Article

Photochemistry of Thymine in Protic Polar Nanomeric Droplets Using Electrostatic Embeding TD-DFT/MM

Miquel Huix-Rotllant [ID]

Aix-Marseille University, CNRS, ICR, 13397 Marseille, France; miquel.huixrotllant@univ-amu.fr

Abstract: Thymine photochemistry is important for understanding DNA photodamage. In the gas phase, thymine undergoes a fast non-radiative decay from S_2 to S_1. In the S_1 state, it gets trapped for several picoseconds until returning to the ground-state S_0. Here, we explore the electrostatic effects of nanomeric droplets of methanol and water on the excited states of thymine. For this purpose, we develop and implement an electrostatic embedding TD-DFT/MM method based on a QM/MM coupling defined through electrostatic potential fitting charges. We show that both in methanol and water, the mechanism is similar to the gas phase. The solvent molecules participate in defining the branching plane of S_0/S_1 intersection and have a negligible effect on the S_1/S_2 intersection. Despite the wrong topology of the ground/excited state intersections, electrostatic embedding TD-DFT/MM allows for a fast exploration of the potential energy surfaces and a qualitative picture of the photophysics of thymine in solvent droplets.

Keywords: electrostatic embedding; QM/MM; conical intersections; nucleobases; thymine

1. Introduction

Photodamage of deoxyribonucleic acid (DNA) originating from ultraviolet (UV) sunlight radiation is one of the main sources of skin cancer [1]. UV light can induce irreparable damage to the genetic code, despite the many natural protection and correction mechanisms that exist [2]. The most basic protection against UV radiation is found at the nucleobase level. Indeed, it has been shown that isolated guanine, adenine, thymine and cytosine undergo ultrafast non-radiative decays back to the ground state [3–7]. Photoexcitation with UV light brings the nucleobase in a singlet excited state of $\pi\pi^*$ character, which relaxes to lower-energy dark states of $n\pi^*$ character (n is a lone-pair electron of a heteroatom in the system) and finally back to the ground state.

The nucleobases decay mechanisms depend strongly on the embedding environment [8–12]. Indeed, when nucleobases are stacked in a DNA polymer, internal non-radiative decay competes with several charge- and proton transfers which ultimately lead to cyclobutane dimers or 6-4 photoadducts [13–18], at the origin of DNA mutations. Electronic structure combined with quantum dynamics simulations are especially suited to determine the mechanisms of deactivation in the excited state. However, there is a lack of methodologies that can treat nucleobases in their native environment at a reasonable computational cost. Especial care has to be taken at the level of theory taken for the electronic structure since the energetic order and the critical points of the photoreaction highly depend on the amount of electron correlation included in the simulations [19]. For example, three different photochemical mechanisms have been inferred from theoretical simulations of thymine: S_1 trapping (fast S_2 and slow S_1 decays) [20], S_2 trapping (slow S_2 and fast S_1 decays) [21], and S_1/S_2 trapping (slow S_1 and S_2 decays) [22].

Recently, the effect of the electronic structure on the photochemistry of thymine in the gas phase has been studied, showing that thymine undergoes a fast non-radiative decay from S_2 to S_1 state thanks to a barrierless access to a conical intersection followed by an S_1-trapping, which appears because of a barrier to reach the conical intersection back to the

ground state [19]. Here, we extend the study of Ref. [19] to determine the effect of protic solvent nanomeric droplets on the photochemistry of thymine. For this purpose, we develop and implement a new quantum mechanics/molecular mechanics (QM/MM) hybrid electrostatic embedding method combining time-dependent density-functional theory (TD-DFT) [23], and classical force fields via an electrostatic potential fitting (ESPF) interaction Hamiltonian [24]. The ESPF method has been recently implemented for the analytic energy, gradient and Hessian for DFT including charge conservation and grid derivatives [25–27]. This method has been shown to successfully represented the electrostatic interaction between complex electrostatic environments such as proteins or solvents [28,29]. The new implementation extends these developments to the excited state, formulating the energy and analytic gradient of TD-DFT/MM required for performing both static (critical point search) and dynamic (semi-classical quantum dynamics) photochemical studies.

We apply the TD-DFT/MM method implemented here to the photochemistry of thymine in water and methanol nanomeric droplets. Despite being only qualitatively correct, TD-DFT has been shown to successfully describe the potential energy surfaces of thymine in the aqueous solution described via continuum models [9]. TD-DFT allows for a fast exploration of potential energy surface (PES) for which excited state mechanisms can be easily extracted. Indeed, we can search for minimum energy structure and approximate conical intersections via the branching update method [30]. Conical intersections are the critical points at the origin of most ultrafast non-radiative decay mechanisms, alike to the transition states in the ground state reactivity. Optimization of conical intersections is thus fundamental for determining the static picture photochemical mechanisms.

This study is organized as follows: in Section 2 the DFT/MM and TD-DFT/MM methods are described, in Section 3 the implementation and computational details are detailed, and in Section 4 the results on the photochemistry and conical intersection topology of thymine in the gas phase and in methanol and water droplets are shown. The conclusions are detailed in Section 5

2. Methodology

The basics of the ESPF QM/MM electrostatic embedding coupling with ground-state self-consistent field single-determinant methods have been published and discussed in previous literature [24,26,27]. Here we followed the notation and method described in Ref. [27], in which a major reformulation of ESPF was established. Here, a formulation is presented based on ESPF that extends the QM/MM coupling to excited states. Without loss of generality, we kept the derivations using the Tamm–Dancoff approximation of the linear response TD-DFT equations (i.e., decoupled excitation and de-excitation parts of the linear response function) [31].

2.1. Electrostatic Embedding Coupling in ESPF DFT/MM

In the electrostatic potential fitting (ESPF) method, the QM and MM subsystems are put in contact through a purely electrostatic interaction, that can be written as

$$\Delta E = \sum_{A}^{N_{QM}} q_A \phi_A \, . \tag{1}$$

In this equation, ϕ_A is the MM electrostatic potential on the QM center A,

$$\phi_A = \sum_{i}^{N_{MM}} \frac{q_i^{FF}}{|\mathbf{R}_A - \mathbf{R}_i|} \, , \tag{2}$$

where q_i^{FF} is the MM partial charge defined from the force field, and $\mathbf{R}_A$ and $\mathbf{R}_i$ are the Cartesian coordinates of QM and MM atoms respectively. Hereafter, we will consider

only classical force fields, that is, q_i^{FF} has a fixed value without polarization. The q_A in Equation (1) is the partial charge on the QM atom A, which is defined as

$$q_A = Z_A - \sum_{\mu\nu}^{N_{\text{AO}}} P_{\mu\nu} Q_{A,\mu\nu} , \tag{3}$$

where Z_A is the atomic charge of atom A, N_{AO} is the number of atomic orbitals, $P_{\mu\nu}$ is the one-electron atomic-orbital density matrix, and $Q_{A,\mu\nu}$ is the electron charge population operator on QM center A. This operator is obtained by a fitting procedure, in which the electrostatic integrals on a grid are fitted into operators on the QM centers. This accounts to solving the equation

$$\sum_{A}^{N_{\text{QM}}} \frac{Q_{A,\mu\nu}}{|\mathbf{r}_k - \mathbf{R}_A|} = \langle \mu | \frac{1}{|\mathbf{r} - \mathbf{r}_k|} | \nu \rangle . \tag{4}$$

Using the short-hand notation $\mathbf{T}_{k,A} = |\mathbf{r}_k - \mathbf{R}_A|^{-1}$ and $V_{k,\mu\nu} = \langle \mu | \frac{1}{|\mathbf{r} - \mathbf{r}_k|} | \nu \rangle = \int d\mathbf{r} \frac{\chi_\mu^*(\mathbf{r})\chi_\nu(\mathbf{r})}{|\mathbf{r} - \mathbf{r}_k|}$, we can formally write the electron charge operators as

$$Q_{A,\mu\nu} = \sum_{k}^{N_k} \left[\left(\mathbf{T}^\dagger \mathbf{T} \right)^{-1} \mathbf{T}^\dagger \right]_{A,k} V_{k,\mu\nu} = \sum_{k}^{N_k} T_{A,k}^+ V_{k,\mu\nu} . \tag{5}$$

Deriving Equation (1) with respect to the density matrix, we arrive at the ESPF Hamiltonian, which is given simply by

$$h_{\mu\nu} = - \sum_{A}^{N_{\text{QM}}} Q_{A,\mu\nu} \phi_A . \tag{6}$$

Due to the finite size of the grid defined in Equation (4), the charge conservation condition,

$$\sum_{A}^{N_{\text{QM}}} \sum_{\mu\nu}^{N_{\text{AO}}} P_{\mu\nu} Q_{A,\mu\nu} = N_{\text{el}} , \tag{7}$$

is not exactly satisfied. In Ref. [27], I showed that this can be imposed at the level of operators by deriving the charge conservation condition with respect to the density matrix,

$$\sum_{A}^{N_{\text{QM}}} Q_{A,\mu\nu} = S_{\mu\nu} , \tag{8}$$

in which $S_{\mu\nu} = \langle \mu | \nu \rangle$ is the atomic-orbital overlap matrix. This charge conservation in operator form can be imposed in the Hamiltonian directly, leading to

$$h'_{\mu\nu} = \sum_{A}^{N_{\text{QM}}} Q_{A,\mu\nu} (\Phi_{\text{av}} - \phi_A) - \Phi_{\text{av}} S_{\mu\nu} , \tag{9}$$

where we introduced the average external potential,

$$\Phi_{av} = \frac{1}{N_{QM}} \sum_{A}^{N_{\text{QM}}} \phi_A . \tag{10}$$

The Hamiltonian defined in Equation (9) can thus be added to the gas phase QM Fock operator (F_0) to define an electrostatic embedding formulation. The total energy of the system then becomes

$$E = \text{Tr}_{\text{AO}} \left[P(F_0 + h') \right] + E_{\text{MM}} + \sum_{A}^{N_{\text{QM}}} Z_A \phi_A , \tag{11}$$

in which E_{MM} is the MM energy contribution and Z_A is the atomic number of atom A.

The gradient of Equation (11) has two blocks with a different expression, depending on the type of atom (QM or MM) with respect to the energy is derived,

$$\nabla E_0 = \left[\begin{array}{c} E_0^x \\ E_0^{\tilde{x}} \end{array} \right] = \left[\begin{array}{c} \text{Tr}_{AO}\left[P(F_0^x + h'^x)\right] - \text{Tr}_{AO}[WS^x] + E_{MM}^x + \sum_A^{N_{QM}} Z_A \phi_A^x \\ \text{Tr}_{AO}\left[Ph'^{\tilde{x}}\right] + E_{MM}^{\tilde{x}} + \sum_A^{N_{QM}} Z_A \phi_A^{\tilde{x}} \end{array} \right], \qquad (12)$$

in which we symbolized the derivatives $E^x = \partial E/\partial x$, in which x and $\tilde{x}$ represent a Cartesian coordinate of a QM or MM atom respectively. In the QM block, we substituted the derivative of the atomic-orbital density matrix with the trace of the derivative of the overlap and the energy-weighted matrix $W = PFP^\dagger$. Since the ESPF charge operator only depends on the coordinates of the grid points and the QM atoms, the derivative of the ESPF operator with respect to MM atom perturbations is simply given by

$$\text{Tr}_{AO}\left[Ph'^{\tilde{x}}\right] = \sum_A^{N_{QM}} Q_A (\Phi_{av} - \phi_A)^{\tilde{x}} - \Phi_{av}^{\tilde{x}} N_{el}, \qquad (13)$$

in which we defined $Q_A = \text{Tr}_{AO}[PQ_A]$. Thus, the gradient becomes

$$\nabla E_0 = \left[\begin{array}{c} E_0^x \\ E_0^{\tilde{x}} \end{array} \right] = \left[\begin{array}{c} \text{Tr}_{AO}\left[P(F_0^x + h'^x)\right] - \text{Tr}_{AO}[WS^x] + E_{MM}^x + \sum_A^{N_{QM}} Z_A \phi_A^x \\ \sum_A^{N_{QM}} Q_A (\Phi_{av} - \phi_A)^{\tilde{x}} + E_{MM}^{\tilde{x}} + \sum_A^{N_{QM}} Z_A \phi_A^{\tilde{x}} - \Phi_{av}^{\tilde{x}} N_{el} \end{array} \right]. \qquad (14)$$

2.2. Electrostatic Embedding Coupling in ESPF TDA-TDDFT/MM

In TDA-TDDFT/MM, the excited state energy of state I is obtained by adding the excitation energy ω_I to Equation (11). The excitation energy is obtained by solving the eigenvalue problem,

$$AX_I = \omega_I X_I, \qquad (15)$$

in which $A_{ai,bj} = \left([F_0 + h']_{ab} - [F_0 + h']_{ij}\right)\delta_{ij}\delta_{ab} + (bj|f_{Hxc}|ai)$, in which $f_{Hxc} = \delta v_H/\delta\rho + \delta v_{xc}/\delta\rho$ is the Hartree-exchange-correlation kernel obtained by deriving the corresponding potentials with respect to the electronic density [32].

The energy gradient is efficiently obtained by the Lagrangian approach by Furche and Ahlrichs [33]. In this approach, TDA-TDDFT/MM excitation energy gradient for state I is written as

$$\nabla E_I = \left[\begin{array}{c} E_0^x + \omega_I^x \\ E_0^{\tilde{x}} + \omega_I^{\tilde{x}} \end{array} \right] = \left[\begin{array}{c} E_0^x + Tr\left[P_I(h_0 + h')^x\right] - Tr[W_I S^x] \\ + Tr\left[f_H^x \Gamma_I\right] + Tr\left[f_{xc}^x X_I X_I^\dagger\right] + \sum_A^{N_{QM}} Z_A \phi_A^x \\ \\ E_0^{\tilde{x}} + \sum_A^{N_{QM}} Q_{A,I}(\Phi_{av} - \phi_A)^{\tilde{x}} + \sum_A^{N_{QM}} Z_A \phi_A^{\tilde{x}} \\ - \Phi_{av}^{\tilde{x}} N_{el} \end{array} \right]. \qquad (16)$$

In this equation, P_I and Γ_I are the one- and two-electron relaxed transition densities to state I, $W_I = P_I F P_I^\dagger$ and $q_{A,I} = Z_A - Tr[P_I Q_A]$ (for the definition of relaxed transition densities, see Ref. [33]). The relaxed densities are obtained after solving a set of coupled-perturbed equations (the Z-vector equations), which only contain two-electron terms and thus not affected by the ESPF one-electron interaction Hamiltonian [33,34].

2.3. Branching Update Method

In the branching update method [30], the optimization is performed by minimizing the average energy $E = (E_I + E_J)/2 + (E_J - E_I)^2$ using the gradient

$$\nabla E = B\nabla E_{av} + 2(E_I - E_J)\nabla E_{diff}, \qquad (17)$$

in which the average gradient is given by $\nabla E_{av} = (\nabla E_I + \nabla E_J)/2$ and the gradient difference is given by $\nabla E_{diff} = (\nabla E_I - \nabla E_2)/2$. The projector T projects out the branching plane vectors from the average gradient, and is defined as,

$$B = 1 - \nabla E'_{diff}\nabla E'^{\mathrm{T}}_{diff} - \nabla E'_{orth}\nabla E'^{\mathrm{T}}_{orth}, \tag{18}$$

in which ∇E_{orth} is an vector orthogonal to ∇E_{diff}. The tilde indicates that these vectors are normalized. The ESPF QM/MM, difference gradient vector between two excited states is defined as

$$\nabla (E_J - E_I) = \begin{bmatrix} \omega^x_J - \omega^x_I \\ \\ \omega^{\tilde{x}}_J - \omega^{\tilde{x}}_I \end{bmatrix} = \begin{bmatrix} Tr\big[(P_J - P_I)(h_0 + h')^x\big] - Tr\big[(W_J - W_I)S^x\big] \\ +Tr\big[f^x_H(\Gamma_J - \Gamma_I)\big] + Tr\big[f^x_{xc}\big(X_J X^\dagger_J - X_I X^\dagger_I\big)\big] \\ \\ \sum^{N_{QM}}_A \Delta Q_{A,IJ}(\Phi_{av} - \phi_A)^{\tilde{x}} \end{bmatrix}, \tag{19}$$

in which we defined $\Delta Q_{A,IJ} = Q_{A,J} - Q_{A,I}$. The particular case of the excited/ground state gradient, the expression for the gradient difference vector is simply the excitation energy of excited state I

$$\nabla (E_I - E_0) = \begin{bmatrix} \omega^x_I \\ \\ \omega^{\tilde{x}}_I \end{bmatrix} = \begin{bmatrix} Tr\big[P_I(h_0 + h')^x\big] - Tr[W_I S^x] \\ +Tr\big[f^x_H\Gamma_I\big] + Tr\big[f^x_{xc}X_I X^\dagger_I\big] + \sum^{N_{QM}}_A Z_A \phi^x_A \\ \\ \sum^{N_{QM}}_A Q_{A,I}(\Phi_{av} - \phi_A)^{\tilde{x}} \end{bmatrix}. \tag{20}$$

As for the orthogonal gradient, it is estimated by a simple update procedure,

$$\nabla E'_{orth,k} = \frac{\left(\nabla E'_{orth,k-1}\nabla E'_{diff,k}\right)\nabla E'_{diff,k-1} - \left(\nabla E'_{diff,k-1}\nabla E'_{diff,k}\right)\nabla E'_{orth,k-1}}{\sqrt{\left(\nabla E'_{diff,k-1}\nabla E'_{diff,k}\right)^2 + \left(\nabla E'_{orth,k-1}\nabla E'_{diff,k}\right)^2}}, \tag{21}$$

in which k indicates the optimization step.

3. Computational Details

The methods described here were implemented in a development version of GAMESS-US (14 FEB 2018 R1) interfaced with a modified version of Tinker 6.3.3 [35,36]. For the computation of ESPF atomic charges, an atom-centred Lebedev grid of 50 angular points and nine radial increments on each atom. Points inside the Van der Waals radii were excluded from the grid. The one-electron ESPF Hamiltonian and the derivative were added accordingly to the one-electron Hamiltonian and derivative parts of the one-electron Hamiltonian of the Fock operator. The gradients of the ground and excited states on MM atoms were computed using the corresponding ESPF atomic charges computed respectively for the density matrix and the relaxed density matrix. The environment was either optimized using the quadratic convergence optimization method of GAMESS-US or using microiterations. All calculations were performed at the TD-BH&HLYP/6-31G* using the Tamm–Dancoff approximation for the thymine atoms [37–39], and Amber94 for the water or methanol atoms [40] .

4. Results and Discussion

In this section, the photochemistry of thymine was studied in gas phase and in nanometic droplets of water and methanol. The purpose was to illustrate the potential of the methodology described in Section 2 rather than performing a full mechanistic study, that would require performing quantum dynamics. The nanomeric droplets were created using the PACKMOL package [41], and subsequently optimized to the minimum energy

structure of the internal QM/MM energy. For methanol and water droplets, 100 and 300 molecules respectively were included in the simulations.

4.1. Photochemistry of Thymine in Gas Phase

The potential energy surface representing the key points of the photochemistry from the lowest singlet energy states of thymine in the gas phase is summarized in Figure 1. The first singlet (S_1) state was described mainly as a single excitation from a lone-pair orbital of oxygen (labeled *n*, with contributions from oxygen 4a mainly and oxygen 2a to a lesser extent) to a π^* orbital delocalized mainly over carbons 4, 5 and 6. The S_2 state was described essentially as a π to π^* transition. The photochemical pathways of deactivation of thymine have been discussed in depth in the literature [5]. Here, we focused on the topology of the potential energy surfaces (PES) and the conical intersection seams found at the TD-BH&HLYP/6-31G* level. The PES in the figure was obtained by a geodesic interpolation between the optimized structures of three conical intersections and two state minima [42]. The geodesic path represents a minimum energy path connecting the different critical points of the photoreaction, for which we can extract an idea of the photochemical mechanism. From the S_0 state, thymine was excited to the bright S_2 state of $\pi\pi^*$ nature. There was no stable minimum for the S_2 state predicted at the TD-DFT level. Rather, the S_2 minimum was found at the seam of conical intersection with S_1, of $n\pi^*$ nature. From the S_1/S_2 intersection, a branching occurs. On the left side of the graph, the path towards the intersection between the ground state and the $\pi\pi^*$ state is shown. This intersection could only be reached by surmounting a barrier of ca. 0.13 eV, and indeed it has been shown as a minor path of deactivation of thymine in the gas phase [5]. Most of the wavepacket would rather follow a minimum energy path towards the $n\pi^*$ state minimum. This minimum was long-lived, since the corresponding conical intersection was 0.8 eV higher in energy.
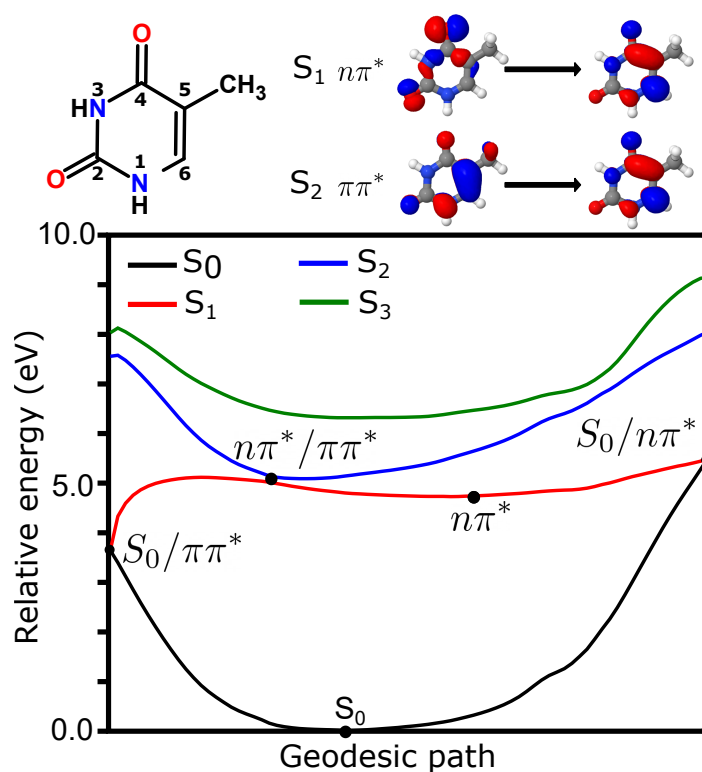
Figure 1. Photochemistry of thymine in the gas phase. (**top**, **left**) Schematic structure of thymine with atom numbering. (**top**, **right**) Main character of the transitions representing the first and second singlet excited states. (**bottom**) Geodesic path connecting the conical intersections between $S_0/n\pi^*$, $S_0/\pi\pi^*$ and $n\pi^*/\pi\pi^*$ as well as S_0 and $n\pi^*$ minima. Energies (in eV) have been computed at the TDA-TDBH&HLYP/6-31G* level.
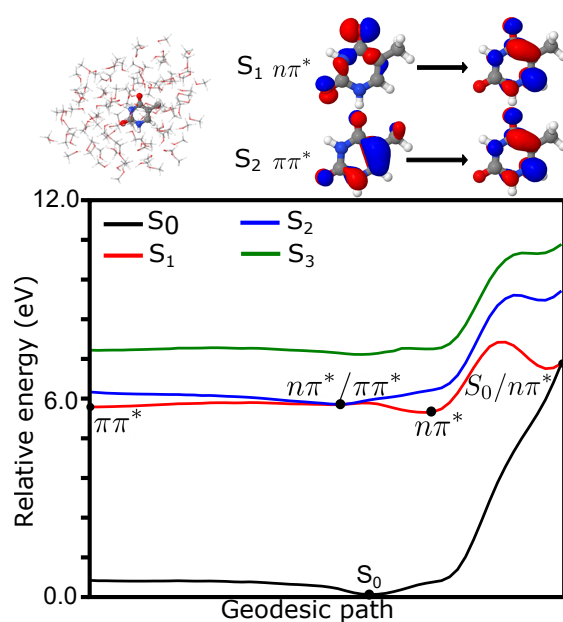
The intersection geometries, vectors and PES around the intersection point for intersections $S_0/n\pi^*$, $S_0/\pi\pi^*$ and $n\pi^*/\pi\pi^*$ are represented in Figure 2. As seen from the PES along with the non-adiabatic coupling, both intersections involving the S_0 state were sloped conical intersections, while the $n\pi^*/\pi\pi^*$ was peaked. In the $S_0/n\pi^*$ intersection, the C_4-O_{4a} bond got perpendicular to the thymine plane, while in the $S_0/n\pi^*$ intersection the C_5-C_{5a} bond was perpendicular to the thymine plane. The $n\pi^*/\pi\pi^*$ was a boat-like structure, in which the bonds N_3-H and C_6-H were parallel but displaced with respect to the plane of thymine. The wrong intersection topology for the $S_0/n\pi^*$ state was observed in the cut along the gradient difference vector, as noted previously for TDA-TD-DFT in the gas phase, in which a negative excitation energy was leading to the "true" ground state [43]. Indeed, TDA-TD-DFT led to diabatic surfaces rather than adiabatic states along the gradient difference vector.
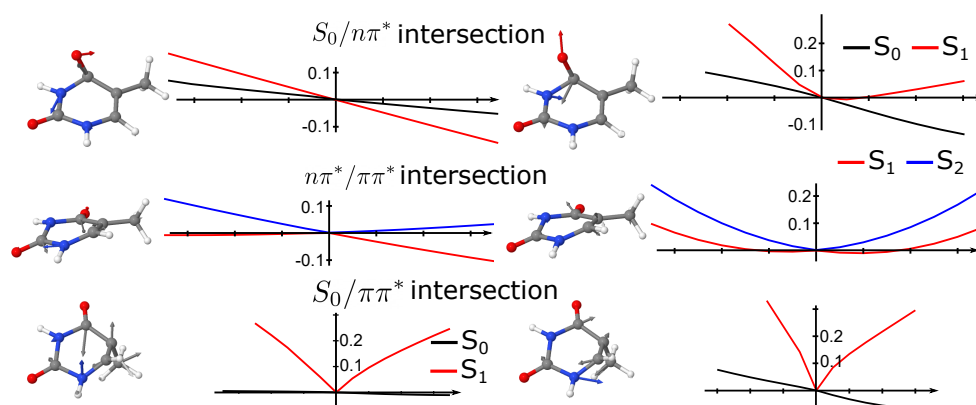
Figure 2. Topology of conical intersections of thymine in the gas phase. (**top**) S0/$n\pi^*$ intersection, (**middle**), $n\pi^*/\pi\pi^*$ intersection and (**bottom**) $n\pi^*/\pi\pi^*$ intersection. Structures and graphs corresponds to the gradient difference vector and the PES along that coordinate (**left**) and non-adiabatic coupling vector (**right**).
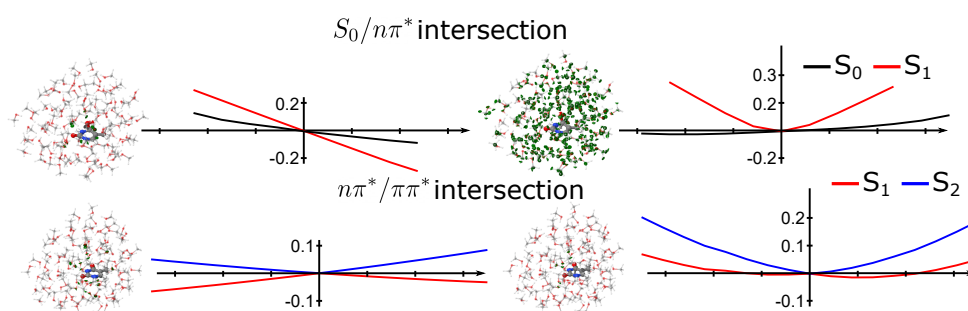
4.2. Photochemistry of Thymine in Methanol Droplet

The photochemistry from the lowest energy singlet states of thymine in a methanol droplet is summarized in Figure 3. The droplet contained 100 methanol molecules treated at the MM level. The state and orbital characters were the same in the gas phase. Indeed, the S_1 is represented by an n to π^* transition while the S_2 state is represented by a π to π^* transition. The PES was computed along the geodesic path connecting the $\pi\pi^*$, $n\pi^*$ and S_0 minimum energy structures as well as the $n\pi^*/\pi\pi^*$ and $S_0/n\pi^*$ intersections. The initial bright state in methanol was still the S_2. From this state, a barrierless access to the $n\pi^*/\pi\pi^*$ intersection was observed, similar to what was observed in the gas phase. From this intersection, a branching towards the $n\pi^*$ and $\pi\pi^*$ minima occurred. From the $\pi\pi^*$ state, the $S_0/\pi\pi^*$ intersection was not found. This could be since this intersection was reached by the motion of a CH_3 group (*vide supra*), which was hindered in solution. Rather, the motions around the $\pi\pi^*$ minimum could reach back to the $n\pi^*/\pi\pi^*$ minimum and thus be trapped in the $n\pi^*$ minimum. From the $n\pi^*$ minimum, a larger barrier was obtained to reach the $S_0/n\pi^*$ than in the case of thymine in the gas phase. The difference between the $n\pi^*$ minimum and the corresponding intersection with the ground state was around 1.45 eV, thus the reaction was expected to be slower than in the gas phase. To reach the conical intersection, a large reorganization of the solvent molecules was needed in the path connecting these two minimum energy structures. As an estimate, the difference between the purely MM interaction energies among the solvent molecules was 0.1 eV higher in $S_0/n\pi^*$ intersection with respect to the $n\pi^*$ minimum. Still, if the solvent was relaxed along this coordinate, the barrier disappeared, but the S_0 and the $n\pi^*$ states intersection disappeared. The two states had an energetic gap of around 0.4 eV with the relaxed solvent. This implies that the solvent had a direct participation in the definition of the conical intersection.

Figure 3. Photochemistry of thymine in methanol. (**top**, **left**) Schematic structure of thymine and methanol QM/MM model. (**top**, **right**) Main character of the transitions representing the first and second singlet excited states. (**bottom**) Geodesic path connecting the conical intersections between $S_0/n\pi^*$, $S_0/\pi\pi^*$ and $n\pi^*/\pi\pi^*$ as well as S_0 and $n\pi^*$ minima. Energies (in eV) have been computed at the TDA-TDBH&HLYP/6-31G* level.

The participation of the solvent on the intersections is clearly shown in Figure 4, in which the intersection geometries, branching plane vectors and PES around the intersection points are represented. In methanol, both the $S_0/n\pi^*$ and $n\pi^*/\pi\pi^*$ intersections peaked. On the one hand, the $n\pi^*/\pi\pi^*$ intersection was completely planar, unlike the boat-like structure in the gas phase. The participation of the solvent in this intersection was negligible. On the other hand, the $S_0/n\pi^*$ intersection was similar to the gas phase, in which the C_4-O_{4a} bond was perpendicular to the thymine bond. While the gradient vector of this intersection did not involve the solvent molecules, the approximate non-adiabatic coupling was completely delocalized over all the solvents. This implies that solvent coordinates were to be considered in the definition of the branching plane of conical intersections. Similar conclusions were extracted from model Hamiltonians of conical intersections of protonated Schiff bases in solution by Burghardt et al [44].


Figure 4. Topology of conical intersections of thymine in methanol. (**top**) $S_0/n\pi^*$ intersection and (**bottom**) $n\pi^*/\pi\pi^*$ intersection. Structures and graphs corresponds to the gradient difference vector and the PES along that coordinate (**left**) and non-adiabatic coupling vector (**right**). The contribution of the solvent on the branching plane vectors is shown in green.

4.3. Photochemistry of Thymine in Water Droplet

The photochemistry from the lowest energy singlet states of thymine in a water droplet is summarized in Figure 5. The droplet contained 300 water molecules treated at the MM

level. The state and orbital characters were the same in the gas phase and methanol droplet. The PES along the geodesic path connecting the $\pi\pi^*$, $n\pi^*$ and S_0 minimum energy structures as well as the $n\pi^*/\pi\pi^*$ and $S_0/n\pi^*$ intersections is shown in the figure. From the bright S_2 state, a barrierless access to the $n\pi^*/\pi\pi^*$ intersection was observed, similar to gas phase and methanol. At variance with methanol and similar to the gas phase, the minimum of the $\pi\pi^*$ state was not found in water, indicating that the minimum was close to the S_1/S_2 intersection. Still, the $S_0/\pi\pi^*$ intersection was not found either, probably for the same reason as in methanol, due to the hindrance of the CH_3 group motion due to the steric interactions with the solvent molecules. From the $n\pi^*/\pi\pi^*$ intersection, all trajectories should go towards the direction of $n\pi^*$ minimum. A double barrier occurs in the path to the $S_0/n\pi^*$ intersection, indicating a larger reorganization of the solvent molecules than in methanol. Indeed, from the $n\pi^*/\pi\pi^*$ to the $n\pi^*$ minimum a first barrier was obtained, which disappeared when the solvent molecules were relaxed along the geodesic path. Additionally, from the $n\pi^*$ minimum a larger barrier was obtained to reach the $S_0/n\pi^*$ than in the case of thymine in the gas phase. The difference between the $n\pi^*$ minimum and the corresponding intersection with the ground state was around 1.34 eV, thus the reaction was expected to be slower than in the gas phase, but slightly faster than in methanol.
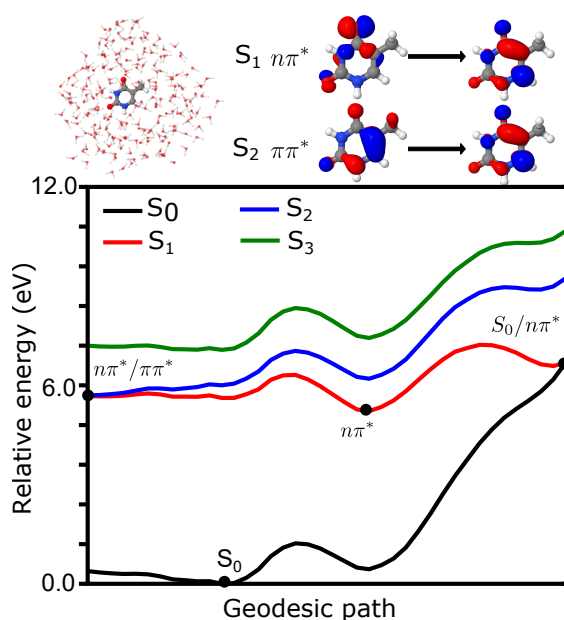
Figure 5. Photochemistry of thymine in water. (**top**, **left**) Schematic structure of thymine and water QM/MM model. (**top**, **right**) Main character of the transitions representing the first and second singlet excited states. (**bottom**) Geodesic path connecting the conical intersections between $S_0/n\pi^*$, $S_0/\pi\pi^*$ and $n\pi^*/\pi\pi^*$ as well as S_0 and $n\pi^*$ minima. Energies (in eV) have been computed at the TDA-TDBH&HLYP/6-31G* level.

The participation of the solvent on the intersections is clearly shown in Figure 6, in which the intersection geometries, branching plane vectors and PES around the intersection points are represented. In water the intersections were similar to the methanol ones, both the $S_0/n\pi^*$ and $n\pi^*/\pi\pi^*$ intersections peaked. On the one hand, the $n\pi^*/\pi\pi^*$ intersection was completely planar, unlike the boat-like structure in the gas phase. The participation of the solvent in this intersection was negligible. On the other hand, the $S_0/n\pi^*$ intersection was similar to the gas phase, in which the C_4-O_{4a} bond was perpendicular to the thymine bond. While the gradient vector of this intersection did not involve the solvent molecules, the approximate non-adiabatic coupling was completely delocalized over all the solvent, as it was observed from methanol. This implies again that solvent coordinates were to be considered in the definition of the branching plane of conical intersections.
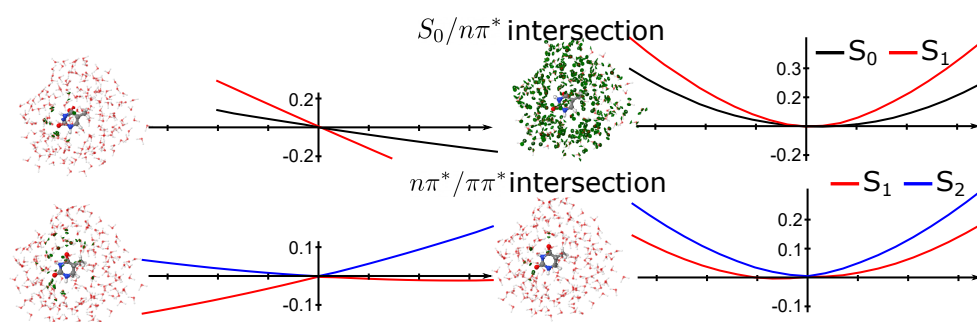
Figure 6. Topology of conical intersections of thymine in water. (**top**) $S_0/n\pi^*$ intersection and (**bottom**) $n\pi^*/\pi\pi^*$ intersection. Structures and graphs correspond to the gradient difference vector and the PES along that coordinate (**left**) and non-adiabatic coupling vector (**right**). The contribution of the solvent on the branching plane vectors is shown in green.

5. Conclusions

A new development of TD-DFT/MM energy and gradient using electrostatic potential fitting interaction Hamiltonian has been developed and implemented. This method has been used to describe the photochemistry of thymine in the gas phase and embedded in methanol and water nanomeric droplets, using critical point search in the excited states and the potential energy surface cut along a geodesic pathway connecting all minima. We show that the gas phase thymine can undergo an ultrafast relaxation pathway from the S_2 state to the ground state via a series of conical intersections. In solution, the barrierless path to the S_2/S_1 intersections is similar to in the gas phase. On the contrary, the S_0/S_1 intersections are highly dependent on the solvent state. Indeed, the approximate non-adiabatic coupling vector defining the branching plane of the $S_0/n\pi^*$ intersection strongly mixes the motions of thymine and the solvent. The $S_0/\pi\pi^*$ intersection was not found in solvent, but rather, a transfer from the $\pi\pi^*$ minimum to the $n\pi^*$ minimum is observed.

In conclusion, TD-DFT/MM is an efficient method to describe the excited state potential energy surface of chromophores embedded in complex media. From this study, we inferred that the $n\pi^*/\pi\pi^*$ intersection is independent of the solvent state and therefore should be an equivalently fast reaction both in gas phase and in solution. On the contrary, to reach the $n\pi^*$ minimum and $S_0/n\pi^*$ intersection requires a relaxation of the solvent and a particular configuration of the solvent molecules to reach the intersection. This implies that the $S_0/n\pi^*$ intersection will be much slower in solution than in the gas phase. It remains to be seen what implication this has for the photochemistry of nucleobases in DNA. Indeed, the ground/excited intersections involve out-of-plane motions that are strongly hindered in solution, but they might still be possible in the electrostatic environment of DNA. In the future, we will use similar techniques to study the degree of delocalization of the non-adiabatic coupling vectors in DNA.

Conflicts of Interest: The author declares no conflict of interest.

References

1. de Gruijl, F.R. Skin cancer and solar UV radiation. *Eur. J. Cancer* **1999**, *35*, 2003–2009. [CrossRef]
2. Chatterjee, N.; Walker, G.C. Mechanisms of DNA damage, repair, and mutagenesis. *Environ. Mol. Mutagen.* **2017**, *58*, 235–263. [CrossRef]
3. Nir, E.; Kleinermanns, K.; Grace, L.; de Vries, M.S. On the Photochemistry of Purine Nucleobases. *J. Phys. Chem. A* **2001**, *105*, 5106–5110. [CrossRef]

4. Merchán, M.; González-Luque, R.; Climent, T.; Serrano-Andrés, L.; Rodríguez, E.; Reguero, M.; Peláez, D. Unified model for the ultrafast decay of pyrimidine nucleobases. *J. Phys. Chem. B* **2006**, *110*, 26471—26476. [CrossRef]

5. Barbatti, M.; Aquino, A.J.A.; Szymczak, J.J.; Nachtigallová, D.; Hobza, P.; Lischka, H. Relaxation mechanisms of UV-photoexcited DNA and RNA nucleobases. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 21453–21458. [CrossRef]

6. Wolf, T.J.A.; Gühr, M. Photochemical pathways in nucleobases measured with an X-ray FEL. *Phil. Trans. R. Soc. A* **2019**, *377*, 20170473. [CrossRef]

7. Arslancan, S.; Martínez-Fernández, L.; Corral, I. Photophysics and Photochemistry of Canonical Nucleobases' Thioanalogs: From Quantum Mechanical Studies to Time Resolved Experiments. *Molecules* **2017**, *22*, 998. [CrossRef]

8. Gustavsson, T.; Bányász, A.; Lazzarotto, E.; Markovitsi, D.; Scalmani, G.; Frisch, M.J.; Barone, V.; Improta, R. Singlet Excited-State Behavior of Uracil and Thymine in Aqueous Solution: A Combined Experimental and Computational Study of 11 Uracil Derivatives. *J. Am. Chem. Soc.* **2006**, *128*, 607–619. [CrossRef]

9. Improta, R.; Barone, V. The excited states of adenine and thymine nucleoside and nucleotide in aqueous solution: A comparative study by time-dependent DFT calculations. *Theor. Chem. Acc.* **2008**, *120*, 491–497. [CrossRef]

10. Lischka, H.; Barbatti, M.; Siddique, F.; Das, A.; Aquino, A.J. The effect of hydrogen bonding on the nonadiabatic dynamics of a thymine-water cluster. *Chem. Phys.* **2018**, *515*, 472–479. [CrossRef]

11. Sun, S.; Brown, A. Effects of hydrogen bonding with H2O on the resonance Raman spectra of uracil and thymine. *Comput. Theor. Chem.* **2017**, *1100*, 70–82. [CrossRef]

12. Janicki, M.J.; Szabla, R.; Šponer, J.; Góra, R.W. Solvation effects alter the photochemistry of 2-thiocytosine. *Chem. Phys.* **2018**, *515*, 502–508. [CrossRef]

13. Johns, H.; Rapaport, S.; Delbrück, M. Photochemistry of thymine dimers. *J. Mol. Biol.* **1962**, *4*, 104–114. [CrossRef]

14. LeClerc, J.E.; Borden, A.; Lawrence, C.W. The thymine-thymine pyrimidine-pyrimidone(6-4) ultraviolet light photoproduct is highly mutagenic and specifically induces 3' thymine-to-cytosine transitions in Escherichia coli. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 9685–9689. [CrossRef]

15. Otoshi, E.; Yagi, T.; Mori, T.; Matsunaga, T.; Nikaido, O.; Kim, S.T.; Hitomi, K.; Ikenaga, M.; Todo, T. Respective Roles of Cyclobutane Pyrimidine Dimers, (6–4)Photoproducts, and Minor Photoproducts in Ultraviolet Mutagenesis of Repair-deficient Xeroderma Pigmentosum A Cells. *Cancer Res.* **2000**, *60*, 1729–1735.

16. Marguet, S.; Markovitsi, D. Time-Resolved Study of Thymine Dimer Formation. *J. Am. Chem. Soc.* **2005**, *127*, 5780–5781. [CrossRef]

17. Rauer, C.; Nogueira, J.J.; Marquetand, P.; González, L. Cyclobutane Thymine Photodimerization Mechanism Revealed by Nonadiabatic Molecular Dynamics. *J. Am. Chem. Soc.* **2016**, *138*, 15911–15916. [CrossRef]

18. Li, J.H.; Zuehlsdorff, T.J.; Payne, M.C.; Hine, N.D.M. Photophysics and Photochemistry of DNA Molecules: Electronic Excited States Leading to Thymine Dimerization. *J. Phys. Chem. C* **2018**, *122*, 11633–11640. [CrossRef]

19. Park, W.; Lee, S.; Huix-Rotllant, M.; Filatov, M.; Choi, C.H. Impact of the Dynamic Electron Correlation on the Unusually Long Excited-State Lifetime of Thymine. *J. Phys. Chem. Lett.* **2021**, *12*, 4339–4346. [CrossRef]

20. McFarland, B.K.; Farrell, J.P.; Miyabe, S.; Tarantelli, F.; Aguilar, A.; Berrah, N.; Bostedt, C.; Bozek, J.D.; Bucksbaum, P.H.; Castagna, J.C.; et al. Ultrafast X-ray Auger probing of photoexcited molecular dynamics. *Nat. Commun.* **2014**, *5*, 4235. [CrossRef]

21. Hudock, H.R.; Levine, B.G.; Thompson, A.L.; Satzger, H.; Townsend, D.; Gador, N.; Ullrich, S.; Stolow, A.; Martínez, T.J. Ab Initio Molecular Dynamics and Time-Resolved Photoelectron Spectroscopy of Electronically Excited Uracil and Thymine. *J. Phys. Chem. A* **2007**, *111*, 8500–8508. [CrossRef]

22. Picconi, D.; Barone, V.; Lami, A.; Santoro, F.; Improta, R. The Interplay between $\pi\pi^*/n\pi^*$ Excited States in Gas-Phase Thymine: A Quantum Dynamical Study. *ChemPhysChem* **2011**, *12*, 1957–1968. [CrossRef]

23. Casida, M.; Huix-Rotllant, M. Progress in Time-Dependent Density-Functional Theory. *Ann. Rev. Phys. Chem.* **2012**, *63*, 287–323. [CrossRef]

24. Ferré, N.; Ángyán, J.G. Approximate electrostatic interaction operator for QM/MM calculations. *Chem. Phys. Lett.* **2002**, *356*, 331–339. [CrossRef]

25. Schwinn, K.; Ferré, N.; Huix-Rotllant, M. Analytic QM/MM atomic charge derivatives avoiding the scaling of coupled perturbed equations with the MM subsystem size. *J. Chem. Phys.* **2019**, *151*, 041102. [CrossRef]

26. Schwinn, K.; Ferré, N.; Huix-Rotllant, M. Efficient Analytic Second Derivative of Electrostatic Embedding QM/MM Energy: Normal Mode Analysis of Plant Cryptochrome. *J. Chem. Theor. Comput.* **2020**, *16*, 3816–3824. [CrossRef]

27. Huix-Rotllant, M.; Ferré, N. Analytic Energy, Gradient, and Hessian of Electrostatic Embedding QM/MM Based on Electrostatic Potential-Fitted Atomic Charges Scaling Linearly with the MM Subsystem Size. *J. Chem. Theor. Comput.* **2021**, *17*, 538–548. [CrossRef]

28. Schwinn, K.; Ferré, N.; Huix-Rotllant, M. UV-visible absorption spectrum of FAD and its reduced forms embedded in a cryptochrome protein. *Phys. Chem. Chem. Phys.* **2020**, *22*, 12447–12455. [CrossRef]

29. Huix-Rotllant, M.; Schwinn, K.; Ferré, N. Infrared spectroscopy from electrostatic embedding QM/MM: local normal mode analysis of infrared spectra of arabidopsis thaliana plant cryptochrome. *Phys. Chem. Chem. Phys.* **2021**, *23*, 1666–1674. [CrossRef]

30. Maeda, S.; Ohno, K.; Morokuma, K. Updated Branching Plane for Finding Conical Intersections without Coupling Derivative Vectors. *J. Chem. Theory Comput.* **2010**, *6*, 1538–1545. [CrossRef]

31. Fetter, A. *Quantum Theory of Many-Particle Systems*; Dover Publications: Mineola, NY, USA, 2003.

32. Casida, M.E. Time-Dependent Density Functional Response Theory for Molecules. In *Recent Advances in Density Functional Methods*; World Scientific: Singapore, 1995; Volume 1, pp. 155–192.
33. Furche, F.; Ahlrichs, R. Adiabatic time-dependent density functional methods for excited state properties. *J. Chem. Phys.* **2002**, *117*, 7433–7447. [CrossRef]
34. Handy, N.C.; Schaefer, H.F. On the evaluation of analytic energy derivatives for correlated wave functions. *J. Chem. Phys.* **1984**, *81*, 5031–5033. [CrossRef]
35. Schmidt, M.W.; Baldridge, K.K.; Boatz, J.A.; Elbert, S.T.; Gordon, M.S.; Jensen, J.H.; Koseki, S.; Matsunaga, N.; Nguyen, K.A.; Su, S.; et al. General atomic and molecular electronic structure system. *J. Comput. Chem.* **1993**, *14*, 1347–1363. [CrossRef]
36. Ponder, J.W. *TINKER: Software Tools for Molecular Design*; Version 6.3.3; Washington University School of Medicine: Saint Louis, MO, USA, 2014.
37. Becke, A.D. A new mixing of Hartree–Fock and local density-functional theories. *J. Chem. Phys.* **1993**, *98*, 1372–1377. [CrossRef]
38. Ditchfield, R.; Hehre, W.J.; Pople, J.A. Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1971**, *54*, 724–728. [CrossRef]
39. Francl, M.M.; Pietro, W.J.; Hehre, W.J.; Binkley, J.S.; Gordon, M.S.; DeFrees, D.J.; Pople, J.A. Self-consistent molecular orbital methods. XXIII. A polarization-type basis set for second-row elements. *J. Chem. Phys.* **1982**, *77*, 3654–3665. [CrossRef]
40. Cornell, W.D.; Cieplak, P.; Bayly, C.I.; Gould, I.R.; Merz, K.M.; Ferguson, D.M.; Spellmeyer, D.C.; Fox, T.; Caldwell, J.W.; Kollman, P.A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197. [CrossRef]
41. Martínez, L.; Andrade, R.; Birgin, E.G.; Martínez, J.M. PACKMOL: A package for building initial configurations for molecular dynamics simulations. *J. Comput. Chem.* **2009**, *30*, 2157–2164. [CrossRef]
42. Zhu, X.; Thompson, K.C.; Martínez, T.J. Geodesic interpolation for reaction pathways. *J. Chem. Phys.* **2019**, *150*, 164103. [CrossRef]
43. Levine, B.G.; Ko, C.; Quenneville, J.; MartÍnez, T.J. Conical intersections and double excitations in time-dependent density functional theory. *Mol. Phys.* **2006**, *104*, 1039–1051. [CrossRef]
44. Burghardt, I.; Cederbaum, L.S.; Hynes, J.T. Environmental effects on a conical intersection: A model study. *Faraday Discuss.* **2004**, *127*, 395–411. [CrossRef] [PubMed]