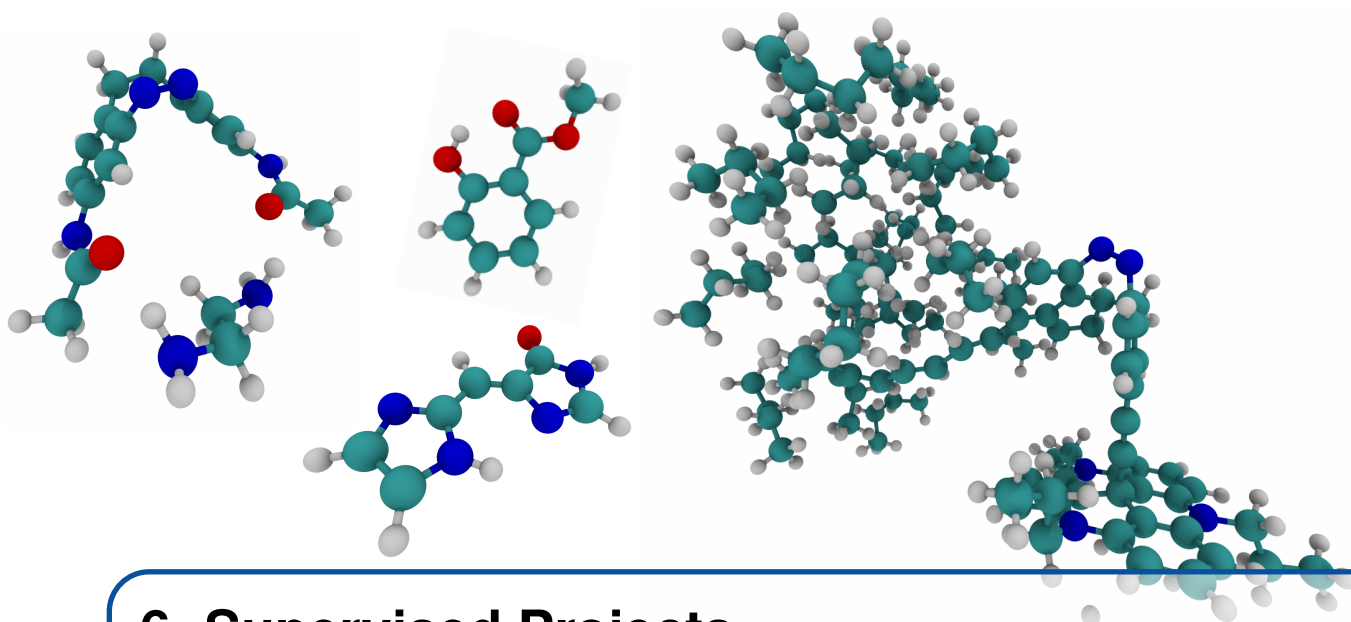
Building a reference data set is the hardest thing in a parametrization process because of many reasons. One example would be overfitting, which means that the parameters may be good in reproducing certain points of the PES – that are part of the reference data – but get substantially worse everywhere else. The presented results for **HMQCA** serve as a perfect example for this: Even though the reference data was reproduced better than before the optimization, the resulting PES, i.e., information that was not part of the reparametrization process, got worse.

The parameter optimization is currently also restricted in the sense that a CI setup – active space and excitation level – needs to be pre-defined and is not optimized alongside the parameters. From the many meta optimizations that were performed during the studies of the thesis, there was no clear indication that a larger active space or a higher CI excitation leads to better results. For example, during the parameter optimization of the azobenzene AM1 parameters, a large [14,13] active space was chosen but with only a CIS excitation level that was extended with an FCI within a [4,3] sub space. Therefore, even though large active spaces and higher excitation levels are advised, more “resource-friendly” settings can get compensated by the optimized parameters.

Because the PSO algorithm presented in this chapter is general and not restricted to optimization of semiempirical parameters, adding an optimization of CI parameters could be realized without a

problem. The only thing that must be done is to ascribe two dimensions of the parameter surface to both excitation level and active space and then convert the position in either of those to any given parameter. For example, if the size of a dimension is $[-1.5;1.5]$ then a particle in the range of $[-1.5;-0.5[$ may be ascribed to CIS, $[-0.5;0.5[$ to Peci and $[0.5;1.5]$ to CISD. But as already mentioned, the more parameters that need to be optimized the harder it is to find a solution.

All in all, the parameter space of the parameter optimization is not only defined by the number of parameters that are to be optimized, but also by the number of parameters that adjust the PSO. On the one hand, more parameters lead to large flexibility of the algorithm, but on the other hand they do not lead to any decisive answer to which PSO setting will yield an optimized set of parameters more efficiently.



6. Supervised Projects

This chapter showcases the projects – Bachelor’s theses and advanced practical courses – that were supervised by the author during the studies for this work. There were three projects on the topic of photoisomerizations of azobenzene and derivatives, and two projects on ESIPT switches. Each project will be presented with its rationale or motivation and a summary of important results.

The following sections on artificial cilia dealt with the same systems from the publication presented in Sec. 3.2 and are therefore not shown here again. Certain acronyms regarding the names of the systems will also be re-used, i.e., azobenzene (**AB**), diazocine (**brAB**), indandiazocine (**ID**), diindandiazocine (**DID**). The cilia are described by the utilized motor followed by **TATA**, e.g, **AB** with with the TATA platform and the tail unit present would correspond to **AB-TATA**.

6.1 Björn Jansen: Photodynamics of Unidirectional Cilia

Björn Jansen performed calculations on static and dynamic properties of a new generation of artificial cilia, as follow-up work of the author’s system.^[108] This new generation used the new bridged-azobenzene derivatives presented in Sec. 3.1 as motor units. Utilizing these new motors should allow for more control over the movement of the cilium which is a crucial design aspect of molecular machines. As for the parent system, ground- and excited-state calculations were performed using a SEQM FOCI setup employing the reparametrized AM1 Hamiltonian of Granucci and Persico^[69] in a QM/MM *ansatz*.

One of the tasks for Björn’s thesis was to update the QM/MM scheme from having the azobenzene motor and tail unit in the QM part and the TATA platform as MM part, to only having the motor in the QM part. This new setup was at first successfully tested on geometry optimizations and dynamics in S_0 . The results from these two types of calculations were compared

to the results of the “naked” (bridged-)azobenzenes from Ref. 141 and experimental data for an indandiazocine on a TATA platform (**ID-TATA**) from Ref. 32. An overall reasonable agreement was achieved for both investigations. Only when the excited-state dynamics were simulated, strange behaviors occurred for the dynamics starting from the *E* isomers, because the produced *Z* isomers all showed extreme bending (Fig. 6.1).

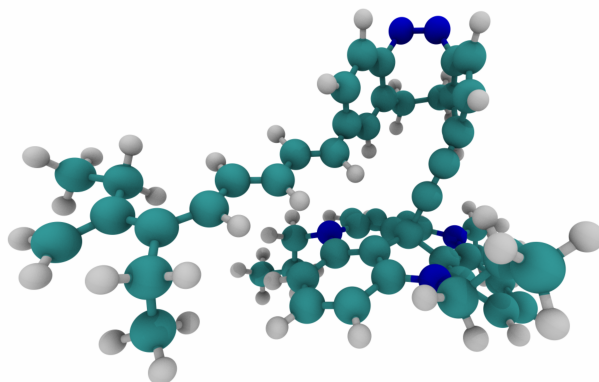


Figure 6.1.: An example for an extremely bent **Z-ID-TATA** during excited-state dynamics.

It was later found out that there were some errors in the input regarding the charge of the connection atom to the tail unit, explaining this behavior of the systems. These were fixed, and all calculations were re-done by the author for the publication presented in Sec. 3.2.

For the scope of Björn’s thesis, all calculations were re-done by him with the old QM/MM setup as it was used in Ref. 108, in which the tail unit was included in the QM part. 250 surface-hopping trajectories were calculated for each system and isomerization direction. Each trajectory was propagated for 5 ps with a time step of 0.1 fs. As it was the case for **AB-TATA**, a larger QM part seems to make the system more prone to stay in S_1 , leading to overall low quantum yields (Table 6.1).

Table 6.1.: Obtained quantum yields for the photodynamics studied in Björn’s Bachelor’s thesis for each system and isomerization direction.

System	$Z \rightarrow E$	$E \rightarrow Z$
brAB-TATA	0.16 ± 0.02	0.09 ± 0.02
ID-TATA	0.02 ± 0.01	0.17 ± 0.03
DID-TATA	0.31 ± 0.03	0.40 ± 0.03

As a side-project Björn also laid the foundation for the Bachelor’s thesis of Dominik Behrens on the transport properties of artificial cilia (see Sec. 6.2), by introducing a small model surface below the system and testing different values for Lennard-Jones (LJ) potential parameters, ϵ and σ (Eq. 6.1).

$$V(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right]. \quad (6.1)$$

6.2 Dominik Behrens: Particle Transport Simulations of Artificial Cilia

Continuing the work of Björn, Dominik Behrens' main task was to simulate the transport properties of the artificial cilia and to optimize the newly introduced model surface to the system.

Regarding the particle transport, Dominik tested several target systems starting with helium as a simple, one-atom target, which proved to be successful (Fig. 6.2).

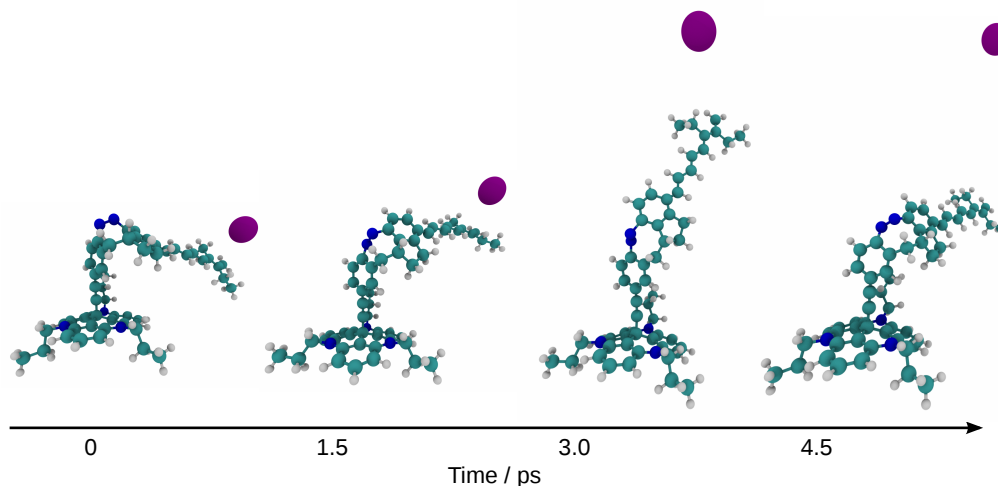


Figure 6.2.: Snapshots of a trajectory of ID-TATA showing particle (He) transport induced by $Z \rightarrow E$ photoisomerization.

Increasing the size of the target to butane, a C_{60} fullerene or a butane cloud with 20 molecules, revealed that the setup has too strong attractive interactions, because the targets kept sticking to the cilium instead of experiencing repulsion. This phenomenon was investigated by changing the parameters of the non-bonding interactions for, again, a single-atom target: Mass of the target, σ , the van der Waals radius, and ϵ , the LJ well-depth.

In short, it was found out that, because there is no transfer of momentum due to collision, the particle transport efficiency is only dependent on the LJ parameters as well as the number of degrees of freedom. The larger the system, i.e., when there is more than one atom, the easier the kinetic energy from the tail of the cilium is dissipated into vibrational instead of translational DOF of the target.

The results of Dominik's Bachelor's thesis corresponded to a large part of the publication presented in Sec. 3.2 and therefore further details are not repeated here.

6.3 Christopher Witt: Photodynamics of a Molecular Pincer

In contrast to the relatively large systems studied in the previous two projects, Christopher Witt dealt with a molecular pincer as another example for a molecule with potential application possibility. The pincer is a **brAB** derivative with added amide groups on either ring that allows grabbing of smaller molecules through hydrogen bonds in the Z form^[142] (Fig. 6.3).

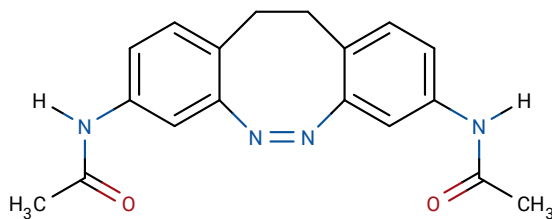


Figure 6.3.: Scheme of the studied pincer system.

To study the efficiency of this process, QM/MM simulations were performed using the same setup as the two previous projects: azobenzene-parametrized FOCI-AM1 as QM method and the OPLSAA-L force field for the MM part. The QM part is the complete pincer system and ethylene diamine (EDA) as target molecule is treated by the force field. This splitting is pretty straightforward because there are no covalent bonds between the two parts, which makes the separation between QM and MM part simple.

The QM/MM setup was benchmarked against DFT reference calculations and crystal structure data available from Ref. 142. A good agreement between the data was found. The pincing process itself was studied with and without a target molecule by photodynamical simulations to investigate the effect of adding a target on the dynamics. It was expected that the complex shows a faster $E \rightarrow Z$ reaction, because of the attractive forces of the hydrogen bonds, and therefore a slower $Z \rightarrow E$ reaction, because of the hydrogen bonds that need to be broken before the isomerization can happen.

The isomerization process was calculated with 200 surface-hopping trajectories for each direction, each propagated for 2 ps, starting in S_1 ($n\pi^*$). The initial structures were sampled from a 10 ps ground-state trajectory including Brownian motion at 298.15 K.

The investigation of the photodynamics of each isomerization direction with and without EDA shows the expected behavior of the systems. As stated in the beginning, the photodynamics of **AB** involve a CoIn, which is reached for CNNC dihedral angles close to 90° . Only through the CoIn, the system is able to return to S_0 . So in addition to looking at the time-evolution of the CNNC dihedral (see below), the population inversion (PI) time, i.e., the time needed to reach 50% S_0 population averaged over the whole trajectory ensemble, is another indicator for the effect of complexation on the dynamics.

Fig. 6.4 shows the ensemble-averaged state populations over time for both isomerization directions with and without EDA. For the $Z \rightarrow E$ reaction without EDA, the PI is reached 28 fs earlier, while for the opposite direction the PI is reached 8 fs later. These numbers are probably within the error margin of the averaging and also probably too small to be measured in any experiment – when there is a collective reaction process of some mmol of molecules. Nevertheless, the effects are as expected.

The same small difference holds true for the QYs. The addition of EDA lowers the QY for the $Z \rightarrow E$ reaction by about 0.04 to 0.70, and increases the QY for the $E \rightarrow Z$ reaction by 0.02 to 0.48

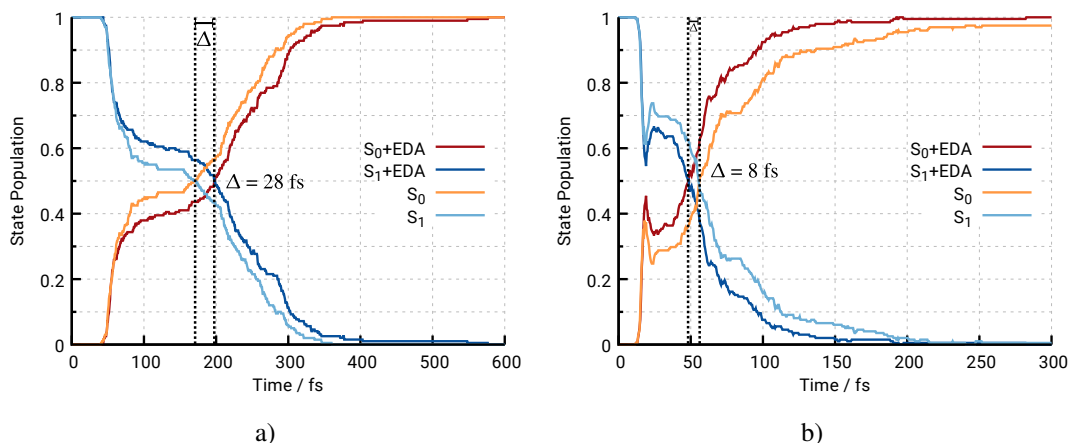


Figure 6.4.: Average ground- (S_0) and excited-state (S_1) populations for the trajectory ensemble starting from a) the Z isomer and b) the E isomer; both for setups with and without an added ethylene diamine (EDA). The dashed lines indicate the population inversion times. The difference for the sets is shown within the plot.

(Table 6.2). Of course, these changes are again within the calculated deviation, but fit the results from the PI rather well.

Looking directly at the degrees of freedom should in theory yield the same results. But due to the oscillative nature of the DOF, such small differences in the low-femtosecond regime may be lost because of the averaging. The only qualitative answer that can be derived from the averaging (Fig. 6.5) might be that, indeed, the $Z \rightarrow E$ reaction with EDA is faster than without EDA, and *vice versa* for $E \rightarrow Z$.

The distance between the two oxygen atoms, which are the two points where molecules are grabbed, seem to give more insight to it. Depending on the isomerization direction, two points of interest may be defined: i) For $Z \rightarrow E$ this can be the first maximum of the plot, because this is where the pincer reached nearly maximum extension so that the amide groups have to rotate to keep the hydrogen bond intact; ii) For $E \rightarrow Z$ this may be the minimum, when the hydrogen bonds are fully formed and any further change in the CNNC dihedral again leads to rotation of the amide groups in order to not “crush” the pinned molecule. The $Z \rightarrow E$ trajectories show a 90 fs increase in reaction time when removing EDA, whereas a 70 fs decrease is seen for the opposite direction in the same case. Even though the numbers are not the same as for the state populations, the qualitative difference that the $E \rightarrow Z$ direction is less affected by the addition of a complexed molecule also can be found here.

Table 6.2.: Quantum yields for the photodynamics of the pincer for both isomerization directions with and without ethylene diamine (EDA).

System	$Z \rightarrow E$	$E \rightarrow Z$
Pincer	0.74 ± 0.03	0.46 ± 0.04
Pincer+EDA	0.70 ± 0.03	0.48 ± 0.04

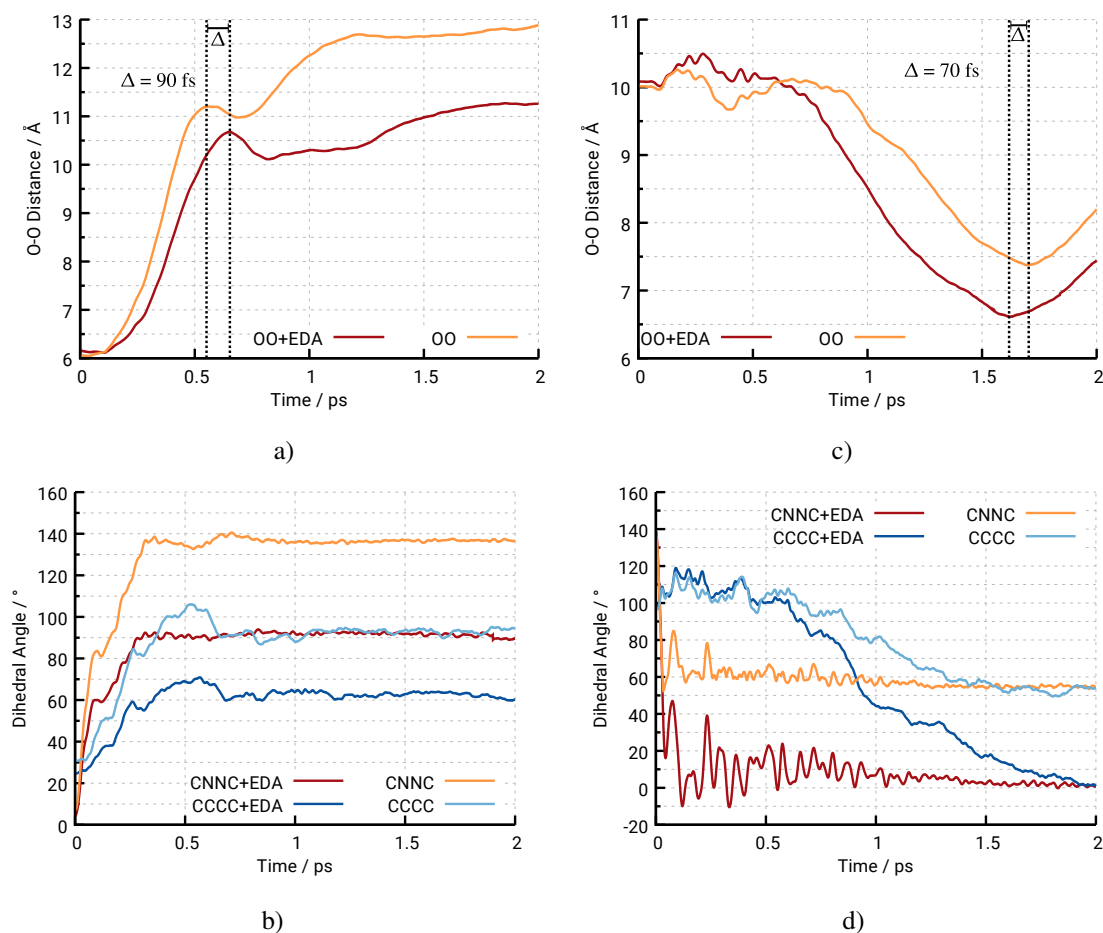


Figure 6.5.: Ensemble-average values for three selected DOF of the pincer during $Z \rightarrow E$ (a, b) and $E \rightarrow Z$ (c, d) dynamics, respectively. Left: The distance between the two carbonyl oxygens; right: The central CNNC and CCCC dihedrals.

In the last part of his Bachelor's thesis, Christopher studied the hydrogen bonds of the Z form with the added EDA. The DFT calculation in Ref. 142 showed hydrogen bonds from the hydrogen of the pincer's amides to the nitrogens of EDA (Fig. 6.6a), whereas the QM/MM calculations show hydrogen bonds from the carbonyl oxygens to the amine protons of EDA (Fig. 6.6b). Comparing the energy of both structures calculated on the FOCI-AM1//OPLSAA-L level reveals a 0.23 eV lower energy for the structure in Fig. 6.6a. Optimizing the DFT structure in the QM/MM setup always resulted in breaking of the hydrogen bonds, because the amide groups always rotated in such a way that the oxygen atoms face "inwards", favoring the $C-O \cdots H$ hydrogen bond.

During visual inspection for other stable hydrogen bonded complexes during the excited-state dynamics, only one additional structure was found that was stable in a subsequent geometry optimization in S_0 . In that case, both carbonyl-oxygen atoms of the pincer bind to each of the two hydrogen atoms of one amine group of EDA (Fig. 6.6c). This complex was found to be only 0.08 eV higher in energy than the structure in Fig. 6.6b. It was argued that the QM/MM optimized structure is chemically more reasonable, because of the electronegativity value of oxygen being

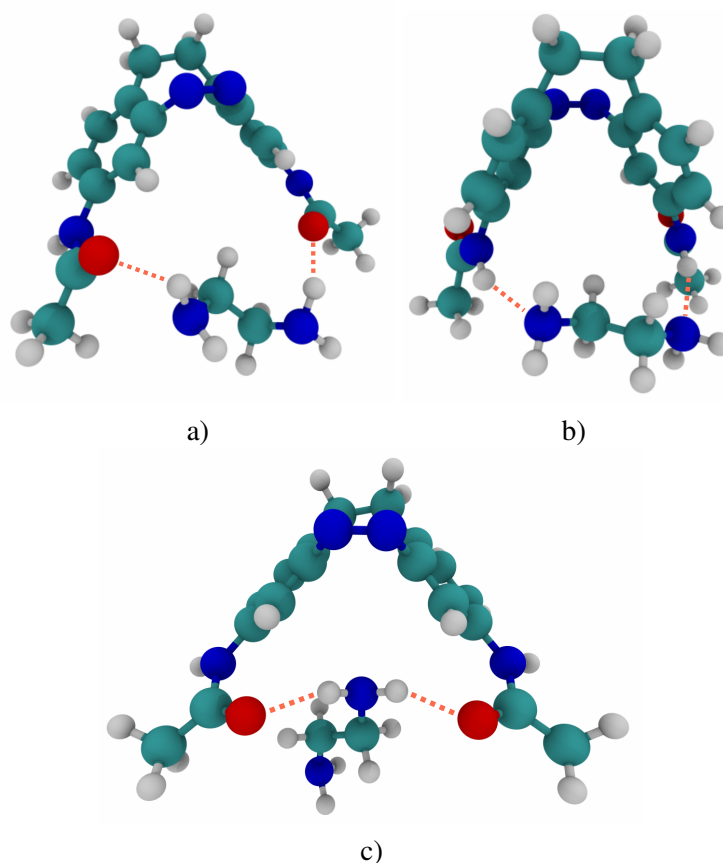


Figure 6.6.: Comparison of the different hydrogen bonds of the optimized FOCI-AM1//OPLSAA-L (a) and B3LYP/6-31G* (b) structures of the pincer with EDA. c) Second stable hydrogen bonded complex of the pincer and EDA found during visual inspection of the trajectories.

greater than the value for nitrogen,^[143] which should result in stronger $O \cdots H$ hydrogen bonds than $N \cdots H$.

As a small design project, Christopher introduced ester groups instead of amide groups (Fig. 6.7), because in that case, the energy difference between the pincer/EDA complex and a “free” pincer is 0.05 eV larger than the same energy with the amide pincer. This could indicate a higher binding energy. Unfortunately, calculating the dynamics revealed the ester groups to be more flexible when it comes to rotation which complicates the formation of hydrogen bonds.

In addition to different binding groups, different molecules to catch were tested for their binding energy. These new molecules were propylene diamine and urea. In both cases, the energy difference between the *Z* and *E* isomers of the pincer is lowered by 0.3 to 0.4 eV (Table 6.3). A destabilized *Z* isomer means that the target molecule is more likely to be set free, thus a higher $Z \rightarrow E$ QY can be expected.

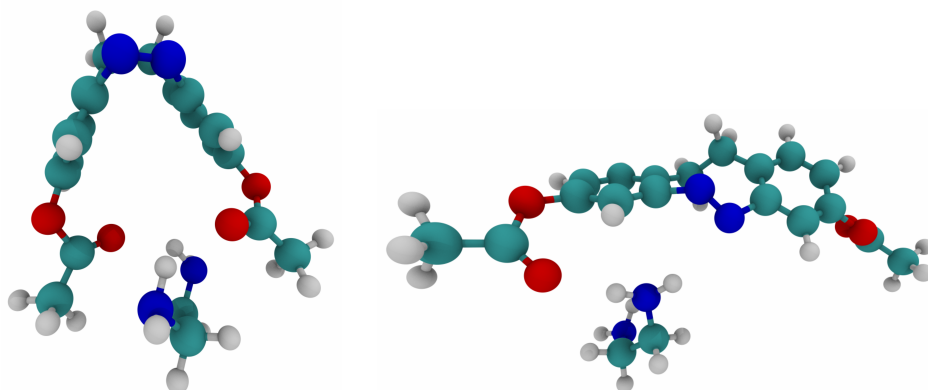


Figure 6.7.: New pincer systems with ester groups instead of amide groups.

Table 6.3.: Energy differences between the *Z* and *E* isomers of the pincer with different target molecules. A positive value means a more stable *Z* isomer. All values in eV.

Target	ΔE_{Z-E}
EDA	0.79
propylene diamine	0.41
urea	0.56
EDA ^a	0.87

^a: When using the pincer with the ester groups

6.4 Jennifer Müller: Photodynamics of Salicylic Acid Derivatives

In her Bachelor's thesis Jennifer Müller investigated different salicylic acid derivatives for their PES and dynamical properties. The first system was methyl salicylate (**SAMe**), in which the esterification disabled the transfer of the "wrong" proton which was an interfering process in the **SAC** dynamics (see Sec. 4.1). The other systems were amide and amine analogues to study the effect of a different proton donor site. The systems are depicted in Fig. 6.8.

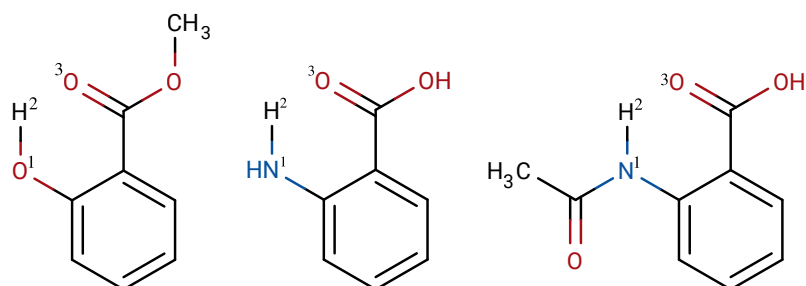


Figure 6.8.: The three different salicylic acid derivatives studied in J. Müllers Bachelor's thesis: Salicylic methylester (**SAMe**), anthranilic acid (**AA**) and N-acetyl anthranilic acid (**AAA**) (from left to right).

The computational setup was given as the same one used for the investigation of the photodynamics of **SAC**, i.e, FOCI-RM1 and an active space of 12 electrons in 12 orbitals, for the sake of comparability.

To study the ESIPT properties, scans of the proton transfer (PT) coordinate and rotation of the carbonyl “crane” were performed, in accordance with the evaluation of **SAC** for straightforward comparison. Also, surface-hopping trajectories were calculated to study the ESIPT timescales, again in comparison to the parent system **SAC**.

Studying the PES at the important points – S_0 and S_1 minima of the two tautomers, the ESIPT transition state and the CoIn – reveals similar properties as **SAC**. Like **SAC**, the proton-transferred forms (II^1) of all three structures are not minima on the S_1 PES and directly evolve to the S_1/S_0 CoIn *via* the rotation of the crane at about $90 \pm 10^\circ$. Surprisingly, the ESIPT energy barrier was found to be higher in the three studied systems: the barrier height between **I-SAMe** and **II-SAMe** is 0.15 eV, in the case of **AA** 0.51 eV and for **AAA** 0.33 eV – for comparison, **SAC** shows a barrier of 0.18 eV. From these values one can expect slightly lower ESIPT quantum yields. All values were benchmarked against TDDFT calculations (PBE0-D3/def2-TZVPP) on S_0 and S_1 . DFT predicted all energy barriers to be about 0.1 eV lower, which is in acceptable agreement.

A 1D scan of the COOH/COOMe rotation in S_1 shows the same behavior for all three system in comparison to **SAC**: There seems to be an extensive CoIn seam spanning from 60° to 120° . 2D scans of the $\text{N}_1\text{-H}_2$ and $\text{N}_1 \cdots \text{O}_3$ distances for all systems in S_1 also do not reveal any qualitative differences: The donor and acceptor atoms still need to come closer to lower the energy barrier during the ESIPT (Fig. 6.9).

To study the dynamical aspects of the ESIPT process of **SAMe**, 400 surface-hopping trajectories were calculated with a duration of 60 ps.

Only 7% of the whole ensemble performed an ESIPT. The reason for this was later found to be the ground-state seeding trajectory. In S_0 , **SAMe** was prone to rotation of the COOMe group leading to an unfavorable conformation for an ESIPT. This behavior is *per se* not problematic, because the rotation of the COOMe group is a known ground-state property of **SAMe**,^[144] unless of course the ESIPT pathway is specifically of interest, as it was the case here.

6.4.1 Additional Information

The computational setup in Jennifer’s Bachelor’s thesis also used the 103 hand-selected determinants as in the publication on the **SAC** excited-state dynamics (cf. Sec. 4.1), which is why the photodynamics of **SAMe** will also be revised using the PECE keyword (cf. Sec. 4.1.3). Additionally, a ground-state seeding trajectory was used in which the system stayed in an ESIPT-favoring rotational conformation by chance.

Applying the former change to a 1D PES scan along the $\text{O}_1\text{-H}_2$ bond of **SAMe**, reveals some changes regarding the proton transferred structure (Fig. 6.10). The ESIPT energy barrier in S_1 is about 0.1 eV smaller than originally. The final ESIPT product at 1.6 Å is energetically favored in the S_1 state, in contrast to the PES constructed with the old CI setup which has the ESIPT minimum 0.07 eV higher in energy than the Franck-Condon minimum (Fig. 6.10b). In both scenarios the

¹ The same nomenclature as for **HMQCA** and its derivatives will be used in this section, i.e., I as the initial structure and II for the proton-transferred form.

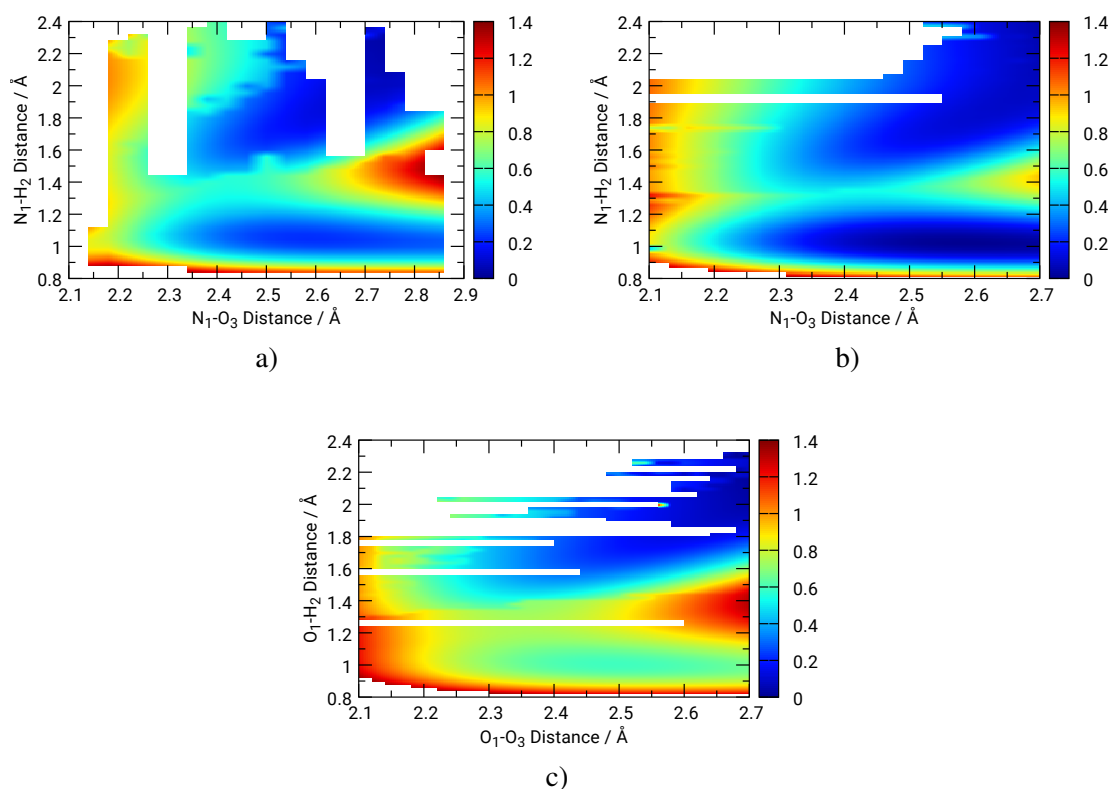


Figure 6.9.: Calculated 2D PES of S_1 for the three studied systems in Jennifer's Bachelor's thesis: a) AAA, b) AA and c) SAME. The white spots correspond to either missing points due to convergence issues or points that are out of the energy range (>2 eV).

COOMe group starts to rotate after the ESIPT is completed, which happens at a larger O_1-H_2 distance in the keyword-based setup.

The smaller barrier and more stabilized ESIPT product were already predicted by the DFT calculations performed in Jennifer's Bachelor's thesis (Fig. 6.10b). Therefore the Peci keyword may be seen as superior.

The S_0 PES is unaffected by the change when looking at the relative energies (Fig. 6.10).

The obtained 2D PES using the keyword-based approach does not show large differences in terms of the energies, because the structures after the ESIPT (upper half of the plots) are both energetically more stabilized than the ones before the ESIPT. But in contrast, the obtained PES utilizing the Peci keyword has converged geometries at every calculated point. In the original approach the white lines indicate missing data points due to problematic/failed convergence of either the SCF or geometry optimization (Fig. 6.9).

Besides the keyword-based Peci approach and ESIPT-favoring starting structures, all other settings, i.e., using the RM1 parametrization, an active space of 12 electrons in 12 orbitals, 200 trajectories and a propagation time of 45 ps, were used in accordance to the original setup for **Sac**. For the sake of comparison, these new trajectories were also recalculated using the original

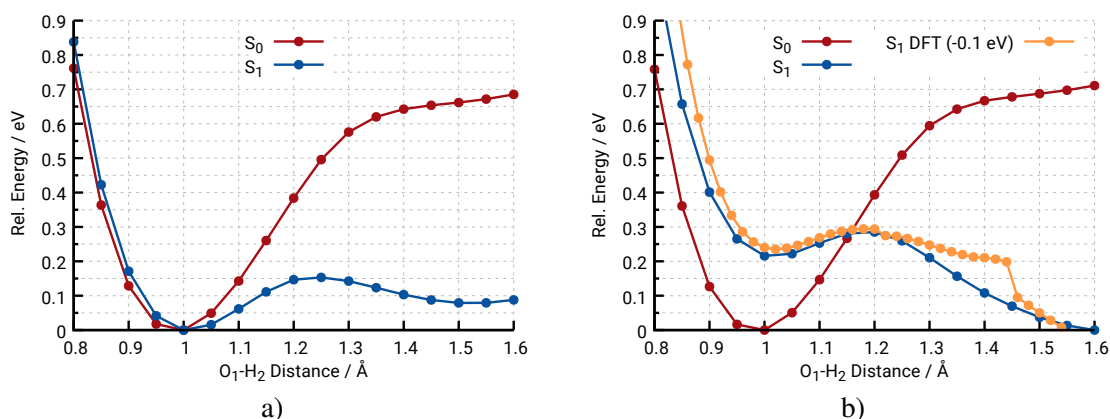


Figure 6.10.: One-dimensional PES scans of the O_1 - H_2 distance of **SAME**, following the S_0 (red) and S_1 gradient (blue), respectively. a) Using the original CI approach; b) using the keyword-based approach. Fig. 6.10b also shows relative DFT energies (shifted by -0.1 eV) taken from Jennifer's Bachelor's thesis.

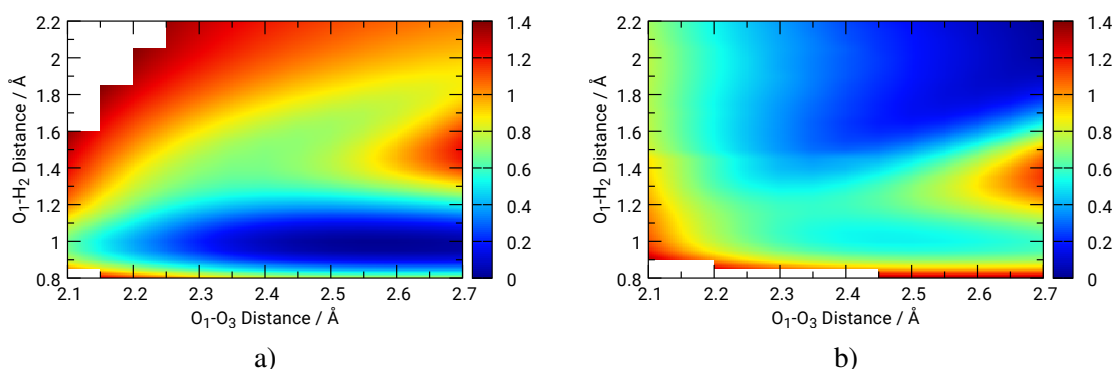


Figure 6.11.: Comparison of the two dimensional PES scans of the O_1 - H_2 distance and O_1 - O_3 distance of **SAME** using the [12,12]-FOPECI-RM1 setup: a) following the S_0 gradient, b) following the S_1 gradient.

hand-selected determinants and also one set using a seed trajectory with a rotated COOMe group, effectively blocking the ability to perform an ESIPT.

The obtained QYs, using the same conditions to define the ESIPT and COOMe rotation as in the publication for **SAC** (cf. Table 4 in the publication shown in Sec. 4.1), are summarized in Table 6.4.

First of all, much higher quantum yields are obtained, regardless of using the PECI keyword or the original setup. Therefore, mindful selection of the initial structures seems to be more important to observe the desired reaction than changing the parameters of the CI part. When starting from a rotated COOMe group, only 1 out of the 200 trajectories performed the ESIPT because the chance of rotating to the ESIPT-favoring position is relatively low with less than 10% for the complete ensemble. The difference between the PECI and original CI setup could be the result of the lower ESIPT barrier, but could also be accounted for by the randomness of the surface hopping method

Table 6.4.: Calculated quantum yields for the ESIPT and COOMe rotation of **SAMe** under different conditions: Using the keyword-based CI approach (PECI), using the 103 hand-selected Slater determinants (original) and starting with a rotated COOMe group (Rot.). The rules that define successful events are the same as used in the publication shown in Sec. 4.1.2 (Table 4 therein).

	PECI	original	Rot.+PECI	Rot.+original
ESIPT	0.96±0.01	0.74±0.03	0.01±0.00	0.01±0.00
COOMe Rotation ^a	0.89±0.02	0.64±0.03	0.09±0.02	0.05±0.02

^a: When starting from the rotated isomer, the target value for the COOMe dihedral changes to 0°.

as explained in Sec. 2.6.1 or the different initial conditions of the trajectories. But the latter two discrepancies can be eliminated when averaging over more trajectories.

The most important result from the photodynamics of **SAMe** is that regardless of using the PECI keyword or the original approach, the system has always reached the electronic ground state within the propagation time (Table 6.5). Only when starting from the rotated isomer, a complete deactivation is not observed.

Table 6.5.: Calculated ensemble-averaged state PI times (50% S_0 population) and time when reaching 100% S_0 population for the four trajectory ensembles of **SAMe** under different conditions: Using the keyword-based CI approach (PECI), using the 103 hand-selected Slater determinants (original) and starting with a rotated COOMe group (Rot.). All values in fs.

	PECI	original	Rot.+PECI	Rot.+original
50% S_0	975	4500	17655	18015
100% S_0	4050	34370	n.a.	n.a.

The two remaining salicylic acid derivatives, **AA** and **AAA**, were also scanned along the ESIPT reaction coordinate using the new setup, but both still exhibited an energy barrier of 0.4 eV and 0.2 eV, respectively (Fig. 6.12). Both values are in accordance to the DFT results obtained by J. Müller.

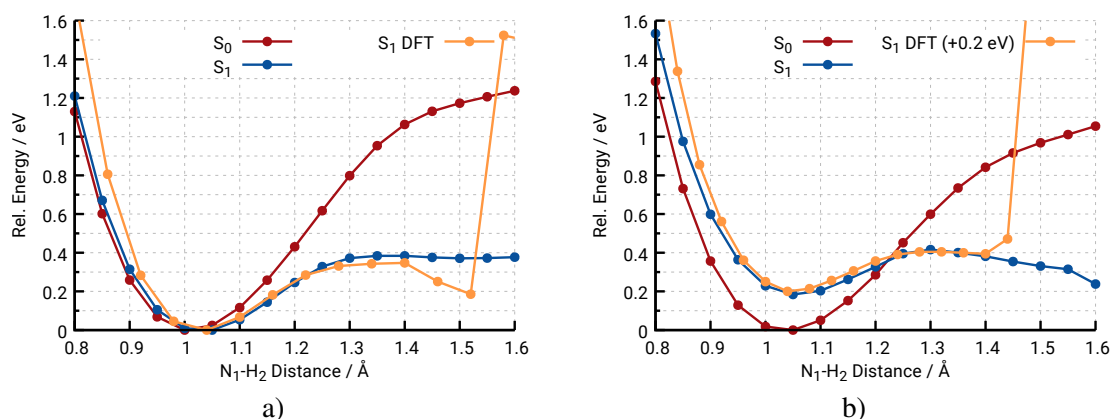


Figure 6.12.: Calculated 1D PES of S_0 and S_1 of the N_1 - H_2 scan of a) AA and b) AAA using the PECI approach. The obtained DFT PES of S_1 from J. Müllers are included for the purpose of comparison. The DFT scan in b) was shifted by +0.2 eV to match the minima. The outliers at the end of the DFT scans are due to larger geometric changes.

6.5 Lasse Janssen: Photodynamics of a Blue Fluorescence Chromophore

Wu *et al.* recently reported theoretical data on the photophysical properties of a blue-fluorescent protein (BFP) analogue.^[138] The BFP in question shows a significantly slower ESIPT rate compared to the widely known green-fluorescent proteins (GFP). GFP analogues were shown to perform an ESIPT within 100 fs or less,^[145–147] whereas the BFP is two orders of magnitude slower.^[148] The report of Wu *et al.* gives detailed insight into the excited state properties and deactivation pathways by means of scans of important DOF on the computationally demanding CASPT2 level of theory.

The scope of Lasse Janssen's practical course was to reproduce the available reference data on the ESIPT and deactivation pathways of the **N** and **T** tautomeric forms of a BFP analogue² (Fig. 6.13) with a SEQM FOCI setup and also to optimize the parameters of the RM1 Hamiltonian to this system. This reference data consisted of vertical excitation energies from the relevant PES minima, and structures and energies from three DOF scans along the ESIPT and rotational deactivation pathways.

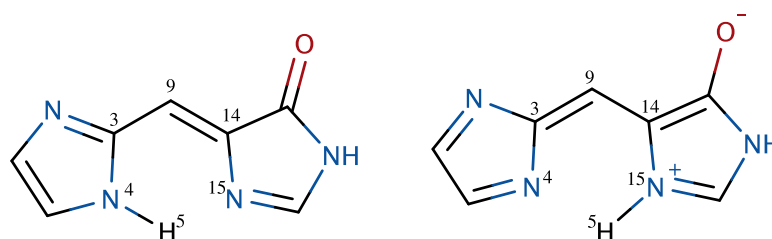


Figure 6.13.: Structures of the two tautomers of the BFP chromophore studied in Ref. 138. Left: **N** tautomer, right: **T** tautomer.

² 5-(1H-imidazol-2-ylmethylidene)-3H-imidazol-4-one.

To find the best fitting semiempirical setup to reproduce this data, a meta optimization was performed. An active space consisting of 12 electrons in 9 orbitals and a CISD excitation level came out as the most suited computational setup. As opposed to, e.g., the meta optimization performed for **SAc**, the RM1 Hamiltonian was predefined. The results are gathered in Table 6.6, with the omitted results for the scans shown in Figures 6.14 and 6.15, for the sake of clarity. For the calculation of the vertical excitation energies and DOF values of the important PES minima, the structures were allowed to relax on the PES of the RM1 Hamiltonian. In contrast, the structures for the scans were kept fixed because they were kindly provided by the authors of Ref. 138.

Table 6.6.: Results of the meta optimization for important points and properties of the PES of the BFP chromophore. All energies in eV.

System	Property	Value	Ref. value	Abs. Error	Rel. Error / %
N^a	$\Delta E_{S_0 \rightarrow S_1}$	3.7201	3.9000	-0.1799	4.61
N^b	$\Delta E_{S_1 \rightarrow S_0}$	3.2678	3.1873	0.0805	2.52
N^c	opt. on S_0	2.7749	3.2089	-0.4340	13.53
N^c	opt. on S_1	2.8971	2.7796	0.1175	4.23
T^a	$\Delta E_{S_0 \rightarrow S_1}$	1.1803	0.8670	0.3133	36.14
T^b	$\Delta E_{S_1 \rightarrow S_0}$	3.2455	3.3087	-0.0632	1.91
T^c	opt. on S_0	2.8327	3.1439	-0.3112	9.90
T^c	opt. on S_1	2.8754	3.1135	-0.2381	7.65

a: Optimized on S_0 , *b:* Optimized on S_1 , *c:* Energy for the structure optimized to the S_0/S_1 CoIn relative to the S_0 minimum

The [12,9]-FOCISD-RM1 setup was able to accurately describe the CASPT2 energies. Differences of up to 4% were obtained for the vertical excitation energies ($\Delta E_{S_0 \rightarrow S_1}$), with the exception of the tautomeric form **T** in S_0 which shows a difference of ~ 0.3 eV, or 36%. Additionally, larger absolute energy differences are found for the structure at the CASPT2 CoIn when optimized on S_0 or S_1 , respectively. But again, with $\sim 10\%$ difference, this can be neglected.

The aforementioned data only show the performance of the RM1 Hamiltonian on six points on the PES. So even though these points fit the CASPT2 PES very well, further insight into the surface is needed to make a more educated statement about the quality of the semiempirical results. DOF scans are the best tool for this. They provide enough insight on the shape of the PES, even though they are – most of the time but not exclusively – restricted to one or two dimensions. The CASPT2 reference data provides 1D information for the ESIPT reaction coordinate (Fig. 6.14), the rotational deactivation *via* the $N_4C_3C_9C_{14}$ dihedral and *via* the $C_3C_9C_{14}N_{15}$ dihedral (Fig. 6.15), respectively.

The scans also exhibit good agreement between the best possible RM1 results and CASPT2 reference. The N_{15} -H₅ scans (Fig. 6.14) both show nearly identical behaviour, except for one qualitative difference: When the N_{15} -H₅ dissociation follows the S_0 gradient, the resulting **T** tautomer seems to be a minimum on the CASPT2 PES at 1 Å, in contrast to the RM1 result. This discrepancy can still be found if the scan allowed the structures to relax on the RM1 PES, rather than using the structures obtained from the CASPT2 scan (Fig. 6.16). Note the different axis

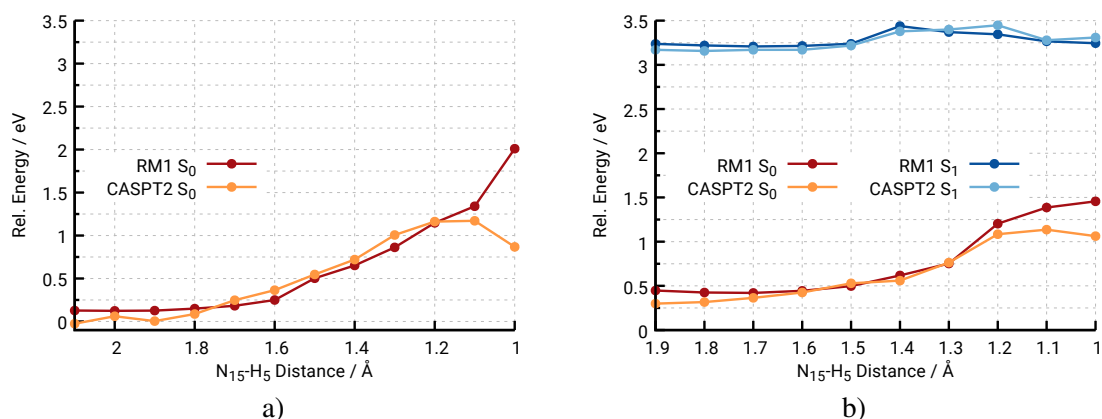


Figure 6.14.: Comparison of the RM1 and CASPT2 energies along a scan of the N_{15} - H_5 distance of the BFP chromophore using the structures from the CASPT2 reference. a) A scan following the ground state and b) following the excited state gradient, respectively. The ground state energies in the right panel are deduced from the vertical excitation energies. All energies for both the RM1 Hamiltonian and CASPT2 are given in reference to the ground state energy of the **N** structure.

label. Whether this is an artifact in the reference calculations or a real qualitative error of the RM1 Hamiltonian, may not be resolved at this point.

The scans of the two important dihedral angles for the **N** tautomer again show good agreement between RM1 and CASPT2 results, in both qualitative behaviour – only the rotation around the $C_3C_9C_{14}N_{15}$ dihedral leads to a CoIn – and relative energies (Figs. 6.15a and 6.15b). For the **T** tautomer qualitative agreement between the two methods exists for the $N_4C_3C_9C_{14}$ dihedral scan, where the energetics are a bit off, but the position of the CoIn at 80° was reproduced (Fig. 6.15c). Artifacts in the form of ground-state energy spikes are present in the scan of the $C_3C_9C_{14}N_{15}$ dihedral (Fig. 6.15d). Without the reference data one could argue that these also correspond to real CoIns, but this is not the case. The position of these spikes at 30° and 50° are present in all tested RM1 setups, so that they may be ascribed to convergence issues of the wave function at the given, fixed structures. These spikes vanish when performing the scans is allowed to relax on either the RM1 or RM1opt surface (Fig. 6.17). The qualitative shape of the 1D PES obtained with both parameter sets shows good agreement with the CASPT2 reference PES. And also the positions of the CoIns were nicely recovered.

This can be seen as another example for the overall satisfying performance of the RM1 parameters for ESIPT systems that perform the reaction exclusively in a $\pi\pi^*$ excited-state.

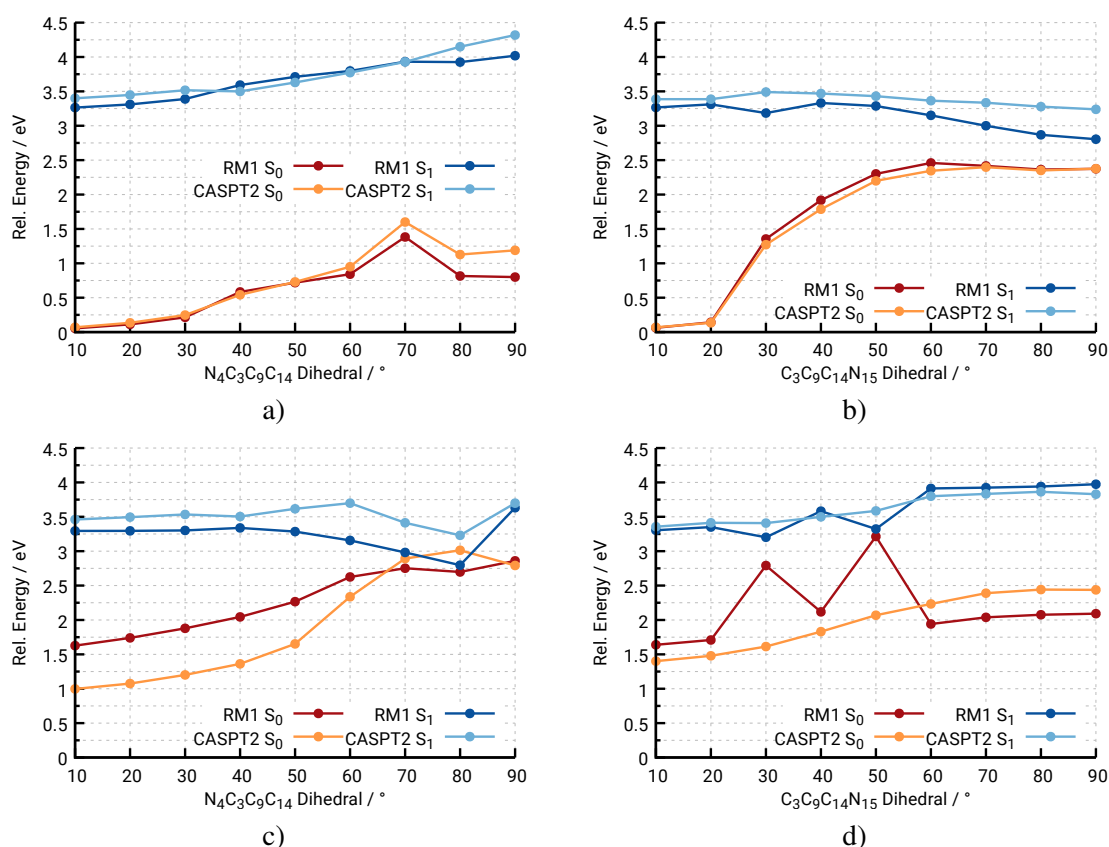


Figure 6.15.: Comparison of the RM1 and CASPT2 energies along a 1D PES scan of the N₄C₃C₉C₁₄ and C₃C₉C₁₄N₁₅ dihedrals of the **N** tautomer (a, b) and **T** tautomer (c, d), respectively, using fixed structures from the CASPT2 reference. The ground state energies are deduced from the vertical excitation energies. All energies for both the RM1 Hamiltonian and CASPT2 are given in reference to the ground state energy of the **N** tautomer. The two spikes for the S₀ scan using RM1 in d) are due to convergence issues.

After the benchmarking, Lasse performed several PSO runs to optimize the RM1 parameters with the given CI setup – only the parameter set with the lowest fitness value will be presented here. Table 6.7 shows the new values for the ones shown in Table 6.6. The new fitness as sum over all absolute differences to the reference, is lowered from 25.04 to 21.39. The important PSO parameter was the scaling interval for the semiempirical parameters, which was set to [0.97;1.03], i.e., allowing a scaling of $\pm 3\%$ of each parameter. This, of course, was by no means an extensive exploration of the parameter surface, but more a local optimization. Larger scaling factors up to $\pm 20\%$ were also tested, but did not lead to a lower fitness value.

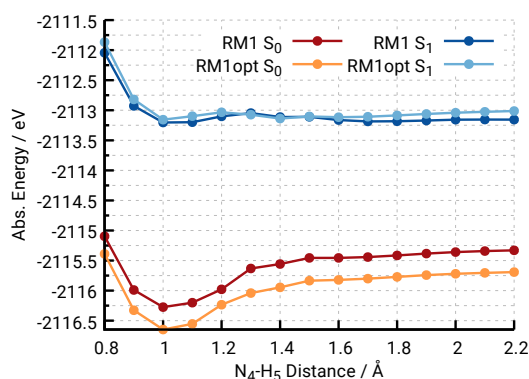
To further compare the two parameter sets, the actual scans as in the reference are performed, instead of just taking the structures from Ref. 138. The ESIPT scan in Fig. 6.16 reveals an energy barrier in the excited state of 0.15 eV, which is about half the size of the reference scan (0.26 eV).

Additionally, the energy difference between the **N**- and **T**-tautomer minima was calculated as 0.05 eV, which again is lower than the reference value of 0.12 eV. Both the energy barrier and

Table 6.7.: Results of the parameter optimization for important points and properties of the PES of the BFP chromophore shown in Table 6.6. All energies in eV.

System	Description	RM1 orig.	Ref. value ^[138]	RM1 opt.
\mathbf{N}^a	$\Delta E_{S_0 \rightarrow S_1}$	3.7201	3.9000	3.7933
\mathbf{N}^b	$\Delta E_{S_0 \rightarrow S_1}$	3.2678	3.1873	3.4658
\mathbf{N}^c	opt. on S_0	2.7749	2.7796	2.7816
\mathbf{N}^c	opt. on S_1	2.8971	3.2089	3.0759
\mathbf{T}^a	$\Delta E_{S_0 \rightarrow S_1}$	1.1803	0.8670	1.1885
\mathbf{T}^b	$\Delta E_{S_0 \rightarrow S_1}$	3.2455	3.3087	3.3935
\mathbf{T}^c	opt. on S_0	2.8327	3.1439	2.8300
\mathbf{T}^c	opt. on S_1	2.8754	3.1135	3.0780

a: Optimized on S_0 , *b:* Optimized on S_1 , *c:* Energy for the structure optimized to the S_0/S_1 CoIn relative to the S_0 minimum

**Figure 6.16.:** Scan of the N_4 - H_5 distance of the BFP chromophore using the RM1 or optimized RM1 gradient, respectively, in the ground and excited state. Note the different DOF in comparison to the scans in Fig. 6.14.

relative energies are increased when the scans are performed with the optimized RM1 parameters, to 0.37 eV for the barrier and 0.19 eV for the relative energy. Smaller barriers are also found for the $C_3C_9C_{14}N_{15}$ dihedral rotations of the \mathbf{N} tautomer and the $N_4C_3C_9C_{14}$ \mathbf{T} tautomer leading to their respective CoIns. This gives a qualitative difference to the reference, because the $C_3C_9C_{14}N_{15}$ dihedral barrier in the both RM1 cases is even smaller than the ESIPT barrier, thus lower a low QY for the ESIPT should be expected. These results are summarized in Table 6.8.

As a third topic, Lasse also performed surface-hopping MD calculations for the given system with the original and optimized RM1 parameters, respectively. Both runs used the same settings for the dynamics, i.e., 200 trajectories with a 8 ps propagation time and a time step of 0.1 fs, including the first five singlet states, starting on the S_1 surface and including a thermostat at 298.15 K. The MD results shall only be reported briefly.

As expected from the scans of the PES, the ESIPT quantum yields are relatively low – only 3 of 200 trajectories show this reaction –, for which the energy barrier of 0.15 eV can be accounted for. The energy barrier for the $C_3C_9C_{14}N_{15}$ dihedral is non-existent (cf. Fig. 6.17b), so therefore

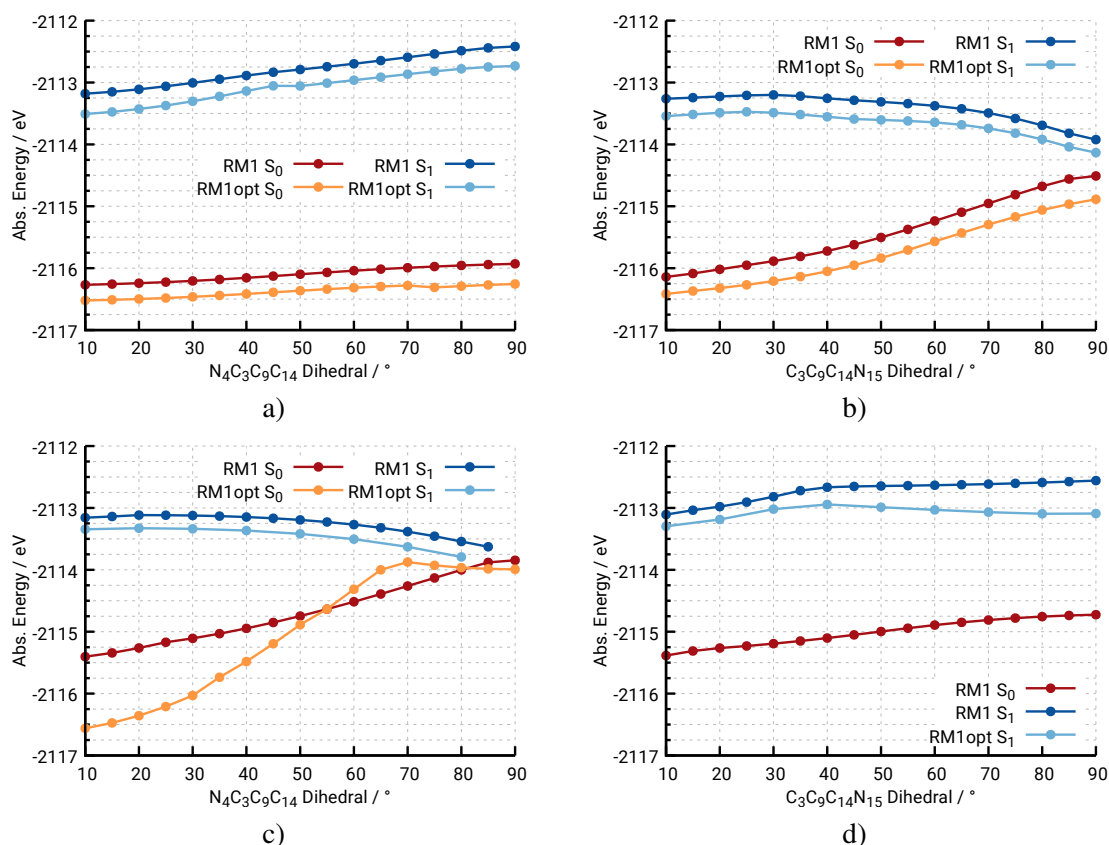


Figure 6.17.: Comparison of the RM1 and RM1opt energies along a 1D PES scan of the $N_4C_3C_9C_{14}$ and $C_3C_9C_{14}N_{15}$ dihedrals of the **N** tautomer (a, b) and **T** tautomer (c, d), respectively. In contrast to the scans shown in Fig. 6.15, these scans are relaxed on the RM1 and RM1opt PES, respectively. The scan for the $C_3C_9C_{14}N_{15}$ dihedral of the **T** tautomer in d) is missing due to convergence issues. The same scans but for a larger range are presented in the appendix (Sec. A.7, Fig. A.5).

the system is more likely to just perform the rotation towards the CoIn without any ESIPT. The missing barrier gives rise to the ultrafast excited-state deactivation, with a half-life of 430 fs for the S_1 population. All trajectories returned to S_0 after 1.5 ps.

Nevertheless, the three “reactive” trajectories subsequently show a deactivation through a CoIn reached by rotation of the $N_4C_3C_9C_{14}$ dihedral angle. All other trajectories also return to the ground state *via* a CoIn, but by the rotation of the $C_3C_9C_{14}N_{15}$ dihedral angle – as expected from the shape of the S_1 PES (cf. Fig. 6.17c). Both pathways are depicted by two trajectories shown in Fig. 6.18.

Since the excited-state PES obtained with the optimized RM1 parameters does not show any significant differences, the trajectories do also. In Fig. 6.17c there are obvious differences for the S_0 surfaces, though. However, this does not matter much for the dynamics initially starting in S_1 .

Table 6.8.: Values for the barrier height and relative energy between the **N** and **T** tautomer of the BFP chromophore as obtained from the scan of the N_4-H_5 distance shown in Fig. 6.16. All values in eV.

Property	Ref.	RM1 orig.	RM1 opt.
ESIPT barrier	0.26	0.15	0.37
ΔE_{N-T}	0.12	0.05	0.19
N $C_3C_9C_{14}N_{15}$ barrier	0.30	n.a.	n.a.
N $N_4C_3C_9C_{14}$ barrier	0.87	0.80	0.83
T $N_4C_3C_9C_{14}$ barrier	0.39	0.06	0.03

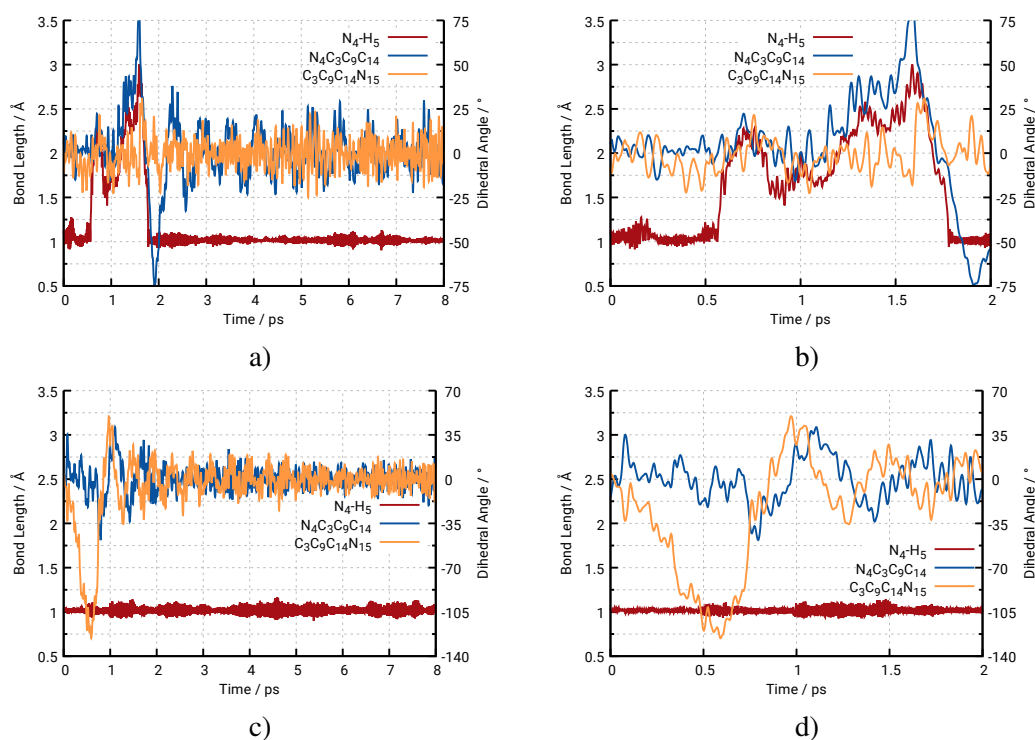
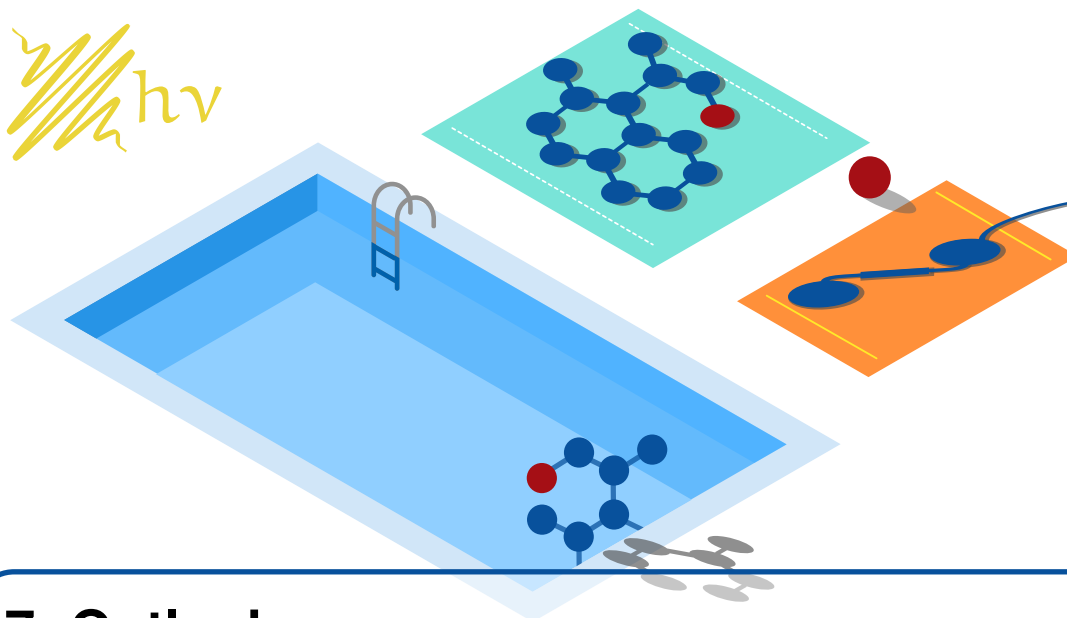


Figure 6.18.: Two trajectories of the excited-state dynamics of the BFP chromophore. a) A trajectory showing an ESIPT and deactivation *via* the $N_4C_3C_9C_{14}$ dihedral. c) A trajectory without an ESIPT and deactivation *via* the $C_3C_9C_{14}N_{15}$ dihedral. The plots on the right-hand side (b, d) show the same corresponding trajectories only in the first 2 ps.

In summary, RM1 is able to accurately reproduce CASPT2 energies of a reference BFP chromophore system. But moving away from the reference structures to performing the same scans reveals some differences, which should be due to different gradients on the PES created by the two methods. Nevertheless, the parameter optimization yielded a better fitness value on paper, but with only minor improvements.

*This page is intentionally left blank; but this text is printed here,
therefore this is not a blank page.*



7. Outlook

Where do we go from here? Regarding the systems in general, solvent was neglected in all ES IPT cases. For ES IPT processes as they are studied at Kiel University this might not be much of a problem, because in order to not influence the reaction, a solvent used in the experiment should be aprotic and apolar. Protic and/or polar solvents would influence the proton during the transfer process. However, for certain use-cases this is actually wanted, for example for the fluorescent probes mentioned in the introduction. To study the ES IPT process itself – as part of fundamental research – these effects just overcomplicate the investigations.

Depending on the size of the “crane” part of ES IPT systems, steric effects of the solvent might be of interest, though. In such a case, explicit treatment of solvent molecules is mandatory. Implicit solvent treatment¹ is not enough, unless, of course, no steric but only (small) electronic effects are expected.

For all of the AB-based systems, explicit solvent is advised, because the photodynamics exhibit a large geometric change that is definitely sterically influenced by the environment, as already shown in the simulations of brAB.^[28] For the even larger AB-based machines, this effect becomes much more crucial. Addition of solvent molecules, though, is most easily done by the QM/MM setup, because usually these molecules do not need to be treated quantum mechanically, especially if only steric effects are of importance. Even simple electronic effects can be realized by the interaction of point charges with the QM wave function. Both steric and electronic effects are easily handled by force fields, therefore these represent a manageable obstacle towards a realistic computational setup.

¹ There are a variety of methods to choose from when it comes to implicit solvent models,^[149,150] with the conductor-like screening model (COSMO)^[151] being one of the popular ones. But basically every program has its own version, which is why the reader should consult the respective program manual to find out which one can be used.

Of course the newest trends in method and hardware development need to be considered for future project planning as well.

Methods based on DFT are the next level after semiempirical methods when it comes to demand of computer resources². The reasons for the limited use of (TD)DFT in photodynamics are its deficient descriptions of conical intersections – because of the mixing of two electronic states which can not be treated accurately by only one SD – or certain kinds of excited state characters.^[152] However, recent developments suggested means of overcoming these deficiencies,^[153, 154] which could make DFT a viable tool in the near future in computational photochemistry.

Research is also devoted to developing low-cost versions of high-level *ab initio* methods like CASPT2, which is the current gold-standard when it comes to multi-reference calculations of excited states. Frank Neese's linear-scaling multireference domain-based pair natural orbital *N*-electron valence perturbation theory (DLPNO-NEVPT2) method is a linear-scaling CASPT2 derivative³, among other approximations, suitable for systems even with even more than 80 atoms.^[155] Hence, the generation of highly-accurate reference data with such methods might also become more accessible in the coming years.

Additionally, because of the simplicity of Tully's surface hopping approach, programs like Sharc^[156] or Newton-X^[157] allow photodynamics with nearly any method that is able to treat electronically excited states and ideally the non-adiabatic coupling vectors (which is usually the missing part). Any upcoming, low-cost electronic structure method can therefore also be (easily) used for excited-state MD calculations.

Hardware-wise, the rise of utilizing graphical processing units (GPUs) in addition to the traditional central processing units (CPUs) in scientific computing may also allow for high-level calculations by means of massive parallelization. One of the pioneers in GPU computing for chemical dynamics is Todd Martínez who has replicated the Miller-Urey experiment *in silico* using HF MD and DFT refinement.^[158] Recent advances also used these techniques for photodynamics, e.g., employing the CASSCF^[159] or TDDFT method,^[160] or to outsource the time-consuming two-electron integrals to the GPU for faster computation.^[161] However, GPU support in QM software packages is still rather limited, and GPUs are not commonly available in traditional computing centers – for instance, there are no servers available at Kiel University that offer GPUs – so that it will take some time before GPU-acceleration of such calculations becomes a standard commodity.

Until then, semiempirical FOCI methods are probably the only resource-friendly way of simulating photodynamics in the multi-picosecond range as well as for extensive design studies. Each SEQM calculation in this thesis, whether it was an energy calculation or a complete trajectory, was run on only one CPU core and used less than 1GB of RAM.

It was shown, though, that semiempirical methods can exhibit deficiencies when representing certain types of reactions or properties. In order to fix those, parameter optimizations or improvements of the underlying Hamiltonian need to be performed. The presented PSO is just one of

² Theoretically it would be HF, but HF-only calculations are nowadays more or less extinct.

³ Assuming the same active space size and basis set.

many possible tools for this purpose. With OGOLEM,^[162,163] there is an in-house program that would also be suited for this task, utilizing genetic algorithms for global optimization problems. However, simply changing the program or algorithm would not magically make the optimization less complex. Any such optimization problem still relies on reference data that must first be obtained. The reference data sets used in this thesis must be seen as the bare minimum in this regard. For example, in the reparametrization process for a reactive force field to a specific type of reaction class in the author's work group, the reference data set consisted of 4877 items.^[164] In contrast, the largest set used here consisted of 108 items⁴. Generating reference data is a project worth its own thesis, though.

Nevertheless, in order to start any optimization process one first needs to show that such a task is both necessary and worth pursuing. Having presented several cases where SEQM FOCI has its strengths and deficiencies, as well as showcasing the success of a system-specific reparametrization, is the contribution of this thesis towards this next step.

⁴ It needs to be noted that the parameter space in Ref. 164 was constructed from >80 parameters per atom, whereas the amount of semiempirical parameters per atom is 28 at most.

*This page is intentionally left blank; but this text is printed here,
therefore this is not a blank page.*



Bibliography

- [1] H. N. Chapman, P. Fromme, A. Barty, T. A. White, R. A. Kirian, A. Aquila, M. S. Hunter, J. Schulz, D. P. DePonte, U. Weierstall, *Nature* **2011**, *470*, 73–77.
- [2] A. Nenov, R. de Vivie-Riedle, *J. Chem. Phys.* **2011**, *135*, 034304.
- [3] G. S. Hartley, *Nature* **1937**, *140*, 281.
- [4] A. Miniewicz, H. Orlikowska, A. Sobolewska, S. Bartkiewicz, *Phys. Chem. Chem. Phys.* **2018**, *20*, 2904–2913.
- [5] J. N. Bull, M. S. Scholz, E. Carrascosa, E. J. Bieske, *Phys. Chem. Chem. Phys.* **2018**, *20*, 509–513.
- [6] P.-M. Su, K.-C. Chang, C.-J. Yang, Y.-C. Liu, W.-S. Chung, *Chem. Commun.* **2017**, *53*, 13241–13244.
- [7] V. Ladanyi, P. Dvorak, J. Al Anshori, L. Vetrakova, J. Wirz, D. Heger, *Photochem. Photobiol. Sci.* **2017**, *16*, 1757–1761.
- [8] C. Nancoz, G. Licari, J. S. Beckwith, M. Soederberg, B. Dereka, A. Rosspeintner, O. Yushchenko, R. Letrun, S. Richert, B. Lang, E. Vauthey, *Phys. Chem. Chem. Phys.* **2018**, *20*, 7254–7264.
- [9] A. Nenov, R. Borrego-Varillas, A. Oriana, L. Ganzer, F. Segatta, I. Conti, J. Segarra-Marti, J. Omachi, M. Dapor, S. Taioli, C. Manzoni, S. Mukamel, G. Cerullo, M. Garavelli, *J. Phys. Chem. Lett.* **2018**, *9*, 1534–1541.
- [10] A. V. Bogdanov, A. K. Vorobiev, *J. Phys. Chem. Lett.* **2018**, *9*, 1372–1376.
- [11] J.-H. Yun, C. Li, S. Kim, M. Cho, *J. Phys. Chem. C* **2018**, *122*, 6310–6317.
- [12] C.-W. Jiang, R.-H. Xie, F.-L. Li, R. E. Allen, *J. Phys. Chem. A* **2010**, *115*, 244–249.

- [13] S. B. Novir, S. M. Hashemianzadeh, *Comp. Theor. Chem.* **2017**, *1102*, 87–97.
- [14] P. Mondal, G. Granucci, D. Rastädter, M. Persico, I. Burghardt **2018**, https://chemrxiv.org/articles/Azobenzene_as_a_Photoregulator_Covalently_Attached_to_RNA_A_Quantum_Mechanics_Molecular_Mechanics-Surface_Hopping_Dynamics_Study/5758665.
- [15] L. Yue, L. Yu, C. Xu, Y. Lei, Y. Liu, C. Zhu, *ChemPhysChem* **2017**, *18*, 1274–1287.
- [16] A. K. Pal, T. J. Duignan, J. Autschbach, *Phys. Chem. Chem. Phys.* **2018**, *20*, 7303–7316.
- [17] A. Muždalo, P. Saalfrank, J. Vreede, M. Santer, *J. Chem. Theory Comput.* **2018**, *14*, 2042–2051.
- [18] T. Cusati, G. Granucci, M. Persico, *J. Am. Chem. Soc.* **2011**, *133*, 5109–5123.
- [19] T. Cusati, Ph.D. thesis, Università degli studi di Pisa, **2009**, group of M. Persico.
- [20] H. Rau, E. Lueddecke, *J. Am. Chem. Soc.* **1982**, *104*, 1616–1620.
- [21] B. K. Pathem, S. A. Claridge, Y. B. Zheng, P. S. Weiss, *Ann. Rev. Phys. Chem.* **2013**, *64*, 605–630.
- [22] S. Monti, G. Orlandi, P. Palmieri, *Chem. Phys.* **1982**, *71*, 87–99.
- [23] T. Ishikawa, T. Noro, T. Shoda, *J. Chem. Phys.* **2001**, *115*, 7503–7512.
- [24] P. Cattaneo, M. Persico, *Phys. Chem. Chem. Phys.* **1999**, *1*, 4739–4743.
- [25] R. Siewertsen, H. Neumann, B. Buchheim-Stehn, R. Herges, C. Näther, F. Renth, F. Temps, *J. Am. Chem. Soc.* **2009**, *131*, 15594–15595.
- [26] R. Siewertsen, J. B. Schönborn, B. Hartke, F. Renth, F. Temps, *Phys. Chem. Chem. Phys.* **2011**, *13*, 1054–1063.
- [27] O. Carstensen, J. Sielk, J. B. Schönborn, G. Granucci, B. Hartke, *J. Chem. Phys.* **2010**, *133*, 124305.
- [28] N. O. Carstensen, *Phys. Chem. Chem. Phys.* **2013**, *15*, 15017–15026.
- [29] R. Liu, A. E. Asato, *Proc. Nat. Acad. Sci.* **1985**, *82*, 259–263.
- [30] W. Fuß, C. Kosmidis, W. E. Schmid, S. A. Trushin, *Angew. Chem. Int. Ed.* **2004**, *43*, 4178–4182.
- [31] J. E. Norton, K. N. Houk, *Mol. Phys.* **2006**, *104*, 993–1008.
- [32] T. Tellkamp, J. Shen, Y. Okamoto, R. Herges, *Eur. J. Org. Chem.* **2014**, *2014*, 5456–5461.
- [33] J. Catalán, J. Palomar, J. L. G. de Paz, *J. Phys. Chem. A* **1997**, *101*, 7914–7921.
- [34] P.-O. Löwdin, *Rev. Mod. Phys.* **1963**, *35*, 724–732.
- [35] L. G. Arnaut, S. a. J. Formosinho, *J. Photochem. Photobiol. A: Chemistry* **1993**, *75*, 1–20.
- [36] S. a. J. Formosinho, L. G. Arnaut, *J. Photochem. Photobiol. A: Chemistry* **1993**, *75*, 21–48.
- [37] T. Mutai, H. Sawatani, T. Shida, H. Shono, K. Araki, *J. Org. Chem.* **2013**, *78*, 2482–2489.
- [38] J. Zhao, S. Ji, Y. Chen, H. Guo, P. Yang, *Phys. Chem. Chem. Phys.* **2012**, *14*, 8803–8817.
- [39] J. E. Kwon, S. Y. Park, *Adv. Mater.* **2011**, *23*, 3615–3642.
- [40] I. P. Pozdnyakov, A. Pigliucci, N. Tkachenko, V. F. Plyusnin, E. Vauthey, H. Lemmetyinen, *J. Phys. Org. Chem.* **2009**, *22*, 449–454.

- [41] H.-C. Lüdemann, F. Hillenkamp, R. W. Redmond, *J. Phys. Chem. A* **2000**, *104*, 3884–3893.
- [42] P. Zhou, M. R. Hoffmann, K. Han, G. He, *J. Phys. Chem. B* **2015**, *119*, 2125–2131.
- [43] L. Lapinski, M. J. Nowak, J. Nowacki, M. F. Rode, A. L. Sobolewski, *ChemPhysChem* **2009**, *10*, 2290–2295.
- [44] M. F. Rode, A. L. Sobolewski, *Chem. Phys.* **2012**, *409*, 41–48.
- [45] M. F. Rode, A. L. Sobolewski, *J. Phys. Chem. A* **2010**, *114*, 11879–11889.
- [46] A. L. Sobolewski, W. Domcke, *Chem. Phys. Lett.* **1999**, *300*, 533–539.
- [47] T. H. van der Loop, F. Ruesink, S. Amirjalayer, H. J. Sanders, W. J. Buma, S. Woutersen, *J. Phys. Chem. B* **2014**, *118*, 12965–12971.
- [48] H. Böhnke, J. Bahrenburg, X. Ma, K. Röttger, C. Näther, M. F. Rode, A. L. Sobolewski, F. Temps, *Phys. Chem. Chem. Phys.* **2018**, *20*, 2646–2655.
- [49] H. Böhnke, University of Kiel, current work.
- [50] S. Lochbrunner, K. Stock, E. Riedle, *J. Mol. Struct.* **2004**, *700*, 13–18.
- [51] F. Jensen, *Introduction to Computational Chemistry*, John Wiley & Sons, **2006**.
- [52] M. Born, R. Oppenheimer, *Ann. d. Phys.* **1927**, *389*, 457–484.
- [53] W. Heitler, F. London, *Z. für Phys.* **1927**, *44*, 455–472.
- [54] L. Piela, *Ideas of Quantum Chemistry 2. Ed.*, Elsevier, Oxford, **2014**.
- [55] T. Helgaker, P. Jørgensen, J. Olsen, *Molecular Electronic-Structure Theory*, John Wiley & Sons, **2000**.
- [56] D. R. Hartree, *Math. Proc. Camb. Phil. Soc.* **1928**, *24*, 89–110.
- [57] D. R. Hartree, *Math. Proc. Camb. Phil. Soc.* **1928**, *24*, 111–132.
- [58] H. Mustroph, *ChemPhysChem* **2016**, *17*, 2616–2629.
- [59] J. C. Slater, *Phys. Rev.* **1928**, *32*, 339–348.
- [60] J. C. Slater, *Phys. Rev.* **1929**, *34*, 1293–1322.
- [61] V. Fock, *Z. für Phys.* **1930**, *61*, 126–148.
- [62] J. E. Lennard-Jones, *Trans. Faraday Soc.* **1929**, *25*, 668–686.
- [63] R. S. Mulliken, *Science* **1967**, *157*, 13–24.
- [64] C. C. J. Roothaan, *Rev. Mod. Phys.* **1951**, *23*, 69–89.
- [65] M. J. S. Dewar, W. Thiel, *J. Am. Chem. Soc.* **1977**, *99*, 4899–4907.
- [66] M. J. S. Dewar, W. Thiel, *Theor. Chim. Acta* **1977**, *46*, 89–104.
- [67] M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, J. J. P. Stewart, *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909.
- [68] G. B. Rocha, R. O. Freire, A. M. Simas, J. J. P. Stewart, *J. Comp. Chem.* **2006**, *27*, 1101–1111.
- [69] T. Cusati, G. Granucci, E. Martínez-Núñez, F. Martini, M. Persico, S. Vázquez, *J. Phys. Chem. A* **2012**, *116*, 98–110.

- [70] J. J. P. Stewart, *J. Mol. Model.* **2013**, *19*, 1–32.
- [71] F. Weigend, R. Ahlrichs, *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
- [72] E. Rossi, G. L. Bendazzoli, S. Evangelisti, D. Maynau, *Chem. Phys. Lett.* **1999**, *310*, 530–536.
- [73] C. Møller, M. S. Plesset, *Phys. Rev.* **1934**, *46*, 618–622.
- [74] G. Granucci, A. Toniolo, *Chem. Phys. Lett.* **2000**, *325*, 79–85.
- [75] H. Koch, P. Jørgensen, *J. Chem. Phys.* **1990**, *93*, 3333–3344.
- [76] X. Ren, P. Rinke, V. Blum, J. Wieferink, A. Tkatchenko, A. Sanfilippo, K. Reuter, M. Scheffler, *New J. Phys.* **2012**, *14*, 053020.
- [77] F. Neese, F. Wennmohs, A. Hansen, U. Becker, *Chem. Phys.* **2009**, *356*, 98–109.
- [78] C.-K. Skylaris, L. Gagliardi, N. C. Handy, A. G. Ioannou, S. Spencer, A. Willetts, *J. Mol. Struct.: THEOCHEM* **2000**, *501*, 229–239.
- [79] C. Hättig, *J. Chem. Phys.* **2003**, *118*, 7751–7761.
- [80] P. Hohenberg, W. Kohn, *Phys. Rev.* **1964**, *136*, B864–B871.
- [81] W. Kohn, L. J. Sham, *Phys. Rev.* **1965**, *140*, A1133–A1138.
- [82] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, M. J. Frisch, *J. Phys. Chem.* **1994**, *98*, 11623–11627.
- [83] C. Adamo, V. Barone, *J. Chem. Phys.* **1999**, *110*, 6158–6170.
- [84] X. Ren, P. Rinke, C. Joas, M. Scheffler, *J. Mater. Sci.* **2012**, *47*, 7447–7471.
- [85] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, *132*, 154104.
- [86] S. Grimme, S. Ehrlich, L. Goerigk, *J. Comp. Chem.* **2011**, *32*, 1456–1465.
- [87] A. K. Rappe, C. J. Casewit, K. S. Colwell, W. A. Goddard, W. M. Skiff, *J. Am. Chem. Soc.* **1992**, *114*, 10024–10035.
- [88] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926–935.
- [89] D. Bakowies, W. Thiel, *J. Phys. Chem.* **1996**, *100*, 10580–10594.
- [90] I. Antes, W. Thiel, *J. Phys. Chem. A* **1999**, *103*, 9290–9295.
- [91] L. Verlet, *Phys. Rev.* **1967**, *159*, 98–103.
- [92] W. C. Swope, H. C. Andersen, P. H. Berens, K. R. Wilson, *J. Chem. Phys.* **1982**, *76*, 637–649.
- [93] N. L. Doltsinis, *Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms* **2002**, *10*, 377–397.
- [94] J. C. Tully, *J. Chem. Phys.* **1990**, *93*, 1061–1071.
- [95] C. Ciminelli, G. Granucci, M. Persico, *Chem. Eur. J.* **2004**, *10*, 2327–2341.
- [96] E. Fabiano, W. Thiel, *J. Phys. Chem. A* **2008**, *112*, 6859–6863.
- [97] M. Barbatti, R. Crespo-Otero, *Surface Hopping Dynamics with DFT Excited States*, Springer International Publishing, **2016**, pp. 415–444.

- [98] M. E. Corrales, V. Lorient, G. Balardi, J. Gonzalez-Vazquez, R. de Nalda, L. Banares, A. H. Zewail, *Phys. Chem. Chem. Phys.* **2014**, *16*, 8812–8818.
- [99] A. W. Jasper, S. N. Stechmann, D. G. Truhlar, *J. Chem. Phys.* **2002**, *116*, 5424–5431.
- [100] G. Granucci, M. Persico, A. Zocante, *J. Chem. Phys.* **2010**, *133*, 134111.
- [101] B. F. E. Curchod, T. J. Martínez, *Chem. Rev.* **2018**, published online, DOI: 10.1021/acs.chemrev.7b00423.
- [102] J. J. P. Stewart, *J. Comp.-Aid. Mol. Des.* **1990**, *4*, 1–103.
- [103] R. Ahlrichs, M. Bär, M. Häser, H. Horn, C. Kölmel, *Chem. Phys. Lett.* **1989**, *162*, 165–169.
- [104] J. J. P. Stewart, *J. Mol. Model.* **2007**, *13*, 1173–1213.
- [105] P. Cattaneo, G. Granucci, M. Persico, *J. Phys. Chem. A* **1999**, *103*, 3364–3371.
- [106] W. van Gunsteren, H. Berendsen, *Mol. Phys.* **1982**, *45*, 637–647.
- [107] G. Bussi, M. Parrinello, *Comp. Phys. Comm.* **2008**, *179*, 26–29.
- [108] T. Raeker, N. O. Carstensen, B. Hartke, *J. Phys. Chem. A* **2012**, *116*, 11241–11248.
- [109] M. Vengris, D. S. Larsen, M. A. van der Horst, O. F. A. Larsen, K. J. Hellingwerf, R. van Grondelle, *J. Phys. Chem. B* **2005**, *109*, 4197–4208.
- [110] M. D. Horbury, L. A. Baker, W.-D. Quan, S. E. Greenough, V. G. Stavros, *Phys. Chem. Chem. Phys.* **2016**, *18*, 17691–17697.
- [111] T. N. V. Karsili, B. Marchetti, M. N. R. Ashfold, W. Domcke, *J. Phys. Chem. A* **2014**, *118*, 11999–12010.
- [112] Y. Miyazaki, K. Yamamoto, J. Aoki, T. Ikeda, Y. Inokuchi, M. Ehara, T. Ebata, *J. Chem. Phys.* **2014**, *141*, 244313.
- [113] M. C. Carreño, I. García, I. Núñez, E. Merino, M. Ribagorda, S. Pieraccini, G. P. Spada, *J. Am. Chem. Soc.* **2007**, *129*, 7089–7100.
- [114] G. Haberhauer, C. Kallweit, C. Wölper, D. Bläser, *Angew. Chem.* **2013**, *125*, 8033–8036.
- [115] K. Takaishi, M. Kawamoto, K. Tsubaki, T. Furuyama, A. Muranaka, M. Uchiyama, *Chem. Eur. J.* **2011**, *17*, 1778–1782.
- [116] K. Takaishi, A. Muranaka, M. Kawamoto, M. Uchiyama, *J. Org. Chem.* **2011**, *76*, 7623–7628.
- [117] G. Despras, J. Hain, S. O. Jaeschke, *Chem. Eur. J.* **2017**, *23*, 10838–10847.
- [118] G. Despras, University of Kiel, current work.
- [119] R. Horbert, Ph.D. thesis, University of Kiel, **2015**, group of C. Peifer.
- [120] R. W. Sparidans, D. Iusuf, A. H. Schinkel, J. H. Schellens, J. H. Beijnen, *J. Chromatogr. B* **2009**, *877*, 4090–4096.
- [121] A. L. Sobolewski, W. Domcke, *Phys. Chem. Chem. Phys.* **2006**, *8*, 3410–3417.
- [122] L. Spörkel, J. Jankowska, W. Thiel, *J. Phys. Chem. B* **2015**, *119*, 2702–2710.
- [123] W. Weber, W. Thiel, *Theor. Chem. Acc.* **2000**, *103*, 495–506.

- [124] A. Koslowski, M. E. Beck, W. Thiel, *J. Comp. Chem.* **2003**, *24*, 714–726.
- [125] G. Cui, Z. Lan, W. Thiel, *J. Am. Chem. Soc.* **2012**, *134*, 1662–1672.
- [126] C.-C. Hsieh, P.-T. Chou, C.-W. Shih, W.-T. Chuang, M.-W. Chung, J. Lee, T. Joo, *J. Am. Chem. Soc.* **2011**, *133*, 2932–2943.
- [127] J. Jankowska, M. F. Rode, J. Sadlej, A. L. Sobolewski, *ChemPhysChem* **2014**, *15*, 1643–1652.
- [128] N. Ole Carstensen, J. M. Dieterich, B. Hartke, *Phys. Chem. Chem. Phys.* **2011**, *13*, 2903–2910.
- [129] A. Toniolo, A. L. Thompson, T. J. Martínez, *Chem. Phys.* **2004**, *304*, 133–145.
- [130] D. R. Armstrong, P. G. Perkins, J. J. P. Stewart, *J. Chem. Soc., Dalton Trans.* **1973**, 838–840.
- [131] M. F. Cardoso, R. Salcedo, S. F. de Azevedo, *Comp. & Chem. Eng.* **1996**, *20*, 1065–1080.
- [132] R. Poli, J. Kennedy, T. Blackwell, *Swarm Intelligence* **2007**, *1*, 33–57.
- [133] M. Dittner, J. Müller, H. M. Aktulga, B. Hartke, *J. Comp. Chem.* **2015**, *36*, 1550–1561.
- [134] K. K. Bejagam, S. Singh, Y. An, C. Berry, S. A. Deshmukh, *J. Phys. Chem. B* **2018**, *122*, 1958–1971.
- [135] H. N. Bhandari, X. Ma, A. K. Paul, P. Smith, W. L. Hase, *J. Chem. Theory Comput.* **2018**, *14*, 1321–1332.
- [136] S. Agrawal, S. Silakari, J. Agrawal, *Mol. Inf.* **2015**, *34*, 725–735.
- [137] T.-D. Liu, T.-E. Fan, G.-F. Shao, J.-W. Zheng, Y.-H. Wen, *Phys. Lett. A* **2014**, *378*, 2965–2972.
- [138] D. Wu, W.-W. Guo, X.-Y. Liu, G. Cui, *ChemPhysChem* **2016**, *17*, 2340–2347.
- [139] K. Schlansker, <https://metacpan.org/pod/AI::PSO>, v0.86.
- [140] P. Jaquierey, <https://metacpan.org/pod/AI::ParticleSwarmOptimization>, v1.006.
- [141] T. Raeker, B. Hartke, *ScienceOpen Research* **2015**, 1–10, DOI: 10.14293/S2199-1006.1.SOR-CHEM.ARDTLN.v1.
- [142] H. Sell, C. Näther, R. Herges, *Beilstein J. Org. Chem.* **2013**, *9*, 1–7.
- [143] W. Gordy, W. J. O. Thomas, *J. Chem. Phys.* **1956**, *24*, 439–444.
- [144] P. Zhou, M. R. Hoffmann, K. Han, G. He, *J. Phys. Chem. B* **2015**, *119*, 2125–2131.
- [145] K.-Y. Chen, Y.-M. Cheng, C.-H. Lai, C.-C. Hsu, M.-L. Ho, G.-H. Lee, P.-T. Chou, *J. Am. Chem. Soc.* **2007**, *129*, 4534–4535.
- [146] C.-C. Hsieh, P.-T. Chou, C.-W. Shih, W.-T. Chuang, M.-W. Chung, J. Lee, T. Joo, *J. Am. Chem. Soc.* **2011**, *133*, 2932–2943.
- [147] Y.-H. Hsu, Y.-A. Chen, H.-W. Tseng, Z. Zhang, J.-Y. Shen, W.-T. Chuang, T.-C. Lin, C.-S. Lee, W.-Y. Hung, B.-C. Hong, S.-H. Liu, P.-T. Chou, *J. Am. Chem. Soc.* **2014**, *136*, 11805–11812.
- [148] X. Fang, Y. Wang, D. Wang, G. Zhao, W. Zhang, A. Ren, H. Wang, J. Xu, B.-R. Gao, W. Yang, *J. Phys. Chem. Lett.* **2013**, *5*, 92–98.

- [149] J. Tomasi, M. Persico, *Chem. Rev.* **1994**, *94*, 2027–2094.
- [150] J. Tomasi, B. Mennucci, R. Cammi, *Chem. Rev.* **2005**, *105*, 2999–3094.
- [151] A. Klamt, G. Schuurmann, *J. Chem. Soc., Perkin Trans. 2* **1993**, 799–805.
- [152] D. J. Tozer, R. D. Amos, N. C. Handy, B. O. Roos, L. Serrano-ANDRES, *Mol. Phys.* **1999**, *97*, 859–868.
- [153] B. G. Levine, C. Ko, J. Quenneville, T. J. Martínez, *Mol. Phys.* **2006**, *104*, 1039–1051.
- [154] S. Pijeu, E. G. Hohenstein, *J. Chem. Theory Comput.* **2017**, *13*, 1130–1146.
- [155] Y. Guo, K. Sivalingam, E. F. Valeev, F. Neese, *J. Chem. Phys.* **2016**, *144*, 094111.
- [156] S. Mai, M. Richter, M. Ruckebauer, M. Oppel, P. Marquetand, L. González **2014**, www.sharc-md.org.
- [157] M. Barbatti, G. Granucci, M. Ruckebauer, F. Plasser, R. Crespo-Otero, J. Pittner, M. Persico, H. Lischka **2016**, www.newtonx.org.
- [158] L.-P. Wang, A. Titov, R. McGibbon, F. Liu, V. S. Pande, T. J. Martínez, *Nature Chem.* **2014**, *6*, 1044–1048.
- [159] J. W. Snyder, B. F. E. Curchod, T. J. Martínez, *J. Phys. Chem. Lett.* **2016**, *7*, 2444–2449.
- [160] B. F. E. Curchod, A. Sisto, T. J. Martínez, *J. Phys. Chem. A* **2017**, *121*, 265–276.
- [161] N. Luehr, I. S. Ufimtsev, T. J. Martínez, *J. Chem. Theory Comput.* **2011**, *7*, 949–954.
- [162] J. M. Dieterich, B. Hartke, *Mol. Phys.* **2010**, *108*, 279–291.
- [163] <https://www.ogolem.org>.
- [164] J. Müller, B. Hartke, *J. Chem. Theory Comput.* **2016**, *12*, 3913–3925.
- [165] B. Hartke, *J. Comp. Chem.* **1999**, *20*, 1752–1759.

All webpages have been accessed on April 10, 2018.

*This page is intentionally left blank; but this text is printed here,
therefore this is not a blank page.*

A. Appendix

A.1 PM6* Parameters

USS H	-11.7199369	BETAP N	-14.8510739
UST H	-11.7199369	BETST N	-18.0520887
BETAS H	-4.8637736	BETPT N	-14.8510739
BETST H	-4.8637736	ZS N	2.2827006
ZS H	0.9696534	ZP N	1.8924715
ALP H	3.7831803	ALP N	3.0063575
GSS H	13.0556814	GSS N	10.7696034
FN11 H	0.1601564	GSP N	11.7425802
FN21 H	1.0951255	GPP N	13.5435421
FN31 H	0.9456355	GP2 N	11.9940423
USS C	-52.1453527	HSP N	1.8706595
UPP C	-39.5919145	FN11 N	0.0701091
UST C	-52.1453527	FN21 N	3.0000163
UPT C	-39.5919145	FN31 N	1.4064711
BETAS C	-14.7446649	USS O	-98.1522233
BETAP C	-6.7458623	UPP O	-76.7874961
BETST C	-14.7446649	UST O	-98.1522233
BETPT C	-6.7458623	UPT O	-76.7874961
ZS C	1.8288179	BETAS O	-29.7468281
ZP C	1.6063756	BETAP O	-26.1198462
ALP C	2.8647964	BETST O	-29.7468281
GSS C	13.0006143	BETPT O	-26.1198462
GSP C	11.2289057	ZS O	3.3218475
GPP C	11.0682246	ZP O	2.5185456
GP2 C	9.8624564	ALP O	3.9798393
HSP C	2.0300783	GSS O	9.6631935
FN11 C	0.0697912	GSP O	16.2638840
FN21 C	3.0005515	GPP O	11.8184957
FN31 C	1.4983387	GP2 O	11.6476900
USS N	-65.7942586	HSP O	4.1271060
UPP N	-58.2141761	FN11 O	0.1876239
UST N	-65.7942586	FN21 O	1.7974688
UPT N	-58.2141761	FN31 O	0.8238496
BETAS N	-18.0520887		

A.2 Carbohydrate Macrocycles

A.2.1 Energies

Table A.1.: Calculated relative energies of all unimproved maltose-containing systems. Each relative energy is given with respect to the *trans* structure of the corresponding group (*cisM* for the open system). All values in eV.

			ΔE				ΔE
M	<i>trans</i>		0	MAc	<i>trans</i>		0
	<i>cisM</i>		-1.189		<i>cisM</i>		-0.251
	<i>cisP</i>		-1.429		<i>cisP</i>		-0.053
MAcNCS	<i>cisM</i>		0				
	<i>cisP</i>		1.268				

A.2.2 Spectra

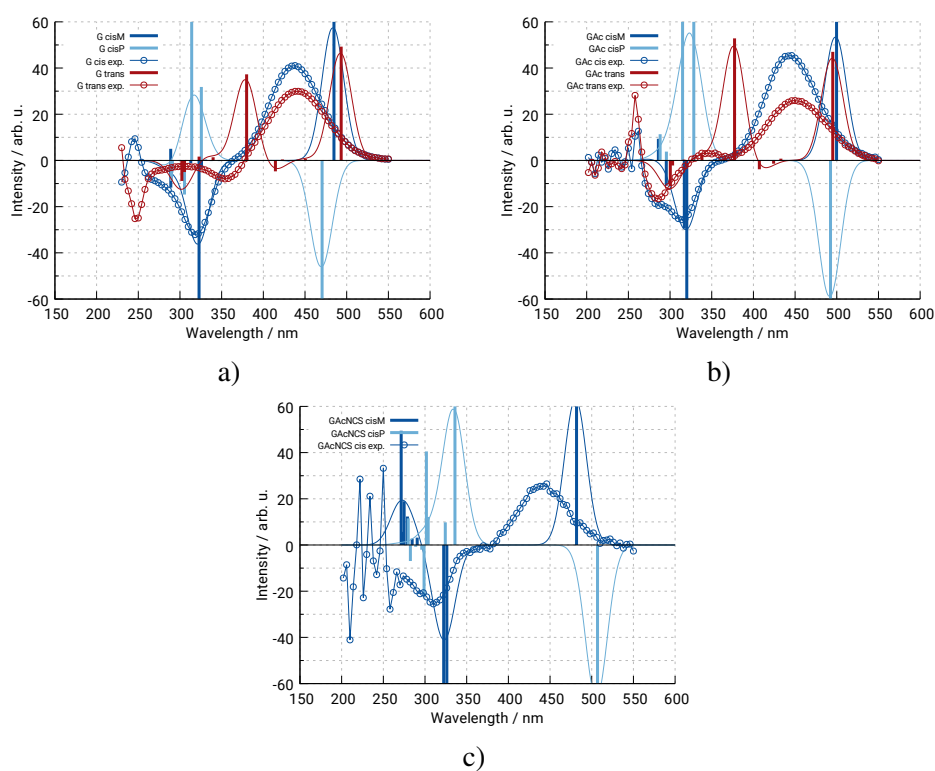


Figure A.1.: Calculated CD spectra (including implicit treatment of solvent) of all three isomers (*trans*, *cisM*, *cisP*) of **G**: a) **G**, b) **GAc**, c) **GAcNCS**. All spectra are compared to obtained experimental spectra.

A.2 Carbohydrate Macrocycles

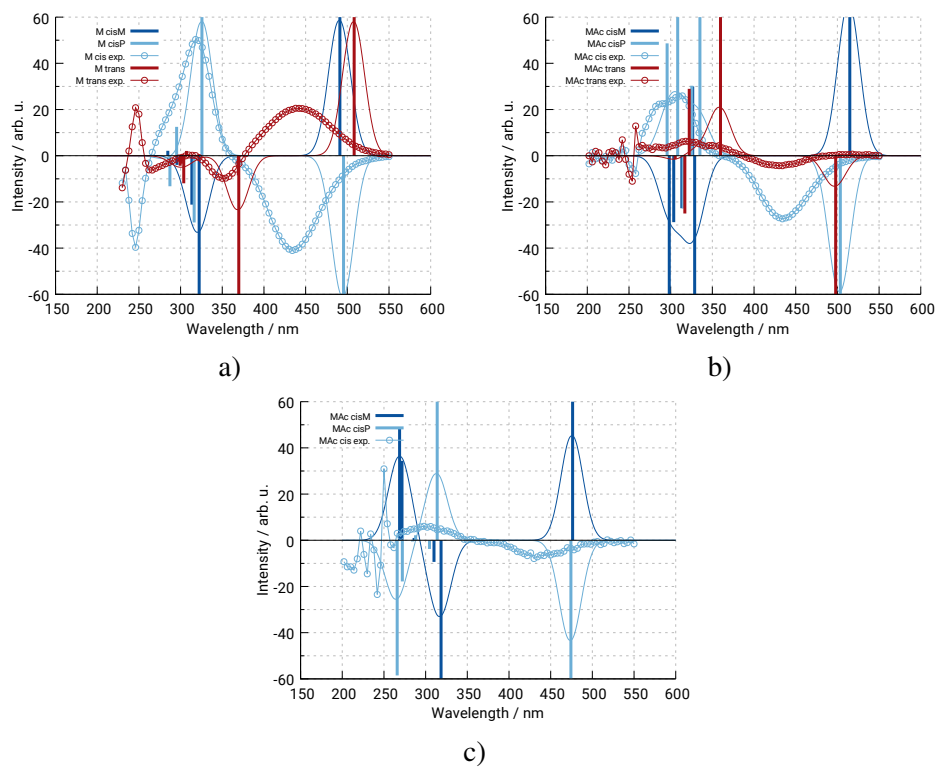


Figure A.2.: Calculated CD spectra (including implicit treatment of solvent) of all three isomers (*trans*, *cisM*, *cisP*) of **M**: a) **M**, b) **MAc**, c) **MAcNCS**. All spectra are compared to obtained experimental spectra.

A.3 Salicylic Acid Derivatives

A.3.1 DFT Scans

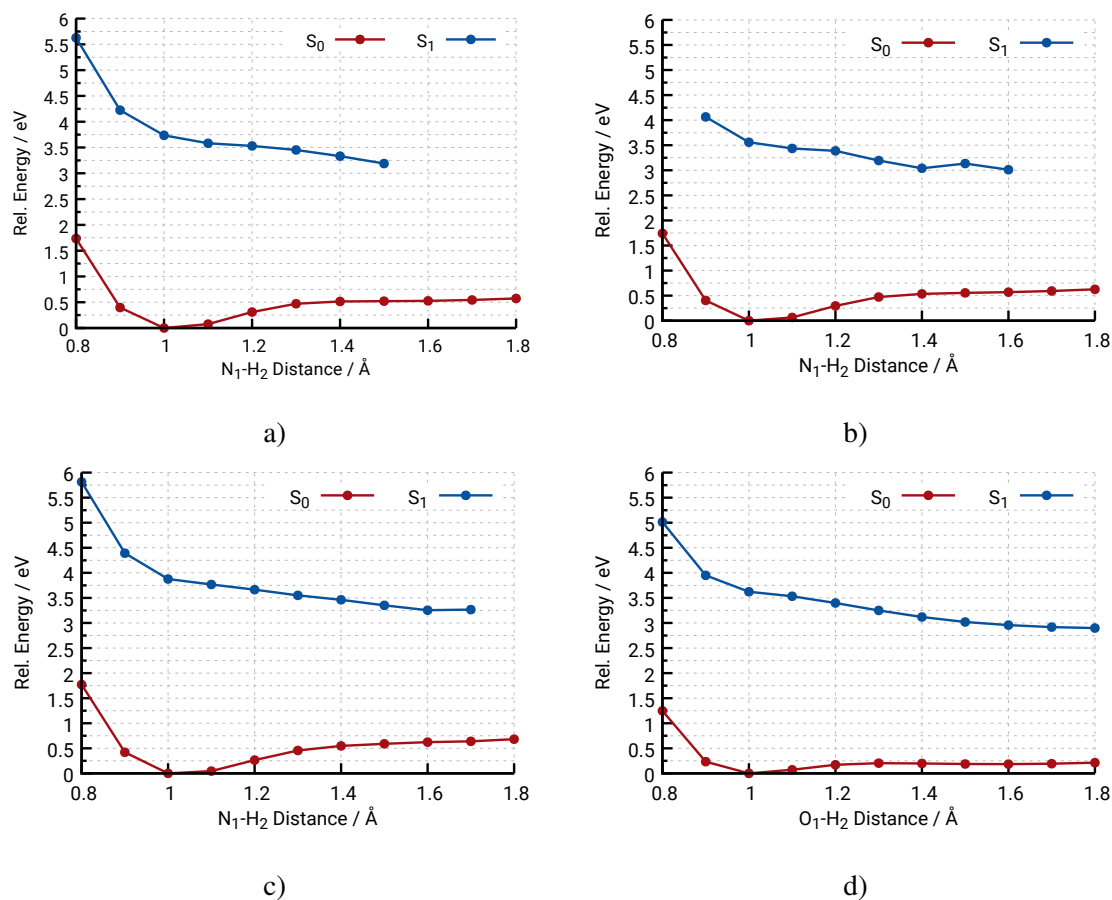
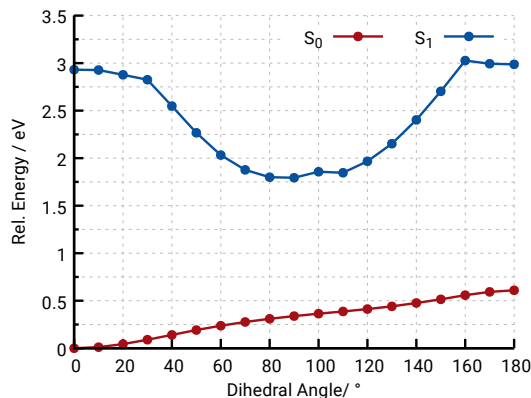


Figure A.3.: Obtained 1D PES for the four studied salicylic acid derivatives on the PBE0-D3BJ/def2-SVP level of theory: a) **MIMA**, b) **MIMAMe**, c) **MIMAac** and d) **MIMPy**. The energies are given relative to the minimum on S_0 .

A.4 Scan of the Crane Rotation of MMIMQO



a)

Figure A.4.: One-dimensional PES scan of the CNMe rotation in S_0 and S_1 of MMIMQO. The energies are given relative to S_0 .

A.5 PSO Routines

In the following the five newly¹ implemented routines in the PSO algorithm shall be presented briefly and how they affect the exploration of the particles.

A.5.1 getNeighborIndices

There are three ways implemented to define the neighbors of a particle:

1. *static*: A particle's neighbors are defined at the initialization stage and are kept throughout all iterations. The only implemented topology for this is "circular", i.e., all particles are arranged in a circle and N_{neigh} neighbors are defined $N_{neigh}/2$ left of the particle and $N_{neigh}/2$ right of it. This requires N_{neigh} to be even and the topology ensures that if particle 1 sees particle $1+x$, particle $1+x$ also sees particle 1.
2. *dynamic*: A particle's neighbors are defined by the N_{neigh} nearest neighbors. An update of the neighbor list is performed at every iteration before the velocity update. So far, there is no restriction with regards to a maximum distance implemented.
3. *dynamic, fitness distance ratio (FDR)*: A particle's neighbors are defined by N_{neigh} nearest neighbors in relation to their own best memorized fitness. This ensures that a particle only sees "good performing" particles even if their geometric distance is larger than other particles. This check is performed at every iteration before the velocity update.

Given the small number of particles used in the parOpts, the *static* setting was used throughout the thesis in which every particle has every other as its neighbor. It is noted that this topology

¹ "New" in terms of these were not implemented in the two available Perl modules in the comprehensive perl archive network (CPAN).^[139,140] The first three subroutines in this section were taken from literature-known derivations of the PSO algorithm discussed in Ref. 132. The last two subroutines are newly invented to the author's best knowledge.

must not necessarily mean that a full-informed particle swarm (FIPS) is used in the classical sense, i.e., as defined in the review article of Ref. 132, because this requires a slightly different equation of the velocity update. However, a FIPS is also possible within the author’s PSO implementation.

When there is a large number of particles, restricting the communication between particles would allow to mimic the concept of “niching”,^[165] i.e., having a subset of particles within separate regions on the parameter surface. Because the “neighbors” of a particle must not be in “close” vicinity – unless the selection was chosen to be dynamic – the particle with the best fitness “pulls” its neighbors towards its best known position. Unless there are better positions found during this migration, this feature helps in converging to a good solution. A deficiency of this approach would be that if the globally best fitness region is not part of the found niches, it can be hard to still find it when all neighboring particles are within their niche. Special care must be taken in this case, e.g., with the following subroutine.

A.5.2 checkParticleDistance

To make sure that there is enough exploration of the given parameter surface a minimum geometrical distance d_{min} between two particles may be set,

$$d = \sqrt{(\Delta x_1)^2 + (\Delta x_2)^2 + \dots + (\Delta x_N)^2} \leq d_{min}. \quad (\text{A.1})$$

If the actual distance between two particles is less than d_{min} , the particle that was moved last is re-initialized randomly on the parameter surface but with kept memory of its best-fitness position.

This routine should mainly help at the end of a PSO run when (ideally) many particles are close at the deepest found minimum (of their neighborhood). Then all except the first particle in this area would give redundant results. To increase the efficiency, these extra particles are then just put somewhere on the parameter surface to explore a new path.

A.5.3 performLocalSearch

Some minima on the parameter surface may be rather shallow. In such areas additional exploration on the currently known deepest minimum could be of help in form of local searches.

At a defined iteration number a specified number of local searches is performed starting from deepest known minimum. The user may specify a search radius, which is basically just the maximum amount of change the particle is allowed in each dimension. If the particle in question finds a deeper minimum, its memory will be updated with this information and depending on the setting regarding the swarm memory, all other particles may know the coordinates as well.

A.5.4 checkForPlateau

There may be a scenario – which is rather likely – that a particle does not find a deeper minimum for a considerable amount of iterations. This is of course not an ideal scenario because it just wastes computational time.

A.6 Optimized Parameters Sets

The user may define a maximum amount of iterations for a particle to not find any deeper point of the parameter surface. If this amount of iterations is reached the particle then gets re-initialized at a random position on the parameter surface. It may also be toggled that this check is applied to the global best fitness. If no deeper point of the surface is found than the deepest so far, the whole PSO may be stopped at once to save resources. The latter setting must be used with caution, though, because at later stages of the propagation finding new global minima gets harder, so ending it prematurely will impede any further optimization.

A.5.5 shrinkSurface

The last additional routine is to force the particles to the globally found minimum by changing the size of the parameter surface. As has been said before, at late stages of the propagation finding new minima gets harder. When assuming many local minima near the global minimum, it could be wise to force all particles near this area. This is done by shrinking the size of each dimension after a specified number of iterations. The size depends on the position of the particle that knows the deepest position of the parameter surface. For example: Imagine a surface where each dimension is defined in a range $[-5;5]$; the only information needed from the deepest known point are the maximum position in all dimension, e.g., 2.5, and the minimal, e.g., -0.75. If the shrinkSurface routine is called it will decrease the size of the parameter surface linearly/exponentially from the current iteration to the last to $[-0.75;2.5]$. If a new best position is found, the final range of course shifts accordingly.

A.6 Optimized Parameters Sets

A.6.1 Minimal Reference Data Set

```
USS H -11.7789635
UST H -12.6316328
ZS H 1.0859402
BETAS H -5.5184515
BETST H -5.6678479
GSS H 14.1902157
ALP H 3.0809533
FN11 H 0.1031719
FN21 H 6.06405
FN31 H 1.1635363
FN12 H 0.0621348
FN22 H 6.2228068
FN32 H 1.9386176
FN13 H -0.035444
FN23 H 2.7544043
FN33 H 1.6038214
USS C -50.0778265
UST C -49.6190345
UPP C -39.7213973
UPT C -38.6831066
ZS C 1.8218722
ZP C 1.7239085
BETAS C -14.7081123
BETST C -15.4403619
BETAP C -8.2950701
BETPT C -8.4385504
GSS C 13.0430841
GSP C 11.293297

GPP C 10.486586
GP2 C 9.2878641
HSP C 1.5152343
ALP C 2.9328736
FN11 C 0.0756506
FN21 C 5.7206687
FN31 C 1.0246369
FN12 C 0.0119209
FN22 C 7.017821
FN32 C 1.749563
FN13 C 0.0367919
FN23 C 6.1950618
FN33 C 1.647172
FN14 C -0.0027265
FN24 C 9.2292092
FN34 C 2.8538259
USS N -67.0596466
UST N -72.8037992
UPP N -56.5032668
UPT N -56.9657211
ZS N 2.4228946
ZP N 1.9478112
BETAS N -21.1396861
BETST N -21.929464
BETAP N -16.6358378
BETPT N -17.4386804
GSS N 13.161623
GSP N 13.0248185

GPP N 14.2377913
GP2 N 11.5336984
HSP N 5.0618183
ALP N 3.0206501
FN11 N 0.0612224
FN21 N 4.6230263
FN31 N 1.377277
FN12 N 0.0246339
FN22 N 4.4456112
FN32 N 2.0003104
FN13 N -0.0230652
FN23 N 2.1535439
FN33 N 1.8259265
USS O -100.2180617
UST O -98.1242696
UPP O -79.308452
UPT O -76.3154405
ZS O 3.1157818
ZP O 2.6502221
BETAS O -31.450568
BETST O -30.2004578
BETAP O -29.0146824
BETPT O -30.2140029
GSS O 14.0605506
GSP O 14.5303637
GPP O 14.1979713
GP2 O 12.7657953
HSP O 3.8680737
```

ALP O 4.0794653
 FN11 O 0.2359824
 FN21 O 5.3035776

FN31 O 0.9486673
 FN12 O 0.059751
 FN22 O 7.6412579

FN32 O 1.4943893

A.6.2 Larger CI Setup

USS H -11.4335901064915
 UST H -12.0471009363815
 ZS H 1.0881480128548
 BETAS H -5.63571952012483
 BETST H -5.48414596455817
 GSS H 13.6111777980475
 ALP H 3.13671972852727
 FN11 H 0.104288815581595
 FN21 H 6.2039803526783
 FN31 H 1.11154957486307
 FN12 H 0.0607827178845102
 FN22 H 6.78514231009997
 FN32 H 2.04621607472028
 FN13 H -0.0368853834073699
 FN23 H 2.85737098904382
 FN33 H 1.75927921031584
 USS C -54.7612277556587
 UST C -52.0872744461548
 UPP C -37.2259225761898
 UPT C -38.0606822070214
 ZS C 1.8261743699767
 ZP C 1.8338742367529
 BETAS C -16.370138955738
 BETST C -16.1810903780144
 BETAP C -7.80938363415853
 BETPT C -8.53129038369678
 GSS C 13.5122760450012
 GSP C 10.5292244136672
 GPP C 10.844067393407
 GP2 C 9.55676371183295
 HSP C 1.65578677603056

ALP C 2.86399210011061
 FN11 C 0.0717447128426341
 FN21 C 5.94358774871847
 FN31 C 1.09120830200692
 FN12 C 0.0117074566985179
 FN22 C 6.99834751403332
 FN32 C 1.64240536259504
 FN13 C 0.0377525399461658
 FN23 C 6.34643751983861
 FN33 C 1.6094744742872
 FN14 C -0.00273860095553426
 FN24 C 8.52613284458914
 FN34 C 2.79109353444265
 USS N -75.3226650358911
 UST N -71.6601149521547
 UPP N -57.8236144140548
 UPT N -59.9297659345707
 ZS N 2.47514679923495
 ZP N 2.02891316131343
 BETAS N -21.0442210137006
 BETST N -20.7923865014626
 BETAP N -17.5902184700535
 BETPT N -16.3510031794987
 GSS N 14.0171192328034
 GSP N 13.3821698381235
 GPP N 14.3333882171498
 GP2 N 11.6940891183963
 HSP N 5.15710199152331
 ALP N 2.837144170111714
 FN11 N 0.0626541113056341
 FN21 N 4.61696238205627

FN31 N 1.33394504261451
 FN12 N 0.0237683603323077
 FN22 N 4.9247455121213
 FN32 N 2.05063970211061
 FN13 N -0.0227832524734404
 FN23 N 2.12643420582781
 FN33 N 1.83000118213893
 USS O -92.3524483391485
 UST O -102.219879317692
 UPP O -84.5044594379046
 UPT O -77.6109297605738
 ZS O 3.0051184524347
 ZP O 2.43855332694247
 BETAS O -29.5927245129389
 BETST O -28.2197565341473
 BETAP O -28.6510838038392
 BETPT O -29.3788277611853
 GSS O 13.7513493747977
 GSP O 15.6702873171911
 GPP O 13.5493629727874
 GP2 O 13.6493100865266
 HSP O 3.92518194452698
 ALP O 3.99736871651075
 FN11 O 0.225351341411567
 FN21 O 5.42534798656265
 FN31 O 0.954118064068487
 FN12 O 0.0610565726854795
 FN22 O 7.03512037084589
 FN32 O 1.62821149445402

A.6.3 Larger Reference Data Set

USS H -11.7455398219427
 UST H -11.7566961295116
 ZS H 1.10420716107291
 BETAS H -5.97060903338928
 BETST H -5.98855945971996
 GSS H 14.2178941393822
 ALP H 3.0749995892279
 FN11 H 0.101969726988834
 FN21 H 6.02488739720408
 FN31 H 1.16157455202179
 FN12 H 0.0635366295488991
 FN22 H 6.39430691138946
 FN32 H 1.98685847295657
 FN13 H -0.0372135142599904
 FN23 H 2.81280185340785
 FN33 H 1.67492606717665
 USS C -50.6772546249653
 UST C -53.346372148906
 UPP C -38.9374358167609
 UPT C -39.8194224297831
 ZS C 1.84187942578584
 ZP C 1.74027963296432
 BETAS C -15.4645071518657
 BETST C -15.2428937448584
 BETAP C -8.48197123205461
 BETPT C -8.27473658149251
 GSS C 13.0720109047644
 GSP C 11.6570469869631
 GPP C 10.9119341598752
 GP2 C 9.70213199789839
 HSP C 1.55795937222321

ALP C 2.77424440072142
 FN11 C 0.0745177376338087
 FN21 C 5.88105466834534
 FN31 C 1.02377818912525
 FN12 C 0.0117653282934468
 FN22 C 6.65142685159252
 FN32 C 1.68214060715311
 FN13 C 0.0385629315181582
 FN23 C 6.20989596013287
 FN33 C 1.60713243721309
 FN14 C -0.00268413908391832
 FN24 C 9.05896942165641
 FN34 C 2.736840043303
 USS N -74.0624351697579
 UST N -69.2013413541164
 UPP N -58.0807596986772
 UPT N -59.3905071266137
 ZS N 2.38234229120612
 ZP N 1.98237160677279
 BETAS N -20.3397760800363
 BETST N -21.6475922322623
 BETAP N -17.0243773744024
 BETPT N -16.3572150487542
 GSS N 13.0013033091873
 GSP N 13.0253001180892
 GPP N 13.860178568046
 GP2 N 12.0550691285332
 HSP N 4.92162347218628
 ALP N 2.93876140460604
 FN11 N 0.0580838151406299
 FN21 N 4.6564224249656

FN31 N 1.38964254936882
 FN12 N 0.0244779839917825
 FN22 N 4.70499157801326
 FN32 N 2.09432927660806
 FN13 N -0.022927362554042
 FN23 N 1.96097244980331
 FN33 N 1.86092776564239
 USS O -95.9863715847098
 UST O -94.1849027114831
 UPP O -78.9170923289229
 UPT O -77.0094963308609
 ZS O 3.14272824936785
 ZP O 2.63933267213423
 BETAS O -30.9290519451228
 BETST O -30.0640907863399
 BETAP O -29.010602808603
 BETPT O -29.640427924289
 GSS O 13.9164133709471
 GSP O 15.2075981375329
 GPP O 14.2311099488818
 GP2 O 12.7859489566756
 HSP O 3.99524017722306
 ALP O 4.24810394411796
 FN11 O 0.22911363380798
 FN21 O 5.24331855192749
 FN31 O 0.94384045196716
 FN12 O 0.0594349391079237
 FN22 O 7.22580295350434
 FN32 O 1.60609241358943

A.6 Optimized Parameters Sets

A.6.4 Including Scan Data

USS H -11.9696813090724	ALP C 2.82220666578781	FN31 N 1.40858589693063
UST H -12.4700953717735	FN11 C 0.0763612558550146	FN12 N 0.0241898932316653
ZS H 1.08906983567538	FN21 C 5.93970423731514	FN22 N 4.56214494995264
BETAS H -6.00415000063566	FN31 C 1.07891165496791	FN32 N 2.09643002549199
BETST H -5.84408028503892	FN12 C 0.0118502153357724	FN13 N -0.0231320949766002
GSS H 13.6609030045147	FN22 C 6.99833555009925	FN23 N 2.10673099772394
ALP H 2.9468960289169	FN32 C 1.65835372082058	FN33 N 1.93534824925988
FN11 H 0.100973216293667	FN13 C 0.0372373627815166	USS O -96.5073030640341
FN21 H 5.84591376872635	FN23 C 6.22017056324631	UST O -96.4281462150823
FN31 H 1.1626765679713	FN33 C 1.65287551746844	UPP O -77.5672244453997
FN12 H 0.0648290751565639	FN14 C -0.00271033104172054	UPT O -74.9384840234469
FN22 H 6.58241266868953	FN24 C 9.04077080980722	ZS O 3.15645168198723
FN32 H 1.9162948609246	FN34 C 2.73496972684357	ZP O 2.49332293388093
FN13 H -0.0367844418085548	USS N -72.81103035368	BETAS O -29.8989346168234
FN23 H 2.79022198390087	UST N -71.1546488884681	BETST O -29.8871637558571
FN33 H 1.6619282171057	UPP N -59.2259529423798	BETAP O -28.0799020977592
USS C -51.8451062883016	UPT N -57.3843076859987	BETPT O -29.4456474199692
UST C -51.8091701995452	ZS N 2.37328554024897	GSS O 14.6010465457969
UPP C -38.5579320970607	ZP N 1.99522703653399	GSP O 15.005104456149
UPT C -39.602736512014	BETAS N -20.66482129974	GPP O 14.1023736136049
ZS C 1.7877006945411	BETST N -21.720671863423	GP2 O 12.2366328787077
ZP C 1.75881250176459	BETAP N -17.1806391641532	HSP O 4.05002579202582
BETAS C -15.7131661523847	BETPT N -16.5846038075377	ALP O 4.11583461256711
BETST C -15.2451342189681	GSS N 13.2799520117246	FN11 O 0.231741160144058
BETAP C -8.27858312215247	GSP N 13.0726651321482	FN21 O 5.23313838386597
BETPT C -8.33770898863264	GPP N 13.9827247402404	FN31 O 0.906911015944829
GSS C 13.1185258225832	GP2 N 11.7021557509812	FN12 O 0.057690667406703
GSP C 11.2881583768392	HSP N 4.88383943397375	FN22 O 7.39174398191767
GPP C 10.9219304569039	ALP N 2.97626463514582	FN32 O 1.45816366122539
GP2 C 9.67246163466832	FN11 N 0.0594768066776018	
HSP C 1.54237679305604	FN21 N 4.61211579899985	

A.7 BFP Chromophore Scans

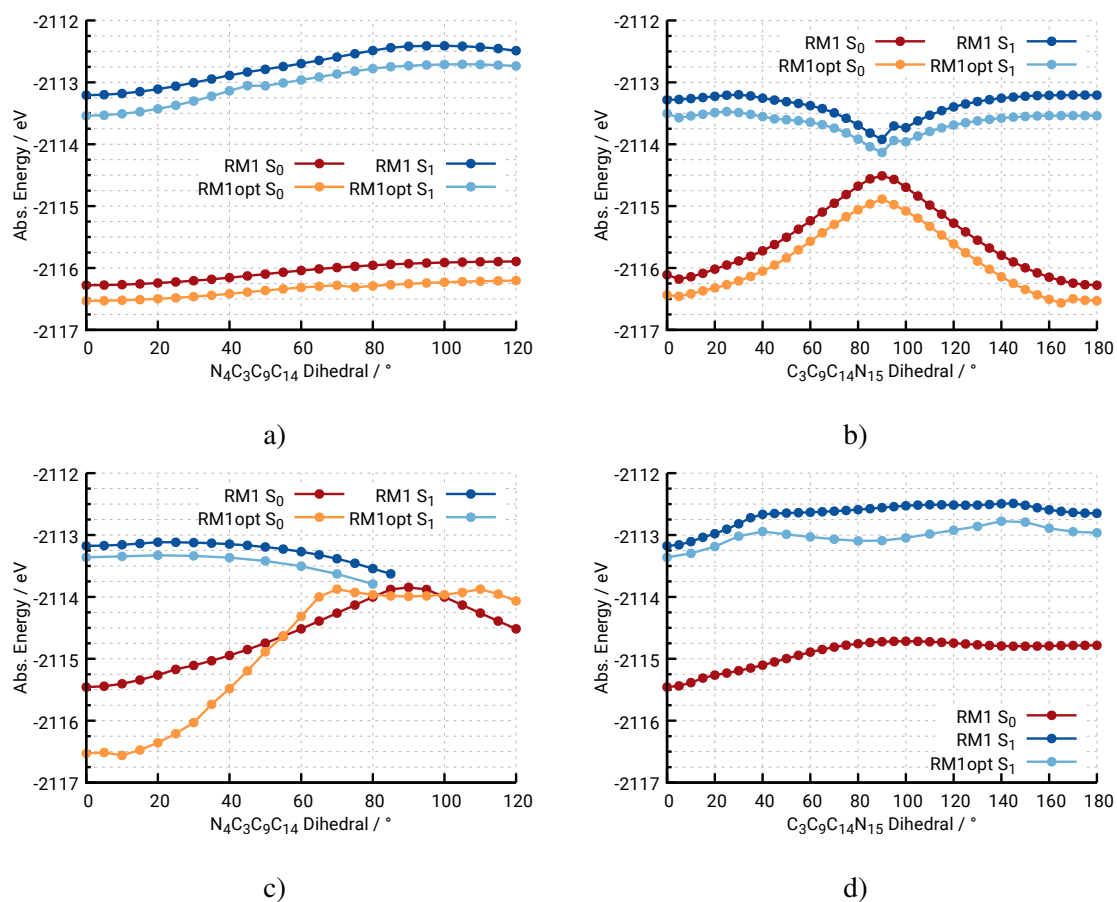


Figure A.5.: Comparison of the RM1 and RM1opt energies along a 1D PES scan of the $N_4C_3C_9C_{14}$ and $C_3C_9C_{14}N_{15}$ dihedrals of the **N** tautomer (a, b) and **T** tautomer (c, d), respectively. In contrast to the scans shown in Fig. 6.15, these scans are relaxed on the RM1 and RM1opt PES, respectively. The scan for the $C_3C_9C_{14}N_{15}$ dihedral of the **T** tautomer in d) is missing due to convergence issues.

A.8 Orbitals

A.8.1 HMQCA

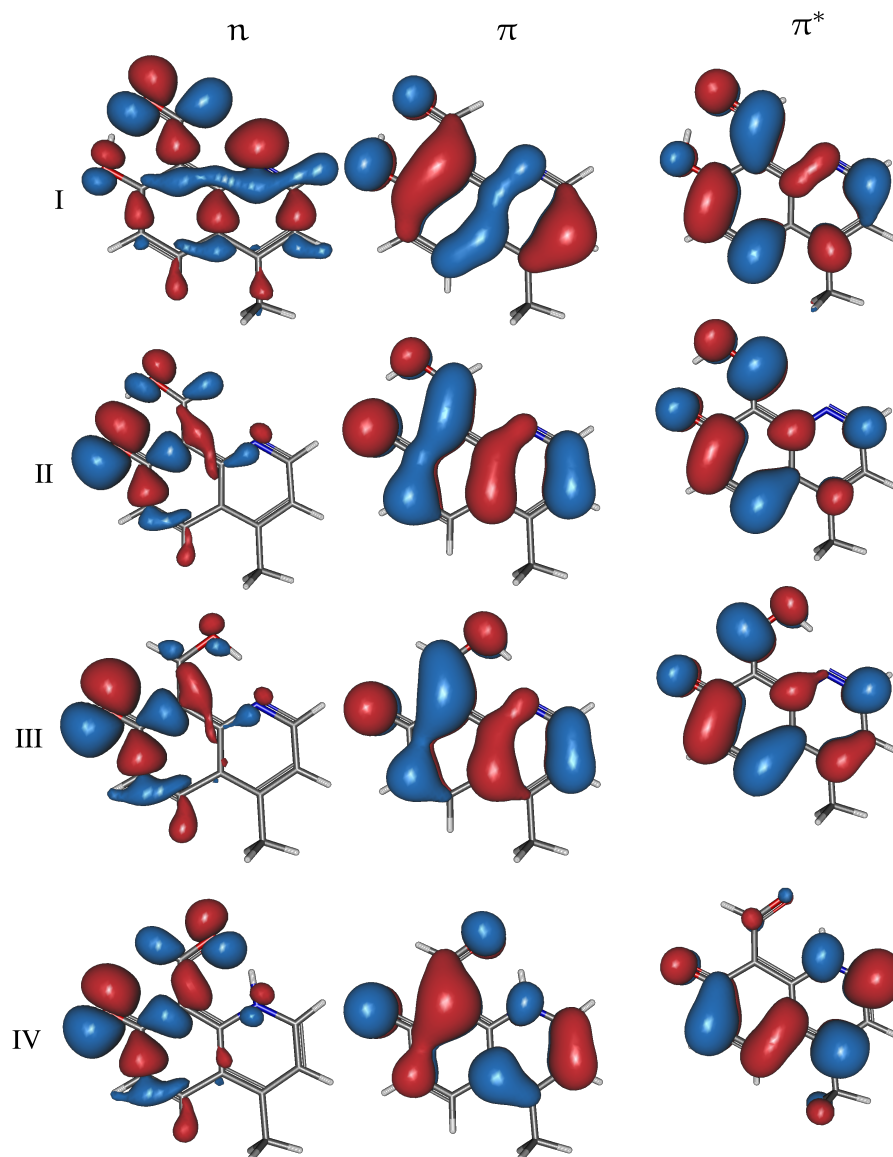


Figure A.6.: Representations of the n , π and π^* orbitals of **HMQCA**. These orbitals were obtained from a PBE0-D3BJ/def2-TZVPP calculations for S_0 (S_1 for **II-HMQCA**). These orbitals are plotted with an isovalue of 0.03.

A.8.2 MMIMQO

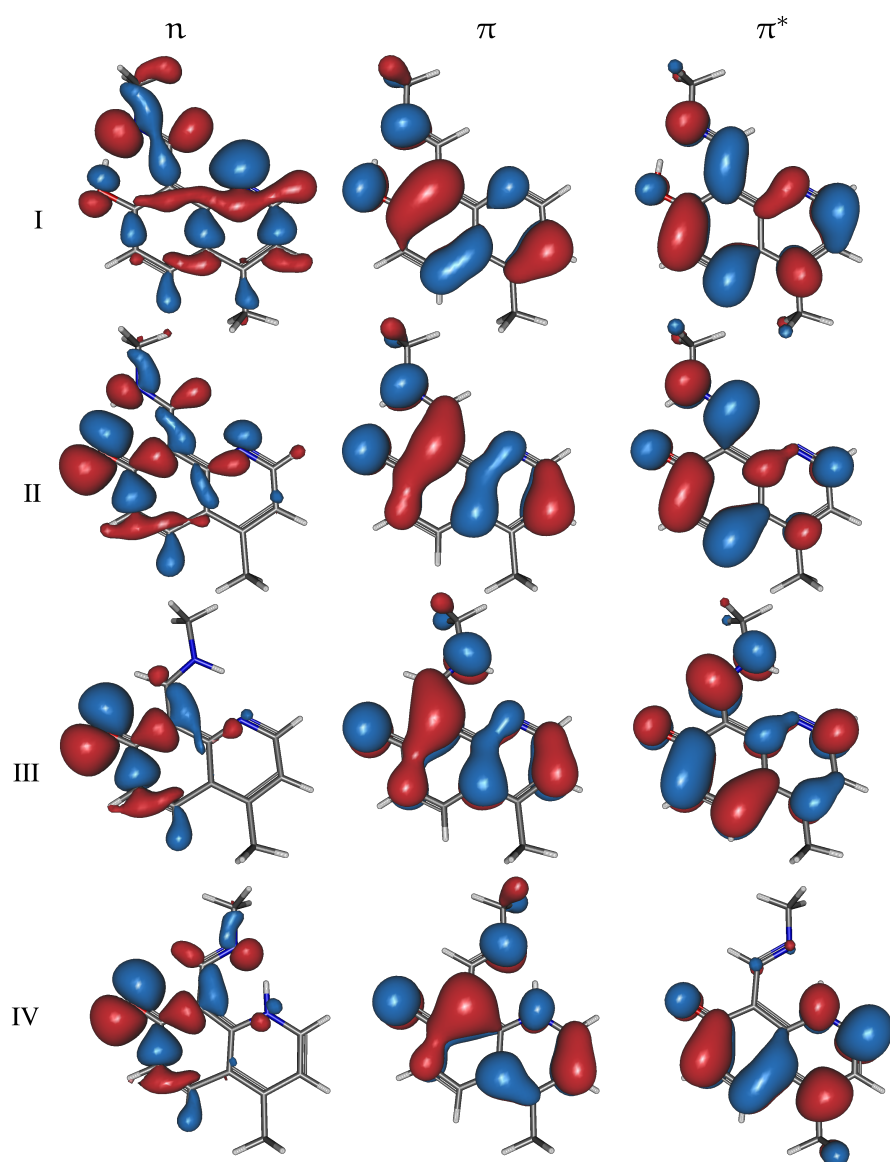


Figure A.7.: Representations of the n , π and π^* orbitals of MMIMQO. These orbitals were obtained from a PBE0-D3BJ/def2-TZVPP calculations for S_0 . These orbitals are plotted with an isovalue of 0.03.

A.8.3 MIMPO

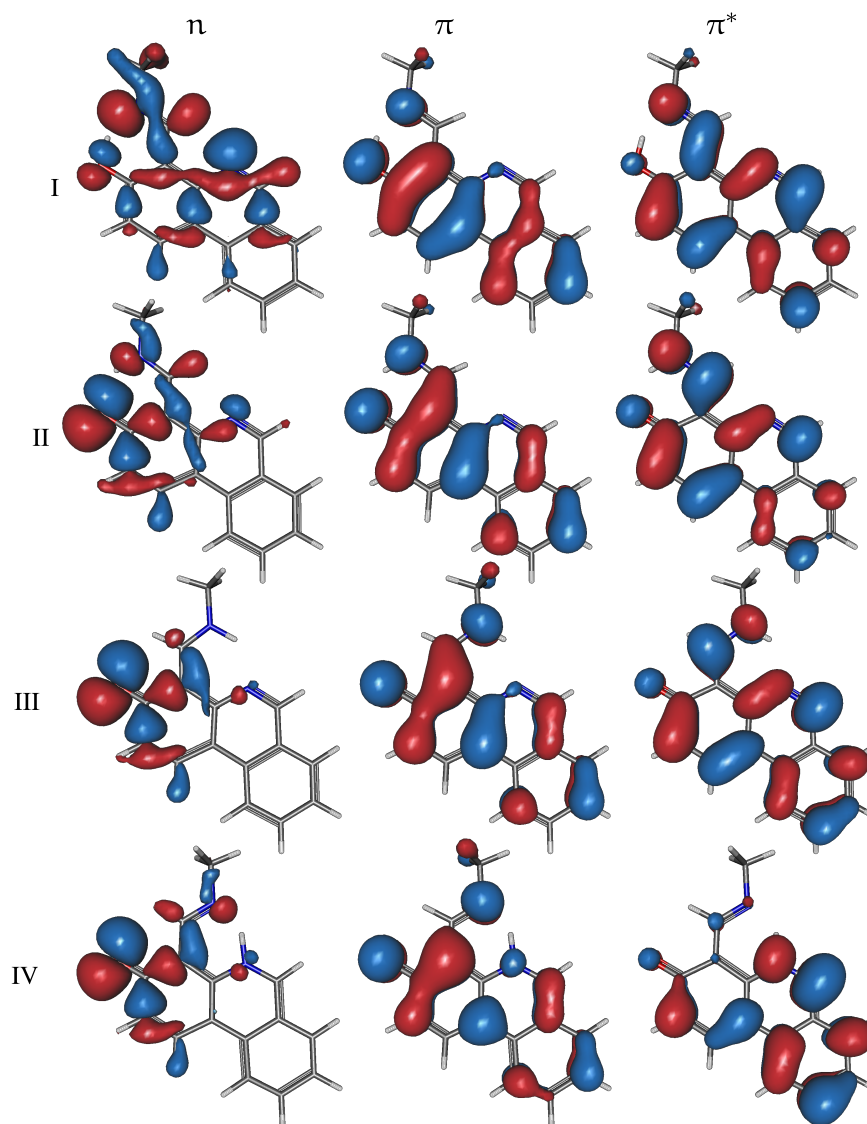


Figure A.8.: Representations of the n , π and π^* orbitals of MIMPO. These orbitals were obtained from a PBE0-D3BJ/def2-TZVPP calculations for S_0 . These orbitals are plotted with an isovalue of 0.03.

Declaration

I hereby declare that the work presented in this thesis was done by me, under the supervision of Prof. Dr. Bernd Hartke, with no other help than the referenced sources in the text.

This is my first dissertation and the work has never been used in any other dissertation attempts.

The dissertation complies to the good scientific practice rules as proposed by the German Research Foundation (DFG).

Tim Raeker

Kiel – April 12, 2018

Curriculum Vitae

Personal Data

DATE AND PLACE OF BIRTH: August 7, 1989 in Buchholz i.d.N.

NATIONALITY: German

ADDRESS: Hanssenstr. 30, 24106 Kiel, Germany

Education

- since 09/2013 Doctoral candidate at **University of Kiel**, Germany
Topic: "Full-Dimensional Photodynamics Simulations From Photoisomerizations to Excited-State Proton Transfer Reactions" with Prof. Dr. B. Hartke, Institute for Physical Chemistry
- 10/2011 – 08/2013 Master of Science (Chemistry) studies at **University of Kiel**, Germany
Master's thesis "Neutrale Wassercluster: Strukturen und Spektren" with Prof. Dr. B. Hartke, Institute for Physical Chemistry
- 10/2008 - 09/2011 Bachelor of Science (Chemistry) studies at **University of Kiel**, Germany
Bachelor's thesis "QM/MM-Simulationen molekularen Schaltens unter sterischer Behinderung durch die Umgebung" with Prof. Dr. B. Hartke, Institute for Physical Chemistry
- 08/2002 - 07/2008 School education at the "Gymnasium am Kattenberge" in Buchholz.
Major courses in chemistry and mathematics.

Awards

- 12/2011 Best Bachelor's thesis in Physical Chemistry

Teaching Experience

10/2017-02/2018	Teaching assistant in Exercises “Computational Chemistry”
10/2017	Teaching assistant in Exercises “Mathematics for Chemists (Pre-Course)”
04/2017-07/2017	Teaching assistant in Exercises “Mathematics for Chemists II”
10/2016-02/2017	Teaching assistant in Exercises “Computational Chemistry”
10/2016-12/2016	Advanced Practical Course in Theoretical Chemistry with Lasse Jannsen
10/2016	Teaching assistant in Exercises “Mathematics for Chemists (Pre-Course)”
06/2016-09/2016	BSc thesis supervision of Jennifer Müller
04/2016-07/2016	Teaching assistant in Lab course “Basic Lab Course in Physical Chemistry for Pharmacists”
10/2015-02/2016	Teaching assistant in Exercises “Computational Chemistry”
10/2015	Teaching assistant in Exercises “Mathematics for Chemists (Pre-Course)”
07/2015-10/2015	BSc thesis supervision of Dominik Behrens
04/2015-07/2015	Teaching assistant in Lab course “Basic Lab Course in Physical Chemistry for Biologists”
10/2014-02/2015	Teaching assistant in Exercises “Mathematics for Chemists I”
10/2014	Teaching assistant in Exercises “Mathematics for Chemists (Pre-Course)”
07/2014-11/2014	BSc thesis supervision of Björn Jansen
04/2014-07/2014	Teaching assistant in Lab course “Basic Lab Course in Physical Chemistry for Biologists”
02/2014-05/2014	BSc thesis supervision of Christopher Witt
10/2013-02/2014	Teaching assistant in Exercises “Computational Chemistry”
09/2013	Teaching assistant in Exercises “Mathematics for Chemists (Pre-Course)”
06/2013-09/2013	BSc thesis supervision (partial) of Julien Steffen
04/2013-07/2013	Student assistant in Exercises “Mathematics for Chemists II”
11/2012-01/2013	BSc thesis supervision of Jana Böckel
10/2012-02/2013	Student assistant in Exercises “Mathematics for Chemists I”
04/2012-07/2012	Student assistant in Exercises “Mathematics for Chemists II”
10/2011-02/2012	Student assistant in Exercises “Physical Chemistry II”
04/2011-09/2011	Student assistant in Exercises “Physical Chemistry I”

Scientific Contributions (Posters and Talks)

03/2018	Poster at the Winterschool of the Integrated Graduate Center of the CRC677, Kiel, Germany
12/2017	Talk at the GDCh Weihnachtskolloquium, Kiel, Germany
09/2017	Poster at the Summerschool of the Integrated Graduate Center of the CRC677, Hohwacht, Germany
08/2017	Poster at the Congress of the World Association of Theoretical and Computational Chemists, Munich, Germany
08/2016	Poster at the Summerschool of the Integrated Graduate Center of the CRC677, Ahrensburg, Germany
03/2016	Poster at the JCF Frühjahrssymposium, Kiel, Germany
11/2015	Poster at the PHENICS Winterschool, Corsica, France
10/2015	Poster at the Zukunftskongress, Kiel, Germany
09/2015	Poster at the Symposium on Theoretical Chemistry, Potsdam, Germany
08/2015	Poster at the Summerschool of the Integrated Graduate Center of the CRC677, Schleswig, Germany
09/2014	Poster at the Symposium on Theoretical Chemistry, Vienna, Austria
08/2014	Talk at the Summerschool of the Integrated Graduate Center of the CRC677, Mölln, Germany
09/2013	Poster at the International Symposium on Photochromism, Berlin, Germany

List of Publications

1. T. Raeker, B. Jansen, D. Behrens, B. Hartke, *J. Comp. Chem.* **2018**, published online, DOI: 10.1002/jcc.25212.
2. S. Wang, S. Schatz, M. C. Stuhldreier, H. Böhnke, J. Wiese, C. Schröder, T. Raeker, B. Hartke, J. K. Keppler, K. Schwarz, F. Renth, F. Temps, *Phys. Chem. Chem. Phys.* **2017**, 45, 30683.
3. T. Raeker, B. Hartke, *J. Phys. Chem. A* **2017**, 32, 5967.
4. T. Raeker, B. Hartke, *ScienceOpen Research*, DOI: 10.14293/S2199-1006.1.SOR-CHEM.ARDTLN.v1
5. T. Raeker, N. O. Carstensen, B. Hartke, *J. Phys. Chem. A* **2012**, 46, 112411.

Acknowledgments

First and foremost I have to thank my supervisor Prof. Bernd Hartke for letting me work on a topic that allowed me to try out many different aspects of computational chemistry. The freedom that I was allowed to choose my own tools is greatly appreciated. Of course it was unavoidable that it would sometimes lead to dead ends, but thank you for always providing me scientific support.

I would also like to thank Prof. Friedrich Temps for being my co-supervisor within the integrated research training group of the CRC677 and accepting to be my second referee.

To the members of his workgroup, especially Hendrik Böhnke, thank you for providing me general information about the experimental side of photochemistry and for keeping me up-to-date on what you guys are working on.

Also to my colleagues, thank you for the great time.

Furthermore I would like to thank my former mentors/supervisors and colleagues throughout the years: Dr. Niss Ole Carstensen, Dr. Jan Boyke Schönborn and Sascha “Sushi” Frick.

Thanks to all proof-readers for lowering the amount of iterations needed to converge to a final version of this manuscript.

I would also like to thank my friends who accompanied me on this journey and especially for your support in the last months.

My parents, Holger and Sonja, and my brother, Jan: Thank you for your invaluable support and encouragement in every way possible throughout all my life.