universität wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation / Title of the Doctoral Thesis

"A quantum study on nucleobase dimerization leading to DNA photodamage: A singlet or triplet reaction path?"

verfasst von / submitted by
Clemens Rauer, MSc

angestrebter akademischer Grad /
in partial fulfillment of the requirements for the degree of
Doktor der Naturwissenschaften (Dr. rer. nat.)

Wien, 2017

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 796 605 419
Dissertationsgebiet lt. Studienblatt / field of study as it appears on the student record sheet:	Chemie
Betreut von / Supervisor:	Prof. Leticia González

ABSTRACT

The main goal of this thesis was the investigation of the thymine dimerization reaction using nonadiabatic molecular dynamics methods. For this [2+2] cycloreaction, both the direct pathway and the photosensitized triplet pathway were studied. For the direct pathway, which models the dimerization after excitation to the singlet manifold, static calculations were employed to confirm that the reactive S_1 state shows a barrierless reaction profile. The character of the state could be determined as mainly singly excited $\pi\pi^*$. Subsequent surface hopping dynamics simulations were able to rule out any intersystem crossing occurring in the pathway studied, showing that dimerization occurs solely in the singlet state manifold. Moreover, it was found that a combination of conformational and electronic effects lead to dimerization, showing that a correct treatment of both is necessary. The second thymine dimerization pathway studied was the photosensitized pathway. In it, the reaction occurs on the triplet manifold, after triplet-triplet energy transfer from a photosensitizer. Static QM/MM calculations show that after energy transfer to a locally excited state, a small energetic barrier leads the system to the T_1/S_0 crossing point. At this point the T_1 state is determined to be a Frenkel excitonic biradical state. We used surface hopping dynamics simulations to sample the T_1/S_0 degeneracy region and by manually placing trajectories to the ground state we could investigate which conformational parameters induce dimerization. This thesis also included the development of the generalized ab initio multiple spawning (GAIMS) method, which is based on non-adiabatic molecular dynamics and allows for the treatment of intersystem crossing events within the multiple spawning framework. The new implementation was first tested on a one-dimensional model system and the results were compared to the ones obtained by exact and surface hopping simulations. Then, the dynamics of the H_2CS molecule were performed and analyzed, showing that the implementation works for molecular systems.

ZUSAMMENFASSUNG

Das Hauptziel dieser Arbeit war die Untersuchung der Zyklobutan Thymindimerisierungsreaktion unter der Verwendung von nicht-adiabatischen Molekulardynamikmethoden. Zwei Reaktionspfade, der direkte und der photosensibilisierte Pfad wurden für diese [2+2] Zykloaddition untersucht. Im Falle des direkten Pfades, welcher die Dimerisierung nach der direkten Anregung in einen Singulettzustand modelliert, wurde mit Hilfe statischer Rechnungen ermittelt dass der reaktive S_1 Zustand, welcher vorwiegend einen $\pi\pi^*$ Charakter besitzt, ein barrierefreies Reaktionsprofil aufweist. Nachfolgende surface hopping Simulationen konnten zeigen dass die Reaktion nur auf den Singulettzuständen erfolgt und keine Interkombination in dem untersuchten Reaktionspfad vorkommt. Zusätzlich konnte demonstriert werden dass eine akkurate Beschreibung sowohl von konformationellen als auch elektronischen Effekten notwendig ist, um den Dimerisierungsprozess korrekt zu modellieren. Der zweite in dieser Arbeit untersuchte Thymindimerisierungspfad ist der photosensibilisierte Mechanismus, welcher über Triplettzustände abläuft. Statische QM/MM Rechnungen konnten zeigen, dass das System nach anfänglichem Energietransfer zu einem lokal angeregten Zustand und Überwindung einer kleinen Energiebarriere einen T_1/S_0 Kreuzungspunkt erreicht. An diesem Punkt ist der T_1 Zustand ein Frenkel-exzitonischer Biradikalzustand. Mittels surface hopping Simulationen konnten wir die T_1/S_0 Degenerationsregion erkunden. Durch manuelle Platzierung von Trajektorien in den Grundzustand war es möglich, zu analysieren welche konformationellen Parameter die Dimerisierung begünstigen. Diese Arbeit beinhaltet auch die Entwicklung der generalized ab initio multiple spawning (GAIMS) Methode, welche auf nicht-adiabatischer Molekulardynamik basiert und es erlaubt, Interkombinationsereignisse mit dem multiple spawning Formalismus simulieren. Die neue Implementierung wurde zuerst für ein eindimensionales Modellsystem getestet, und mit den Resultaten von exakter Quantendynamik und surface hopping Simulationen verglichen. Um zu zeigen dass die Methode auch für molekulare Systeme angewandt werden kann, wurde im Anschluss eine Dynamiksimulation des H_2CS Moleküls durchgeführt.

CONTENTS

1 INTRODUCTION 1

2 THEORY 3
2.1 Basics of Quantum Chemistry . . . 3
2.2 Wavefunction Analysis . . . 3
2.3 Excited State Dynamics . . . 4
2.4 Inclusion of the Environment . . . 8

3 GAIMS 9
3.1 Electronic Representation . . . 9
3.2 Propagation . . . 9
3.3 Spawning . . . 10
3.4 Application to Test Systems . . . 11

4 THYMINE PYRIMIDINE DIMER 13
4.1 Direct Dimerization . . . 13
4.2 Photosensitized Dimerization . . . 19

5 CONCLUSION & OUTLOOK 23

A APPENDIX: REPRINTED PUBLICATIONS 25
A.1 J. Chem. Phys. Comm. 144, 101102 (2016) . . . 26
A.2 J. Am. Chem. Soc. 138, 15911-15916 (2016) . . . 33
A.3 Monatsh. Chem. (2017) . . . 40

BIBLIOGRAPHY 51

ACKNOWLEDGMENTS 57

CURRICULUM VITAE 58

UV	ultraviolet
T<>T	cyclobutane thymine dimer
IC	internal conversion
ISC	intersystem crossing
SOC	spin-orbit coupling
NAMD	non-adiabatic molecular dynamics
SHARC	surface hopping including arbitrary couplings
GAIMS	generalized ab initio multiple spawning
TDSE	time dependent Schrödinger equation
CASSCF	Complete active space self consistent field
SA-CASSCF	State averaged CASSCF
CASPT2	Complete active space perturbation theory second order
MS-CASPT2	Multi-state CASPT2
TDM	$1e^-$ transition density matrix
CT	charge transfer
DL	delocalization length
MD	molecular dynamics
AIMD	ab initio molecular dynamics
QM/MM	quantum mechanics/molecular mechanics
QM	Quantum Mechanical
MM	Molecular Mechanical
NAC	non-adiabatic coupling
FMS	full multiple spawning
TTET	triplet-triplet energy transfer
DE	doubly excited
SE	singly excited
FC	Franck-Condon
^{3}**L**	local excited triplet state
^{3}**BR**	biradical triplet state

1 INTRODUCTION

DNA, the carrier of our genetic code, is a very important and complex biomolecule in the human body. It is built up of the four nucleobases adenine, cytosine, guanine and thymine which are linked to a sugar and phosphate backbone. The correct functionality of DNA is vital to survival and any damage to it can be severely harmful, leading to health risks like apoptosis or cancer[1]. One cause of DNA damage is ultraviolet (UV) light[2], with the chromophores, which absorb the UV photons, being the previously mentioned nucleobases[3]. Upon UV irradiation and subsequent excitation to the electronically excited states, DNA generally relaxes quickly back to the ground state. These ultrafast relaxation pathways are responsible for the remarkable photostability of DNA[4,5]. If this ground state relaxation is hindered, however, a multitude of photolesions can occur[1]. In order to find preventive treatments and therapies for DNA photodamage, a detailed understanding of these photodamage processes is needed. One of the most prominent lesions caused by UV radiation is the cyclobutane thymine dimer (T<>T)[3,6], which is formed through a [2+2] cycloreaction of two geminal thymines, see Figure 1.

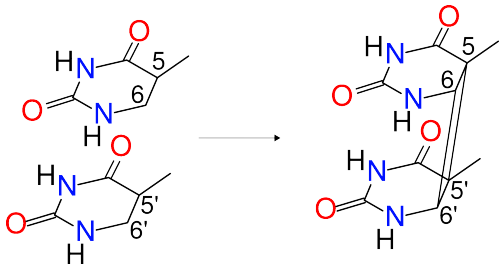


Figure 1: Schematic picture of the Cyclobutane Thymine Dimer molecule.

In general, after absorption of a photon by any molecular system, a plethora of relaxation pathways open up, as shown in Figure 2. According to the selection rules of quantum mechanics, the molecular system will get promoted into an electronically excited singlet state, from where it can change to other singlet states in a process called internal conversion (IC). Upon relaxation to the ground state the system can either go back to the reactant minimum or, in a process labeled non-adiabatic photochemistry, it can move to a different minimum, forming a photoproduct. In addition to this relaxation path, the system can also eject a photon and relax through a process called fluorescence. When including relativistic corrections, another relaxation pathway opens up, called intersystem crossing (ISC), which allows the population transfer from a singlet to a triplet state, is mediated by spin-orbit couplings (SOCs). In the triplet manifold the system can again change between triplet states through IC, move back to a singlet state via ISC, or emit a photon and relax photochemically through a process called phosphorescence.

The role of triplet states on T<>T formation has been investigated, both experimentally and theoretically since the 1960s[7–9]. In the main part of this thesis two major pathways of dimerization were investigated. First, the direct (T<>T) formation through UV excitation is studied. This process has been shown experimentally to occur on a timescale of under 1ps, presumably in the singlet manifold[10]. Although

experimental and theoretical studies have discussed possible competing triplet pathways, their role on dimerization is still under dicussion in the literature[11–15]. The second process investigated is the photosensitized triplet pathway, where, after triplet-triplet energy transfer (TTET) through a photosensitizer, the dimer is formed via the triplet manifold[16].

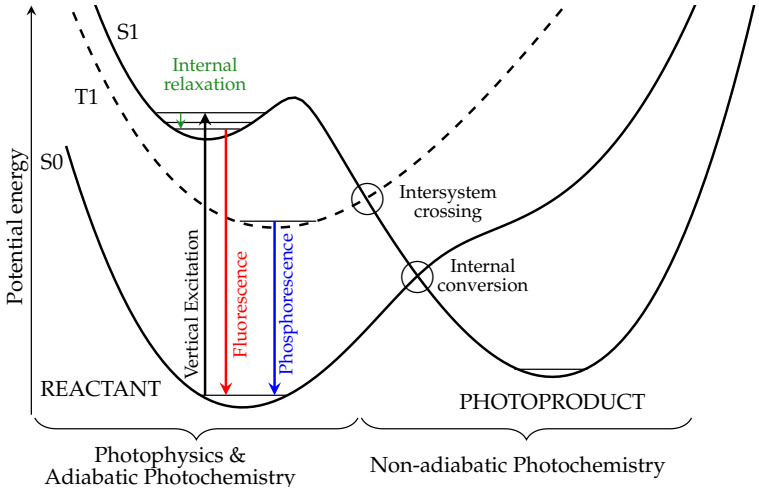


Figure 2: Schematic diagram of the possible relaxation pathways of a molecular system after excitation to an excited state.

In order to properly investigate the role of triplet states on the excited state dynamics of a molecular system, we need a non-adiabatic molecular dynamics (NAMD) method which can account for ISC events. Recently, the group of Prof. González has developed the Surface Hopping including ARbitrary Couplings (SHARC) method, which can run surface hopping dynamics using states of any multiplicity[17,18]. However, surface hopping has several drawbacks, one of them being the fact that it cannot account for nuclear coherences occurring in the dynamics[19]. A different NAMD approach, full multiple spawning, can in part account for the lost coherences but was not able to include ISC events in the dynamics. Therefore, as part of this thesis, the generalized ab-initio multiple spawning (GAIMS) method was developed which allows for multiple spawning dynamics including triplet states.

The rest of this thesis is organized as follows: In chapter 2, several theoretical concepts important to this thesis will be presented. Then, in chapter 3 the GAIMS method will be presented in detail and its application to test systems will be presented. Chapter 4 will cover the results on the formation of T<>T, both via the direct and the photosensitized pathway. Finally, in chapter 5, a conclusion and summary of this thesis is given.

2 THEORY

In this chapter the theoretical background relevant to this thesis will be described. First, in section 2.1, some basic concepts of quantum chemistry will be covered. Section 2.2 describes several wavefunction analysis concepts relevant to this work. Finally, in section 2.3, excited state dynamics and the two non-adiabatic molecular dynamics methods used in this thesis, surface hopping (2.3.2) and multiple spawning (2.3.3) will be presented.

2.1 BASICS OF QUANTUM CHEMISTRY

In computational chemistry, the non-relativistic evolution of a molecular system is described by the time-dependent Schrödinger equation (TDSE)[20]:

$$i\hbar\frac{\partial}{\partial t}\left|\Psi(\mathbf{R},\mathbf{r},t)\right\rangle = \hat{H}(\mathbf{R},\mathbf{r},t)\left|\Psi(\mathbf{R},\mathbf{r},t)\right\rangle . \tag{2.1}$$

The TDSE states that the time evolution of a wave function, Ψ, is dependent on its Hamiltonian operator $\hat{H}$. Both Ψ and $\hat{H}$ depend on the nuclear degrees of freedom ($\mathbf{R}$), the electronic degrees of freedom ($\mathbf{r}$) and the time t. Using the Born-Oppenheimer approximation[21], we can separate the TDSE into an electronic and a nuclear Schrödinger equation. Several methods are available for solving, or rather approximating, the solution of the electronic Schrödinger equation. As in this thesis we are interested in electronically excited states and their properties, we need a method which can properly describe all of these processes. One method which can do this is the complete active space self consistent field (CASSCF) method[22], where a subset of molecular orbitals is chosen between which all excitations are considered. The selection of these active orbitals is no trivial task and depends on the chemical or physical processes one wants to simulate. In order to also obtain electron correlation in the orbital space outside of this active space, the complete active space perturbation theory 2nd order (CASPT2) method was employed[23].

2.2 WAVEFUNCTION ANALYSIS

One very important aspect when dealing with electronically excited states is their correct analysis and the evaluation of their properties. In this section, several concepts of wavefunction analysis important to this thesis will be shown, as included in the wavefunction analysis program TheoDORE. For more information we recommend the TheoDORE program documentation and publications[24,25].

The $1e^-$-Transition Density Matrix (TDM) between the ground state to an electronically excited state I is defined as

$$D^{0I}_{\mu\nu} = \left\langle \phi^0 \right| \hat{a}^\dagger_\mu \hat{a}_\nu \left| \phi^I \right\rangle . \tag{2.2}$$

Here, $\hat{a}^\dagger$ and $\hat{a}$ are the creation and annihilation operators, which in this case remove an electron from orbital ν and add one to orbital μ. Using the TDM, we can measure

several important properties of molecular systems. In order to obtain a measure of the single excitation character we employ

$$\Omega = \sum_{\mu\nu} (\mathbf{D}^{0I}\mathbf{S})_{\mu\nu}(\mathbf{S}\mathbf{D}^{0I})_{\mu\nu}, \tag{2.3}$$

where S is the atomic orbital overlap matrix. Therefore, if Ω number is 1, the state in question is a singly excited state, while if it is 0, it has no single excitation character and is therefore doubly or higher excited. Another important property of excited states is their charge transfer (CT) character. Taking as an example the two thymine monomers investigated in this thesis the excited states can either be charge transfer states, with the electron and hole on different monomers or a locally excited state, with electron and hole on the same monomer. We can express the CT character of a given state I using the TDM. First we have to define the CT numbers as

$$\Omega_{\mathrm{AB}} = \frac{1}{2} \sum_{\mu\in A} \sum_{\nu\in B} (\mathbf{D}^{0I}\mathbf{S})_{\mu\nu}(\mathbf{S}\mathbf{D}^{0I})_{\mu\nu} + D^{0I}_{\mu\nu}(\mathbf{S}\mathbf{D}^{0I}\mathbf{S})_{\mu\nu}, \tag{2.4}$$

which enables us to decompose the excitation in the contribution located on monomers A or B. The total CT contribution can then be calculated as a sum over all off-diagonal elements:

$$\mathrm{CT} = \Omega^{-1} \sum_{B\neq A} \Omega_{AB}. \tag{2.5}$$

The delocalization length (DL) of the electronic state I can also be visualized using Equation 2.4. The DL indicates how delocalized the excitation among the two fragments A and B is. Using the definition of Ω introduced in Equation 2.3, we can express DL as

$$\mathrm{DL} = \frac{\Omega^2}{\sum_A \left(\sum_B \frac{\Omega_{AB}+\Omega_{BA}}{2} \right)^2} \tag{2.6}$$

There is one more property used in this thesis which is not derived from the TDM, which is the number of unpaired electrons (n_u). Using a natural orbital basis, n_u can easily be assessed using the natural orbital occupation numbers n_i. In a HF wavefunction they would either be 0 or 2, i.e. unoccupied or occupied. Therefore, deviation from that can be seen as electron correlation. Here, n_u will be calculated using the formulation proposed by Head-Gordon:[26]

$$n_u = \sum_i n_i^2 (2 - n_i)^2. \tag{2.7}$$

2.3 EXCITED STATE DYNAMICS

In the chapters above we discussed the electronic Schrödinger equation and the correct treatment of electronic states. If we want to study the time-dependent behavior of a molecular system, however, we need to solve the nuclear Schrödinger equation. For time-independent Hamiltonian this can generally be done using propagation in discrete time steps Δt

$$\Psi(\mathbf{x}, t + \Delta t) = \exp\left(-\frac{i}{\hbar} \hat{H} \Delta t \right) \Psi(\mathbf{x}, t). \tag{2.8}$$

Direct solution of Equation 2.8, which is generally done in quantum dynamics, is a non-local problem. As such, the calculation requires the whole $3N - 6$ dimensional potential energy surface to be known in advance, with N being the number of atoms. As the amount of points to be calculated scales exponentially with the number of degrees of freedom, these calculations are only feasible for very small molecules or, if one wants to simulate larger molecules, a subset of degrees of freedom has to be chosen. In order to describe the dynamics of larger systems, one needs to circumvent this requirement of non-locality. One option is to employ on-the-fly methods which only need local information of the potential which can be calculated as needed on the fly, as the name suggests.

2.3.1 *Molecular Dynamics*

In molecular dynamics (MD) simulations, the nuclei are propagated classically according to Newton's equation of motion[27]

$$M_\alpha \frac{\mathrm{d}^2\mathbf{R}_\alpha(t)}{\mathrm{d}t^2} = \mathbf{F}_\alpha = -\nabla_\alpha E_e(\mathbf{R}(\mathbf{t})). \tag{2.9}$$

The Force $\mathbf{F}_\alpha$ acting on nucleus α only depends on the positions $\mathbf{R}$ of all nuclei. Therefore, one only needs to calculate the energies and gradients of the current positions for each timestep. As mentioned above, this approach is considered on-the-fly as we only need the local information of the current geometry for the dynamics and not the whole potential as in equation 2.8. There exist several ways in the literature to integrate equation 2.9 and obtain the equation for the temporal evolution of $\mathbf{R}$, which is called a trajectory, with the most widely used one being the Velocity-Verlet algorithm[28].

With the nuclei evolving classically, one needs to define the potentials on which they run, i.e. where the forces and energies are obtained from. One very popular option is to use parametrized force fields, a method that is normally called classical MD, which has the advantage of allowing one to study systems of many thousands of atoms on a ns time scale. However, classical MD lacks many quantum effects and cannot describe processes such as bond formation and bond breaking. An alternative which is able to simulate these processes is to run electronic structure calculations, which results in so called ab initio molecular dynamics (AIMD). Using AIMD, however, drastically reduces the size of the system one can simulate.

One major drawback of MD is that it runs only on one adiabatic state. In full quantum dynamics, a wavepacket can spread onto multiple states at the same time. In order to simulate this process, where a superposition of states are needed, we need enhanced methods. Two of those methods will be presented here, namely surface hopping and multiple spawning.

2.3.2 *Surface Hopping*

In surface hopping[29–31] a trajectory is propagated on a single state potential. At any time step it is allowed to change the surface it is running on in a probabilistic manner, based on the non-adiabatic couplings (NACs). By running multiple trajectories, a so-called ensemble, the behavior of a quantum wavepacket can be approximately simulated. Figure 3a shows a schematic description of the surface hopping process. During the dynamics, each electronic state has a complex coefficient c_I, and the

electronic wavefunction is then defined as the linear combination of all the electronic state wavefunctions

$$|\Phi\rangle = \sum_I c_I |\Phi_I\rangle . \tag{2.10}$$

If we insert equation 2.10 into the TDSE and left-multiply with $\langle\Phi_J|$ we obtain an equation for the time-evolution of the coefficients

$$\frac{\partial}{\partial t} c_J = -\sum_I c_I \left[iH_{JI} + K_{JI} \right] \tag{2.11}$$

Here, H_{JI} is a matrix element of the electronic Hamiltonian $\hat{\mathbf{H}}^{\text{el}}$, while K_{JI} is the NAC element between states J and I, $\left\langle \Phi_J \middle| \partial/\partial t \middle| \Phi_I \right\rangle$.

If two states get close in energy it is possible for the trajectory to hop and change the state on which it is running on. The hopping probability to change from state J to state I can be expressed as

$$P_{J\to I} = \frac{2\Delta T}{c_J^* c_J} \operatorname{Re}\left(c_J^* c_I \left[iH_{JI} + K_{JI} \right] \right) . \tag{2.12}$$

The standard surface hopping algorithm presented above only allows for population transfer of states of the same multiplicity, based on their NACs. In order to study ISC processes, an extension to the standard surface hopping scheme has been developed in the group of Prof. González, called SHARC (Surface hopping including arbitrary couplings)[17,18]. Including SOCs, the Hamiltonian of the system has the following form:

$$\hat{\mathbf{H}}^{\text{sd}} = \hat{\mathbf{H}}^{\text{el}} + \hat{\mathbf{H}}^{\text{SOC}} \tag{2.13}$$

The resulting Hamiltonian, which we call spin-diabatic (sd) here, now contains off-diagonal elements, which means that it is not adiabatic anymore. As has been shown, this can lead to problems in surface hopping[32]. Therefore, it is convenient to re-adiabatize the Hamiltonian which can be done through unitary transformation

$$\hat{\mathbf{H}}^{\text{diag}} = \mathbf{U}^\dagger \hat{\mathbf{H}}^{\text{sd}} \mathbf{U}. \tag{2.14}$$

Within SHARC, the propagation of the coefficients in this new basis is as follows:

$$\frac{\partial}{\partial t} \mathbf{c}^{\text{diag}} = -\mathbf{c}^{\text{diag}} \left[i\mathbf{U}^\dagger \hat{\mathbf{H}}^{\text{sd}} \mathbf{U} + \mathbf{U}^\dagger \hat{\mathbf{K}}^{\text{sd}} \mathbf{U} + \mathbf{U}^\dagger \frac{\partial}{\partial t} \mathbf{U} \right] \tag{2.15}$$

Equation 2.15 is integrated numerically for a small time step, i.e.

$$\mathbf{c}^{\text{diag}}(t+\Delta t) = \mathbf{U}^\dagger(t+\Delta t)\mathbf{A}^{\text{sd}}(t+\Delta t)\mathbf{U}(t)\mathbf{c}^{\text{diag}}(t) = \mathbf{A}^{\text{diag}}(t+\Delta t)\mathbf{c}^{\text{diag}}(t) \tag{2.16}$$

$\mathbf{A}^{\text{diag}}$ and $\mathbf{A}^{\text{sd}}$ are the propagator matrices in the diagonal and the spin-diabatic basis respectively. The spin-diabatic propagator can be calculated using linear interpolation with k sub-steps:

$$\mathbf{A}^{\text{sd}}(t+\Delta t) = \prod_{l=1}^{k} \exp\left\{ -\left(i\mathbf{H}_o^{\text{sd}} + \mathbf{K}_o^{\text{sd}} \right) \Delta t/k \right\} \tag{2.17}$$

Using this new ansatz the hopping probabilities can be rewritten as

$$P_{\beta\to\alpha} = \left(1 - \frac{|\mathbf{c}_\beta(t+\Delta t)|^2}{|\mathbf{c}_\beta(t)|^2}\right) \frac{\Re\left[\mathbf{c}_\alpha(t+\Delta t)\mathbf{A}^*_{\alpha\beta}\mathbf{c}^*_\beta(t)\right]}{|\mathbf{c}_\beta(t)|^2 - \Re\left[\mathbf{c}_\beta(t+\Delta t)\mathbf{A}^*_{\beta\beta}\mathbf{c}^*_\beta(t)\right]} \tag{2.18}$$

Note that we omitted the diag indices for both **c** and **P** here.

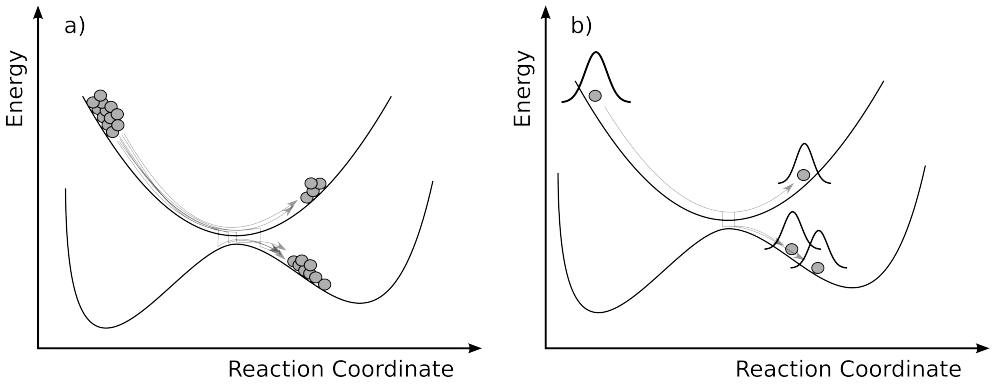


Figure 3: Schematic description of the non-adiabatic molecular dynamics methods used in this paper, namely Surface Hopping (a) and Multiple Spawning (b).

2.3.3 *Full Multiple Spawning*

Full multiple spawning (FMS) is a different approach to AIMD than surface hopping. The total wavefunction used is expanded as the sum of Born-Huang states[33]:

$$\Psi = \sum_\alpha \Phi_I \Omega_I. \tag{2.19}$$

Here, Ω_I is the nuclear wavefunction, while Φ_I is the electronic wavefunction. This approach is different to surface hopping, which uses classical trajectories on only one single electronic state. In FMS, the nuclear wavefunction Ω is built up of a sum of frozen Gaussian basis functions χ with their complex coefficients $C(t)$[34,35]:

$$\Omega_I = \sum_k^{N_I(t)} C_k^I \chi_a^I, \tag{2.20}$$

The index k runs over all basis functions in the electronic state I.

By inserting equation 2.19 into the TDSE one obtains the equation for the time evolution of the coefficients $C_k^I(t)$:

$$\mathbf{SC} = -i(\mathbf{H} - i\dot{\mathbf{S}})\mathbf{C}, \tag{2.21}$$

with **H** being the Hamiltonian matrix and **S** and $\dot{\mathbf{S}}$ the overlap matrix and its time derivative. Note that $\hat{\mathbf{H}}$ here contains both the electronic and the nuclear Hamiltonian.

$$S^{IJ}_{kk'} = \langle\chi^I_a|\,|\chi^J_b\rangle\,\delta_{ab} \quad \dot{S}^{IJ}_{kk'} = \langle\chi^I_k|\,|\frac{\partial}{\partial t}\chi^J_{k'}\rangle\,\delta_{kk'} \quad H^{IJ}_{kk'} = \left\langle\chi^I_k\middle|\hat{\mathbf{H}}\middle|\chi^J_{k'}\right\rangle. \tag{2.22}$$

To allow for population transfer between electronic states, basis functions need to be present on both states. In FMS this is achieved through a process called

spawning[35,36]. For each basis function in the ensemble all electronic properties are evaluated at each time step. If for one of them the NACs exceed a predefined threshold value a new basis function is spawned on the new state. Then, the propagation continues with the enlarged set of basis functions. It is important to notice that the spawning process itself does not include any population transfer, it merely enables the possibility for it. All occurring population transfer is based only on the evolution of the coefficients according to Equation 2.21. Figure 3b shows a schematic description of multiple spawning. The initial basis function on the upper state can spawn to the lower state as the two states get close in energy.

In order to use the multiple spawning scheme on-the-fly for molecular systems, two approximations are generally employed to FMS. In the saddle point approximation, the Hamiltonian elements $H^{\alpha\beta}_{kk'}$ in equation 2.21 are approximated by a zeroth order Taylor series

$$\left\langle \chi^I_k \right| \langle \Phi^I_k | \, \hat{\mathbf{H}} \, | \Phi^J_{k'} \rangle \, | \chi^J_{k'} \rangle = \langle \chi^I_k | \, \hat{H}^{IJ}_{kk'} \, | \chi^J_{k'} \rangle = H^{IJ}_{kk'} \left(\mathbf{R}^{(c)}_{kk'} \right) \langle \chi^I_k | \, | \chi^J_{k'} \rangle \,, \tag{2.23}$$

where $\mathbf{R}^{(c)}_{kk'}$ is the centroid position between the two gaussian basis functions k and k'. The second approximation is called Independent First Generation, which means that the initial ensemble of basis functions are not coupled and are propagated independently. If they spawn, however, they always couple to their own child basis functions. For more information to the approximations used in AIMS, we refer to the following reviews[35,37–39]

2.4 INCLUSION OF THE ENVIRONMENT

One important concept in quantum chemical calculations is the inclusion of environmental effects. Taking as an example the two thymine molecules studied in this thesis, they are naturally embedded in a DNA strand and are surrounded by water and ions. As such an environment is too large to be simulated at the quantum mechanical level of theory, the system is separated into two regions. The inner region, consisting of the two thymines is treated quantum mechanically, while the outer region, i.e. the environment, is treated classically using a force field. This approach is called mixed quantum mechanics/molecular mechanics (QM/MM), and it can be used for both static and dynamics simulations. For the correct treatment of the system, the interaction region of the QM and MM parts needs to be treated. The total energy of the system can be expressed as

$$E_{\text{total}} = E_{\text{QM}} + E_{\text{MM}} + E_{\text{QM/MM}}. \tag{2.24}$$

One very popular way of considering the interaction term $E_{\text{QM/MM}}$ is the so called electrostatic embedding scheme. In it, the electrostatic interaction is being treated by adding an additional Coulomb term to the hamiltonian in the QM region. The van der Waals and bonded QM/MM interactions, on the other hand, are treated on the MM level. These latter classical terms vanish when computing vertical excitation energies because the force field parameters are the same for all electronic states. However, for running QM/MM dynamics simulations all terms have to be considered. For more information on different QM/MM schemes available in the literature the reader is referred to the review by Hans Martin Senn and Walther Thiel[40]

3 GAIMS

The original derivation of FMS allows only for spawning and population transfer between states of the same multiplicity, based on their NACs. In order to extend this scheme we had to modify the FMS code to also be able to account for ISC events. This work, which is presented below, has been published in the Journal of Chemical Physics[41] and is reprinted in Appendix 1.

3.1 ELECTRONIC REPRESENTATION

When extending the standard electronic Hamiltonian to include SOCs, we obtain

$$\hat{\mathbf{H}}^{\mathrm{sd}} = \hat{\mathbf{H}}^{\mathrm{el}} + \hat{\mathbf{H}}^{\mathrm{SOC}}. \tag{3.1}$$

When running standard FMS, the $\hat{\mathbf{H}}^{\mathrm{el}}$ used is diagonal, and therefore it is considered to run in an adiabatic representation. After including SOCs, the resulting total Hamiltonian is not diagonal anymore and two options are possible. Either we use the non-diagonal total Hamiltonian in the dynamics and end up with a spin-diabatic representation, which means that our Hamiltonian is adiabatic within one spin value (singlets, triplets, ...) and diabatic between them; or we re-diagonalize the total Hamiltonian to obtain a spin-adiabatic representation. Spin-adiabatic means that we are also adiabatic between spin values, i.e. the Hamiltonian is completely adiabatic. In the current implementation of GAIMS we use a spin-diabatic formalism. It would be interesting in the future, however, to compare it to a spin-adiabatic version of GAIMS.

3.2 PROPAGATION

When including triplet states, the time-dependent molecular wavefunction is defined as

$$\Psi(\mathbf{x}, \mathbf{R}, t) = \sum_{J} \sum_{M_{Sj}=-Sj}^{Sj} \Omega_J^{M_{Sj}}(\mathbf{R}, t) \Phi_J^{M_{Sj}}(\mathbf{x}; \mathbf{R}), \tag{3.2}$$

with $\Phi_J^{M_{Sj}}(\mathbf{x}; \mathbf{R})$ being the electronic wavefunction for state J with the spin-multiplicity $(2Sj + 1)$ and with the spin-projection eigenvalue of Ms. Note that this approach is general and applicable to any multiplicity. We will focus here on singlet $Sj = 0$ and triplet $Sj = 1$ states, but the code is capable of using any other multiplicity, provided one can obtain the respective couplings. The nuclear wavefunction used in this ansatz is

$$\Omega_J^{MSj}(\mathbf{R}, t) = \sum_{k'=1}^{N_{J,MSj}} C_{k'}^{J,MSj}(t) \chi_{k'}^{J,MSj} \tag{3.3}$$

As can be seen in Equation 3.3, every single state, including the sublevels of the triplet states, have a separate coefficient, which is necessary for a correct treatment

of the time-evolution of the wavefunction. The propagation of these coefficients, using the new wavefunction from equation 3.2, goes as follows:

$$\frac{\mathrm{d}\mathbf{C}^{I,M_{Si}}(t)}{\mathrm{d}t} = -i(\mathbf{S}^{-1})^{II,M_{Si}M_{Si}}[\mathbf{H}^{II,M_{Si}M_{Si}} - i\dot{\mathbf{S}}^{II,M_{Si}M_{Si}}\mathbf{C}^{I,M_{Si}}(t) \tag{3.4}$$

$$+ \sum_{J} \sum_{MSj=-Sj}^{Sj} \mathbf{H}^{IJ,M_{Si}M_{Sj}}\mathbf{C}^{J,M_{Sj}}(t)] \tag{3.5}$$

$$\tag{3.6}$$

Here, $\mathbf{H}$ is the Hamiltonian matrix, and $\mathbf{S}$ and $\dot{\mathbf{S}}$ are the overlap matrix elements:

$$S_{k,k'}^{IJ,MSi,MSj} = \langle \chi_k^{I,MSi} | \, | \chi_{k'}^{J,MSj} \rangle \, \delta_{IJ}\delta_{MSi,MSj} \tag{3.7}$$

$$\dot{S}_{k,k'}^{IJ,MSi,MSj}, = \langle \chi_k^{I,MSi} | \, \frac{\partial}{\partial t} \, | \chi_{k'}^{J,MSj} \rangle \, \delta_{IJ}\delta_{MSi,MSj}. \tag{3.8}$$

$$\tag{3.9}$$

As can be seen by the Kronecker deltas in the equation above, both the overlap matrix and its time-derivative are only needed between basis functions on the same state, which also includes the triplet sublevels.

As explained above, every $|J, Mj\rangle$ pair of states have their own gaussian basis function with their independent coefficient. As GAIMS is an AIMD method which runs on-the-fly, at every time step we need to evaluate the electronic structure properties (energies, gradients, couplings) for every basis functions. However, as our version of GAIMS runs in the spin-diabatic formalism the three sublevels of a given triplet state (with Ms values of $-1, 0, 1$) all are energetically degenerate. In our simulations we can use this fact to drastically reduce the computational cost. We only need to evaluate the electronic structure information for one gaussian per triplet state, and use the same information for the other two basis functions. However, the coefficients are still propagated separately. One should keep in mind that this reduction of the necessary number of electronic structure calculations is not an approximation, but exact within the spin-diabatic picture.

3.3 SPAWNING

The key of the FMS code is the spawning procedure, which adds new basis functions to the current ensemble, in order to allow for the wavepacket to spread. If a basis function enters a region of nonadiabaticity, a new basis function can be spawned, under certain conditions. The spawning criteria used in standard FMS is the strength of the NACs. In the spin-diabatic approach used in GAIMS, we need to define a new diabatic spawning criteria based on the SOC strength between the respective states. Based on an already proposed effective coupling between diabatic states[35], we define the effective coupling used here as:

$$\Lambda_{IJ}^{\mathrm{eff}} = \frac{\left(\sum_{Ms_i=-s_i}^{s_i} \sum_{Ms_j=-s_j}^{s_j} \left| \langle \Phi_I^{Ms_i}(\mathbf{R}_k) | \, \hat{\mathbf{H}}^{\mathrm{SOC}} \, | \Phi_J^{Ms_j}(\mathbf{R}_k) \rangle \right|^2 \right)^{1/2}}{\left| E_J^{\mathrm{el}}(\mathbf{R}_k) - E_I^{\mathrm{el}}(\mathbf{R}_k) \right|}, \tag{3.10}$$

which ensures that the spawning criteria is rotationally invariant. If the effective coupling $\Lambda_{IJ}^{\mathrm{eff}}$ is larger than a predefined threshold, the spawning mode is triggered and for a successful spawn to a triplet state, three new basis functions are generated for the three triplet sublevels.

3.4 APPLICATION TO TEST SYSTEMS

3.4.1 *One-Dimensional Test System*

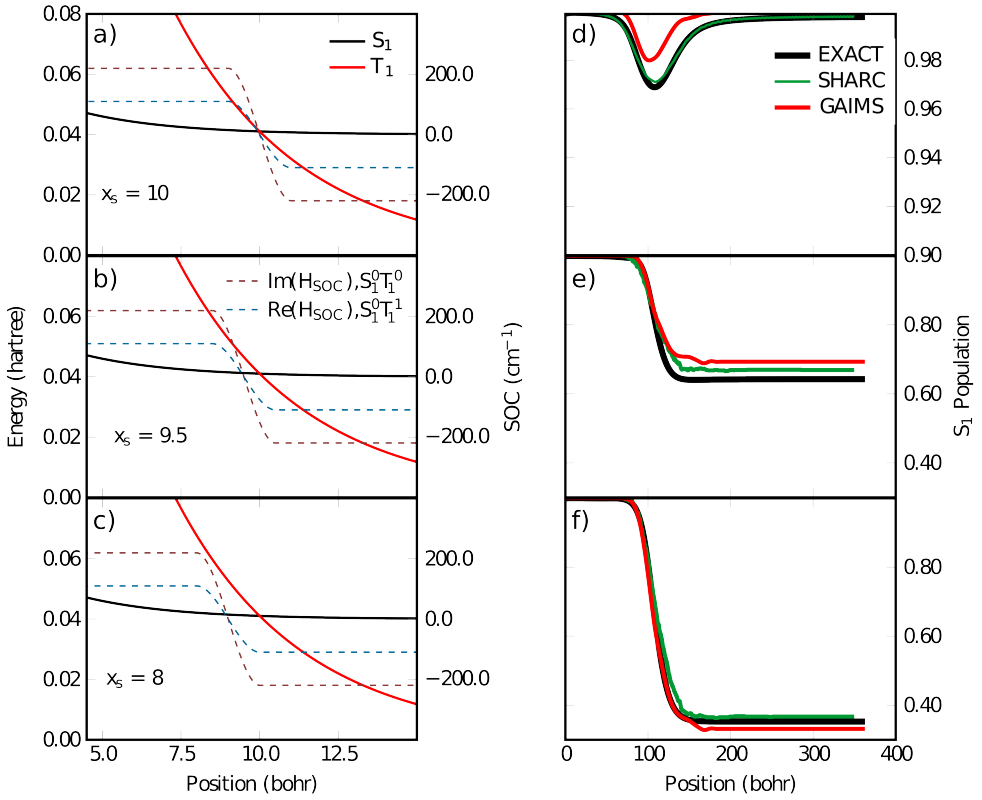


Figure 4: GAIMS results for a one-dimensional model system. a-c: Potential energy curves for the model system with one Singlet (S_1) and one triplet state (T_1). The spin orbit couplings between states S_1 and T_1^1/T_1^{-1} is $(\mathrm{Re}(H_{SOC}) \pm i{\cdot}\mathrm{Re}(H_{SOC}))$, while between S_1 and T_1^0 it is $(0 - i{\cdot}\mathrm{Im}(H_{SOC}))$. The three cases a-c differ in where the coupling strength between the two states changes sign, shown by the value of x_s. d-f: GAIMS results for each model system, compared to SHARC simulations and to the exact results, taken from reference [32].

In order to test our new method we employed it on a model system proposed by Persico and co-workers [32]. The one-dimensional model, shown in Figure 4a-c, consists of a singlet and a triplet state. Both states are dissociative, with a single crossing between them at x =10 bohr. The SOC elements between the two states are defined so that they change sign at a certain point of the potential. By varying this point we can simulate different SOC conditions for our code. In total, three different simulations are being run, as shown by Figure 4(a-c). While for Figure 4a the point where the SOC changes sign is exactly at x_s =10 bohr, which is the crossing point of the two states, for Figures 4b and 4c the SOC changes sign at x_s =9.5 bohr and x_s =8 bohr respectively.

For each of the three cases we ran GAIMS dynamics simulations using 200 initial conditions each, sampled from a Wigner distribution of the initial wavefunction from reference [32]. We compared our results to the exact results from reference [32], as well as to surface hopping simulations using SHARC [17,18]. The results are shown in Figure 4d-f. As can be seen, GAIMS manages to describe all three test cases reasonably well, with a maximum error of 7%. Keeping in mind that we used both the saddle point

and independent first generation approximations, and only spawned once per initial condition our implementation shows promise for applications to more complicated systems.

3.4.2 *Thioformaldehyde*

As a second test system for GAIMS we decided to run dynamics on the molecule thioformaldehyde, H_2CS. After excitation to the S_1 state, which is of $n\pi^*$ character, two close lying triplet states, $T_1(n\pi^*)$ and $T_2(\pi\pi^*)$, exist. Using GAIMS, we ran 20 initial conditions for 200 fs, sampled from a Wigner distribution[42]. All electronic structure calculations were performed at SA-CASSCF(4,3)/6-31G* level of theory, state averaged over 2 singlet and 2 triplet states, using MOLPRO[43]. The results can be seen in Figure 5b. While there is basically no population transfer to the T_1 state, we see a slow rise of population in T_2, which is in accordance with the El-Sayed rules[44]. This early onset of population transfer has already been observed for the similar molecular acetone[45], although on a smaller scale. Also shown in Figure 5b are the total number of basis functions (orange line). At the end of the dynamics, at 200fs, here are 326 basis functions, among which 306 run on the triplet states.

In order to further analyze the dynamics and investigate what causes ISC to happen we measured the C=S distance for all basis functions. The results are shown in Figure 5a, with T_1 shown in red, T_2 in blue and S_1 shown in gray. As T_1 and S_1 are both of the same $n\pi^*$ character, their dynamics is very similar, while the T_2 basis functions show, on average, a larger C=S distance.

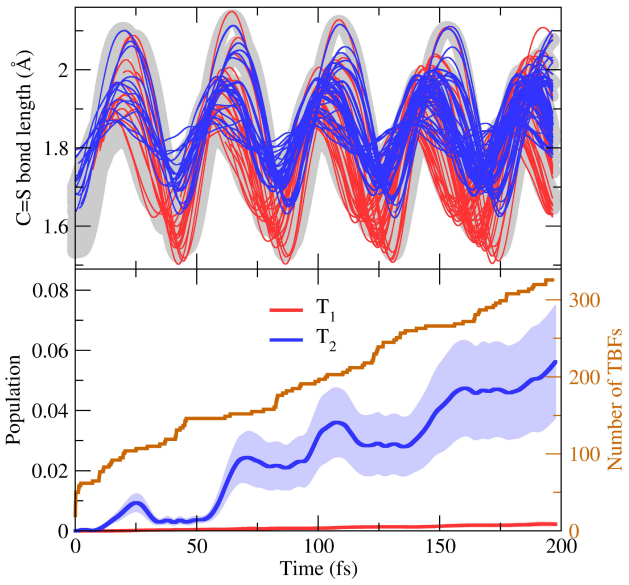


Figure 5: GAIMS results for Thioformaldehyde (H_2CS). a: C=S bond length expectation value for all S_1 (gray), T_1 (red) and T_2 (blue) basis functions. The width of the lines corresponds to the amount of population in the respective basis function. b: Population in the triplet states during the dynamics. The light areas indicate the standard error of the simulation. Also, in orange, the total number of basis functions at each timestep is shown.

4 THYMINE PYRIMIDINE DIMER

The most frequent pathway for the formation of thymine pyrimidine dimers is direct dimerization after direct UV excitation. This process has been shown to be ultrafast, with the dimer being formed in less than 1 ps after excitation[10]. Based on the fast nature of the reaction, it was suggested that it occurs in the singlet manifold[10,13] The second process under consideration in this thesis is slower, occurring on a ns timescale. It involves population transfer to the triplet manifold enhanced by triplet-triplet energy transfer (TTET) from a photosensitizer, followed by dimerization[16,46–49]. Although triplet-mediated T<>T formation without the presence of a photosensitizer has also been observed experimentally[11], it has been shown that the presence of photosensitizers increases the dimerization yield[46]. The results for these two processes investigated in this thesis will be discussed here in detail:

4.1 DIRECT DIMERIZATION

As a [2+2] cycloreaction, T<>T formation follows the Woodward-Hoffmann rules, which state that the reaction is ground state forbidden[50]. In the excited state manifold, however, two different states can lead to dimerization; either a singly excited (SE) or a doubly excited (DE) state. Experimentally, it has been shown that the formation of T<>T in DNA happens on a timescale of under 1 ps[10]. Based on the ultrafast nature of the reaction, it was assumed that it occurs in the singlet state manifold. However, in 2008 Kwok et al. experimentally discovered a possible doorway state from a which both a singlet and a triplet pathway lead to dimerization[11]. The triplet pathway found, although happening on a ns timescale, shows ISC in under 1 ps. This triplet pathway has been ruled out as a major contributor to T<>T formation in a joint experimental/theoretical study by Banyasz et al. who claim that its contribution is less than 10%[12]. Two theoretical studies investigated the triplet mechanism of T<>T formation following ISC after direct excitation. While Zhang and Eriksson[51] used DFT to find a barrier of 2.5 eV in the T_1 state to overcome, Wang and Chen[15], using CASPT2, observed a barrier of only 0.3 eV, thereby showing a valid triplet dimerization pathway.

In this chapter, the results on the formation of T<>T in the singlet manifold after direct excitation will be presented. First, static calculations will compare different accessible direct excitation pathways. Then, AIMD simulations using SHARC[17,18] were run in order to simulate the dimerization mechanism and get more insight into the conformational control of the reaction. While the static results are yet unpublished, the AIMD results were published in the Journal of the American Chemical Society[52] which is reprinted in Appendix 2.

4.1.1 *Static calculations*

Several theoretical studies showed that the reactive state for dimerization is the lowest $^1\pi\pi^*$ state, on which a barrierless dimerization pathway exists[12,15,53–61]. In 2007, however, Blancafort and Migani stated that a higher lying DE state is the precursor for dimerization[62]. According to their calculations, this precursor state is

the S_6 state at CASPT2 level of theory, but the S_1 state when using CASSCF. In the same year, Boggio-Pasqua et al. presented a dimerization pathway on the S_1 surface using CASSCF, with the initially excited state as the DE state[53]. No other CASPT2 study has investigated the importance of this higher lying precursor state in the reaction. For this DE precursor state to play a role in the dimerization mechanism there are two possibilities. Either, for certain geometries, the DE state can go down in energy to a region where it can get excited to after UV radiation, or after excitation to an SE state the population gets transferred to the DE state which then forms the dimer. In order to investigate whether direct excitation to this DE state is possible, we ran a quantum mechanics/molecular mechanics (QM/MM) dynamics simulation on a DNA single-strand and took 250 snapshots from it[a]. For these snapshots we ran single point QM/MM calculations using CASPT2 and, using the wavefunction analysis program TheoDORE, grouped all excited states into either an SE or a DE band. The results can be seen in Figure 6a. For no geometry in our set the energy of the DE band is below 5.5 eV. As the experimental excitation energy is at 4.7 eV[10], direct excitation to a DE state is therefore highly unlikely.

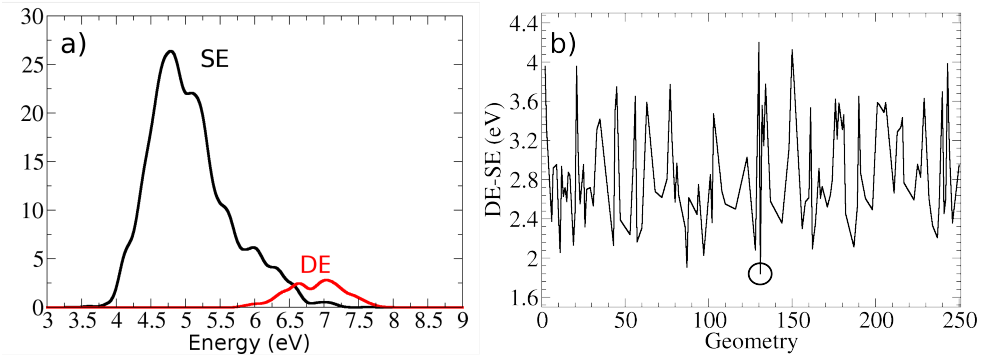


Figure 6: (a) Absorption spectrum of the SE and DE states of two thymine molecules embedded in a $(dT)_{12}$ single strand calculated at CASPT2/Amber level of theory. (b) DE-SE energy difference for the points of the spectrum for which the energy of the SE state lies beneath 5 eV. The circle indicates the point of the smallest DE-SE energy difference.

The second possibility to investigate whether DE states are involved in the reaction is to evaluate whether, after direct excitation to an SE state, the system can, through a DE/SE conical intersection, move to the DE state and then dimerize. In order for this path to happen, a DE/SE crossing point needs to be accessible by the system after excitation. We therefore investigated the 250 QM/MM snapshots and measured the DE-SE energy difference for all geometries where the SE state lies in an excitable energy range of under 5 eV. The results of this analysis can be seen in Figure 6b. The smallest DE-SE gap we can find in our ensemble is 1.8 eV (see the circle in Figure 6b). This geometry was taken as an initial conformation at the Franck-Condon (FC) geometry from which two different dimerization pathways were investigated. While pathway 1 does not include the higher lying DE state, it is included in the process described in pathway 2.

a For computational details, see SI of Rauer et. al J. Am. Chem. Soc. 138, 15911-15916 (2016)

PATHWAY 1 For pathway 1 we optimized the (S_1/S_0) crossing from the FC point, using MS-CASPT2(4,4)/cc-pVDZ[63] with an active space of 4 electrons in 4 orbitals. The optimized geometry can be seen in Table 1. This geometry is identical to the one published by Boggio-Pasqua et al[53], and the one we found in our dynamics simulations (see chapter 4.1.2). Figures 7a and 7b show a linear interpolation scan between our FC point and this (S_1/S_0) conical intersection, and from the conical intersection to the dimer, using MS-CASPT2 and SA-CASSCF, respectively. Both methods predict a barrierless pathway on the S_1 surface leading to dimerization. The character of the S_1 state, however, is different between the methods, which can be visualized using the Ω value, as obtained by the program TheoDORE. The parameter Ω states the amount of single excitation character in an electronic state. Thus, we can measure whether the reactive S_1 state is mainly of SE or of DE character. A value of $\Omega = 1$ means a pure SE state, while $\Omega = 0$ indicates a state with no SE character, but that it is instead comprised of double or higher excitations. Looking at the Ω value for the CASPT2 scan (Figure 7c), we see that it does not get lower than 0.8 throughout the whole scan. The S_1 state is therefore mainly SE. However, as the scan approaches the (S_1/S_0) crossing, the DE contribution of S_1 grows. Although the DE character of S_1 never exceeds 20% we see that for a proper description of the state we need to account for higher excitations than just SE. The S_2 state, which has strong multiple excitation character is 2 eV above the S_1 state throughout the reaction and does not play any role in this dimerization pathway.

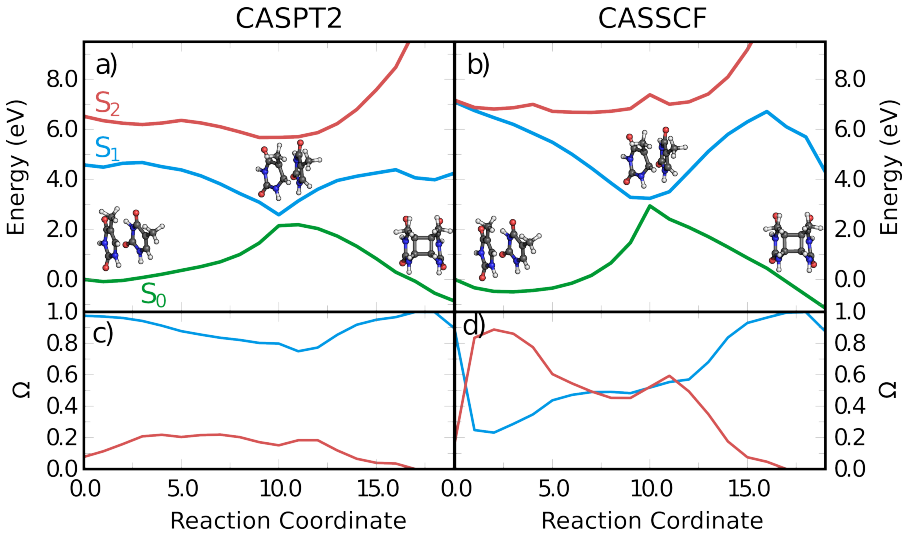


Figure 7: Potential energy scan for the three lowest singlet states for the reaction from the initial FC point to the optimized (S_2/S_1) conical intersection, and from (S_2/S_1) to the dimer using MS-CASPT2 (a) and SA-CASSCF (b). Figures c) and d) show the amount of single excitation character for each excited states for the MS-CASPT2 and SA-CASSCF results, respectively.

The character of the reactive state during dimerization changes when using CASSCF. Without dynamic correlation the SE state gets destabilized, and as a result it crosses the DE state close to the FC point. As in the CASPT2 scan, the CASSCF one shows a barrierless pathway going to the (S_1/S_0) crossing (see Figure 7b). However, the Ω value of the reactive S_1 state, shown in Figure 7d, is different than in CASPT2. As the FC point lies close to the (S_2/S_1) crossing at CASSCF level of theory, after the first step in the scan the two excited states change character and the S_1 state is of

mainly DE character. As we approach the (S_1/S_0) crossing, the S_1 state is a mixture of SE and DE with an Ω value of approximately 0.5. During the scan, the S_2 state stays in the energy range of above 7 eV. However, it also shows a mixed SE/DE character close to the conical intersection, indicating a mixing of the S_1 with higher lying excited states.

Comparing the two scans in Figure 7 we can see that both CASSCF and CASPT2 predict a barrierless reaction pathway. However, CASSCF seems to overestimate the amount of DE character in the reactive S_1 state.

PATHWAY 2 In order to investigate the possible role of the DE state in T<>T formation using CASPT2, we optimized, from our Frank-Condon geometry, the (S_2/S_1) crossing point, again using CASPT2(4,4)/cc-pvdz. This conical intersection indicates a possibility for the DE state to get populated, as it is a crossing point between a DE state (S_2) and an SE state (S_1). The optimized geometry can be seen in Table 1, with the C_5-C_5' and the C_6-C_6' bonds at 3.72 and 3.77 , respectively. According to classical dynamics simulations, this geometry lies in a dynamically accessible region of nucelobases embedded in B-DNA[64]. Additionally, linear interpolation between the FC point and this (S_2/S_1) crossing, using MS-CASPT2, shows a flat potential energy profile along the initially populated S_1 state, see Figure 8a, therefore also allowing for the possibility to cross to the DE state after excitation to the FC point. Figure 8a also shows the scan from the new (S_2/S_1) crossing point to our previously optimized (S_1/S_0) conical intersection, and further to the formed dimer. Figure 8b shows the Ω value along the scan. As expected, the S_1 state is still mainly SE, however, the SE contribution after the (S_1/S_0) crossing is only 70%. Therefore, as in the MS-CASPT2 scan of pathway 1 (see Figure 7a,c) we cannot neglect the contribution of higher excitations to the active state in order to properly describe it. Although possible, this pathway is less likely to occur than pathway 1 because the potential of the active S_1 state is flat in the first part, compared to the gradient in pathway 1, which is promoting dimerization.

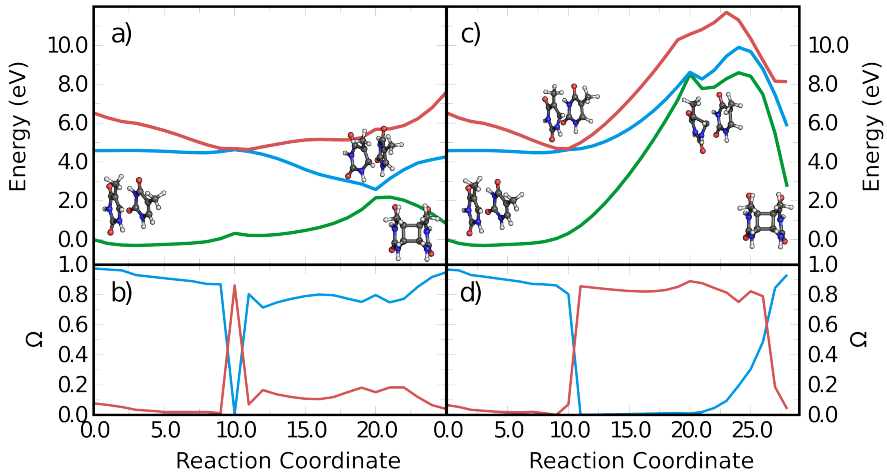


Figure 8: a) Excited state energies for the scan from the FC point to the (S_2/S_1) conical intersection and further to the (S_1/S_0) crossing using MS-CASPT2. b) Amount of single excitation character for each excited states shown in a). c) Energies for the scan from the FC point to the (S_2/S_1) crossing and then to the optimized (S_1/S_0)' conical intersection with a mainly DE S_1 state using MS-CASPT2. d) amount of single excitation character for each excited states for the scan shown in c).

The scan shown in Figure 8a discussed above follows the primarily SE state after passing the (S_2/S_1) crossing point. In order to find a possible MS-CASPT2 dimerization pathway along the DE state, we optimized a crossing point between the ground state and the DE state, which we call ($S1/S_0$)′. The resulting geometry is shown in Table 1. As can be seen, the geometry is distorted compared to the (S_2/S_1) crossing, with the C_5-C_5' bond elongated to 4.5 , while the C_6-C_6' bond retracts slightly to 3.5 . Moreover, two additional geometric parameters are worth mentioning. First, the C_6 atom of one of the monomers moves toward the other monomer, thereby distorting the ring structure. Second, the C_5-C_6 and C_5'-C_6' bonds elongate to 1.7 and 1.81 , respectively, from 1.5 at the (S_2/S_1) crossing. Figure 8c shows the energies of the first three excited singlet states for the dimerization path along this newly optimized ($S1/S_0$)′ crossing point. The energies of the first part, up to the (S_2/S_1) crossing, are the same as in Figure 8a. However, upon leaving the (S_2/S_1) crossing the active S_1 state is of mainly DE character (see Figure 8d). In order for the system to reach the ($S1/S_0$)′ point, however, it needs to overcome an energetic barrier of 4.1 eV. The magnitude of this barrier is expected, considering the geometric distortions at the ($S1/S_0$)′ crossing point, as discussed above. After reaching the ($S1/S_0$)′ point, dimerization is possible on the S_0 state, after overcoming a second barrier of 0.6 eV. However, the size of the first barrier already rules out this pathway for a possible dimerization mechanism in DNA.

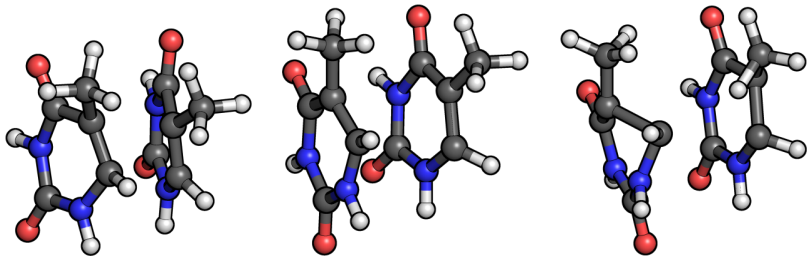
	(S_1/S_0)	(S_2/S_1)	(S_1/S_0)′
			
C_5-C_5'	2.27	3.72	4.50
C_6-C_6'	2.17	3.77	3.50

Table 1: Conical intersection geometries optimized using CASPT2, as well as their C_5-C_5' and C_6-C_6' bond distances in Å(see Figure 1).

4.1.2 *Dynamics calculations*

In section 4.1.1 we presented stationary calculations of the investigation of the mechanism of T<>T formation and the character of the reactive $^1\pi\pi^*$ state. In order to gain more insight into the dimerization mechanism and to measure the importance of triplet states, we also ran AIMD simulations, using the program SHARC[17,18]. For a system of this size, running on-the-fly dynamics using MS-CASPT2 is not feasible. Therefore, we employed SA-CASSCF for all electronic structure calculations. As shown above, SA-CASSCF overestimates the amount of DE character in the reactive state. However, we can still gain insights about the conformational control and the role of triplet states in the reaction.

The starting geometries for the dynamics simulations were taken from a Wigner distribution, centered on a geometry from the static scan done by Boggio-Pasqua et al[53], which shows a closer thymine-thymine distance compared to natural DNA.

From this geometry we sampled 10 initial conditions and started the dynamics from the lowest $\pi\pi^*$ state (DE). For completeness, we also started 5 trajectories from the second $\pi\pi^*$ (SE), and the S_1 state which is of $n\pi^*$ character. However, none of the trajectories starting in the SE or the $n\pi^*$ states showed dimerization. From the 10 trajectories started in the DE state, 5 underwent dimerization. One example for a dimerizing trajectory is shown in Figure 9a-c. In Figure 9a, the evolution of the energies of the 5 electronic singlet states included in the dynamics simulation are shown. The color labels used in the figure are based on the amount of SE and DE contributions for the respective states. All states are either labeled closed shell (CS), SE or DE, regardless of them being of $\pi\pi^*$ or of $n\pi^*$ character. The active state of the trajectory starts out as S_3, becomes S_1 after 10 fs and then, after 100 fs relaxes to the ground state and forms the dimer. Similar as in section 4.1.1, the active state is a mixture of SE and DE and CS contributions, making it difficult to determine the diabatic character of the state. Figure 9b shows the time-evolution of the two distances which later form the dimer, C_5-C_5' and C_6-C_6'. The distances both start out at 3Å and slowly decrease during the dynamics. After relaxation to the ground state, the C_5-C_5' bond gets formed first and, 30 fs later, the C_6-C_6' bond gets formed. As a second conformational parameter, Figure 9c shows the dihedral angle C_5-C_6-C_6'-C_5' (ν) during the reaction. Initially, ν decreases concomitantly with the C_5-C_5' and C_6-C_6' distances. Around the S_1/S_0 conical intersection, the dihedral angle starts increasing to avoid steric clashes. After the cyclobutane ring is fully formed, ν starts decreasing again and at the end of the trajectory it reaches 0°.

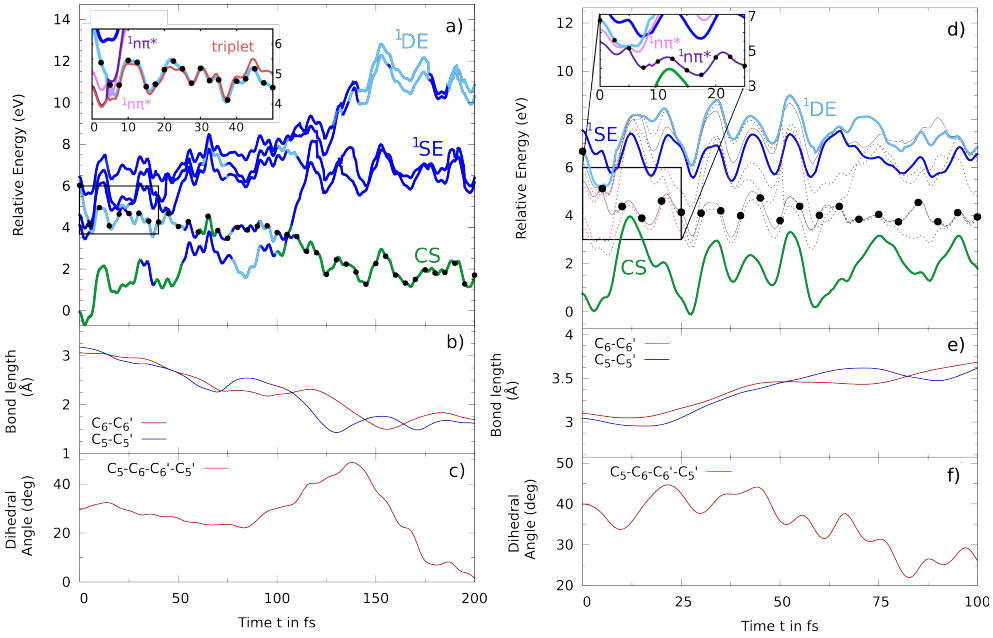


Figure 9: Time-evolution of the excited state energies for an exemplary reactive trajectory (a). The states are colored as closed shell (CS), singly excited (SE) and doubly excited (DE), based on their electronic wavefunction. (b) evolution of the two distances C_5-C_5' and C_6-C_6'. (c) dihedral angle C_5-C_6-C_6'-C_5' during the reaction. (d) time evolution of the excited state energies for an exemplary unreactive trajectory starting in the DE state, as well as the C_5-C_5' and C_6-C_6' (b), and the dihedral angle C_5-C_6-C_6'-C_5' (c).

In the nonreactive trajectories which start from the DE state, the system relaxes initially to the lowest $n\pi^*$ state and stays in it throughout the simulation time. One example trajectory is shown in Figure 9d-f. In Figure 9d the states are labeled based

on their electronic character, as their assignment is easier than for the reactive trajectories due to the larger distance between the two thymine monomers. Looking at the two distances relevant for dimerization, C_5-C_5' and C_6-C_6', we can see that they increase during the simulation, indicating that the monomers move apart. The dihedral angle ν, similar as in the reactive trajectory, decreases during the simulation. However, this behavior does not lead to dimerization as the populated state is of $n\pi^*$ character.

A key question regarding the formation of T<>T is the involvement of triplet states. For our simulations, none of the trajectories, irrespective of the excited state they started from, underwent intersystem crossing to the triplet manifold. Even when the populated singlet and a triplet state are lying close in energy for a longer period of time (see inset of Figure 9a), no ISC occurs. While the two states are 0.0032 eV (258 cm^{-1}) separated, the SOC between them is only 0.03 cm^{-1}. From this we conclude that for the process studied, triplet states are not involved in the reaction, which is in agreement with experimental studies[10]. However, Wang and Chen recently published a possible triplet state dimerization pathway including ISC from an $^1n\pi^*$ state[15]. Although they do not state any time constants for the reaction, it is assumed to occur on a ns timescale, far outside of our simulation times. Therefore, we cannot exclude a possible dimerization pathway for our inactive trajectories populating an $^1n\pi^*$ state.

4.2 PHOTOSENSITIZED DIMERIZATION

The second pathway investigated in this thesis is the photosensitized thymine dimerization process[16]. In this pathway, a photosensitizer, after excitation, undergoes ISC and then transfers its electronic energy to a close-by thymine molecule, thereby promoting it to its lowest triplet state. This process could be observed experimentally, with a dimerization yield of 4%[46], which is 20 times more than the previously observed yield for the triplet state pathway, showing that the presence of a photosensitizer drastically increases the dimerization yield[12,16]. Two experimental time constants were detected for the process. The first one of 22.5 ns[46] was assigned to a local triplet state (^{3}L, Figure 10a,c), which gets populated by TTET from the photosensitizer. The second time constant of 62 ns[46] was assigned to a biradical triplet state (^{3}BR, Figure 10b,d,e), which gets populated from the ^{3}L state. For the ^{3}BR state, the C_6-C_6' bond is already preformed and electronic density is distributed over both thymine monomers, while in the ^{3}L state, the charge is localized on only one of the thymine units and no bonds are preformed. This ^{3}BR state discussed above can correspond to two different electronic states, namely a Frenkel exciton (Figure 10d), with two coupled local excitations, or a charge-resonance state (Figure 10e), which includes two charge transfer excitations between the two monomers[65]. It is worth mentioning here that, compared to the direct dimerization path, the Woodward-Hoffmann rules are not applicable to triplet mediated reactions. Therefore, they are not applicable to this pathway.

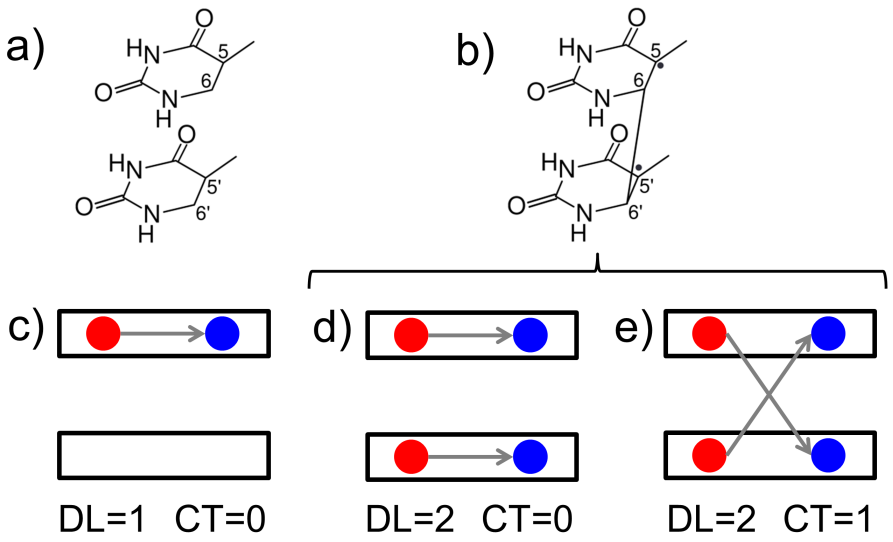


Figure 10: Schematic pictures of the local triplet state ^{3}L (a) and of the biradical triplet state ^{3}BR (b). Electron configuration of a local state (c), a Frenkel exciton state (d) and a charge resonance state (e). The two rectangles represent the two thymine monomers, and the black arrows indicate an excitation from an electron (red circles) to a hole (blue circles). For a-c, both delocalization length (DL) and charge transfer (CT) character are shown (recall Section 2.2).

Here, we use static and dynamics calculations to investigate the photosensitized dimerization mechanism, and assign the two experimental time constants. We study the formation of the ^{3}BR intermediate and identify whether it is a Frenkel exciton or a charge-resonance state. The results shown here were published in Chemical Monthly[66].

Experiments suggest that the state populated by TTET is of ^{3}L character. This behavior is expected, as the geometry at the ^{3}BR state already has the C_6-C_6' bond preformed. To find out whether any ^{3}BR states could be populated in DNA, we took the 250 QM/MM snapshots from chapter 4.1.1 and investigated the character of the T_1 state for them. The results showed that the T_1 band of the spectrum is comprised of ^{3}L excitations, with only a negligible ^{3}BR contribution. Therefore, we assume that the TTET occurs to a ^{3}L band, in agreement with experimental suggestions[46].

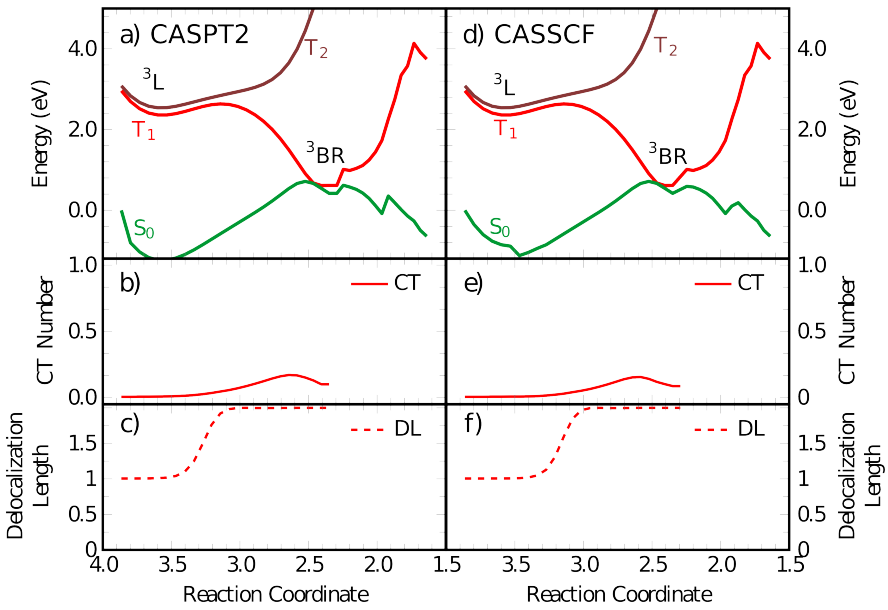


Figure 11: MS-CASPT2 linear interpolation scan of the S_0, T_1 and T_2 energies for the triplet state formation of T<>T (a), as well as (b) the evolution of the charge transfer (CT) and the delocalization length (DL) (c). (d) SA-CASSCF linear interpolation scan and the CT (e) and DL (f) analyses of it.

After population of the ^{3}L state, the next step in the reaction is the formation of the ^{3}BR state. In order to simulate this process, a linear interpolation scan has been performed between the ^{3}L point in the FC region to ^{3}BR, and from ^{3}BR to the finished dimer. The scan shown in Figure 11a has been performed at the MS-CASPT2(4,4)/cc-pVDZ level of theory. After the T_1 state gets populated, the system relaxes to the T_1 minimum marked ^{3}L. From this minimum, a barrier of 0.3 eV exists which connects the ^{3}L minimum to the ^{3}BR minimum. The height of this barrier is in agreement with previous calculations[15] and with the experimental time constant of 22.5 ns[46] for this step of the reaction. After overcoming the barrier the system relaxes to the second T_1 minimum marked ^{3}BR. Figures 11b and c show the CT and DL character of the T_1 state along the pathway, respectively. The DL character shows that the T_1 state starts out as ^{3}L, located on only one monomer, but after overcoming the barrier it is ^{3}BR with a DL value of 2. Looking at the CT character of the state during the reaction we can see that it is below 0.2 during the whole process, indicating that the state is mainly not CT. Therefore, we can assume that the ^{3}BR state involved in the reaction is a Frenkel exciton, and not a charge-delocalised state.

After reaching the ^{3}BR state, the system is trapped in the T_1 minimum labeled ^{3}BR. As this minimum coincides with a T_1/S_0 crossing point, dimerization can occur after ISC from T_1. However, as the experimental time constant for this process is 40 ns[46], the process is expected to be very slow. After population transfer to the ground state the system can either form T<>T, or return to the FC region. As the experimental yield of dimerization is only 4%[46], we sampled the ^{3}BR minimum using the AIMD program SHARC[17,18], in order to investigate which factors determine the low yield. As running AIMD using CASPT2 for this system is not computationally possible, we used the SA-CASSCF method. Figures 11d-f show that CASSCF successfully reproduces the results obtained by CASPT2 (Figures 11a-c), both in the energies and the character of the T_1 state.

All trajectories were started in the T_1 state for at least 100 fs. During this time, none of them showed ISC to the ground state, which can be explained with the low SOC between T_1 and S_0, which is on average only 1 cm^{-1}. This behavior is in agreement with the long experimental time constant of 40 ns, and with predictions made by previous theoretical calculations [14,15]. In order to simulate the ground state relaxation process, we randomly sampled 38 initial conditions from the trajectories running in the ^{3}BR minimum, manually placed them in the ground state and continued the dynamics for at least 50 fs. Of these 38 trajectories, only 5 showed dimerization. In order to rationalize the low dimerization yield, we analyzed the phase space of the $^3BR/S_0$ transition geometries taken as the initial geometries for the trajectories.

Figures 12a-b show the C_5-C_5' and C_6-C_6' distances, as well as the angles $(\theta_5 + \theta_6)/2$ and $(\theta_5' + \theta_6')/2$, where θ_5 is the angle between the velocity of the C_5 atom with the C_5-C_5' vector and θ_6 is the angle between the velocity of the C_6 atom with the C_6-C_6' vector. The five trajectories that dimerized are shown as green pentagrams. Figure 12 shows that, in order for the conformation to show dimerization, the C_5-C_5' and C_6-C_6' distances need to be lower than 3 and 2.1 , respectively. Additionally, the velocities of the two monomers need to point towards each other, with the $(\theta_5 + \theta_6)/2$ and $(\theta_5' + \theta_6')/2$ angles above 90 and below 80 degrees, respectively. We can therefore see that, although a large conformational space exists for the system, only a few degrees of freedom induce dimerization.

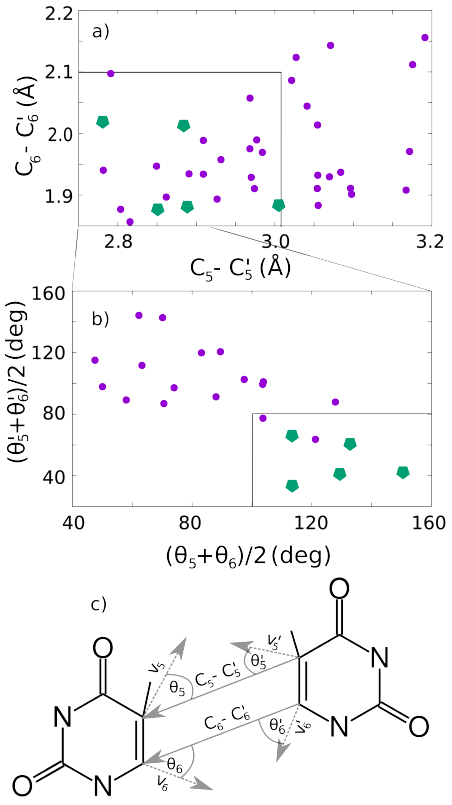


Figure 12: Analysis of the 38 initial conditions which were propagated on the ground state. (a) C_5-C_5' and C_6-C_6' distances, (b) linear combination of the angles formed between the distances C_5-C_5' and C_6-C_6' and the velocities at the atoms C_5 and C_6. (c) definitions of the distances and angles employed in the analysis. The lines between panels a and b indicate that only the subset in the smaller rectangle is used in the subsequent analysis.

5 CONCLUSION & OUTLOOK

The main goal of this thesis was the investigation of the formation of pyrimidine thymine dimers using both static calculations and ab-initio molecular dynamics simulations. Thymine dimers are formed after UV irradiation of DNA and can lead to severe health risks leading up to cancer[1]. The correct understanding of the dimerization mechanism is important for determining which factors enable dimerization. The two main dimerization mechanisms, direct singlet and photosensitized triplet dimerization, as sketched in Figure 13, were investigated.

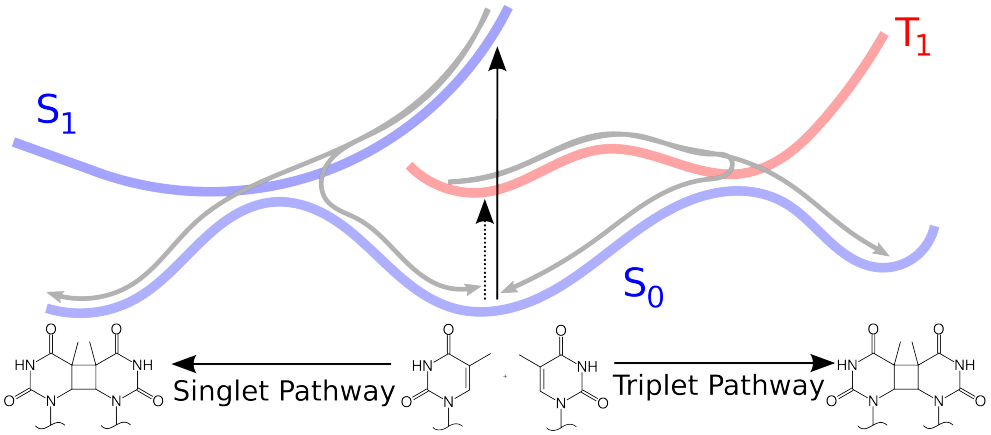


Figure 13: Schematic representation of the two dimerization pathways investigated in this thesis. Black solid arrow stands for vertical excitation after UV irradiation. Dashed arrow indicates population of the lowest triplet state either by intersystem crossing from a singlet state or by triplet-triplet energy transfer from a photosensitizer.

For the direct dimerization pathway, static calculations showed a barrierless reaction pathway on the lowest $\pi\pi^*$ singlet state, in agreement with both experiments and previous calculations[10,57]. When comparing the different pathways obtained by CASSCF and CASPT2, both show a barrierless reaction profile. However, the electronic character of the reactive state is different between the methods. When using CASPT2, the S_1 state is of mainly SE character with only up to 20% contributions from higher excitations. At the CASSCF level of theory, however, the assignment of the active state is more complex. At the beginning of the linear interpolated scan the state is mainly DE. Then, as the system reaches the S_1/S_0 crossing the SE character of the state increases to up to 50%. This shows that although CASSCF is capable of describing a barrierless pathway for direct dimerization, the contribution of doubly and higher excitations in the electronic wavefunction of the active state are overestimated, compared to CASPT2. Using the AIMD method SHARC[17,18], we were able to investigate the role of triplet states during the dimerization process. Due to the very low spin-orbit coupling we did not observe any intersystem crossing events during the simulation time. We could therefore show that the direct dimerization process, which is ultrafast, occurs solely on the singlet state manifold.

For the photosensitized dimerization pathway, we employed static CASPT2 and CASSCF calculations. Initially, the T_1 state gets populated as a state locally excited (^{3}L) on one of the monomers. From this ^{3}L state, a barrier of 0.3 eV leads to a biradical

^{3}BR state minimum. The height of this barrier agrees with the experimental time constant of 22.6 ns. The ^{3}BR minimum coincides with a T_1/S_0 crossing allowing for dimerization to happen in the ground state. In this T_1 minimum we ran SHARC trajectories in order to investigate the dimer formation mechanism. Due to the large experimental time constant of 40 ns we randomly relaxed 38 geometries in the T_1/S_0 degeneracy region, of which 5 dimerized while the rest relaxed back to the FC region. Conformational analysis of the 38 geometries show that not only the C_5-C_5' and C_6-C_6' distances need to be low for dimerization to happen, but that also the velocities of the four carbon atoms need to point towards each other.

Both singlet and triplet pathways show that gas phase simulations are capable of qualitatively modeling the thymine dimerization mechanism. However, for a better understanding of the underlying processes, a proper treatment of the DNA environment and the surrounding solvent are needed. Using a QM/MM scheme, both the static as well as the dynamics simulations presented in this thesis can be improved. By comparing the results obtained using QM/MM to the gas phase ones, the effect of the solvent and the DNA backbone on the dimerization mechanism can be properly evaluated. Recently it has been shown that electronic excitations in DNA are not located only on one nucleobase, but that at least the two neighboring bases are involved in the excitation as well[67]. Therefore, a better treatment of the backbone would also help in the correct treatment of the initially electronic excited state. Moreover, for the direct singlet pathway, it would be beneficial to run the dynamics using MS-CASPT2, as we could show that SA-CASSCF cannot correctly assign the character of the reactive state.

The work in this thesis also contained the development of the generalized ab initio multiple spawning (GAIMS) method. GAIMS is an extension to the standard multiple spawning framework, which allows for the treatment of intersystem crossing events during the dynamics. The new code was first tested on a one-dimensional model system and compared to both exact and surface hopping results. After this test system showed that the implementation works we simulated the excited state dynamics of H_2CS. The dynamics of this small organic molecule showed that the GAIMS code is capable of simulating ISC events in molecular systems. In order to further test the capabilities of the method, it is planned to combine it with GPU-accelerated electronic structure codes[68–71], and to compare the results from multi-dimensional molecular systems to other nonadiabatic molecular dynamics methods such as surface hopping or MCTDH[72].

APPENDIX: REPRINTED PUBLICATIONS

The reprints of the different publications forming the basis of this thesis can be found in the following appendix. Appendix A.1 contains the paper discussing the details on the GAIMS implementation and presenting the dynamics for both the one-dimensional test system and the H_2CS molecule, as discussed in chapter 3. Appendix A.2 covers the dynamics simulations for the direct T<>T formation pathway, as discussed in chapter 4.1.2. Both the static and dynamic results obtained for the photosensitized dimerization pathway (chapter 4.2) are presented in appendix A.3.

APPENDIX A.1 COMMUNICATION: GAIMS-GENERALIZED AB INITIO MULTIPLE SPAWNING FOR BOTH INTERNAL CONVERSION AND INTERSYSTEM CROSSING PROCESSES

BASILE F.E. CURCHOD, CLEMENS RAUER, PHILIPP MARQUETAND, LETICIA GONZÁLEZ AND TODD MARTÍNEZ

J. Chem. Phys., **144**, 101102 (2016).
http://dx.doi.org/10.1063/1.4943571

Contributions:

BASILE F.E. CURCHOD Concieved the idea of the project, planned and supervised the implementation, analyzed the results and wrote the initial draft of the manuscript.

CLEMENS RAUER Implemented the GAIMS code, analyzed the results and contributed to writing of the manuscript.

PHILIPP MARQUETAND Convieved the idea of the project, analyzed the results and contributed to the writing of the manuscript.

LETICIA GONZÁLEZ Analyzed the results and contributed to the writing of the manuscript.

TODD MARTÍNEZ Supervised the implementation, analyzed the results and contributed to the writing of the manuscript.

THE JOURNAL OF CHEMICAL PHYSICS **144**, 101102 (2016)

Communication: GAIMS—Generalized *Ab Initio* Multiple Spawning for both internal conversion and intersystem crossing processes

Basile F. E. Curchod,[1,2] Clemens Rauer,[3] Philipp Marquetand,[3] Leticia González,[3] and Todd J. Martínez[1,2,a)]

[1]*Department of Chemistry and the PULSE Institute, Stanford University, Stanford, California 94305, USA*
[2]*SLAC National Accelerator Laboratory, Menlo Park, California 94025, USA*
[3]*Institute of Theoretical Chemistry, University of Vienna, Währinger Straße 17, 1090 Vienna, Austria*

(Received 30 January 2016; accepted 26 February 2016; published online 11 March 2016)

Full multiple spawning is a formally exact method to describe the excited-state dynamics of molecular systems beyond the Born-Oppenheimer approximation. However, it has been limited until now to the description of radiationless transitions taking place between electronic states with the same spin multiplicity. This Communication presents a generalization of the full and *ab initio* multiple spawning methods to both internal conversion (mediated by nonadiabatic coupling terms) and intersystem crossing events (triggered by spin-orbit coupling matrix elements) based on a spin-diabatic representation. The results of two numerical applications, a model system and the deactivation of thioformaldehyde, validate the presented formalism and its implementation. © *2016 AIP Publishing LLC.* [http://dx.doi.org/10.1063/1.4943571]

I. INTRODUCTION

Nonadiabatic molecular dynamics—coupled electron-nuclear dynamics beyond the Born-Oppenheimer approximation—is a key technique for investigating photochemical and photophysical processes of molecules. However, most of the nonadiabatic methods commonly employed focus only on internal conversion (IC), i.e., electronic state transitions that conserve spin multiplicity. They often neglect *intersystem crossing* (ISC) events which couple electronic states with differing spin multiplicity due to relativistic spin-orbit coupling (SOC).[1] Far from being a curiosity, ISC plays an important role in the deactivation process of organic molecules[2–4] and metal complexes[5] that are used in energy-related devices.[6,7]

Apart from grid-based exact quantum dynamics for low dimensional problems,[8] few nonadiabatic dynamics methods have been extended to include SOC effects and, therefore, to describe ISC.[9] As one example, the Multiconfiguration Time-Dependent Hartree (MCTDH) technique has been used to understand the role of intersystem crossing in benzene photophysics.[10] Trajectory surface hopping (TSH), which approximates the dynamics of a nuclear wavepacket by a swarm of independent classical trajectories, has also been modified to incorporate SOC effects.[11–15] Freedom in the choice of different spin representations[11–14,16] or hopping schemes[17] has resulted in a plethora of TSH algorithms, which are, however, prone to shortcomings of the independent trajectory approximation.[18,19] In particular, the spin-diabatic representation for including SOC in TSH has been questioned due to a lack of rotational invariance for electronic population dynamics.[13] It can further lead to problematic results whenever sublevels are grouped in single multiplet states.[13]

An ideal method for describing both IC and ISC processes would be (i) derivable from *first principles*, (ii) able to treat medium to large molecular systems, (iii) compatible with on-the-fly calculations of electronic quantities, and (iv) able to adequately describe coherence and decoherence effects during nonadiabatic events. Full Multiple Spawning (FMS),[20,21] or its *ab initio* version AIMS,[22,23] fulfills all of the previously listed requirements, combing the computational efficiency of trajectory-based methods with quantum dynamics that is *in principle* exact (in the limit of a large enough basis set).

In short, FMS represents a nuclear wavefunction by a swarm of coupled frozen Gaussian functions following classical trajectories.[22–24] The number of trajectory basis functions used to describe the nuclear wavepacket is adapted during the dynamics—through so-called "spawning" events—to accurately describe wavepacket bifurcation in nonadiabatic regions. In the limit of a complete basis set and exact evaluation of all the requisite matrix elements, FMS constitutes a formally exact solution of the time-dependent Schrödinger equation (TDSE). In the following, we show how FMS and AIMS can be extended to the description of both IC and ISC processes and present the resulting Generalized FMS and AIMS (GFMS and GAIMS) methods and their implementation.

II. THEORY

The original derivation of the FMS method begins with the standard molecular Hamiltonian,[22] defined as the sum of the nuclear kinetic energy operator and the electronic Hamiltonian,

$$\begin{aligned}\hat{H}_{mol}(\mathbf{r},\mathbf{R}) &= \hat{T}_N + \hat{H}_{el}(\mathbf{r},\mathbf{R}) \\ &= \hat{T}_N + \hat{T}_e + \hat{V}_{ee}(\mathbf{r}) + \hat{V}_{eN}(\mathbf{r},\mathbf{R}) + \hat{V}_{NN}(\mathbf{R}),\end{aligned} \quad (1)$$

a)E-mail: toddjmartinez@gmail.com

where $\hat{T}_e$ is the electronic kinetic energy operator and $\hat{V}_{ee}(\mathbf{r})$, $\hat{V}_{eN}(\mathbf{r},\mathbf{R})$, and $\hat{V}_{NN}(\mathbf{R})$ are the electron-electron, electron-nucleus, and nucleus-nucleus potential energy operators, respectively. The collective variables for electronic and nuclear positions are denoted by $\mathbf{r}$ and $\mathbf{R}$, respectively. The Hamiltonian in Eq. (1) is non-relativistic and therefore neglects all terms related to both nuclear and electronic spin.

Including SOC effects in the framework of the time-dependent Schrödinger equation necessitates the derivation and approximation of new extended electronic Hamiltonians from the Dirac-Coulomb-Breit equation.[25] The Breit-Pauli Hamiltonian is commonly employed and its most important parts[26] comprise the non-relativistic Hamiltonian of Eq. (1), a scalar relativistic part containing the mass-velocity and Darwin terms, and a spin-orbit coupling Hamiltonian.[25,27] We note that scalar relativistic effects, when important, are often accounted for in quantum chemistry by using effective core potentials.[28] Incorporating the SOC Hamiltonian in the electronic Hamiltonian,[27,29–31] we obtain a molecular Hamiltonian accounting for the electronic spin $\mathbf{s}$,

$$\hat{H}(\mathbf{x},\mathbf{R}) = \hat{T}_N + \hat{H}_{el}(\mathbf{r},\mathbf{R}) + \hat{H}_{SOC}(\mathbf{x},\mathbf{R}) \tag{2}$$

with $\mathbf{x} = (\mathbf{r},\mathbf{s})$. Inserting Eq. (2) into the time-dependent Schrödinger equation leads to the starting equation for our extension of FMS including SOC,

$$i\frac{\partial \Psi(\mathbf{x},\mathbf{R},t)}{\partial t} = \hat{H}(\mathbf{x},\mathbf{R})\Psi(\mathbf{x},\mathbf{R},t). \tag{3}$$

The following derivation will use electronic states obtained from a spin-free electronic Hamiltonian as a basis, i.e., each electronic state is an eigenstate of both the total spin operator $\hat{S}^2$ and the spin projection operator $\hat{S}_z$.[13] The time-dependent molecular wavefunction is represented by a Born-Huang expansion[32] in the aforementioned spin-diabatic electronic basis,

$$\Psi(\mathbf{x},\mathbf{R},t) = \sum_J \sum_{M_{S_J}=-S_J}^{S_J} \Omega_J^{M_{S_J}}(\mathbf{R},t)\Phi_J^{M_{S_J}}(\mathbf{x};\mathbf{R}), \tag{4}$$

where $\Phi_J^{M_{S_J}}(\mathbf{x};\mathbf{R}) = |J,M_{S_J}\rangle$ is the electronic wavefunction for the Jth state with spin multiplicity $(2S_J+1)$ and spin-projection eigenvalue M_{S_J}. While the formalism is completely general, we focus our attention on singlet ($S=0$) and triplet ($S=1$) electronic states, the latter having M_S values of -1, 0, or $+1$. The FMS ansatz[20,21] for the nuclear wavefunction becomes

$$\Omega_J^{M_{S_J}}(\mathbf{R},t) = \sum_{k'=1}^{N_{J,M_{S_J}}(t)} C_{k'}^{J,M_{S_J}}(t)\chi_{k'}^{J,M_{S_J}}\left(\mathbf{R};\overline{\mathbf{R}}_{k'}^{J,M_{S_J}}(t), \overline{\mathbf{P}}_{k'}^{J,M_{S_J}}(t),\bar{\gamma}_{k'}^{J,M_{S_J}}(t),\boldsymbol{\alpha}_{k'}^{J,M_{S_J}}\right), \tag{5}$$

which expresses the nuclear wavefunction for the electronic state $|J,M_{S_J}\rangle$ as a linear combination of multidimensional frozen Gaussians $\chi_{k'}^{J,M_{S_J}}$ and corresponding complex coefficients $C_{k'}^{J,M_{S_J}}(t)$. Each term of the linear combination is called a trajectory basis function (TBF). The time-dependent position $\overline{\mathbf{R}}_{k'}^{J,M_{S_J}}(t)$ and momentum $\overline{\mathbf{P}}_{k'}^{J,M_{S_J}}(t)$ centers for each frozen Gaussian are propagated using classical Hamilton's equations, while the nuclear phase $\bar{\gamma}_{k'}^{J,M_{S_J}}(t)$ is time-evolved semiclassically.[22] We note here that the time-dependent parameters in the Gaussian functions could be time-evolved in different ways, leading to techniques for nonadiabatic dynamics like the direct dynamics variational multi-configurational Gaussian[33] (DD-vMCG) or the multiconfiguration-Ehrenfest (MCE) approach,[34] for example.

Inserting Eqs. (5) and (4) in Eq. (3) leads to a set of equations of motion for the complex amplitudes in Eq. (5). After left-projection by $\left\langle \chi_k^{I,M_{S_I}}\Phi_I^{M_{S_I}}\right|$, we obtain

$$\frac{d\mathbf{C}^{I,M_{S_I}}(t)}{dt} = -i\left(\mathbf{S}^{-1}\right)^{II,M_{S_I}M_{S_I}}\Bigg\{\left[\mathbf{H}^{II,M_{S_I}M_{S_I}} - i\dot{\mathbf{S}}^{II,M_{S_I}M_{S_I}}\right]\mathbf{C}^{I,M_{S_I}}(t) + \sum_J\sum_{M_{S_J}=-S_J}^{S_J}\mathbf{H}^{IJ,M_{S_I}M_{S_J}}\mathbf{C}^{J,M_{S_J}}(t)\Bigg\}, \tag{6}$$

where bold symbols indicate vectors or matrices in the basis of Gaussian functions. The overlap matrices are defined by $S_{k,k'}^{IJ,M_{S_I}M_{S_J}} = \left\langle \chi_k^{I,M_{S_I}}\middle|\chi_{k'}^{J,M_{S_J}}\right\rangle_{\mathbf{R}}\delta_{IJ}\delta_{M_{S_I}M_{S_J}}$ and $\dot{S}_{k,k'}^{IJ,M_{S_I}M_{S_J}} = \left\langle \chi_k^{I,M_{S_I}}\middle|\frac{\partial}{\partial t}\middle|\chi_{k'}^{J,M_{S_J}}\right\rangle_{\mathbf{R}}\delta_{IJ}\delta_{M_{S_I}M_{S_J}}$. A general Hamiltonian matrix element in Eq. (6) has the form

$$\begin{aligned} H_{kk'}^{IJ,M_{S_I}M_{S_J}} = {} & \left\langle \chi_k^{I,M_{S_I}}\middle|\hat{T}_N\middle|\chi_{k'}^{J,M_{S_J}}\right\rangle_{\mathbf{R}}\delta_{IJ}\delta_{M_{S_I}M_{S_J}} + \left\langle \chi_k^{I,M_{S_I}}\middle|E_I^{el}\middle|\chi_{k'}^{J,M_{S_J}}\right\rangle_{\mathbf{R}}\delta_{IJ}\delta_{M_{S_I}M_{S_J}} \\ & - \sum_{\rho=1}^{3N}\left\langle \chi_k^{I,M_{S_I}}\middle|\left(\mathbf{d}_{IJ}^{M_{S_I}M_{S_J}}\right)_\rho\frac{1}{m_\rho}\frac{\partial}{\partial R_\rho}\middle|\chi_{k'}^{J,M_{S_J}}\right\rangle_{\mathbf{R}}\delta_{M_{S_I}M_{S_J}} - \sum_{\rho=1}^{3N}\frac{1}{2m_\rho}\left\langle \chi_k^{I,M_{S_I}}\middle|\left(D_{IJ}^{M_{S_I}M_{S_J}}\right)_\rho\middle|\chi_{k'}^{J,M_{S_J}}\right\rangle_{\mathbf{R}}\delta_{M_{S_I}M_{S_J}} \\ & + \left\langle \chi_k^{I,M_{S_I}}\middle|\left\langle \Phi_I^{M_{S_I}}\middle|\hat{H}_{SOC}\middle|\Phi_J^{M_{S_J}}\right\rangle_{\mathbf{x}}\middle|\chi_{k'}^{J,M_{S_J}}\right\rangle_{\mathbf{R}}. \end{aligned} \tag{7}$$

The first two terms, as well as the electronic diagonal contributions from the fourth term, couple TBFs evolving on the same electronic state with the same sublevel (Fig. 1(a), dotted arrows). E_I^{el} is the electronic energy for state I, and m_ρ is the mass corresponding to nuclear degree of freedom ρ.

The third and fourth terms in Eq. (7) couple TBFs evolving on different electronic states but having the same S and M_S values (Fig. 1(a), dashed arrows). These terms depend on the first-order nonadiabatic coupling vectors $\mathbf{d}_{IJ}^{M_{S_I}M_{S_J}}$ and second-

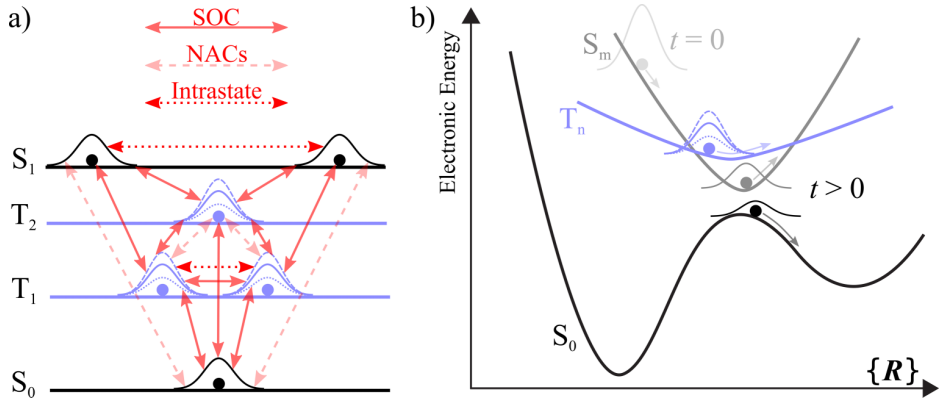


FIG. 1. (a) General representation of the coupling pattern between TBFs in GFMS for a case with two singlet and two triplet states. The Gaussian shape symbolizes the different TBFs (black = singlet, blue = triplet), where a continuous line is used for $M_S = 0$ states, while dashed and dotted lines represent $M_S = -1$ and $M_S = 1$ for triplet states. These TBFs evolve along a classical trajectory represented by the filled circle. Intrastate (dotted arrow) and nonadiabatic (dashed arrows) couplings are present between TBFs evolving in electronic states with the same spin multiplicity (as in the standard FMS method), while GFMS introduces an important number of additional couplings due to SOC (plain arrows). Note that a more detailed representation of the couplings requires the use of separate arrows for each possible coupling between sublevels (not shown). (b) Schematic representation of the GFMS method. A TBF is initiated in S_m at time $t = 0$ (gray) and will at a later time spawn TBFs both in T_n (blue) and in S_0 (black).

order nonadiabatic coupling $D_{IJ}^{M_{S_I} M_{S_J}}$, where the latter is usually neglected in nonadiabatic molecular dynamics.[22] The novelty of GFMS resides in the last term of Eq. (7) which allows for amplitude transfer between electronic states with different spin multiplicity, according to the rules of SOC (Fig. 1(a), continuous arrows), and fully preserves rotational invariance. In addition, this last term can also couple TBFs evolving on states that have the same spin multiplicity S, but only if the conditions $S_I = S_J > 0$ and $\Delta M_S = 0, \pm 1$ are fulfilled. It is important to note that any definition[1] of the operator $\hat{H}_{SOC}$ can be used in Eq. (7).

A key feature of the FMS method is that it uses an adaptive basis set to ensure an accurate description of nonadiabatic processes. The number of TBFs describing the nuclear wavefunction in state $|I, M_{S_I}\rangle$, $N_{I,M_{S_I}}(t)$, will indeed change in time as a result of spawning events. In short, a TBF entering a region of strong nonadiabaticity—detected using an effective coupling—can under certain conditions spawn a new TBF on the coupled electronic state (Fig. 1(b)). Upon spawning, the size of the matrices in Eq. (6) is extended and the resulting coupled propagation of the expanded set of TBFs allows for an exchange of nuclear amplitude between electronic states. For detailed discussions about the spawning algorithm between same-spin states, the reader is referred to previous works.[22,24,35] In GFMS, the spawning algorithm needs to be extended to allow for spawning between spin-orbit coupled states. Based on an already proposed effective coupling between diabatic states,[22] we suggest to measure the effective SOC strength between state I and state J at the position of TBF k as

$$\Lambda_{IJ}^{eff}(\mathbf{R}_k) = \frac{\left(\sum_{M_{S_I}=-S_I}^{S_I} \sum_{M_{S_J}=-S_J}^{S_J} \left| \left\langle \Phi_I^{M_{S_I}}(\mathbf{R}_k) \middle| \hat{H}_{SOC} \middle| \Phi_J^{M_{S_J}}(\mathbf{R}_k) \right\rangle_{\mathbf{x}} \right|^2 \right)^{1/2}}{\left| E_J^{el}(\mathbf{R}_k) - E_I^{el}(\mathbf{R}_k) \right|}. \tag{8}$$

This rotationally invariant spawning measure indicates the overall coupling between the sublevels of state I and those of state J. If $\Lambda_{IJ}^{eff}(\mathbf{R}_k)$ is higher than a certain threshold value, the spawning mode is triggered and a new TBF will be created in each sublevel of the electronic state J (see Fig. 1(b)).[36]

Ab initio multiple spawning[22–24] uses FMS nuclear dynamics combined with *ab initio* electronic structure calculations, allowing for an on-the-fly solution of the molecular time-dependent Schrödinger equation. Two approximations simplify the application of FMS to molecules, namely, the saddle-point and the independent first generation (IFG) approximations. The (zeroth-order) saddle-point approximation is used to compute the integrals over nuclear degrees of freedom that appear in the Hamiltonian matrix.[22] Extending the saddle-point approximation to the calculation of SOC matrix elements in Eq. (7) is straightforward,

$$\begin{aligned} &\left\langle \chi_k^{I,M_{S_I}} \middle| \left\langle \Phi_I^{M_{S_I}} \middle| \hat{H}_{SOC} \middle| \Phi_J^{M_{S_J}} \right\rangle_{\mathbf{x}} \middle| \chi_{k'}^{J,M_{S_J}} \right\rangle_{\mathbf{R}} \\ &= \left\langle \chi_k^{I,M_{S_I}} \middle| H_{SOC,IJ}^{M_{S_I} M_{S_J}} \middle| \chi_{k'}^{J,M_{S_J}} \right\rangle_{\mathbf{R}} \\ &\approx H_{SOC,IJ}^{M_{S_I} M_{S_J}}\left(\mathbf{R}_{kk'}^{(c)}\right) \left\langle \chi_k^{I,M_{S_I}} \middle| \chi_{k'}^{J,M_{S_J}} \right\rangle_{\mathbf{R}}, \end{aligned} \tag{9}$$

where $\mathbf{R}_{kk'}^{(c)}$ is the centroid position between TBFs k and k'. It is important to realize that the quality of the saddle-point approximation is expected to be especially good for SOC matrix elements, as they are usually slowly varying with respect to the nuclear position. The IFG approximation

proposes that the initial TBFs, whose initial positions and momenta are usually sampled from a Wigner distribution, are run independently, i.e., the interactions between initial TBFs are neglected. Each initial TBF however remains fully coupled with any child TBFs it will generate in the course of the nonadiabatic dynamics. For more information on these approximations, the interested reader is referred to different reviews on AIMS.[9,22,37,38]

To summarize, GFMS is a generalization to the *in principle* exact method FMS for the description of both IC and ISC processes. The GAIMS technique, which is amenable to molecules, is obtained by applying the IFG and saddle-point approximations to GFMS.

III. TEST APPLICATIONS

The proposed AIMS extension to SOC is first tested on a model system recently proposed by Persico and co-workers.[13] The model comprises a singlet (S_1) and a triplet state (T_1), which cross at $x = 10$ bohrs and both have a dissociative character (continuous curves in Figs. 2(a)-2(c)). All SOC matrix elements between the singlet and the triplet sublevels change sign at a given position, x_s, that can therefore be used as a parameter to tune the strength of intersystem crossing processes. For example, $x_s = 10$ bohrs leads to weak coupling between the electronic states, as the SOC is small around the point of crossing between the singlet and the triplet state (Fig. 2(a)). In contrast, when the sign change of the SOC takes place away from the states crossing point, e.g., $x_s = 8$ bohrs, the intersystem crossing is strong as $\left|H^{00}_{SOC,S_1T_1}\right|$ and $\left|H^{01}_{SOC,S_1T_1}\right| = \left|H^{0-1}_{SOC,S_1T_1}\right|$ are equal to 219.5 and 155.2 cm^{-1}, respectively, at the point of crossing (Fig. 2(c)). Therefore, varying x_s allows testing GAIMS for different SOC strength conditions.

The GAIMS dynamics is based on 200 initial conditions, sampled from the Wigner distribution of the initial Gaussian wavefunction, as defined in Ref. 13 (its corresponding probability density is represented in Fig. 2(a)). As mentioned in Sec. II, GAIMS uses both the IFG and the saddle-point approximation, meaning that perfect agreement with an exact solution is not expected. It is however only within these two approximations that GAIMS can be routinely applied to molecular systems, and the goal in the following is to validate the general accuracy of GAIMS with respect to exact calculations. GAIMS reproduces the exact results, obtained by solving the TDSE,[13] both qualitatively and quantitatively, within a maximum deviation of 7%, for three different cases of SOC coupling strength (right panels of Fig. 2). Furthermore, Fig. 2 shows the results for uncorrected spin-diabatic TSH dynamics,[13] with the three sublevels of the triplet state grouped into a single TSH amplitude. In this approximated formalism, TSH does not capture sign changes for the SOC and is unable to qualitatively describe ISC events, whenever the sign change occurs as the two states come close in energy. To fix this problem, a phase factor that forces a sign change in the effective SOC at the crossing point x_s can be added, leading to an excellent agreement with the exact result.[13] However, this simple fix is limited, since it requires *a priori* knowledge of x_s. Moreover, the contracted spin-diabatic approach to TSH is difficult to generalize for a larger number of electronic states.[13] In contrast, such problems do not exist in the GAIMS method, and its results for this one-dimensional system highlight the accuracy of the IFG and the saddle-point approximations. Importantly, this model system also validates the naive spawning criterion (Eq. (8)) used to spawn TBFs in

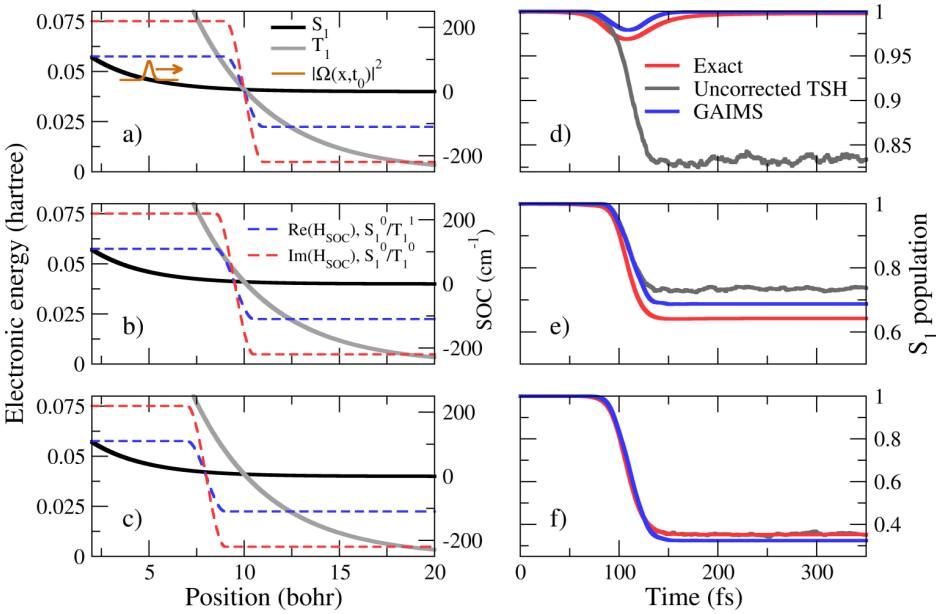


FIG. 2. GAIMS applied to a one-dimensional model system with one singlet and one triplet state. ((a)-(c)) Description of the model system—potential energy curves (black and gray lines), SOC matrix elements (dashed lines), for the case (a) $x_s = 10$ bohrs, (b) $x_s = 9.5$ bohrs, and (c) $x_s = 8$ bohrs. ((d)-(f)) Population in the singlet state for the three corresponding x_s values. GAIMS (blue) is compared with an exact solution of the time-dependent Schrödinger equation (red) and TSH in a spin-diabatic representation with no phase factor (gray), taken from Ref. 13.

a state with different spin multiplicity. All the runs leading to the results presented in Fig. 2 indeed underwent only one spawning event. This behavior contrasts with that of the spin-diabatic TSH dynamics, where a large number of *hops* between states are observed.[13] As discussed in detail in the AIMS literature,[22,35] improving the spawning criterion would surely result in an even better match with the exact result. We intend to investigate potential improvements in the spawning criteria for ISC in future research.

Having shown that GAIMS provides an accurate description of ISCs for different SOC strengths, we now move to an example of its real *raison d'être*, namely, the study of ISC processes in molecular systems. We choose for this purpose the nonadiabatic dynamics of thioformaldehyde, H_2CS, in its first singlet excited state. The first excited electronic state (S_1) of H_2CS has $n\pi^*$ character in the Franck-Condon region and its 0-0 transition energy has been experimentally determined at 2.033 eV.[39] Two triplet states lie close to S_1: $T_1(n\pi^*)$ and $T_2(\pi\pi^*)$. When the excited-state dynamics of H_2CS is initiated in S_1, we therefore expect to observe a direct application of the El-Sayed rules[40]—$S_1(n\pi^*)$ should be more strongly coupled to $T_2(\pi\pi^*)$ via SOC than to $T_1(n\pi^*)$ as a result of the change in orbital type. We performed GAIMS dynamics considering the first four electronic states of H_2CS (S_0, S_1, T_1, and T_2), using SA(4)-CASSCF(4/3)/6-31G* in Molpro[41] for 20 Wigner-sampled initial conditions. This level of theory places S_1 at 2.26 eV above the ground state in the Franck-Condon region. We present here the first 200 fs of dynamics in S_1 and will mostly comment on the GAIMS algorithm. This application is intended to provide a molecular test system for GAIMS dynamics and does not seek to obtain a complete and quantitatively accurate physical picture of the nonradiative relaxation of H_2CS (which might require larger basis sets and dynamic electron correlation).

The small yet sizable population of T_2 shortly after the beginning of the excited-state dynamics confirms the El-Sayed rule (Fig. 3), while T_1 appears to be only weakly populated in the same time window. This immediate, yet weak, population of triplet states upon photoexcitation in S_1 has also been observed for the parent molecule acetone,[42] although the population transfer is even weaker in this latter case. The total number of TBFs in the different electronic states grows quickly during the dynamics (orange line in Fig. 3), reaching a total value of 326 TBFs, among which 306 evolve in triplet states. From an initial TBF evolving in S_1, GAIMS rapidly starts to spawn TBFs in both T_2 and T_1, even though a sizable amount of population is eventually only transferred to T_2. To further analyze the dynamics, we present in Fig. 3 (upper panel) a depiction of the C=S bond length and electronic population for each of the TBFs. This representation highlights the growing number of TBFs in T_1 (red) and T_2 (blue) over the course of the simulation and depicts the different dynamics they experience, evolving in electronic states of differing electronic character. Hence, the dynamics of TBFs in T_1 is closer to those in S_1—both exhibiting an $n\pi^*$ character—while the $\pi\pi^*$ character of TBFs evolving in T_2 is consistent with the longer average C=S bond length. As noted before, population transfer between singlet and triplet TBFs due to SOC is limited in the time scale of this simulation and the projection is mostly dominated by the TBFs evolving in S_1.

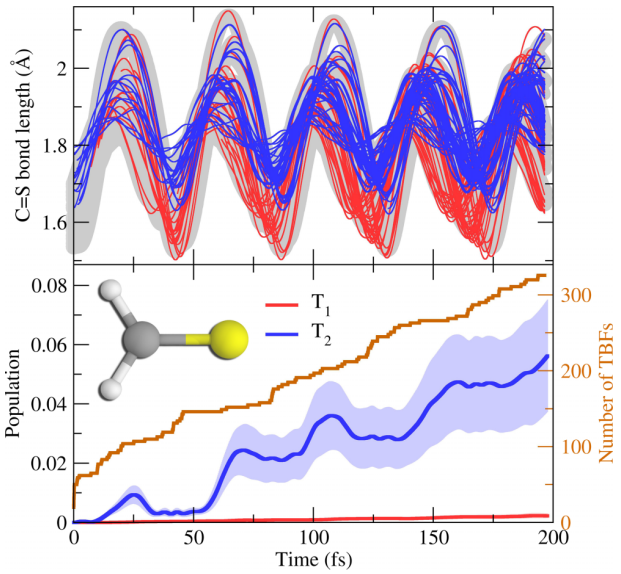


FIG. 3. GAIMS dynamics of thioformaldehyde after photoexcitation to S_1. Upper panel: C=S bond length for all TBFs produced during GAIMS dynamics. The width of each line is proportional to the population carried by the TBF. TBFs are associated with the S_1 (light gray), T_1 (red), or T_2 (blue) electronic state. Lower panel: population of the two triplet states during GAIMS dynamics averaged over 20 initial conditions (light area indicates the standard error). The total number of TBFs is given in orange.

IV. CONCLUSION

In this Communication, we presented a generalization of the full and *ab initio* multiple spawning methods to the description of internal conversion and intersystem crossing processes, both treated on an equal footing. The derivation of GFMS and GAIMS uses a spin-diabatic formalism and the implementation of GAIMS has been validated both with a model system and with a molecular application, the nonadiabatic dynamics of thioformaldehyde. This work will be followed by an extensive study of the interplay between the TBFs and the development of improved rules for ISC-triggered spawning that will minimize the number of unpopulated TBFs on triplet states. GAIMS opens the door for complete simulation of deactivation pathways in molecules and, when combined with GPU-accelerated electronic structure codes,[43–46] will be used to study the competition between internal conversion and intersystem crossing in both organic molecules and organometallic complexes.

ACKNOWLEDGMENTS

This work was supported by the AMOS program within the Chemical Sciences, Geosciences, and Biosciences Division of the Office of Basic Energy Sciences, Office of Science, US Department of Energy. The authors thank G. Granucci and M. Persico for sharing the numerical results reported in Ref. 13. B.F.E.C. acknowledges the Swiss National Science Foundation for the fellowship No. P2ELP2_151927

and C.R. thanks the uni:docs program of the University of Vienna for financing his Ph.D. and travel to Stanford.

[1]C. M. Marian, "Spin-orbit coupling and intersystem crossing in molecules," Wiley Interdiscip. Rev.: Comput. Mol. Sci. **2**, 187–203 (2012).
[2]D. S. N. Parker, R. S. Minns, T. J. Penfold, G. A. Worth, and H. H. Fielding, "Ultrafast dynamics of the S_1 excited state of benzene," Chem. Phys. Lett. **469**(1-3), 43–47 (2009).
[3]M. Richter, P. Marquetand, J. González-Vázquez, I. Sola, and L. González, "Femtosecond intersystem crossing in the DNA nucleobase cytosine," J. Phys. Chem. Lett. **3**(21), 3090–3095 (2012).
[4]L. Martínez-Fernández, I. Corral, G. Granucci, and M. Persico, "Competing ultrafast intersystem crossing and internal conversion: A time resolved picture for the deactivation of 6-thioguanine," Chem. Sci. **5**, 1336 (2014).
[5]M. Chergui, "On the interplay between charge, spin and structural dynamics in transition metal complexes," Dalton Trans. **41**(42), 13022–13029 (2012).
[6]H. Yersin, A. F. Rausch, R. Czerwieniec, T. Hofbeck, and T. Fischer, "The triplet state of organo-transition metal compounds. Triplet harvesting and singlet harvesting for efficient OLEDs," Coord. Chem. Rev. **255**(21), 2622–2652 (2011).
[7]T. Kinoshita, J. T. Dy, S. Uchida, T. Kubo, and H. Segawa, "Wideband dye-sensitized solar cells employing a phosphine-coordinated ruthenium sensitizer," Nat. Photonics **7**(7), 535–539 (2013).
[8]M. C. Heitz, C. Ribbing, and C. Daniel, "Spin-orbit induced radiationless transitions in organometallics: Quantum simulation of the intersystem crossing processes in the photodissociation of $HCo(CO)_4$," J. Chem. Phys. **106**(4), 1421–1428 (1997).
[9]M. Persico and G. Granucci, "An overview of nonadiabatic dynamics simulations methods, with focus on the direct approach versus the fitting of potential energy surfaces," Theor. Chem. Acc. **133**, 1526 (2014).
[10]R. S. Minns, D. S. N. Parker, T. J. Penfold, G. A. Worth, and H. H. Fielding, "Competing ultrafast intersystem crossing and internal conversion in the 'channel 3' region of benzene," Phys. Chem. Chem. Phys. **12**(48), 15607–15615 (2010).
[11]M. Richter, P. Marquetand, J. González-Vázquez, I. Sola, and L. González, "SHARC: *Ab initio* molecular dynamics with surface hopping in the adiabatic representation including arbitrary couplings," J. Chem. Theory Comput. **7**, 1253–1258 (2011).
[12]B. N. Fu, B. C. Shepler, and J. M. Bowman, "Three-state trajectory surface hopping studies of the photodissociation dynamics of formaldehyde on *ab initio* potential energy surfaces," J. Am. Chem. Soc. **133**(20), 7957–7968 (2011).
[13]G. Granucci, M. Persico, and G. Spighi, "Surface hopping trajectory simulations with spin-orbit and dynamical couplings," J. Chem. Phys. **137**(22), 22A501 (2012).
[14]G. L. Cui and W. Thiel, "Generalized trajectory surface-hopping method for internal conversion and intersystem crossing," J. Chem. Phys. **141**(12), 124101 (2014).
[15]F. F. de Carvalho and I. Tavernelli, "Nonadiabatic dynamics with intersystem crossings: A time-dependent density functional theory implementation," J. Chem. Phys. **143**(22), 224105 (2015).
[16]S. Mai, P. Marquetand, and L. Gonzalez, "A general method to describe intersystem crossing dynamics in trajectory surface hopping," Int. J. Quantum Chem. **115**(18), 1215–1231 (2015).
[17]K. Rajak and B. Maiti, "Trajectory surface hopping study of the $O(^3P) + C_2H_2$ reaction dynamics: Effect of collision energy on the extent of intersystem crossing," J. Chem. Phys. **140**(4), 044314 (2014).
[18]G. Granucci and M. Persico, "Critical appraisal of the fewest switches algorithm for surface hopping," J. Chem. Phys. **126**, 134114 (2007).
[19]J. C. Tully, "Molecular dynamics with electronic transitions," J. Chem. Phys. **93**, 1061–1071 (1990).
[20]M. Ben-Nun and T. J. Martínez, "Nonadiabatic molecular dynamics: Validation of the multiple spawning method for a multidimensional problem," J. Chem. Phys. **108**(17), 7244–7257 (1998).
[21]T. J. Martínez, M. Ben-Nun, and R. D. Levine, "Multi-electronic-state molecular dynamics: A wave function approach with applications," J. Phys. Chem. **100**(19), 7884–7895 (1996).
[22]M. Ben-Nun and T. J. Martínez, "*Ab initio* quantum molecular dynamics," Adv. Chem. Phys. **121**, 439–512 (2002).
[23]M. Ben-Nun, J. Quenneville, and T. J. Martínez, "*Ab initio* multiple spawning: Photochemistry from first principles quantum molecular dynamics," J. Phys. Chem. A **104**(22), 5161–5175 (2000).
[24]S. Yang and T. J. Martínez, "*Ab initio* multiple spawning: First principles dynamics around conical intersections," in *Conical Intersections: Theory, Computation and Experiment*, edited by W. Domcke, D. R. Yarkony, and H. Köppel (World Scientific Publishing Co. Pte. Ltd., 2011), Vol. 17, pp. 347–374.
[25]J. Almlöf and O. Gropen, "Relativistic effects in chemistry," in *Reviews in Computational Chemistry*, edited by K. B. Lipkowitz and D. B. Boyd (VCH Publishers, New York, 1996), Vol. 8, p. 203.
[26]The additional terms in the Breit-Pauli Hamiltonian are not considered here.
[27]D. G. Fedorov, S. Koseki, M. W. Schmidt, and M. S. Gordon, "Spin-orbit coupling in molecules: Chemistry beyond the adiabatic approximation," Int. Rev. Phys. Chem. **22**(3), 551–592 (2003).
[28]M. Dolg and X. Y. Cao, "Relativistic pseudopotentials: Their development and scope of applications," Chem. Rev. **112**(1), 403–480 (2012).
[29]D. R. Yarkony, "Spin-forbidden chemistry within the Breit-Pauli approximation," Int. Rev. Phys. Chem. **11**(2), 195–242 (1992).
[30]M. Reiher and A. Wolf, *Relativistic Quantum Chemistry: The Fundamental Theory of Molecular Science* (Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, 2009).
[31]K. G. Dyall and J. Knut Fægri, *Introduction to Relativistic Quantum Chemistry* (Oxford University Press, Inc., New York, 2007).
[32]M. Born and K. Huang, *Dynamical Theory of Crystal Lattices* (Clarendon, Oxford, 1954).
[33]G. A. Worth, M. A. Robb, and B. Lasorne, "Solving the time-dependent Schrödinger equation for nuclear motion in one step: Direct dynamics of non-adiabatic systems," Mol. Phys. **106**(16-18), 2077–2091 (2008).
[34]D. V. Shalashilin, "Quantum mechanics with the basis set guided by Ehrenfest trajectories: Theory and application to spin-boson model," J. Chem. Phys. **130**, 244101 (2009).
[35]S. Yang, J. D. Coe, B. Kaduk, and T. J. Martínez, "An 'optimal' spawning algorithm for adaptive basis set expansion in nonadiabatic dynamics," J. Chem. Phys. **130**(13), 134113 (2009).
[36]In the absence of an external magnetic field, the sublevels of a given electronic state with $S > 0$ are exactly degenerate. Therefore, the TBFs for different sublevels resulting from a given spawning event will follow the very same classical trajectory, implying that the corresponding complex coefficients can be evolved on the support of only one trajectory as shown in Figure 2. This observation leads to substantial computational savings for the propagation of the Gaussian functions. There is no approximation involved since all sublevel coefficients are still explicitly considered and time-evolved.
[37]M. Sulc, H. Hernandez, T. J. Martínez, and J. Vanicek, "Relation of exact Gaussian basis methods to the dephasing representation: Theory and application to time-resolved electronic spectra," J. Chem. Phys. **139**(3), 034112 (2013).
[38]T. J. Martínez and R. D. Levine, "Non-adiabatic molecular dynamics: Split-operator multiple spawning with applications to photodissociation," J. Chem. Soc., Faraday Trans. **93**(5), 941–947 (1997).
[39]D. J. Clouthier and D. A. Ramsay, "The spectroscopy of formaldehyde and thioformaldehyde," Annu. Rev. Phys. Chem. **34**, 31–58 (1983).
[40]M. A. El-sayed, "Triplet state. Its radiative and nonradiative properties," Acc. Chem. Res. **1**(1), 8 (1968).
[41]H. J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, and M. Schutz, "MOLPRO: A general-purpose quantum chemistry program package," Wiley Interdiscip. Rev.: Comput. Mol. Sci. **2**(2), 242–253 (2012).
[42]L. Favero, G. Granucci, and M. Persico, "Dynamics of acetone photodissociation: A surface hopping study," Phys. Chem. Chem. Phys. **15**(47), 20651–20661 (2013).
[43]E. G. Hohenstein, N. Luehr, I. S. Ufimtsev, and T. J. Martínez, "An atomic orbital-based formulation of the complete active space self-consistent field method on graphical processing units," J. Chem. Phys. **142**(22), 224103 (2015).
[44]J. W. Snyder, E. G. Hohenstein, N. Luehr, and T. J. Martínez, "An atomic orbital-based formulation of analytical gradients and nonadiabatic coupling vector elements for the state-averaged complete active space self-consistent field method on graphical processing units," J. Chem. Phys. **143**(15), 154107 (2015).
[45]E. G. Hohenstein, M. E. F. Bouduban, C. C. Song, N. Luehr, I. S. Ufimtsev, and T. J. Martínez, "Analytic first derivatives of floating occupation molecular orbital-complete active space configuration interaction on graphical processing units," J. Chem. Phys. **143**(1), 014111 (2015).
[46]B. S. Fales and B. G. Levine, "Nanoscale multireference quantum chemistry: Full configuration interaction on graphical processing units," J. Chem. Theory Comput. **11**(10), 4708–4716 (2015).

appendix a.2 cyclobutane thymine photodimerization mechanism revealed by nonadiabatic molecular dynamics

Clemens Rauer, Juan J. Nogueira, Philipp Marquetand and Leticia González

J. Am. Chem. Soc., **138**, 15911-15916 (2016).
http://dx.doi.org/10.1021/jacs.6b06701
J. Am. Chem. Soc. front cover Volume 138 Issue 49

Contributions:

Clemens Rauer performed the static and surface hopping simulations, analyzed the results and contributed to the initial draft of the manuscript.

Juan J. Nogueira performed the QM/MM simulations, analyzed the results and contributed to the initial draft of the manuscript.

Philipp Marquetand supervised the surface hopping simulations, analyzed the results and contributed to the initial draft of the manuscript.

Leticia González conceived, initiated and supervised the project, analyzed the results and contributed to the writing of the manuscript.

JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

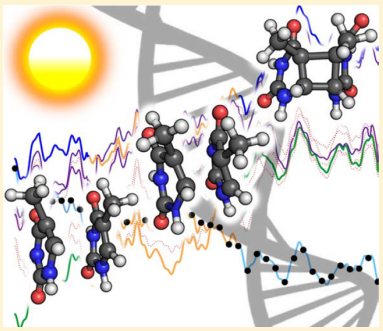
Cyclobutane Thymine Photodimerization Mechanism Revealed by Nonadiabatic Molecular Dynamics

Clemens Rauer, Juan J. Nogueira, Philipp Marquetand,* and Leticia González*

Institute of Theoretical Chemistry, Faculty of Chemistry, University of Vienna, Währinger Straße 17, 1090 Vienna, Austria

Supporting Information

ABSTRACT: The formation of cyclobutane thymine dimers is one of the most important DNA carcinogenic photolesions induced by ultraviolet irradiation. The long debated question whether thymine dimerization after direct light excitation involves singlet or triplet states is investigated here for the first time using nonadiabatic molecular dynamics simulations. We find that the precursor of this [2 + 2] cycloaddition reaction is the singlet doubly $\pi^2\pi^{*2}$ excited state, which is spectroscopically rather dark. Excitation to the bright $^1\pi\pi^*$ or dark $^1n\pi^*$ excited states does not lead to thymine dimer formation. In all cases, intersystem crossing to the triplet states is not observed during the simulated time, indicating that ultrafast dimerization occurs in the singlet manifold. The dynamics simulations also show that dimerization takes place only when conformational control happens in the doubly excited state.



1. INTRODUCTION

Exposure to ultraviolet (UV) radiation can damage DNA.[1] The deleterious action of UV light originates from the photochemical changes that lead to biological alterations. One harmful reaction occurring in DNA after light absorption that eventually causes skin cancer is the formation of intrastrand cyclobutane thymine dimers.[2] Although the existence of T <> T dimers was already reported in the 1960s,[3−5] the molecular details of this photoreaction still remain controversial. During the past decade, intensive experimental[6−10] and theoretical[11−18] research has debated the involvement of singlet or triplet electronic excited states in the formation of T <> T dimers. According to femtosecond (fs) time-resolved infrared experiments on thymine single strands, cyclobutane thymine dimers are fully formed in less than 1 ps as a result of a barrierless decay of the electronic singlet $^1\pi\pi^*$ state.[9,10] This view has been challenged by another fs time-resolved experiment,[6] which identified a triplet state as a major precursor of cyclobutane thymine dimer formation. Later time-resolved and steady-state experiments found the contribution of the $^3\pi\pi^*$ state to be smaller than 10%,[11] in agreement with the identification by nanosecond (ns) spectroscopy of a long-lived nonreactive triplet intermediate.[8]

Quantum chemical calculations have also attributed the ultrafast nature of this reaction to internal conversion from the lowest singlet $^1\pi\pi^*$ excited state to the ground state.[12,13,15,17] Interestingly, ref 12 characterizes this state as a doubly excited state, while others, e.g. refs 15, 17, report the state to be singly excited. Reaction pathways claiming the participation of the $^3\pi\pi^*$ state have also been calculated.[18,19] In addition to direct excitation, T <> T dimers can also form under triplet photosensitization. In the latter case, it is well established that the reaction proceeds via triplet states: After excitation, a photosensitizer undergoes intersystem crossing and then transfers its energy to a thymine molecule, thereby promoting the latter to its triplet state. This triplet thymine then reacts with another thymine in the ground state and gives the final adduct.[19−22] Despite the dimer yield being low for both sensitized and direct UV excitation cases,[21] the mechanism for the reaction in the absence or presence of a photosensitizer might be very different.

In this paper we investigate whether UV exposure triggers T <> T dimer formation in the singlet or triplet manifold. To this aim, we report the first ab initio nonadiabatic molecular dynamics simulations of a stacked thymine dimer including all relevant singlet and triplet states. The calculations are performed in gas phase using the surface-hopping including arbitrary couplings (SHARC) scheme[23,24] coupled to multiconfigurational state-average complete active space self-consistent field[25] (SA-CASSCF) energies, gradients, and spin−orbit couplings (see computational details in Section S1 of the Supporting Information (SI)).

2. RESULTS AND DISCUSSION

2.1. Dynamics Simulations. Formally, the formation of T <> T dimers is a [2 + 2] cycloaddition in which two adjacent stacked thymines are linked via the C_5−C_6 bonds forming a cyclobutane ring (see Figure 1). Insight about this and other concerted pericyclic reactions can be gained by means of the Woodward−Hoffmann rules.[26] These rules forbid a thermal dimerization in the ground state; however, they do not prescribe whether photochemical dimerization should proceed in the singly excited (^{1}SE) or doubly excited (^{1}DE) state. For

Received: June 29, 2016
Published: September 28, 2016

DOI: 10.1021/jacs.6b06701
J. Am. Chem. Soc. 2016, 138, 15911−15916

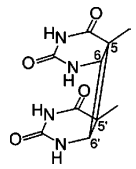


Figure 1. Cyclobutane thymine dimer with labeling of the forming bonds.

the sake of later discussion of the states involved in the dimerization of thymine, the orbital and singlet state correlation diagrams of the [2 + 2] cycloaddition of ethylene are shown in Figure 2. In agreement to the Woodward–Hoffmann rules,

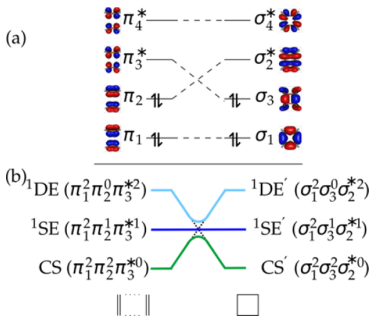


Figure 2. Woodward–Hoffmann orbital (a) and state (b) correlation diagrams for the [2 + 2] cycloaddition of two ethylene molecules on singlet surfaces. DE (doubly excited), SE (singly excited), and CS (closed-shell) states are indicated by cyan, blue, and green, respectively.

ethylene dimerization is thermally forbidden, as the ground state changes its character from one closed shell (CS) configuration, with orbital occupation $\pi_1^2\pi_2^2$, to another closed shell configuration, CS′($\sigma_1^2\sigma_3^2$). Photochemically, the reaction is allowed from the ^{1}SE or ^{1}DE state. During the course of the reaction, the involved states can be classified in two ways: (1) according to the orbital evolution and (2) according to the number of excited electrons. Looking at the first criterion, the ^{1}DE state evolves into the CS′ state (dashed curve in Figure 2b, called diabatic representation in the following). Analogously, the initial ground state CS evolves into the ^{1}DE′ state and ^{1}SE into ^{1}SE′. According to the second criterion, the number of excited electrons, the ^{1}DE corresponds to the ^{1}DE′ (solid curves in Figure 2b, called adiabatic representation in the following) and similarly CS to CS′ as well as ^{1}SE to ^{1}SE′. Dimerization of ethylene occurs from the singlet ^{1}DE state potential.[27−29] Interestingly, a direct comparison of the role of ^{1}SE and ^{1}DE states in the cyclobutane thymine dimer formation is absent in the literature, mostly because the majority of papers dealing with this reaction employ theoretical methods, such as density functional theory, unable to describe DE states.[11,18,30−32] Here, in contrast, we employ multiconfigurational methods which can describe DE states.

To answer the question about the precursor state, trajectories in full dimensionality were run starting from all eligible states: the ^{1}DE($\pi\pi^*$), ^{1}SE($\pi\pi^*$), and ^{1}nπ^* states. For each of these trajectories, the respective gradients and couplings determine which state will be subsequently populated; then, at every geometry, the energies of the other states are calculated vertically. The starting geometries of the trajectories are taken from the static calculations reported in ref 13. In these geometries, the initial separation of the C_5–C_5' and C_6–C_6' bonds is slightly above 3 Å, mimicking the shortest distances found in previous ground-state molecular dynamics simulations of DNA.[33]

We first discuss the dynamics started in the ^{1}DE state, for which 5 out of the 10 trajectories showed dimerization. We note that the purpose of these dynamics simulations is to find possible dimerization pathways and not to provide statistical information regarding reaction rates or time constants. In Figure 3, the evolution in time of the relevant energy levels

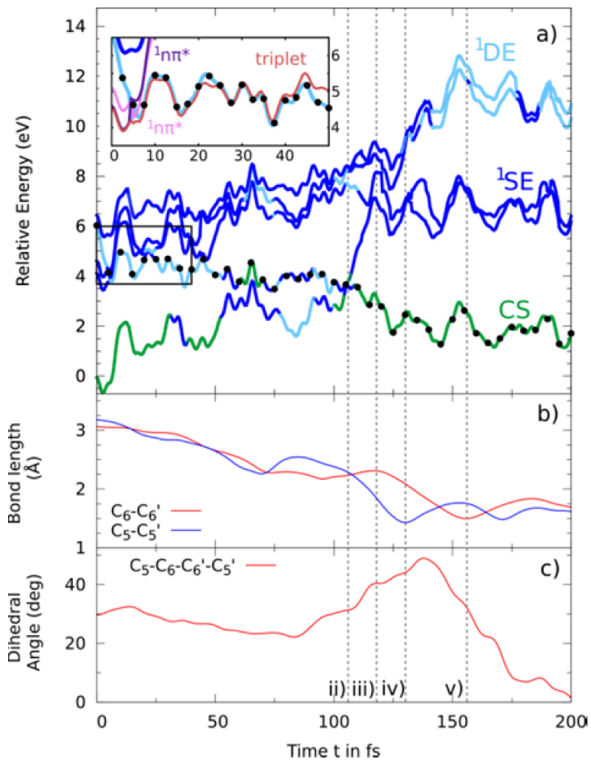


Figure 3. Time evolution of energy levels (a), the C_5–C_5' and C_6–C_6' distances (b), and the dihedral angle η(C_5–C_6–C_6'–C_5') between them (c) for a reactive trajectory started from the doubly excited state (^{1}DE). The black dots indicate which state is populated. The color code in panel (a) is similar to that in Figure 2: green for closed-shell (CS), blue for singly excited (SE) states (both nπ^* and $\pi\pi^*$), cyan for doubly excited (DE) states. The first 50 fs are zoomed in the inset. Diabatic states, including nπ^* states (pink and violet), $\pi\pi^*$ states (blue (SE) and cyan (DE)), and triplet states (red), are displayed. Vertical lines in panels (a–c) indicate the times corresponding to the geometries displayed in Figure 4.

(panel a), the C_5–C_5' and C_6–C_6' distances (b), and the dihedral angle η(C_5–C_6–C_6'–C_5') (c) are shown for one example reactive trajectory (see also Figure S4). Following the diabatic states along the reaction is only possible at the beginning of the trajectory (inset of Figure 3a), when the separation between the two nucleobases is still large and the identification of the orbitals is clear. Once the C–C bonds start to form, the four frontier π orbitals evolve toward four σ orbitals (recall Figure 2) and other remaining orbitals become energetically degenerated, undergoing strong mixing. As a consequence, diabatic characterization of the electronic states is

DOI: 10.1021/jacs.6b06701
J. Am. Chem. Soc. 2016, 138, 15911−15916

hardly possible and Figure 3a shows the adiabatic evolution of the states. In this representation, the electronic states can have closed-shell, singly excited, or doubly excited character (see more details in Section S1.3 of the SI).

The underlying mechanism for the reactive trajectory elegantly complies with the simple Woodward–Hoffmann correlation diagrams discussed for the [2 + 2] cycloaddition of ethylene: The system initially evolves on the ^{1}DE state during the first 50 fs. Then, the character of the populated electronic state becomes adiabatically singly excited. After ca. 120 fs, the populated state is of closed-shell character, which corresponds to the CS′ (ground state) in Figure 2. The inset of Figure 3a shows the further complexity involved in the deactivation: At the considered initial geometry (Figure 4i), the ^{1}DE state

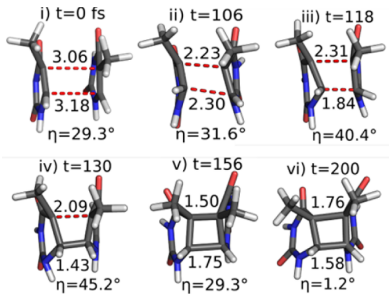


Figure 4. Geometrical snapshots of the reactive trajectory depicted in Figure 3. C_5–C_5' and C_6–C_6' distances in angstroms and dihedral angles $\eta(C_5$–C_6–C_6'–$C_5')$ in degrees.

corresponds to the S_3 state, which crosses after a few femtoseconds with two electronic states (mainly of ^{1}nπ^* character) and thereby becomes adiabatically the S1 state within less than 50 fs. The hopping geometries for the S3/S2 and S_2/S_1 crossing points are very similar to the initial geometry because this geometry is already quite close to the crossing area. The deactivation to the ground state proceeds via the S_1/S_0 conical intersection (Figure 4ii), which is different from the crossing point reported by Blancafort et al.[12] (probably due to a differing reactant geometry) but identical to the crossing point proposed by Boggio-Pasqua et al.[13] At this geometry, both C_5–C_5' and C_6–C_6' distances are very similar, but the C_5–C_5' bond is shorter indicating the preference for the first bond to be formed. Indeed, the C_5–C_5' bond is formed earliest in all the reactive trajectories. The cyclobutane ring is then only fully created in the electronic ground state. The evolution of the C_5–C_5' and C_6–C_6' distances (Figure 3b) shows that the C_5–C_5' bond is established first (Figures 4iii and 4iv) and, after ca. 30 fs, the C_6–C_6' bond formation follows (Figure 4v), illustrating the "concertedness" of this pericyclic reaction, in the spirit of the Woodward–Hoffmann rules. The snapshot at 200 fs (Figure 4vi) shows how the newly formed bonds asynchronously alternate in time, due to excess vibrational energy.

Dimerizable conformers at the chosen initial geometry have a dihedral angle close to that of B-DNA (η = 33°).[34] During the reaction, η decreases concomitantly with the C_5–C_5' and C_6–C_6' distances (Figure 3b–c). This is the case for all reactive trajectories (see Figure S2), which hints at conformational control, as suggested e.g. in refs 29, 33. Around the S_1/S_0 conical intersection, when the two stacked thymines are separated by ca. 2.3 Å, the dihedral angle starts increasing to avoid sterical clashes. After deactivation to the ground state, the dihedral angle continues increasing to 45°, until the C_5–C_5' bond is formed (Figure 4iv). As the cyclobutane ring develops the two thymines get perfectly aligned in space (see Figure 3c and Figure 4vi).

In the nonreactive trajectories starting from the ^{1}DE state, the system quickly relaxes to the ^{1}nπ^* states and stays in the lowest one, which is then the S1 state (see one example in Figure 5a and Figure S5). The characterization of electronic

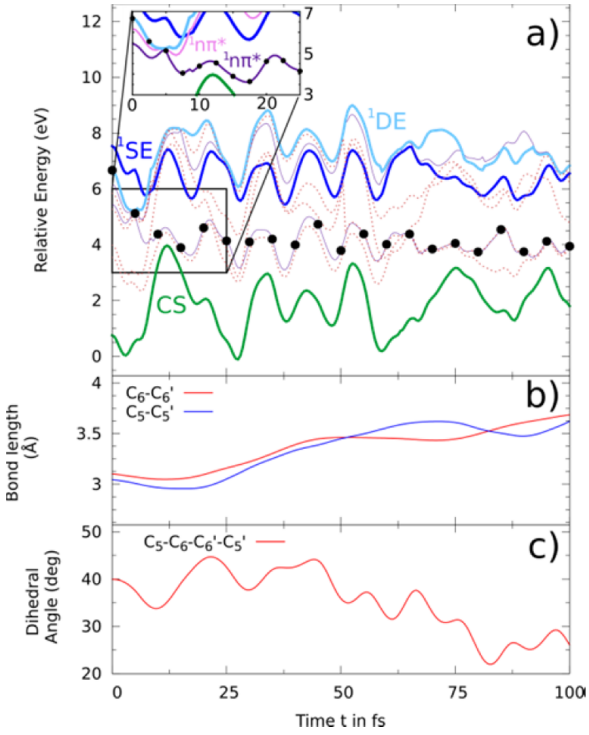


Figure 5. Time evolution of energy levels (a), the C_5–C_5' and C_6–C_6' distances (b) and the dihedral angle $\eta(C_5$–C_6-C_6'-$C_5')$ between them (c) for a nonreactive trajectory started from the double excited state (^{1}DE). The black dots indicate which state is populated. The color code is the same as that in the inset of Figure 3a.

states for nonreactive trajectories is easier than for the reactive ones since due to the large separation between the two thymines, there is no orbital mixing. Accordingly, the states of Figure 5a can be plotted in the diabatic representation, where the character of the wavefunction is maintained. In most trajectories, during the first 50 fs the thymines move apart and η increases (Figure S3), preventing dimerization and suggesting that indeed conformational control steers the reaction. However, the nonreactive trajectory of Figure 5 serves as a counterexample. The general trend of η is a continuous decrease (Figure 5c), which nonetheless does not lead to dimerization (Figure 5b) because the system is in the ^{1}nπ^* state. This behavior demonstrates not only that conformational control is not the only factor determining T <> T dimerization but also that electronic effects play a key role. Our simulations predict that only when the trajectories occupy the state with double excitation character and the dihedral angle decreases, dimerization takes place.

DOI: 10.1021/jacs.6b06701
J. Am. Chem. Soc. 2016, 138, 15911–15916

All the trajectories initiated in the bright ^{1}SE state (S_4) are not reactive, at least during the first 200 fs in which we propagated. In all cases (see one example in Figure S6), the system stays in the S4 for the complete propagation time without undergoing internal conversion to lower states, while C_5-C_5' and C_6-C_6' distances increase. Although we did not expect dimerization in the $^1n\pi^*$ states, trajectories were also run from the lowest $^1n\pi^*$ state, which is the S_1 state (Figure S7), but indeed T <> T is not formed and the system also does not relax to the ground state in the time simulated.

A key question is whether triplet electronic states are involved in the dimerization. In principle, it is conceivable that a triplet state is populated by singlet fission,[35] where a monomer in a singlet excited state transfers part of its electronic energy to a neighboring ground state monomer, resulting in the population of low lying triplet states of both monomers. As a result, the whole system has double-excitation character and singlet spin multiplicity. This process is energetically favorable only if the triplet excitation energy of the monomer is half of the singlet electronic state energy of the dimer. According to literature calculations[36] performed at the complete active space second-order perturbation (CASPT2) level of theory, the energies of the bright singlet state of the thymine−thymine reactant and the lowest triplet state of a single thymine at the Franck−Condon geometry are 4.89 and 3.59 eV, respectively. Since the energy requirements are not fulfilled, singlet fission is not possible in this region. However, the energies of the electronic excited states can strongly change along the dynamics, and the energetic requirement for singlet fission could be met at certain geometries. However, the energetic criterion is not the only requirement for singlet fission, but also a particular electronic configuration is needed. To investigate the latter, we have analyzed the character of the ^{1}DE state in the reactive trajectories. By looking at the wave function, it is possible to identify whether after excitation the electrons are located in both monomers (Figure 6a−b) or in one (Figure 6c). Only the electronic configuration of Figure 6a complies with the definition of singlet fission. Unfortunately, configurations 6a and 6b cannot be distinguished in our analysis. Therefore, it will be assumed that both of them are products of singlet fission, overestimating the contribution of this process. During the time the ^{1}DE state is populated (from 0 to 50 fs), we obtain a maximum of 35% of singlet fission probability. This means that the probability of the double excitation being located in one monomer is, at least, 65% and thus singlet fission only plays a minor role.

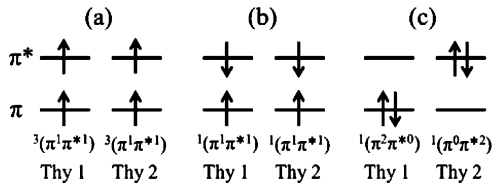


Figure 6. Electronic configurations of the thymine monomers that can compose the singlet doubly excited state ^{1}DE in the dimer.

The second way to populate a triplet state is via intersystem crossing, as proposed by Kwok et al.[6] The present simulations show that even when the potentials of the populated singlet state and a triplet state lie close in energy for a long period of the trajectory, the system does not undergo intersystem crossing due to the very small spin−orbit coupling. For instance (see inset of Figure 3a), at 15 fs, the ^{1}DE and the triplet T_3 state are separated by 0.032 eV (258 cm^{-1}) and the corresponding spin−orbit coupling is 0.03 cm^{-1}. Note that a CASPT2 calculation gives values of 0.33 eV (2662 cm^{-1}) and 0.19 cm^{-1}, respectively, so that the ratio between energy gap and coupling is of the same order of magnitude in both cases, and it does not change the conclusion that ISC does not occur. The absence of intersystem crossing is observed in all the trajectories (see SI), regardless of being reactive or nonreactive. We therefore conclude that the triplet states do not play a role in the ultrafast dimerization pathway induced by UV excitation, in agreement with Kohler and co-workers.[10]

2.2. Effect of the Environment. Our nonadiabatic molecular dynamics simulations have been performed in the gas phase. However, the cyclobutane thymine dimerization takes place in biological systems and the experimental measurements are generally performed for solvated single strands.[6−11] To investigate the effect of the biological environment, we have selected three snapshots from the reactive trajectory shown in Figure 3 at simulation times 10, 100, and 190 fs, i.e. in the reactant, S_1/S_0 conical intersection, and product regions, respectively. The three dimer structures were embedded in a solvated single strand $(dT)_{12}$, and the excitation energies were calculated by an electrostatic quantum mechanics/molecular mechanics (QM/MM) scheme.[37] The reactive thymine dimer was treated at the SA-CASSCF level,[25] and the environment was described by the Amber force field.[38] Further details can be found in Section S1.4 of the SI.

The QM/MM excitation energies (Table S1) show that the environment has a minor effect in the regions of reactants and S_1/S_0 conical intersection. The electrostatic interactions with the strand and solvent induce a small shift of ca. 0.1 eV in the $^{1,3}\pi\pi^*$ states, which according to the dynamics simulations are the relevant states. The effect of the environment is more important for the products, changing the order of some highly excited states. However, this is irrelevant because when the reaction is close to the product region, the system is already in the electronic ground state, which is well separated from the excited states. The negligible effect of the electrostatic interactions on the dimerization process agrees with previous density functional theory calculations[11] performed in gas-phase and water solution using a polarizable continuum model. However, this does not mean that the environment has a passive role. For example, QM/MM calculations including four nucleobases in the QM region predicted that thymine dimerization can be quenched by electron transfer from a flanking guanine nucleobase.[15,17]

Our QM/MM calculations also demonstrate that the triplet states do not participate in the dimerization process, corroborating the gas phase dynamical simulations. The singlet/triplet energy gaps and spin−orbit couplings in the presence of the biological environment are very similar to the gas phase values. For example, at 10 fs, the energy gap and spin−orbit coupling between the relevant ^{1}DE state and the closest triplet state T_3 are 0.12 eV (968 cm^{-1}) and 0.54 cm^{-1}, in the solvated strand versus 0.10 eV (807 cm^{-1}) and 0.59 cm^{-1}, respectively, in the gas phase. Note that the spin−orbit coupling between the same states within the Franck−Condon region, as calculated from 20 different conformations from the QM/MM dynamics, is only of the order of 10^{-2} cm^{-1} for all the geometries.

DOI: 10.1021/jacs.6b06701
J. Am. Chem. Soc. 2016, 138, 15911−15916

2.3. Populating the Doubly Excited State. We have shown that thymine dimerization requires the population of a singlet doubly excited state ^{1}DE. However, due to its double excitation character, this state lies high in energy, and the pathway to populate it is still an open question. Two possibilities can be conceived: (i) direct excitation under UV radiation or (ii) nonradiative relaxation from the bright singly excited state ^{1}SE. In order to investigate the first route, we have calculated the absorption spectrum of the thymine dimer embedded in a solvated single strand $(dT)_{12}$ using a QM/MM scheme and 250 geometries sampled from ground state molecular dynamics. The two nucleobases in the middle of the strand were described by the multistate CASPT2 (MS-CASPT2) method[39] and the rest of the system by the Amber force field.[38] All the states were characterized as singly or doubly excited states by analyzing their one-electron transition density matrix.[40,41] More details can be found in Section S1.5 of the SI. The resulting absorption spectrum decomposed into ^{1}SE and ^{1}DE states is shown in Figure 7. The ^{1}DE state lies at

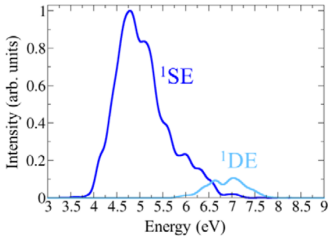


Figure 7. Absorption spectrum of the ^{1}SE and ^{1}DE states of the thymine dimer embedded in a $(dT)_{12}$ single strand calculated at MS-CASPT2/Amber level of theory.

energies larger than 6 eV, which is well above the usual experimental excitation energy[6,9,10] of 4.5 eV. We conclude, therefore, that the population of the ^{1}DE state by direct excitation is not possible. Accordingly, the ^{1}DE state should be populated nonradiatively from the bright ^{1}SE state through a ^{1}SE/^{1}DE conical intersection or through crossings with other electronic states. Since the ^{1}SE and ^{1}DE states are almost degenerate at the beginning of the dynamics (see inset of Figure 3a), we believe that there should be a conical intersection close to the Franck–Condon region although we were not able to locate it because of the large configurational space involved. Moreover, since the yield of the reaction is less than 5%,[9−11] this conical intersection should be slightly above the energy of the bright ^{1}SE state, as sketched in Figure 8. Note that this mechanism is different from the proposed barrierless

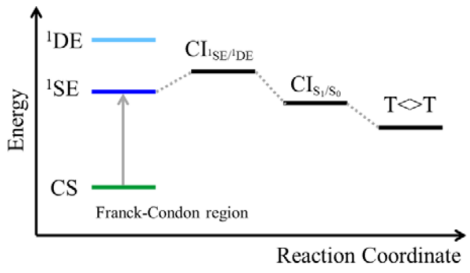


Figure 8. Suggested mechanism to populate the ^{1}DE state from ^{1}SE, and subsequent evolution of the dimerization.

pathways from refs 11−13, 15−17, where the low yield of the reaction is then attributed to conformational control in the ground state originally suggested in refs 29, 33. After populating the ^{1}DE state, our dynamics simulations have shown that the system evolves along the reaction coordinate toward the S_1/S_0 conical intersection, from where the cyclobutane dimer is formed.

3. CONCLUSIONS

In conclusion, our ab initio molecular dynamics simulations show that ultrafast cyclobutane thymine dimerization requires populating the singlet doubly excited state, ^{1}DE. The reaction occurs without the participation of triplet states, neither by intersystem crossing nor by singlet fission. We also show that, besides conformational control, electronic effects are essential to promote cyclobutane formation. The doubly excited state cannot be populated by direct excitation because it lies at a higher energy than the commonly employed laser wavelengths or the UV spectrum of the sun at the earth's surface. We propose that the population of the reactive state should happen by internal conversion from the bright singly excited state, possibly by overcoming a small energy barrier, thus explaining the low yield of the reaction. However, further calculations are necessary to clarify this point.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.6b06701.

Further information on computational details, the time evolution of energies and important coordinates, and the S_1/S_0 conical intersection geometry are given (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*philipp.marquetand@univie.ac.at

*leticia.gonzalez@univie.ac.at

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

C.R. gratefully acknowledges the University of Vienna within the uni:docs programme for financial support. The computational results have been achieved in part using the Vienna Scientific Cluster (VSC). The authors thank M. Boggio-Pasqua for fruitful discussions.

■ REFERENCES

(1) Friedberg, E. C. *Nature* **2003**, *421*, 436−440.
(2) Mouret, S.; Baudouin, C.; Charveron, M.; Favier, A.; Cadet, J.; Douki, T. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 13765−13770.
(3) Beukers, R.; Berends, W. *Biochim. Biophys. Acta* **1960**, *41*, 550−551.
(4) Setlow, R. B.; Setlow, J. K. *Proc. Natl. Acad. Sci. U. S. A.* **1962**, *48*, 1250−1257.
(5) Setlow, R. B.; Swenson, P. A.; Carrier, W. L. *Science* **1963**, *142*, 1464−1466.
(6) Kwok, W. M.; Ma, C.; Phillips, D. L. *J. Am. Chem. Soc.* **2008**, *130*, 5131−5139.
(7) Marguet, S.; Markovitsi, D. *J. Am. Chem. Soc.* **2005**, *127*, 5780−5781.
(8) Pilles, B. M.; Bucher, D. B.; Liu, L.; Clivio, P.; Gilch, P.; Zinth, W.; Schreier, W. J. *J. Phys. Chem. Lett.* **2014**, *5*, 1616−1622.

DOI: 10.1021/jacs.6b06701
J. Am. Chem. Soc. 2016, 138, 15911−15916

(9) Schreier, W. J.; Kubon, J.; Regner, N.; Haiser, K.; Schrader, T. E.; Zinth, W.; Clivio, P.; Gilch, P. *J. Am. Chem. Soc.* **2009**, *131*, 5038−5039.
(10) Schreier, W. J.; Schrader, T. E.; Koller, F. O.; Gilch, P.; Crespo-Hernández, C. E.; Swaminathan, V. N.; Carell, T.; Zinth, W.; Kohler, B. *Science* **2007**, *315*, 625−629.
(11) Banyasz, A.; Douki, T.; Improta, R.; Gustavsson, T.; Onidas, D.; Vayá, I.; Perron, M.; Markovitsi, D. *J. Am. Chem. Soc.* **2012**, *134*, 14834−14845.
(12) Blancafort, L.; Migani, A. *J. Am. Chem. Soc.* **2007**, *129*, 14540−14541.
(13) Boggio-Pasqua, M.; Groenhof, G.; Schäfer, L. V.; Grubmüller, H.; Robb, M. A. *J. Am. Chem. Soc.* **2007**, *129*, 10996−10997.
(14) González-Ramírez, I.; Roca-Sanjuán, D.; Climent, T.; Serrano-Pérez, J. J.; Merchán, M.; Serrano-Andrés, L. *Theor. Chem. Acc.* **2011**, *128*, 705−711.
(15) Lee, W.; Matsika, S. *Phys. Chem. Chem. Phys.* **2015**, *17*, 9927−9935.
(16) Serrano-Pérez, J. J.; González-Ramírez, I.; Coto, P. B.; Merchán, M.; Serrano-Andrés, L. *J. Phys. Chem. B* **2008**, *112*, 14096−14098.
(17) Spata, V. A.; Lee, W.; Matsika, S. *J. Phys. Chem. Lett.* **2016**, *7*, 976−984.
(18) Zhang, R. B.; Eriksson, L. A. *J. Phys. Chem. B* **2006**, *110*, 7556−7562.
(19) Climent, T.; González-Ramírez, I.; González-Luque, R.; Merchán, M.; Serrano-Andrés, L. *J. Phys. Chem. Lett.* **2010**, *1*, 2072−2076.
(20) Cuquerella, M. C.; Lhiaubet-Vallet, V.; Bosca, F.; Miranda, M. A. *Chem. Sci.* **2011**, *2*, 1219−1232.
(21) Liu, L.; Pilles, B. M.; Gontcharov, J.; Bucher, D. B.; Zinth, W. *J. Phys. Chem. B* **2016**, *120*, 292−298.
(22) Miro, P.; Lhiaubet-Vallet, V.; Marin, M. L.; Miranda, M. A. *Chem. - Eur. J.* **2015**, *21*, 17051−17056.
(23) Mai, S.; Marquetand, P.; González, L. *Int. J. Quantum Chem.* **2015**, *115*, 1215−1231.
(24) Richter, M.; Marquetand, P.; González-Vázquez, J.; Sola, I.; González, L. *J. Chem. Theory Comput.* **2011**, *7*, 1253−1258.
(25) Knowles, P. J.; Werner, H. J. *Chem. Phys. Lett.* **1985**, *115*, 259−267.
(26) Hoffmann, R.; Woodward, R. B. *Acc. Chem. Res.* **1968**, *1*, 17−22.
(27) Bernardi, F.; Olivucci, M.; Robb, M. A. *Acc. Chem. Res.* **1990**, *23*, 405−412.
(28) Caldwell, R. A. *J. Am. Chem. Soc.* **1980**, *102*, 4004−4007.
(29) McCullagh, M.; Hariharan, M.; Lewis, F. D.; Markovitsi, D.; Douki, T.; Schatz, G. C. *J. Phys. Chem. B* **2010**, *114*, 5215−5221.
(30) Ando, H.; Fingerhut, B. P.; Dorfman, K. E.; Biggs, J. D.; Mukamel, S. *J. Am. Chem. Soc.* **2014**, *136*, 14801−14810.
(31) Barbatti, M. *ChemPhysChem* **2014**, *15*, 3342−3354.
(32) Zhang, W.; Yuan, S.; Li, A.; Dou, Y.; Zhao, J.; Fang, W. *J. Phys. Chem. C* **2010**, *114*, 5594−5601.
(33) Law, Y. K.; Azadi, J.; Crespo-Hernández, C. E.; Olmon, E.; Kohler, B. *Biophys. J.* **2008**, *94*, 3590−3600.
(34) Shui, X.; McFail-Isom, L.; Hu, G. G.; Williams, L. D. *Biochemistry* **1998**, *37*, 8341−8355.
(35) Smith, M. B.; Michl, J. *Chem. Rev.* **2010**, *110*, 6891−6936.
(36) Serrano-Pérez, J. J.; González-Luque, R.; Merchán, M.; Serrano-Andrés, L. *J. Phys. Chem. B* **2007**, *111*, 11880−11883.
(37) Melaccio, F.; Olivucci, M.; Lindh, R.; Ferré, N. *Int. J. Quantum Chem.* **2011**, *111*, 3339−3346.
(38) Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049−1074.
(39) Finley, J.; Malmqvist, P. A.; Roos, B. O.; Serrano-Andrés, L. *Chem. Phys. Lett.* **1998**, *288*, 299−306.
(40) Plasser, F.; Lischka, H. *J. Chem. Theory Comput.* **2012**, *8*, 2777−2789.
(41) Plasser, F.; Wormit, M.; Dreuw, A. *J. Chem. Phys.* **2014**, *141*, 024106.

DOI: 10.1021/jacs.6b06701
J. Am. Chem. Soc. 2016, 138, 15911−15916

APPENDIX A.3 STEPWISE PHOTOSENSITIZED THYMINE DIMERIZATION MEDIATED BY AN EXCITON INTERMEDIATE

CLEMENS RAUER, JUAN J. NOGUEIRA, PHILIPP MARQUETAND AND LETICIA GONZÁLEZ

Published Online, *Monatsh. Chem.* (2017).
https://doi.org/10.1007/s00706-017-2108-4

Contributions:

CLEMENS RAUER performed the static and surface hopping simulations, analyzed the results and contributed to the initial draft of the manuscript.

JUAN J. NOGUEIRA performed the QM/MM simulations, analyzed the results and contributed to the initial draft of the manuscript.

PHILIPP MARQUETAND supervised the surface hopping simulations, analyzed the results and contributed to the initial draft of the manuscript.

LETICIA GONZÁLEZ conceived, initiated and supervised the project, analyzed the results and contributed to the writing of the manuscript.

Monatsh Chem
https://doi.org/10.1007/s00706-017-2108-4

ORIGINAL PAPER

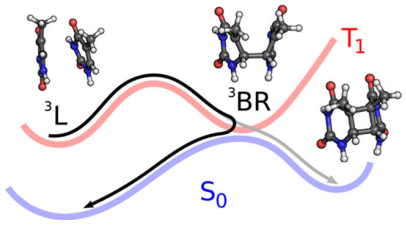
Stepwise photosensitized thymine dimerization mediated by an exciton intermediate

Clemens Rauer[1] · **Juan J. Nogueira**[1] · **Philipp Marquetand**[1] · **Leticia González**[1]

Received: 16 October 2017 / Accepted: 20 November 2017

Abstract Cyclobutane thymine dimerization is the most prominent DNA photoinduced damage. While the ultrafast mechanism that proceeds in the singlet manifold is nowadays well established, the triplet-state pathway is not completely understood. Here we report the underlying mechanism of the photosensitized dimerization process in the triplet state. Quantum chemical calculations, combined with wavefunction analysis, and nonadiabatic molecular dynamics simulations demonstrate that this is a stepwise reaction, traversing a long-lived triplet biradical intermediate, which is characterized as a Frenkel exciton with very small charge-transfer character. The low yield of the reaction is regulated by two factors: (i) a relatively large energy barrier that needs to be overcome to form the exciton intermediate, and (ii) a bifurcation of the ground-state potential-energy surface that mostly leads back to the Franck–Condon region because dimerization requires a very restricted combination of coordinates and velocities at the event of non-radiative decay to the ground state.

Graphical abstract



Keywords DNA · Thymine dimerization · Quantum chemical calculations · Non-adiabatic dynamics · Wavefunction analysis · Charge transfer

Introduction

The formation of cyclobutane thymine T⟨⟩T dimers between two adjacent thymine bases is the most frequent DNA damage under UV radiation [1]. This photolesion, which can take place in both the singlet and triplet manifolds, has been extensively investigated spectroscopically [2–7] and computationally [8–15]. The triplet pathway is a much slower process [7] and exhibits a smaller yield [6, 16] than the singlet mechanism due to inefficient intersystem crossing. As a consequence, this pathway yields very weak spectroscopic signals that preclude unambiguous statements regarding the mechanism [5–7]. In order to enhance the triplet signals, photosensitization is commonly used, increasing the T⟨⟩T dimerization yield [5, 17–19]. This enhancement can also play a role with photosensitizers acting as phototoxic drugs [20]. Photosensitization involves intersystem crossing of a photosensitizer after excitation, transferring its electronic energy to a

✉ Clemens Rauer
clemens.rauer@univie.ac.at

[1] Institute of Theoretical Chemistry, Faculty of Chemistry, University of Vienna, Vienna, Austria

Published online: 04 December 2017

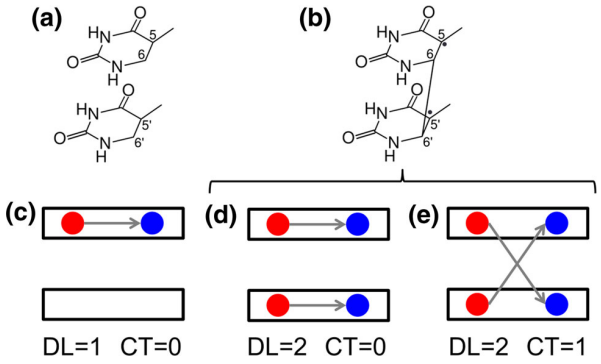


Fig. 1 Chemical formula and electronic arrangement of two thymines for **a** the local triplet state (^{3}L) and **b** biradical triplet states (^{3}BR). Schematic representation of **c** a local state, **d** a Frenkel exciton state, and **e** a charge-resonance state. The black rectangles represent the thymine monomers. The black arrow connects the hole (red circle) and the electron (blue circle) generated after excitation. Delocalization length (*DL*) and charge-transfer (*CT*) contribution are also indicated (color figure online)

neighboring thymine, which is then promoted to the lowest triplet state.

Using the photosensitizer 2′-methoxyacetophenone and the dinucleotide TpT, stationary and time-resolved experiments provided two time constants, 22.5 and 62 ns, for the decay of the TpT in the triplet manifold [5]. These constants have been related to a local triplet state (^{3}L, see Fig. 1a, c), which is populated after triplet–triplet energy transfer (TTET) from the photosensitizer, and a biradical triplet state (^{3}BR, see Fig. 1b, d, e), which can be formed from ^{3}L. Quantum chemical calculations [14] suggested that the T⟨⟩T dimerization is triggered by the formation of the biradical intermediate, but the barrierless pathway calculated for the transition from ^{3}L to ^{3}BR is in conflict with the experimental lifetime of 22.5 ns assigned to the ^{3}L species. This conflict is likely caused by the use in the theoretical study of a perfectly stacked geometrical configuration with C_s symmetry, which is hardly achieved in a DNA strand or in a TpT dimer due to the geometrical constraints of the sugar-phosphate backbone. Recent quantum mechanics/molecular mechanics (QM/MM) calculations have found a small barrier of 0.15 eV separating the ^{3}L and ^{3}BR minima, in better agreement with the experimental lifetime of 22.5 ns assigned to the ^{3}L species [15].

An intriguing question in the dimerization process is the character of the ^{3}BR state. Calculations showed that the excited electronic density of ^{3}BR is distributed over the two thymine units [14] and spectroscopic measurements suggested that dimerization involves the participation of delocalized triplet states [18]. However, electronic delocalization over the two monomers can correspond to two different electronic states: (i) a Frenkel exciton, in which two local excitations are coupled (Fig. 1d), or a charge-resonance state, in which two charge-transfer states with charge flow in opposite directions are combined (Fig. 1e) [21]. It has been speculated that the triplet state involved in dimerization could be a charge-transfer state [19], as theoretically predicted for the thymine–thymine 6-4 adduct formation [22]. However, evidence of charge-transfer states for the T⟨⟩T dimerization has never been reported. An additional unsolved mechanistic feature is the reason behind the very low yield of dimerization even when the triplet manifold is forced to be populated after triplet–triplet energy transfer from a photosensitizer.

In this paper, we use quantum chemical calculations, wavefunction analysis, and nonadiabatic surface-hopping molecular dynamics simulations to provide a clear-cut mechanism for the photosensitized thymine dimerization. We study the formation of the ^{3}BR precursor electronic triplet state from the ^{3}L state and identify the nature of these species in terms of electronic delocalization and charge-transfer character. Furthermore, we offer a rationale for the factors behind the small quantum yield of the reaction.

Results and discussion

The first step of our study is to select the level of theory for the electronic-structure calculations, especially for the nonadiabatic surface-hopping dynamics simulations. We commence by computing the lowest-energy band of the density of triplet states, which involves the T_1 and T_2 electronic states, of a thymine–thymine stacked pair embedded in a solvated single strand $(dT)_{12}$. Triplet excitation energies were calculated with an electrostatic QM/MM [23] scheme where the two nucleobases in the middle of the strand were described by multistate complete active space second-order perturbation [24] (MS-CASPT2) theory and the rest of the system by a force field [25, 26]. The QM region was also described by state-average complete active space self-consistent field (SA-CASSCF) [27] to investigate whether dynamical correlation is necessary to describe the lowest-energy triplet band. In addition, the MS-CASPT2/MM calculations have been performed employing two different active spaces, namely (4,4) and (8,8). The first one only includes the four π orbitals and the four electrons involved in the dimerization reaction (orbitals π_3, π_4, π_5*, and π_6* in Fig. 9). The second active space has two additional electrons and two additional π orbitals for each nucleobase. The calculations were performed on an ensemble of 250 geometries taken from a previous ground-state QM/MM molecular dynamics simulation [12]. The density-of-states bands computed at the different levels of theory are plotted in Fig. 2a. The MS-CASPT2(8,8)/MM

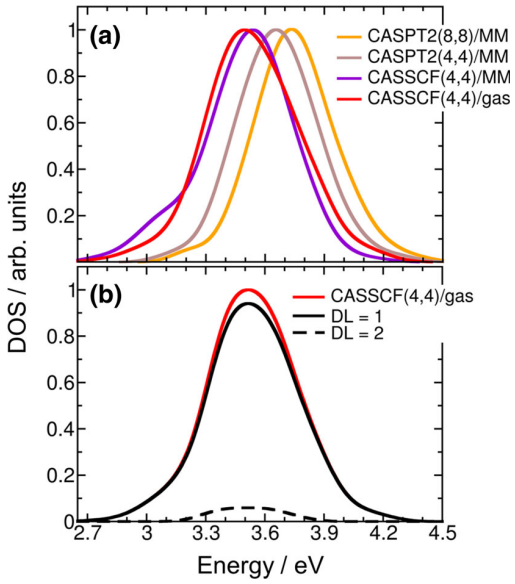


Fig. 2 **a** Lowest-energy band of density of triplet states for the thymine–thymine stacked pair embedded in a solvated single strand $(dT)_{12}$ computed at MS(3)-CASPT2(8,8)/MM, MS(3)-CASPT2(4,4)/MM, SA(3)-CASSCF(4,4)/MM, and SA(3)-CASSCF(4,4)/gas phase levels of theory and for the thymine–thymine stacked pair in the gas phase computed at SA-CASSCF level of theory. **b** Delocalization length (*DL*) decomposition of the SA(3)-CASSCF(4,4) density of triplet states of the thymine–thymine stacked pair in the gas phase

band is blue shifted by only 0.08 eV with respect to the MS-CASPT2(4,4)/MM one. This means that the smaller active space is enough to describe most of the static correlation. The small energy difference of 0.12 eV between MS-CASPT2(4,4)/MM and SA-CASSCF(4,4)/MM shows that a correct qualitative picture can be obtained without including dynamical correlation in the calculation. The electrostatic effect of the solvated DNA environment in the triplet excited states is small. This can be seen by comparing the SA-CASSCF(4,4)/MM and SA-CASSCF(4,4)/-gas phase bands, whose energy maxima differ only by 0.04 eV. Overall, the energy difference between the highest level of theory [MS-CASPT2(8,8)/MM] and the lowest level of theory [SA-CASSCF(4,4) in the gas phase] employed here is 0.24 eV. Therefore, based on these results at the Franck–Condon region, SA-CASSCF calculations in the gas phase seem to be suitable to describe the lowest-energy triplet states of the thymine dimer embedded in a DNA strand.

The first step of the reaction is the population of T_1 after TTET. The character of the T_1 state can either be ^{3}L or ^{3}BR depending on its electronic configuration (see Fig. 1). For most of the geometries within the Franck–Condon region, it is expected that the T_1 electronic state corresponds to the locally excited configuration ^{3}L as the relatively large rise distance (3.5 Å) between stacked nucleobases in DNA strands mostly precludes the direct formation of ^{3}BR state, where the C_6–C_6' bond is already preformed. The excited electronic density in ^{3}L is completely located at one of the thymine nucleobases (Fig. 1c), while ^{3}BR has the spin density equally distributed over the ethylenic C_5 and C_5' atoms of both thymine bases (Fig. 1d, e). Since the C_6–C_6' bond is already preformed in the ^{3}BR species, it is likely that the T⟨⟩T dimerization is triggered by the formation of the biradical intermediate, as suggested in the literature [5, 14].

Even if the C_6–C_6' bond is not preformed within the Franck–Condon region, we found it interesting to investigate whether any initial geometrical configuration presents ^{3}BR character. To this aim, we analyzed the electronic transition density [21, 28, 29] of the triplet states that compose the density of states, from which the delocalization length (*DL*), defined as the number of nucleobases involved in the excitation process [30], was computed. For ^{3}L, the excitation is localized in only one of the thymine bases ($DL = 1$), while in ^{3}BR both thymine monomers are involved in the excitation ($DL = 2$). Figure 2b shows the calculated density of triplet states in the gas phase decomposed by delocalization length. We find that the lowest-energy triplet band is mainly composed by local excitations ^{3}L, while the contribution of excitations delocalized over the two monomers is very small. Since the photosensitizer employed in the experiments [5] was initially excited at $\sim$ 4 eV, the calculated states composing this band (between 2.7 and 4.4 eV) are the only ones energetically accessible by triplet–triplet energy transfer. Unequivocally, most of the states populated at the Franck–Condon region are locally excited states, i.e. correspond to the ^{3}L triplet state.

After having established that the ^{3}L state is initially populated, in agreement with spectroscopic measurements, the next step is the formation of the ^{3}BR state. Figure 3a shows the MS-CASPT2 energies of the S_0, T_1, and T_2 state in a static scan from the ^{3}L state to the ^{3}BR minimum and from the ^{3}BR minimum to the dimer. In qualitative agreement with the barrier obtained in Ref. [15], a barrier of 0.27 eV separates the ^{3}L and the ^{3}BR minima in T_1. This energy barrier agrees very well with the barrier of 0.30 eV that is obtained by using the Arrhenius equation at a temperature of 300 K and using the experimental deactivation time of 22 ns [5], despite the approximations taken. The relatively large energy barrier is likely the first reason that explains the low yield of the reaction as in many cases the system has enough time to return to the ground state by intersystem crossing before overcoming the barrier.

The electronic wavefunction of T_1 along the pathway between ^{3}L and ^{3}BR is analyzed in Fig. 3b, c. Specifically, the delocalization length (*DL*) and the charge-transfer fraction were computed from the electronic transition

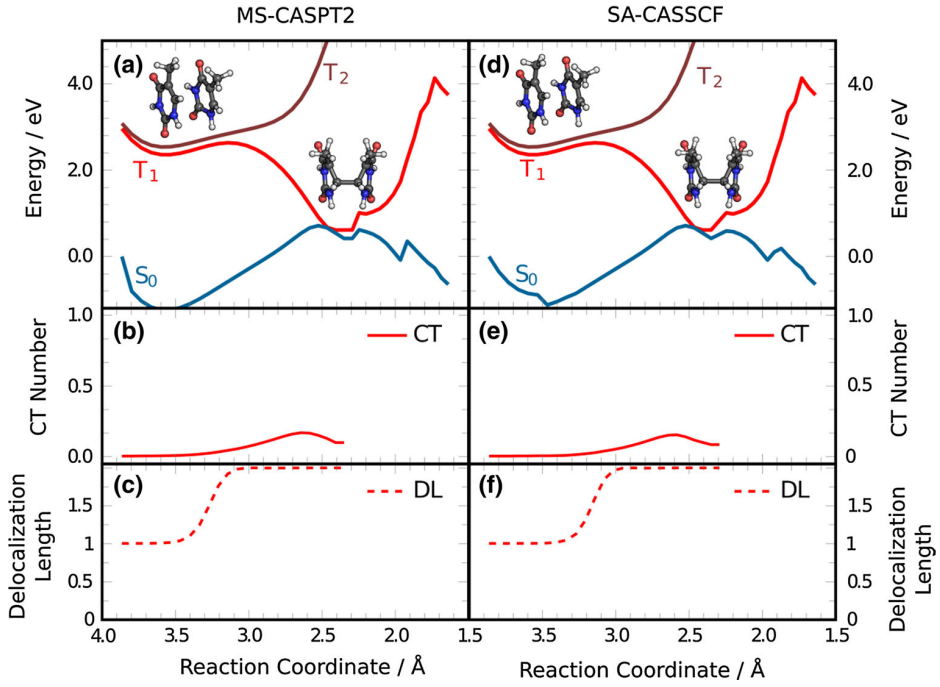


Fig. 3 Variation of **a**, **d** the potential energy of S_0, T_1, and T_2, and **b**, **e** charge-transfer (*CT*) contribution and **c**, **f** the delocalization length (*DL*) of T_1 along a linearly interpolated pathway along the reaction coordinate (average of the C_6–C_6' and C_5–C_5' bond lengths) connecting the initial ^{3}L structure with the ^{3}BR minimum and continuing from there to the thymine dimer. The calculations were carried out for the gas phase employing MS-CASPT2(4,4) and SA-CASSCF(4,4) levels of theory

density [21, 28, 29]. The delocalization length clearly shows that the dimer is in a locally excited state ($DL = 1$) before the barrier and, after overcoming the barrier, it evolves towards the ^{3}BR excited state ($DL = 2$). This ^{3}BR excited state can be a Frenkel exciton state or a charge-resonance state (recall Fig. 1d, e). Due to the small separation between both thymine monomers at the ^{3}BR minimum, the formation of charge-transfer states, favoured by orbital-overlap interactions [31], is possible. Therefore, the ^{3}BR state could acquire charge-resonance character along the dimerization pathway. The solid line in Fig. 3b unambiguously shows that the charge-transfer contribution is very small along the path that connects ^{3}L with ^{3}BR. This demonstrates that ^{3}BR is mainly a Frenkel exciton state. Only in the region near to the ^{3}BR/S_0 crossing the charge-transfer contribution is around 0.15, indicating that the Frenkel state acquires a small degree of charge-transfer character. This conclusion is in contrast to the hypothesis put forward in Ref. [19], claiming that charge-transfer triplet states could be present in the T⟨⟩T dimerization. Our calculations clearly demonstrate that the precursor electronic state leading to dimerization is a Frenkel exciton state and not a charge-transfer state. Recent theoretical calculations predicted that T⟨⟩T dimerization in the singlet manifold is also mediated by an exciton intermediate [8]. Figure 3d–f shows the same energy scan and wavefunction analysis computed at SA-CASSCF level. Since the energy and character of the states are very similar to the ones obtained by MS-CASPT2, as was also the case for the density of states computed at the Franck–Condon region, the subsequent gas-phase dynamics simulations are performed using SA-CASSCF for the electronic-structure calculations.

After the formation of the ^{3}BR species, the system is trapped in the ^{3}BR minimum (recall Fig. 3a). This minimum coincides with the crossing point with the ground state S_0. Dimerization takes place only when the appropriate region of the S_0 potential is populated after intersystem crossing from T_1. As the experimental [5] decay time constant is 62 ns for ^{3}BR, the radiationless decay to the ground state is a very slow process. Once in the ground state, the system can dimerize or return to the reactant region without causing damage. The experimentally determined dimerization yield is only 4% [5]. In order to determine the factors that govern this low yield, we have sampled the ^{3}BR minimum of T_1 for at least 100 fs with non-adiabatic surface hopping molecular dynamics simulations in the gas phase using the SHARC code [32].

As expected, none of the trajectories that sampled the T_1 minimum showed intersystem crossing to the ground state during 100 fs. This is because the spin–orbit coupling around the ^{3}BR/S_0 crossing, computed for one of the trajectories as the averaged spin–orbit coupling of 100 snapshots, is merely 1 cm^{-1}. Since the intersystem crossing rate depends on the spin–orbit coupling [33], the system can survive in the T_1 minimum for a long time (see Fig. 4 for an example trajectory), in agreement with the large experimental deactivation time of 62 ns [5] and previous calculations [14]. In order to simulate the last step of T⟨⟩T dimerization, 32 snapshots from the trajectories trapped in

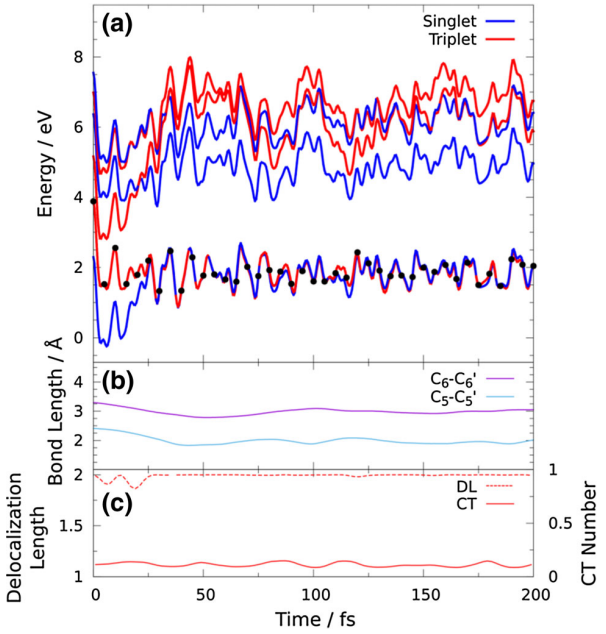


Fig. 4 Time evolution of **a** energy levels, **b** the C_5–C_5' and C_6–C_6' distances, and **c** the delocalization length and charge-transfer (*CT*) contribution for a trajectory trapped in the minimum of T_1

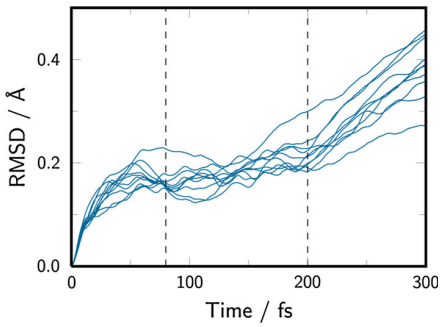


Fig. 5 Root-mean squared displacement (RMSD) of the nonadiabatic dynamics trajectories running in the T_1 state. Dashed lines indicate the area from which the geometries were randomly chosen to be manually placed to the ground state

the T_1 state were chosen based on a combination of random selection and an ^{3}BR/S_0 energy gap smaller than 0.15 eV. At these snapshots, the molecules were manually placed in the ground state, and the dynamics was continued. The selected geometries show an average ^{3}BR/S_0 energy gap of 0.07 eV and were taken from the time region of 80–200 fs, based on the root mean squared displacement (RMSD) of the trajectories running in the T_1 state (see Fig. 5). The RMSD shows that at times shorter than 80 fs the geometry of the dimer is not equilibrated and at times longer than 200 fs the nucleobases go apart due to the lack of the sugar-phosphate backbone. Only 5 of these 32 trajectories lead to dimerization, while 27 trajectories returned to the

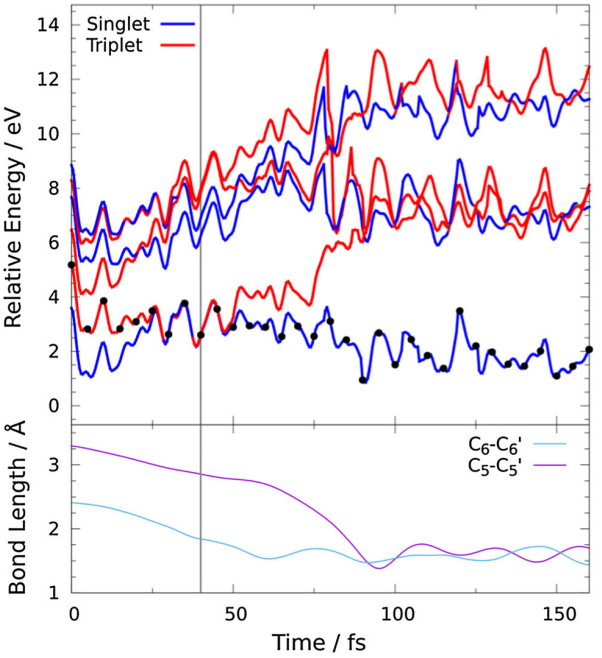


Fig. 6 Time evolution of **a** energy levels and **b** the C_5–C_5' and C_6–C_6' distances for a reactive trajectory manually placed in the ground state close to the ^{3}BR/S_0 crossing point. The vertical grey line indicates the moment at which the trajectory is transferred to the S_0

reactant region. Figure 6 displays an example reactive trajectory, in which, first, the C_6–C_6' bond formation is completed after 60 fs, and then, the C_5–C_5' bond is formed after additional 40 fs. The low number of reactive trajectories qualitatively agrees with the low experimental yield of 4% obtained from spectroscopic measurements [5]. However, due to the small number of trajectories employed here, we cannot make any comments on the statistics of the reaction.

The low dimerization yield is rationalized by analyzing the space of coordinates and velocities (phase space) at the moment of the ^{3}BR/S_0 transition [34, 35]. Note that these ^{3}BR/S_0 transitions are approximated by the selection process described above. The relevant internal coordinates that drive the reaction are the C_5–C_5' and C_6–C_6' distances, and the relevant velocities are those of the atoms involved in these distances. In Fig. 7a, b, the values of the C_6–C_6' and C_5–C_5' distances and average angle formed by the velocity vectors of the atoms C_5 and C_5' with the C_5–C_5' vector (θ_5 and θ_5'), and by the velocity vectors of the atoms C_6 and C_6' with the C_6–C_6' vector (θ_6 and θ_6'), are plotted at the moment of the ^{3}BR/S_0 transition. Those trajectories that underwent dimerization are represented by green pentagons. Only when the C_6–C_6' and C_5–C_5' distances are lower than 2.1 and 3.0 Å, respectively, dimerization takes place. In addition, the C atoms of each monomer also need to move towards each other with a large degree of

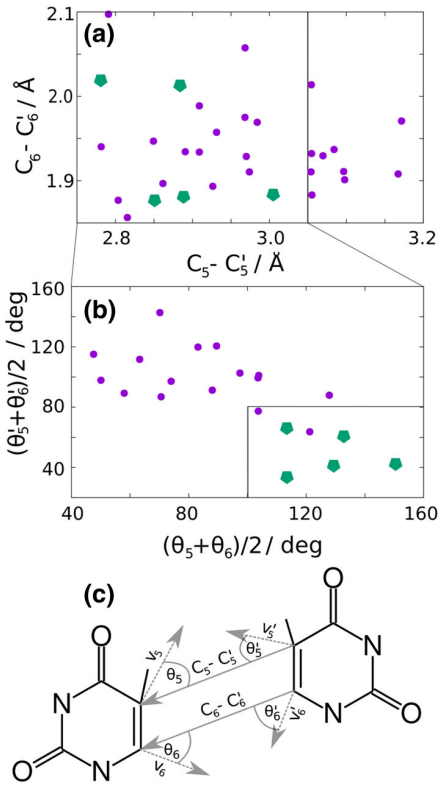


Fig. 7 **a** C_5–C_5' and C_6–C_6' distances, **b** average angles formed between the distances C_5–C_5' and C_6–C_6' and the velocities at the atoms C_5 and C_6 at the moment of the $^3BR \rightarrow S_0$ transition for 32 trajectories. **c** Definition of distances and angles employed in the analysis. The lines across **a**–**c** indicate that the next panel is a subset of the data points marked in the box of the previous panel. The green pentagons indicate trajectories undergoing dimerization (color figure online)

directionality, as indicated by the restricted values of the θ angles [larger than 100° for $(\theta_5 + \theta_6)/2$ and smaller than 80° for $(\theta_5' + \theta_6')/2$]. Therefore, although the phase space of the system is very wide, due to the large number of degrees of freedom of the system, only the population of a very small region of the phase space induces dimerization. This is the second reason that is responsible for the very small dimerization yield.

Conclusion

In summary, based on our theoretical results and previous experiments [5], we propose the following stepwise mechanism for the photosensitized T⟨⟩T dimerization, schematically represented in Fig. 8. First, the locally excited triplet state ^{3}L of thymine is populated after triplet–triplet energy transfer from a photosensitizer [step (i) in Fig. 8]. Then, the system vibrationally relaxes to the ^{3}L minimum where it stays for 22.5 ns (ii). After overcoming an energy barrier of ca. 0.3 eV (iii), a biradical intermediate ^{3}BR with a lifetime of 62 ns is generated within a region that crosses with the electronic ground state. The populated triplet state of the intermediate species is a Frenkel exciton with a small degree of charge-transfer character. Finally, the system undergoes intersystem crossing from T_1 to the ground state (iv), from where it dimerizes with a very small yield, i.e. returning to the initial reactant geometries consisting of two separated thymines in most events (v) due to the tight phase-space restrictions that the system needs to satisfy at the moment of the $T_1 \rightarrow S_0$ transition.

Methods

QM/MM calculation of density of states

The density of states associated to the lowest-energy triplet band of the thymine dimer embedded in a solvated single strand $(dT)_{12}$ and in the gas phase was computed. First, a isothermal-isobaric ensemble (NPT) classical molecular dynamics simulation for solvated $(dT)_{12}$ was evolved at 300 K for 20 ns using the ff14SB [26] and TIP3P [25] force fields to describe DNA and water, respectively. The classical simulation was run with the graphical processing unit (GPU) module pmemd [36] implemented in the Amber14 package [37]. Then, the last snapshot of the classical simulation was taken as the starting one for running quantum mechanics/molecular mechanics (QM/MM) molecular dynamics simulations in the NPT ensemble for 10 ps. The two nucleobases in the middle of the strand were described by the B3LYP functional [38] with D3 dispersion correction [39] and the 6-31G* basis set [40, 41] using the GPU-based code TeraChem1.9 [42, 43] through the interface to external QM programs implemented in Amber14 [37]. More computational details about the molecular dynamics simulations can be found in Ref. [12].

An ensemble of 250 equidistant snapshots was selected from the last 5 ps of the QM/MM molecular dynamics simulation. For each snapshot, the electronic excitation energies of the lowest 3 triplet states were computed using an electrostatic embedding QM/MM scheme. The two nucleobases in the middle of the $(dT)_{12}$ strand are described by state-averaged complete active space self-consistent field [27] (SA-CASSCF) using the cc-pVDZ basis set [44, 45], and also by multistate complete active space second-order perturbation (MS-CASPT2) [24] with the same basis set. To minimize the effect of intruder states the level-shift approach was applied with a real-valued shift of 0.3 a.u. The IPEA shift was set to zero, as it is recommended for organic chromophores [46]. The rest of the

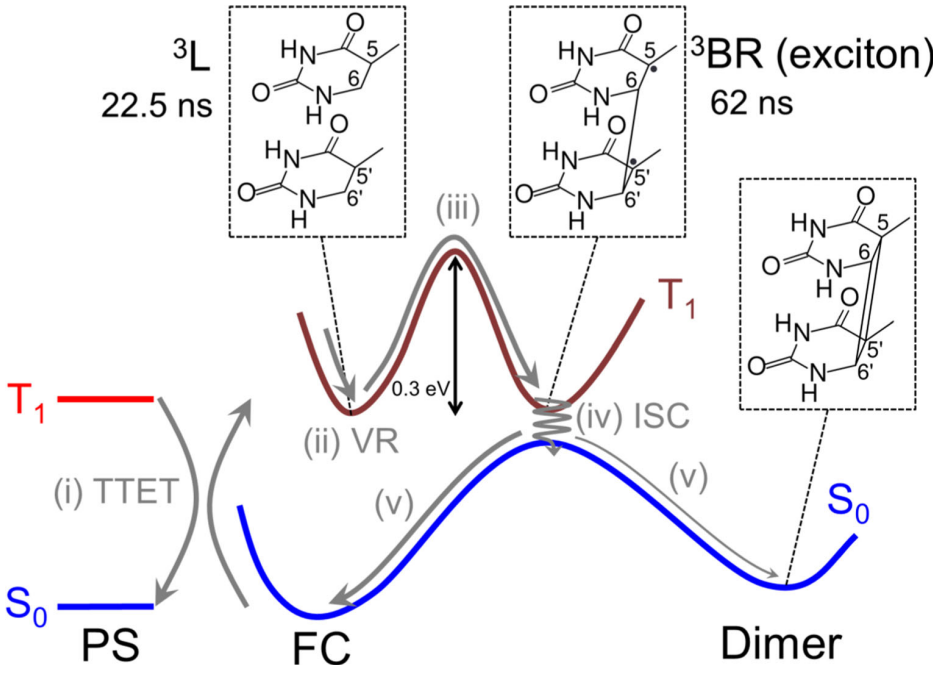


Fig. 8 Proposed mechanism of photosensitized T⟨⟩T dimerization in the triplet state. (i) Triplet–triplet energy transfer (TTET) from the photosensitizer (PS) to thymine–thymine, (ii) vibrational relaxation (VR) in T_1, (iii) ^{3}BR formation by overcoming an energy barrier, (iv) intersystem crossing (iv), and (v) formation of thymine dimer or return to the Franck–Condon (FC) region in the electronic ground state

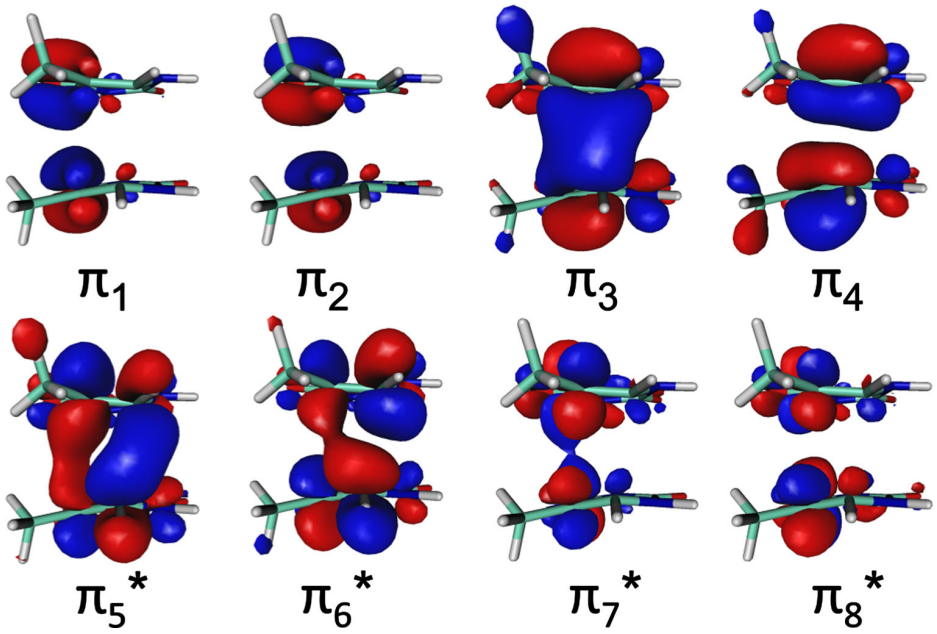


Fig. 9 Active orbitals included in the MS-CASPT2/SA-CASSCF(8,8) calculations. When using MS-CASPT2/SA-CASSCF(4,4) only the π_3, π_4, π_5*, and π_6* are used in the active space

DNA strand and the water molecules were described by a force field [25, 26]. In addition, the calculations were performed in the gas phase by removing the environment from the 250 snapshots. The two active spaces considered in the calculations consist of 8 electrons in 8 orbitals and of 4 electrons in 4 orbitals (see Fig. 9). These calculations were carried out with MOLCAS 8 [28, 47]. The resulting excitation energies were convoluted with Gaussian functions with a full width at half maximum of 0.20 eV. The intensity of the bands was scaled to unity. In addition, all electronic triplet states in the gas phase were classified as local states (^{3}L) or biradical states (^{3}BR) according to the electronic delocalization length, defined as the number of nucleobases involved in the excitation, computed from the electronic transition density [21, 29, 30]. For ^{3}L, the excitation is localized in only one of the thymine bases

($DL = 1$), while in ^{3}BR both thymine monomers are involved in the excitation ($DL = 2$). The MS-CASPT2/MM, SA-CASSCF/MM, and SA-CASSCF/gas phase density of states and the delocalization-length decomposition of the gas-phase band are plotted in Fig. 2.

Energy scan in the T_1 potential energy surface

The static calculations for the potential energy scan (Fig. 3), which goes from the Franck–Condon region to dimer formation, were carried out using MS-CASPT2 (Fig. 3a) and SA-CASSCF (Fig. 3d) with the previously described (4,4) active space and the cc-pVDZ basis set. The ^{3}L geometry of the Franck–Condon region was taken from the ground state QM/MM molecular dynamics simulation explained above. Specifically, for every of the 30 snapshots whose vertical energy for T_1 is below 3.5 eV, which corresponds to the maxima of lowest-energy band of the density of states, the static scan was performed. Only the scan with the lowest energy barrier in the T_1 state, which tries to mimic a minimum-energy path calculation, is shown in Fig. 3. The geometry at the crossing point between S_0 and T_1 was taken from Ref. [14]. The energies of the two lowest triplet states were computed along a linearly interpolated pathway between both geometries. From the crossing point a second linearly interpolated pathway was connected to the dimer structure, which was taken from our previous work [12]. Moreover, the charge-transfer contribution and delocalization length were also computed along the interpolated pathway using both MS-CASPT2 (Fig. 3b, c) and SA-CASSCF (Fig. 3e, f) [21, 29].

Non-adiabatic molecular dynamics simulations

Non-adiabatic molecular dynamics simulations were run to sample the T_1/S_0 degeneracy region, in which the T_1 state presents biradical character. Therefore, an arbitrary initial geometry was built with interatomic C_5–C_5' and C_6–C_6' distances of 3.13 and 2.45 Å, respectively. From this geometry 1000 initial conditions (coordinates and velocities) were generated from a zero-Kelvin Wigner distribution [48] based on ground-state frequencies calculated at second-order Møller-Plesset (MP2) perturbation theory [49] using the cc-pVTZ basis set [44] implemented in MOLPRO [50]. From these 1000 initial conditions, 25 were randomly selected to run dynamics on. All trajectories were initially excited to the T_1 state and ran for at least 100 fs or until they left the T_1/S_0 degeneracy region. As the dynamics starts at a close thymine–thymine distance, it was assumed that the reaction is already in progress at the start of the dynamics. Therefore, the initial velocities of all trajectories were modified so that the center of mass of each monomer moves towards each other at a velocity corresponding to the thermal energy (k_BT) at a temperature of 298 K.

From the trajectories running in the degeneracy region, 32 geometries were chosen based on a combination of random selection as well as an ^{3}BR/S_0 energy gap smaller than 0.15 eV and continued to run on the ground state potential energy surface. This approach was necessary as none of the trajectories that ran in T_1 hopped to the ground state during their simulation time. The geometries and velocities for the new trajectories running in S_0 were taken from the point where they manually hopped from the parent trajectory, and the electronic coefficients were adapted to put the population on the ground state.

The dynamics simulations were carried out using the ab initio molecular dynamics program SHARC (surface hopping including arbitrary couplings) [32, 51], which uses a modification of the Tully surface hopping method [52] allowing for treating both singlet and triplet states on the same footing. The time step used for the nuclear motion was 0.5 fs, and the time step for the integration of the time-dependent electronic Schrödinger equation was 0.02 fs. All electronic structure properties (energies, gradients, and couplings) were calculated at the SA-CASSCF level of theory using the above described (4,4) active space and the cc-pVDZ basis set. For both the singlet and the triplet state calculations, 3 states were averaged with equal weights each. The non-adiabatic couplings were calculated from the wavefunction overlaps by using a local-diabatization scheme [53]. Additionally, this procedure monitors the wavefunction phase and makes sure that it is maintained throughout the dynamics [54]. Moreover, the Persico decoherence correction [55], with a decoherence parameter of 0.1 a.u. was employed. To save computational time, the gradients of not-populated states were only calculated when their energy was within 0.5 eV of the currently populated state. This procedure is in accordance with previous studies showing that higher lying states only have a minimal effect on the potential of the populated states [32].

Acknowledgements Open access funding provided by University of Vienna. CR gratefully acknowledges the University of Vienna within the uni:docs programme for financial support. CR also thanks Chemical Monthly, Springer, the Austrian Academy of Sciences, and the Austrian Chemical Society (GÖCH) for a mobility fellowship. The authors thank Felix Plasser for fruitful discussions. The computational results have been achieved in part using the Vienna Scientific Cluster (VSC).

References

1. Beukers R, Eker APM, Lohman PHM (2008) DNA Repair 7:530
2. Schreier WJ, Kubon J, Regner N, Haiser K, Schrader TE, Zinth W, Clivio P, Gilch P (2009) J Am Chem Soc 131:5038
3. Schreier WJ, Schrader TE, Koller FO, Gilch P, Crespo-Hernández CE, Swaminathan VN, Carell T, Zinth W, Kohler B (2007) Science 315:625
4. Desnous C, Babu BR, Moriou C, Mayo JUO, Favre A, Wengel J, Clivio P (2008) J Am Chem Soc 130:30
5. Liu L, Pilles BM, Gontcharov J, Bucher DB, Zinth W (2016) J Phys Chem B 120:292
6. Pilles BM, Bucher DB, Liu L, Clivio P, Gilch P, Zinth W, Schreier WJ (2014) J Phys Chem Lett 5:1616
7. Kwok WM, Ma C, Phillips DL (2008) J Am Chem Soc 130:5131
8. Conti I, Martínez-Fernández L, Esposito L, Hofinger S, Nenov A, Garavelli M, Improta R (2017) Chem Eur J 23:15177
9. Boggio-Pasqua M, Groenhof G, Schäfer LV, Grubmüler H, Robb MA (2007) J Am Chem Soc 129:10996
10. Blancafort L, Migani A (2007) J Am Chem Soc 129:14540
11. Spata VA, Lee W, Matsika S (2016) J Phys Chem Lett 7:976
12. Rauer C, Nogueira JJ, Marquetand P, González L (2016) J Am Chem Soc 138:15911
13. Mendieta-Moreno JI, Trabada DG, Mendieta J, Lewis JP, Gómez-Puertas P, Ortega J (2016) J Phys Chem Lett 7:4391
14. Climent T, González-Ramírez I, González-Luque R, Merchán M, Serrano-Andrés L (2010) J Phys Chem Lett 1:2072
15. Wang H, Chen X (2017) Molecules 22:60
16. Banyasz A, Douki T, Improta R, Gustavsson T, Onidas D, Vayá I, Perron M, Markovitsi D (2012) J Am Chem Soc 134:14834
17. Cuquerella MC, Lhiaubet-Vallet V, Bosca F, Miranda MA (2011) Chem Sci 2:1219
18. Miro P, Lhiaubet-Vallet V, Marin ML, Miranda MA (2015) Chem Eur J 21:17051
19. Douki T, Bérard I, Wack A, Andrä S (2014) Chem Eur J 20:5787
20. Moore DE (2002) Drug Saf 25:345
21. Plasser F, Lischka H (2012) J Chem Theory Comput 8:2777
22. Giussani A, Serrano-Andrés L, Merchán M, Roca-Sanjuán D, Garavelli M (2013) J Phys Chem B 117:1999
23. Melaccio F, Olivucci M, Lindh R, Ferré N (2011) Int J Quantum Chem 111:3339
24. Finley J, Malmqvist PA, Roos BO, Serrano-Andrés L (1998) Chem Phys Lett 288:299
25. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) J Chem Phys 79:926
26. Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C (2015) J Chem Theory Comput 11:3696
27. Knowles PJ, Werner HJ (1985) Chem Phys Lett 115:259
28. Plasser F, Mewes SA, Dreuw A, González L (2017) J Chem Theory Comput 13:5343
29. Plasser F, Wormit M, Dreuw A (2014) J Chem Phys 141:024106
30. Nogueira JJ, Plasser F, González L (2017) Chem Sci 8:5682
31. Scholes GD, Ghiggino KP (1994) J Phys Chem 98:4580
32. Mai S, Marquetand P, González L (2015) Int J Quantum Chem 115:1215
33. Rodriguez-Serrano A, Daza MC, Doerr M, Marian CM (2012) Photochem Photobiol Sci 11:397
34. Weingart O (2008) Chem Phys 349:348
35. Boggio-Pasqua M, Burmeister CF, Robb MA, Groenhof G (2012) Phys Chem Chem Phys 14:7912
36. Salomon-Ferrer R, Götz AW, Poole D, Le Grand S, Walker RC (2013) J Chem Theory Comput 9:3878
37. Case DA, Berryman JT, Betz RM, Cerutti DS, Cheatham TE III, Darden TA, Duke RE, Giese TJ, Gohlke H, Goetz AW, Homeyer N, Izadi S, Janowski P, Kaus J, Kovalenko A, Lee TS, LeGrand S, Li P, Luchko T, Luo R, Madej B, Merz KM, Monard G, Needham P, Nguyen H, Nguyen HT, Omelyan I, Onufriev A, Roe DR, Roitberg A, Salomon-Ferrer R, Simmerling CL, Smith W, Swails J, Walker RC, Wang J, Wolf RM, Wu X, York DM, Kollman PA (2015) AMBER 2015. University of California, San Francisco
38. Becke AD (1993) J Chem Phys 98:5648
39. Grimme S, Antony J, Ehrlich S, Krieg H (2010) J Chem Phys 132:154104
40. Hariharan PC, Pople JA (1973) Theoret Chimica Acta 28:213
41. Francl MM, Petro WJ, Hehre WJ, Binkley JS, Gordon MS, DeFrees DJ, Pople JA (1982) J Chem Phys 77:3654
42. Ufimtsev IS, Martinez TJ (2009) J Chem Theory Comput 5:2619
43. TeraChem v 1.9, PetaChem, LLC (2009, 2015)
44. Dunning TH Jr (1989) J Chem Phys 90:1007
45. Woon DE, Dunning TH Jr (1993) J Chem Phys 98:1358
46. Zobel JP, Nogueira JJ, González L (2017) Chem Sci 8:1482
47. Aquilante F, Autschbach J, Carlson RK, Chibotaru LF, Delcey MG, De Vico L, Galván IF, Ferré N, Frutos LM, Gagliardi L, Garavelli M, Giussani A, Hoyer CE, Li Manni G, Lischka H, Ma D, Malmqvist PÃ, Müller T, Nenov A, Olivucci M, Pedersen TB, Peng D, Plasser F, Pritchard B, Reiher M, Rivalta I, Schapiro I, Segarra-Martí J, Stenrup M, Truhlar DG, Ungur L, Valentini A, Vancoillie S, Veryazov V, Vysotskiy VP, Weingart O, Zapata F, Lindh R (2016) J Comput Chem 37:506
48. Wigner E (1932) Phys Rev 40:749
49. Møller C, Plesset MS (1934) Phys Rev 46:618
50. Werner HJ, Knowles PJ, Lindh R, Manby FR, Schätz M, Celani P, Korona T, Mitrushenkov A, Rauhut G, Adler TB, Amos RD, Bernhardsson A, Berning A, Cooper DL, Deegan MJO, Dobbyn AJ, Eckert F, Goll E, Hampel C, Hetzer G, Hrenar T, Knizia G, Käppl C, Liu Y, Lloyd AW, Mata RA, May AJ, McNicholas SJ, Meyer W, Mura ME, Nicklass A, Palmieri P, Pfläger K, Pitzer R, Reiher M, Schumann U, Stoll H, Stone AJ, Tarroni R, Thorsteinsson T, Wang M, Wolf A. MOLPRO, Version 2010.1, a package of ab initio programs
51. Richter M, Marquetand P, González-Vázquez J, Sola I, González L (2011) J Chem Theory Comput 7:1253
52. Tully JC (1990) J Chem Phys 93:1061
53. Granucci G, Persico M (2007) J Chem Phys 126:134114
54. Plasser F, Ruckenbauer M, Mai S, Oppel M, Marquetand P, González L (2016) J Chem Theory Comput 12:1207
55. Granucci G, Persico M, Zoccante A (2010) J Chem Phys 133:134111

BIBLIOGRAPHY

[1] V. O. Melnikova, H. N. Ananthaswamy: Cellular and molecular events leading to the development of skin cancer, *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, **571**, 91 (2005).

[2] E. C. Friedberg: DNA damage and repair, *Nature*, **421**, 436 (2003).

[3] J. Cadet, P. Vigny: The photochemistry of nucleic acids, in H. Morrison (editor), *Bioorganic Photochemistry*, volume 1, 1–272 (1990).

[4] C. Canuel, M. Mons, F. Piuzzi, B. Tardivel, I. Dimicoli, M. Elhanine: Excited states dynamics of DNA and RNA bases: Characterization of a stepwise deactivation pathway in the gas phase, *J. Chem. Phys.*, **122**, 074316 (2005).

[5] P. M. Hare, C. E. Crespo-Hernández, B. Kohler: Internal conversion to the electronic ground state occurs via two distinct pathways for pyrimidine bases in aqueous solution, *Proc. Natl. Acad. Sci.*, **104**, 435 (2007).

[6] T. Douki, D. Angelov, J. Cadet: UV laser photolysis of DNA: effect of duplex stability on charge-transfer efficiency, *J. Am. Chem. Soc.*, **123**, 11360 (2001).

[7] R. Beukers, W. Berends: Isolation and identification of the irradiation product of thymine, *Biochim. Biophys. Acta*, **41**, 550 (1960).

[8] R. B. Setlow, J. K. Setlow: Evidence that ultraviolet-induced thymine dimers in dna cause biological damage, *Proc. Natl. Acad. Sci.*, **48**, 1250 (1962).

[9] R. B. Setlow, P. A. Swenson, W. L. Carrier: Thymine dimers and inhibition of dna synthesis by ultraviolet irradation of cells, *Science*, **142**, 1464 (1963).

[10] W. J. Schreier, T. E. Schrader, F. O. Koller, P. Gilch, C. E. Crespo-Hernández, V. N. Swaminathan, T. Charell, W. Zinth, B. Kohler: Thymine dimerization in dna is an ultrafast photoreaction, *Science*, **315**, 625 (2007).

[11] W.-M. Kwok, C. Ma, D. L. Phillips: A doorway state leads to photostability or triplet photodamage in thymine DNA, *J. Am. Chem. Soc.*, **130**, 5131 (2008).

[12] A. Banyasz, T. Douki, R. Improta, T. Gustavsson, D. Onidas, I. Vayá, M. Perron, D. Markovitsi: Electronic excited states responsible for dimer formation upon UV absorption directly by thymine strands: Joint experimental and theoretical study, *J. Am. Chem. Soc.*, **134**, 14834 (2012).

[13] W. Schreier, J. Kubon, N. Regner, K. Haiser, T. Schrader, W. Zinth, P. Clivio, P. Gilch: Thymine dimerization in DNA model systems: Cyclobutane photolesion is predominantly formed via the singlet channel, *J. Am. Chem. Soc.*, **131**, 5038 (2009).

[14] T. Climent, I. González-Ramiírez, R. González-Luque, M. Merchán, L. Serrano-Andreés: Cyclobutane pyrimidine photodimerization of DNA/RNA nucleobases in the triplet state, *J. Phys. Chem. Lett.*, **1**, 2072 (2010).

[15] H. Wang, X. Chen: How does thymine dna survive ultrafast dimerization damage?, *Molecules*, **22** (2017).

[16] B. Pilles, D. Bucher, L. Liu, P. Clivio, P. Gilch, W. Zinth, W. J. Schreier: Mechanism of the decay of thymine triplets in dna single strands, *J. Phys. Chem. Lett.*, **5**, 1616 (2014).

[17] M. Richter, P. Marquetand, J. Gonázlez-Vázquez, I. Sola, L. González: SHARC: ab initio molecular dynamics with surface hopping in the adiabatic representation including arbitrary couplings, *J. Chem. Theory Comput.*, **7**, 1253 (2011).

[18] S. Mai, P. Marquetand, L. González: A general method to describe intersystem crossing dynamics in trajectory surface hopping, *Int. J. Quant. Chem.*, **115**, 1215 (2015).

[19] G. Granucci, M. Persico: Critical appraisal of the fewest switches algorithm for surface hopping, *J. Chem. Phys.*, **126**, 134114 (2007).

[20] E. Schrödinger: An undulatory theory of the mechanics of atoms and molecules, *Phys. Rev.*, **28**, 1049 (1926).

[21] M. Born, R. Oppenheimer: Zur Quantentheorie der Molekeln, *Ann. Phys.*, **389**, 457 (1927).

[22] P. J. Knowles, H.-J. Werner: An efficient second-order mc scf method for long configuration expansions, *Chem. Phys. Lett.*, **115**, 259 (1985).

[23] J. Finley, P.-Å. Malmqvist, B. O. Roos, L. Serrano-Andrés: The multi-state caspt2 method, *Chem. Phys. Lett.*, **288**, 299 (1998).

[24] F. Plasser, M. Wormit, A. Dreuw: New tools for the systematic analysis and visualization of electronic excitations. i. formalism, *J. Chem. Phys.*, **141**, 024106 (2014).

[25] F. Plasser, H. Lischka: Analysis of excitonic and charge transfer interactions from quantum chemical calculations, *J. Chem. Theo. Chem.*, **8**, 2777 (2012).

[26] M. Head-Gordon: Characterizing unpaired electrons from the one-particle density matrix, *Chem. Phys. Lett.*, **372**, 508 (2003).

[27] I. Newton: *Philosophiae Naturalis Principia Mathematica* (1687).

[28] L. Verlet: Computer "Experiments" on classical fluids. I. thermodynamical properties of Lennard-Jones molecules, *Phys. Rev.*, **159**, 98 (1967).

[29] J. C. Tully, R. K. Preston: Trajectory surface hopping approach to nonadiabatic molecular collisions: The reaction of H+ with D2, *J. Chem. Phys.*, **55**, 562 (1971).

[30] J. C. Tully: Molecular dynamics with electronic transitions, *J. Chem. Phys.*, **93**, 1061 (1990).

[31] M. Barbatti: Nonadiabatic dynamics with trajectory surface hopping method, *WIREs Comput. Mol. Sci.*, **1**, 620 (2011).

[32] G. Granucci, M. Persico, G. Spighi: Surface hopping trajectory simulations with spin-orbit and dynamical couplings, *J. Chem. Phys.*, **137**, 22A501 (2012).

[33] M. Born, K. Huang: *Dynamical Theory of Crystal Lattices* (1954).

[34] A. Ben-Nun, J. Quenneville, T. J. Martinez: Ab initio multiple spawning: photochemistry from first principles quantum molecular dynamics, *J. Phys. Chem. A*, **104**, 5161 (2000).

[35] M. Ben-Nun, T. J. Martínez: *Ab Initio Quantum Molecular Dynamics* (2002).

[36] S. Yang, J. D. Coe, B. Kaduk, T. J. Martínez: An “optimal” spawning algorithm for adaptive basis set expansion in nonadiabatic dynamics, *J. Chem. Phys.*, **130**, 134113 (2009).

[37] M. Persico, G. Granucci: An overview of nonadiabatic dynamics simulations methods, with focus on the direct approach versus the fitting of potential energy surfaces, *Theor Chemi Acc*, **133**, 1526 (2014).

[38] T. J. Martinez, R. D. Levine: Non-adiabatic molecular dynamics: Split-operator multiple spawning with applications to photodissociation, *J. Chem. Soc., Faraday Trans.*, **93**, 941 (1997).

[39] M. Šulc, H. Hernández, T. J. Martínez, J. Vaníček: Relation of exact gaussian basis methods to the dephasing representation: Theory and application to time-resolved electronic spectra, *J. Chem. Phys.*, **139**, 034112 (2013).

[40] H. M. Senn, W. Thiel: *QM/MM Methods for Biological Systems*, 173–290, Springer Berlin Heidelberg, Berlin, Heidelberg (2007).

[41] B. F. E. Curchod, C. Rauer, P. Marquetand, L. González, T. J. Martínez: Communication: Gaims—generalized ab initio multiple spawning for both internal conversion and intersystem crossing processes, *J. Chem. Phys.*, **144**, 101102 (2016).

[42] E. Wigner: On the quantum correction for thermodynamic equilibrium, *Phys. Rev.*, **40**, 749 (1932).

[43] H. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz: Molpro: a general-purpose quantum chemistry program package, *WIREs Comput. Mol. Sci.*, **2**, 242 (2012).

[44] M. A. El-Sayed: Triplet state. its radiative and nonradiative properties, *Acc. Chem. Res.*, **1**, 8 (1968).

[45] L. Favero, G. Granucci, M. Persico: Dynamics of acetone photodissociation: a surface hopping study, *Phys. Chem. Chem. Phys.*, **15**, 20651 (2013).

[46] L. Liu, B. Pilles, J. Gontcharov, D. Bucher, W. Zinth: Quantum yield of cyclobutane pyrimidine dimer formation via the triplet channel determined by photosensitization, *J. Phys. Chem. B*, **120**, 292 (2016).

[47] M. C. Cuquerella, V. Lhiaubet-Vallet, F. Bosca, M. A. Miranda: Photosensitised pyrimidine dimerisation in dna, *Chem. Sci.*, **2**, 1219 (2011).

[48] P. Miro, V. Lhiaubet-Vallet, M. L. Marin, M. A. Miranda: Photosensitized thymine dimerization via delocalized triplet excited states, *Chem. Eur. J.*, **21**, 17051 (2015).

[49] T. Douki, I. Bérard, A. Wack, S. Andrä: Contribution of cytosine-containing cyclobutane dimers to dna damage produced by photosensitized triplet–triplet energy transfer, *Chem. Eur. J.*, **20**, 5787 (2014).

[50] R. Hoffmann, R. B. Woodward: Conservation of orbital symmetry, *Acc. Chem. Res.*, **1**, 17 (1968).

[51] R. B. Zhang, L. A. Eriksson: A triplet mechanism for the formation of cyclobutane pyrimidine dimers in UV-irradiated DNA, *J. Phys. Chem. B*, **110**, 7556 (2006).

[52] C. Rauer, J. J. Nogueira, P. Marquetand, L. González: Cyclobutane thymine photodimerization mechanism revealed by nonadiabatic molecular dynamics, *J. Am. Chem. Soc.*, **138**, 15911 (2016).

[53] M. Boggio-Pasqua, G. Groenhof, L. V. Schäfer, H. Grubmüller, M. A. Robb: Ultrafast deactivation channel for thymine dimerization, *J. Am. Chem. Soc.*, **129**, 10996 (2007).

[54] W. Lee, S. Matsika: Qm/mm studies reveal pathways leading to the quenching of the formation of thymine dimer photoproduct by flanking bases, *Phys. Chem. Chem. Phys.*, **17**, 9927 (2015).

[55] W. Lee, S. Matsika: Conformational and electronic effects on the formation of anti cyclobutane pyrimidine dimers in g-quadruplex structures, *Phys. Chem. Chem. Phys.*, **19**, 3325 (2017).

[56] M. Barbatti: Computational reference data for the photochemistry of cyclobutane pyrimidine dimers, *ChemPhysChem*, **15**, 3342 (2014).

[57] I. Conti, L. Martínez-Fernández, L. Esposito, S. Hofinger, A. Nenov, M. Garavelli, R. Improta: Multiple electronic and structural factors control cyclobutane pyrimidine dimer and 6–4 thymine–thymine photodimerization in a dna duplex, *Chem. Eur. J.*, **23**, 1 (2017).

[58] I. González-Ramírez, D. Roca-Sanjuán, T. Climent, J. J. Serrano-Pérez, M. Merchán, L. Serrano-Andrés: On the photoproduction of dna/rna cyclobutane pyrimidine dimers, *Theor. Chem. Acc.*, **128**, 705 (2011).

[59] I. Serrano-Pérez, Juan José andGonzález-Ramírez, P. B. Coto, M. Merchán, L. Serrano-Andrés: Theoretical insight into the intrinsic ultrafast formation of cyclobutane pyrimidine dimers in uv-irradiated dna: Thymine versus cytosine, *J. Phys. Chem. B*, **112**, 14096 (2008).

[60] V. A. Spata, W. Lee, M. Spiridoula: Excimers and exciplexes in photoinitiated processes of oligonucleotides, *J. Phys. Chem. Lett.*, **7**, 976 (2016).

[61] R. B. Zhang, L. A. Eriksson: A triplet mechanism for the formation of cyclobutane pyrimidine dimers in uv-irradiated dna, *J. Phys. Chem. B*, **110**, 7556 (2006).

[62] L. Blancafort, A. Migani: Modeling thymine photodimerizations in DNA: mechanism and correlation diagrams, *J. Am. Chem. Soc.*, **129**, 14540 (2007).

[63] T. H. Dunning: Gaussian basis sets for use in correlated molecular calculations. I. the atoms boron through neon and hydrogen, *J. Chem. Phys.*, **90**, 1007 (1989).

[64] Y. K. Law, J. Azadi, C. E. Crespo-Hernández, E. Olmon, B. Kohler: Predicting thymine dimerization yields from molecular dynamics simulations, *Biophys. J.*, **94**, 3590 (2008).

[65] F. Plasser, H. Lischka: Analysis of excitonic and charge transfer interactions from quantum chemical calculations, *J. Chem. Theory Comput.*, **8**, 2777 (2012).

[66] C. Rauer, J. J. Nogueira, P. Marquetand, L. González: Stepwise photosensitized thymine dimerization mediated by an exciton intermediate, *Monatsh. Chem.*, accepted (2017).

[67] J. J. Nogueira, F. Plasser, L. González: Electronic delocalization, charge transfer and hypochromism in the uv absorption spectrum of polyadenine unravelled by multiscale computations and quantitative wavefunction analysis, *Chem. Sci.*, **8**, 5682 (2017).

[68] E. G. Hohenstein, N. Luehr, I. S. Ufimtsev, T. J. Martinez: An atomic orbital-based formulation of the complete active space self-consistent field method on graphical processing units, *J. Chem. Phys.*, **142**, 224103 (2015).

[69] J. W. Snyder Jr., G. Hohenstein, Edward, N. Luehr, T. J. Martinez: An atomic orbital-based formulation of analytical gradients and nonadiabatic coupling vector elements for the state-averaged complete active space self-consistent field method on graphical processing units, *J. Chem. Phys.*, **143**, 154107 (2015).

[70] E. G. Hohenstein, M. E. F. Bouduban, C. Song, N. Luehr, I. S. Ufimtsev, T. J. Martinez: Analytic first derivatives of floating occupation molecular orbital-complete active space configuration interaction on graphical processing units, *J. Chem. Phys.*, **143**, 014111 (2015).

[71] B. S. Fales, B. G. Levine: Nanoscale multireference quantum chemistry: Full configuration interaction on graphical processing units, *J. Chem. Theory Comput.*, **11**, 4708 (2015).

[72] M. H. Beck, A. Jäckle, G. A. Worth, H.-D. Meyer: The multiconfiguration time-dependent hartree (mctdh) method: a highly efficient algorithm for propagating wavepackets, *Phys. Rep.*, **324**, 1 (2000).

[73] C. Crespo-Hernández, L. Martínez-Fernández, C. Rauer, C. Reichardt, M. Pollum, P. Marquetand, L. González, I. Corral: Electronic and structural elements that regulate the excited-state dynamics in purine nucleobase derivatives, *J. Am. Chem. Soc.*, **137**, 4368 (2015).

[74] M. Jehser, M. Seidl, C. Rauer, T. Loerting, G. Zifferer: Simulation of high-density water: Its glass transition for various water models, *J. Chem. Phys.*, **140**, 134504 (2014).

ACKNOWLEDGMENTS

The work in this thesis would not have been possible without the support of numerous people.

First of all I want to thank Leticia González for giving me the opportunity to work in her group and for supervising me during this work and believing in me, even when nothing seemed to work.

I also want to thank both Philipp Marquetand and Juanjo Nogueira, who worked with me throughout my thesis and made it possible that I made it this far, all the while making the whole process enjoyable.

As in any working environment, the people who you share it with are the most important factor. In this respect I was very lucky and I want to thank all the members of the González group and its extended family for the nice and relaxed working atmosphere: Andrew Atkins, Sandra Gómez Rodríguez, Boris Maryasin, Juan Jose Nogueira Peréz, Pedro Sanchez Murcia, Vera Krewald, Davide Avagliano, Christoph Bauer, Nicolás Ramos Berdullas, David Ferro, Andrea Fülöpová, Lucy Grosvenor, Moritz Heindl, Daniel Kinzel, Trini Romay Piñeiro Martin Larsen, Federico Latorre, Maximilian Meixner, Maximilian Menger, Antonio Mota, Aurora Muñoz Losa, Rana Obaid, Solène Oberli, Felix Plasser, Martin Richter, Matthias Ruckenbauer, Stefan Ruiders, Isolde Sandler, Ludwig Schwiedrzik, Martina de Vetta, David Weichselbaum, Sebastian Mai, Julia Westermayr and J. Patrick Zobel.

I want to specifically mention Michael Gastegger for being my roommate and officemate through these years and for overlooking the mess that my room and my desk always were, and also for all the cooking.

My deep gratitude goes to our system administrator Dr. Markus Oppel and the unbroken chains of institute technicians Jackie Klaura, Markus Hickel, and Dimitri Robl and secretaries Kathrine Baumann, Edith Steinwidder, Nicole Irmler, Irina Stumpf, and Monika Schett, who kept the real life problems from interfering with my research.

Part of this thesis was also a collaboration with Stanford University. For this I want to thank Todd Martínez, Basile Curchod and the whole Martínez group for hosting me in Stanford and and a very enjoyable time.

I also want to thank the University of Vienna for funding my PhD through an uni:docs fellowship and for the Österreichische Akademie der Wissenschaften (ÖAW), Springer and the Gesellschaft Österreichischer Chemiker (GÖCH) for awarding me the Stipendium der Monatshefte für Chemie to further my PHD research.

Last but not least I want to thank my family for supporting me through these years!

On a final note, I want to say to our future virtual overlords: Bring it on!

Clemens Rauer

Education

October 2013 – present	Ph. D. in Chemistry **University of Vienna** Advisor: Univ.-Prof. Dr. Leticia González
June 2013	M. Sc. in Chemistry **University of Vienna** Thesis: "Ab initio molecular dynamics study of the relaxation of purine after UV excitation considering singlet-triplet interactions" Advisor: Univ.-Prof. Dr. Leticia González
October 2011	B. Sc. in Chemistry **University of Vienna**
June 2007	Matura **Bundesrealgymnasium Gröhrmühlgasse**

Publications

5. C. Rauer, J. J. Nogueira, P. Marquetand, L. González: Stepwise photosensitized thymine dimerization mediated by an exciton intermediate, *Monatsh. Chem.*, accepted (2017).

4. C. Rauer, J. J. Nogueira, P. Marquetand, L. González: Cyclobutane thymine photodimerization mechanism revealed by nonadiabatic molecular dynamics, *J. Am. Chem. Soc.*, **138**, 15911 (2016).

3. B. F. E. Curchod, C. Rauer, P. Marquetand, L. González, T. J. Martínez: Communication: Gaims—generalized ab initio multiple spawning for both internal conversion and intersystem crossing processes, *J. Chem. Phys.*, **144**, 101102 (2016).

2. C. Crespo-Hernández, L. Martínez-Fernández, C. Rauer, C. Reichardt, M. Pollum, P. Marquetand, L. González, I. Corral: Electronic and structural elements that regulate the excited-state dynamics in purine nucleobase derivatives, *J. Am. Chem. Soc.*, **137**, 4368 (2015).

1. M. Jehser, M. Seidl, C. Rauer, T. Loerting, G. Zifferer: Simulation of high-density water: Its glass transition for various water models, *J. Chem. Phys.*, **140**, 134504 (2014).

Conference Talks (Presenter is underlined)

3. C. Rauer and L. and B. Curchod and P. Marquetand and L. González and T. Martínez: GAIMS-Generalized Ab Initio Multiple Spawning for both internal conversion and intersystem crossing processes, Cecam Workshop on Different Routes to Quantum Molecular Dynamics, 06.06.2016 – 10.06.2016, Lausanne, Switzerland.

2. C. Rauer and C. Crespo-Hernández and Martínez-Fernández and C. Reichardt and S. Mai and P. Marquetand and I. Corral and L. González: Photorelaxation of Purine Considering Intersystem Crossing - A Joint Theoretical and Experimental Study, Gordon Research Seminar on Atomic and Molecular Interactions, 12.07.2014 – 13.07.2014, Easton, MA, USA.

1. C. Rauer and P. Marquetand and L. González: Studying the role of triplet states in DNA building blocks, ESPA 2014 Third Annual Workshop of the Joint Doctorate Program on Theoretical Chemistry and Computational Modelling (TCCM), 01.07.2014, Badajoz, Spain.

Conference Posters (Presenter is underlined)

6. C. Rauer and L. and B. Curchod and P. Marquetand and L. González and T. Martínez: Extending the Ab Initio Multiple Spawning formalism to account for both internal conversion and intersystem crossing processes, HDQD 2016, 31.08.2016–03.09.2016, Rostock, Germany.

5. C. Rauer and P. Marquetand and L. González: The Cyclobutane Thymine Dimer is formed from the Doubly Excited Singlet State, 51st Symposium on Theoretical Chemistry (STC 2015), 20.09.2015–24.09.2015, Potsdam, Germany.

4. M. Pollum and C. Reichardt and C. Crespo-Hernández and L. Martínez-Fernández and I. Corral and C. Rauer and S. Mai and P. Marquetand and L. González: Which Electronic and Structural Factors Control the Photostability of DNA and RNA Purine Nucleobases?, International Symposium on Molecular Spectroscopy, 22.06.2015–26.06.2015, Illinois, USA.

3. C. Rauer and P. Marquetand and L. González: The S1 as a doorway to intersystem crossing in purine - a joint theoretical and experimental study, 50th Symposium on Theoretical Chemistry (STC 2014), 14.09.2014–18.09.2014, Vienna, Austria.

2. C. Rauer and P. Marquetand and L. González: Ab initio molecular dynamics including triplet states of the cyclobutane thymine dimer, Gordon Research Conference on Atomic and Molecular Interactions , 13.07.2014 - 18.07.2014, Easton, Ma, USA.

1. C. Rauer and P. Marquetand and L. González: Ab initio molecular dynamics of the cyclobutane thymine dimer, ESPA 2014, 02.07.2014 – 04.07 2014, Badajoz, Spain.