THE JOURNAL OF PHYSICAL CHEMISTRY Letters

Letter

pubs.acs.org/JPCL

Quantum Mechanics/Molecular Mechanics Free Energy Maps and Nonadiabatic Simulations for a Photochemical Reaction in DNA: Cyclobutane Thymine Dimer

Jesús I. Mendieta-Moreno,[†,‡] Daniel G. Trabada,[†] Jesús Mendieta,[†,‡,§] James P. Lewis,[¶] Paulino Gómez-Puertas,[‡] and José Ortega*[,†]

[†]Departamento de Física Teórica de la Materia Condensada and Condensed Matter Physics Center (IFIMAC), Universidad Autónoma de Madrid, ES-28049 Madrid, Spain

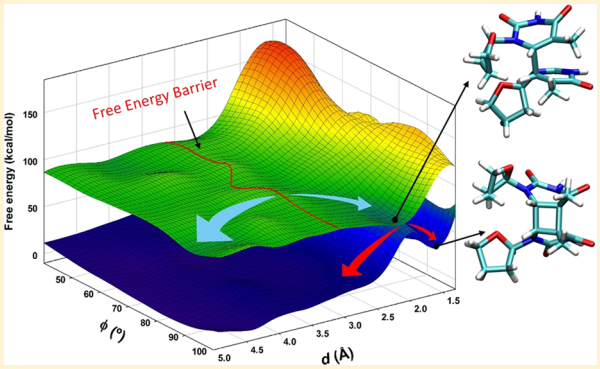
[‡]Molecular Modelling Group, Center of Molecular Biology Severo Ochoa (CSIC-UAM), ES-28049 Madrid, Spain

[¶]Department of Physics, West Virginia University, Morgantown, West Virginia 26506-6315, United States

[§]Departamento de Biotecnología, Universidad Francisco de Vitoria, ctra. Pozuelo-Majadahonda, km 1,800, 28223 Pozuelo de Alarcón, Madrid, Spain

Supporting Information

ABSTRACT: The absorption of ultraviolet radiation by DNA may result in harmful genetic lesions that affect DNA replication and transcription, ultimately causing mutations, cancer, and/or cell death. We analyze the most abundant photochemical reaction in DNA, the cyclobutane thymine dimer, using hybrid quantum mechanics/molecular mechanics (QM/MM) techniques and QM/MM nonadiabatic molecular dynamics. We find that, due to its double helix structure, DNA presents a free energy barrier between nonreactive and reactive conformations leading to the photolesion. Moreover, our nonadiabatic simulations show that most of the photoexcited reactive conformations return to standard B-DNA conformations after an ultrafast nonradiative decay to the ground state. This work highlights the importance of dynamical effects (free energy, excited-state dynamics) for the study of photochemical reactions in biological systems.



The absorption of ultraviolet (UV) radiation by DNA may trigger different photochemical reactions that lead to the formation of covalent bonds between adjacent nucleobases.[1,2] DNA presents remarkable stability against this photodamage: more than 99.9% of the absorbed photons are transformed into heat, which is transferred to the solvent without causing any lesion.[3] The most frequent DNA photolesion produced by sunlight is the cyclobutane thymine dimer (CTD) that is characterized by the formation of a ring of four carbon atoms that includes two C−C bonds between adjacent thymine bases.[4]

The development of new experimental techniques has provided very interesting information regarding these photochemical reactions.[1,2,5−13] Due to the intricacy of cellular DNA, most of these investigations have been carried for DNA bases or short single-stranded synthetic DNA molecules in solution. It has been shown that after photoexcitation, these systems present ultrafast relaxation to the electronic ground state. In particular, studies on all thymine oligonucleotides have shown that CTD formation is an ultrafast reaction along the singlet state that is completed in 1 ps.[7,8] In light of these experiments, it has been suggested that the conformation of the thymine bases at the moment of photon absorption determines if CTD will be formed and that the low quantum yield for the reaction is due to the small number of reactive conformations in the thermal ensemble at room temperature (RT).[7,14−16] Experiments on genomic DNA suggest, however, that the formation of photolesions is a very complex problem in which different mechanisms (interaction with adjacent bases, collective excited states, etc.) play an important role.[17−19]

Despite these important advances in the experimental techniques, theoretical modeling is essential to understand the photochemistry of DNA.[10,20,21] Due to the complexity of this problem, most of the theoretical studies have focused on the analysis of the photochemical reaction for simpler systems such as two nucleobases in the gas phase[22−29] or thymine dinucleoside monophosphate (TpT) in water,[30,31] using accurate quantum chemistry or density functional theory (DFT) methods. In these studies, the reaction takes place via

Received: September 22, 2016
Accepted: October 21, 2016
Published: October 21, 2016

DOI: 10.1021/acs.jpclett.6b02168
J. Phys. Chem. Lett. 2016, 7, 4391−4397

a nonadiabatic mechanism on a barrierless singlet excited-state (S_1) pathway to a conical intersection[32,33] between the S_1 and S_0 potential energy surfaces (PESs), which is the funnel for ultrafast decay into the ground state (S_0), leading to the photolesion. However, the dynamics of these photochemical reactions should be quite different for the case of cellular DNA with the bases immersed in the complex DNA environment. Notice also that the biomolecular system presents a high number of degrees of freedom and a large number of accessible conformations for a given state of the system at RT.[20,34] Due to the computational cost, this problem is usually ignored, assuming that a single energy-minimized structure represents each state. Moreover, in order to model these photochemical reactions, we must go beyond the Born−Oppenheimer approximation and take into account the excited-state dynamics and nonadiabatic couplings of the electronic and nuclear degrees of freedom. Theoretical analysis of photochemical reactions in biomolecules is a very challenging problem that requires mixing different simulation strategies.[19−21,35]

In this work, we analyze theoretically CTD formation in a realistic model for DNA at physiological conditions. We use quantum mechanics/molecular mechanics (QM/MM) techniques in which the QM method is used to describe the region where the photochemical reaction is taking place, while the MM method is used to properly take into account the rest of the DNA double helix structure, as well as the water molecules and counterions in the solvent.[19,36,37] In a first step, we determine the S_0 and S_1 free energy maps for the reaction as a function of the reaction coordinates (see below) that define the state of the system. In these free energy maps, each state is associated with an ensemble of different structures and energy values, taking into account in this way the plethora of accessible structures at RT.[34] The calculation of these free energy maps allows us to define, in a precise way, the region of reactive conformations,[7] that is, the region of conformational space for which electron photoexcitation may lead to CTD formation in an ultrafast reaction; this region is closely related to the S_1/S_0 conical intersection zone (CIZ) in which de-excitation from S_1 to S_0 actually takes place. In a second step, we perform nonadiabatic[33,38−41] QM/MM molecular dynamics (MD) to further analyze the atomic mechanisms of this reaction. These simulations allow us to obtain the probability of CTD formation or, alternatively, internal conversion back to initial DNA conformations, once the system reaches the CIZ. This work highlights the importance of the dynamical effects (free energy, excited-state motion, and nonadiabatic dynamics) for the study of photochemical reactions in biomolecules.

Our system consists of a fragment of solvated B-DNA with 10 DNA base pairs (base sequence: CGAATTAAGC)[42] and a water cap of 30 Å that also includes the corresponding Na^+ counterions. The QM region contains 52 atoms: the two adjacent thymine nucleobases as well as the two corresponding deoxyriboses; see Figure 1. The two thymine bases are located in the same DNA strand, in the middle of the DNA sequence. The rest of the DNA atoms in the two strands (phosphate backbones and the rest of the nucleosides) and all of the water molecules and counterions (Na^+) are included in the MM region. In total, this MM region contains ∼11 000 atoms.

Figure 2a,b shows typical atomic structures for the two thymine nucleosides before and after the photochemical reaction; the two C double bonds C5═C6 and C5′═C6′ in adjacent thymine bases give rise to a cyclobutane ring with the formation of the two covalent bonds C5−C5′ and C6−C6′

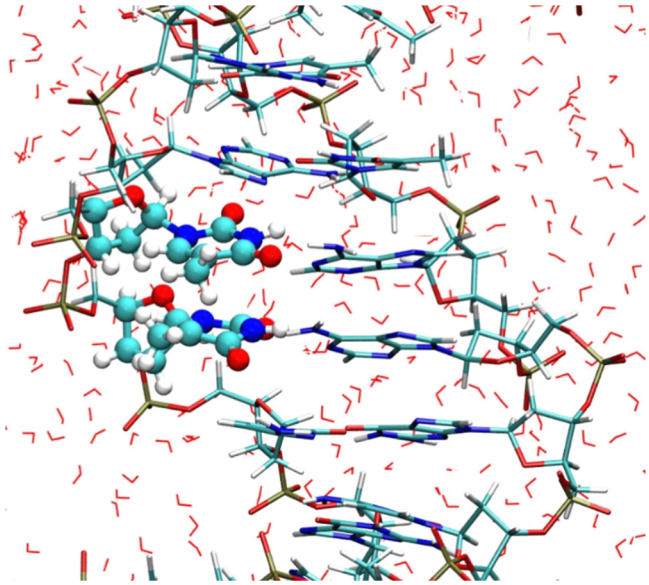


Figure 1. View of the central part of the DNA double helix showing the QM region and the surrounding DNA and solvent environment. The QM region (highlighted in the figure) consists of two adjacent thymine nucleosides in the middle of a sequence of 10 DNA base pairs, while the MM region comprises the rest of the DNA atoms and all of the water molecules and counterions in the solvent.

(see the notation in Figure 2b). Figure 2a also shows the two reaction coordinates that we have used in our study: the distance, d, between the midpoints of the two C5−C6 bonds and the angle ϕ for atoms C5′−C6′−C6. Using these reaction coordinates, we have determined the free energy maps for the S_0 and S_1 electronic states, following the methodology described in the Computational Methods section.

Figure 3a shows the ground-state free energy map as a function of the coordinates d and ϕ. In standard DNA conformations, the distance d between the two thymine bases is ∼4−5 Å, and the angle ϕ is ∼50−100°. The local minimum that appears for values of $\phi \approx 90°$ and $d \approx 1.6$ Å corresponds to the CTD photolesion (Figure 2b). Between this minimum and standard DNA conformations, there is a high free energy barrier (red dashed line in Figure 3a) of 60−70 kcal/mol; the system cannot overcome this barrier at RT. This result is in agreement with previous calculations for simpler systems, for example, thymine dimers in the gas phase[22−29] or TpT in water,[30,31] as well as with the energy difference between the conical intersection and ground state found in a QM/MM study for B-DNA, 77 kcal/mol[19] (in this QM/MM calculation, the difference between the energy-minimized CTD and the ground state of B-DNA is 33 kcal/mol, in excellent agreement with the free energy value shown in Figure 3a).

Figure 3b shows the free energy map for the first excited singlet state S_1. The region corresponding to CTD configurations ($d \approx 1.6$ Å and $\phi \approx 90°$) is clearly unstable now, that is, in the S_1 state, the two thymine bases cannot get too close to each other at RT. The main feature in this map is the minimum found for $d \approx 2.2$ Å and $\phi \approx 95°$; as discussed below, this minimum determines the region in conformational space where de-excitation from S_1 to S_0 takes place, that is, the CIZ. Notice that, at variance with calculations for two thymine bases[22−29] or TpT,[30,31] in double-helix DNA there is an energy barrier between standard DNA conformations and the CIZ (red dashed line in Figure 3b). This barrier defines in a precise way

DOI: 10.1021/acs.jpclett.6b02168
J. Phys. Chem. Lett. 2016, 7, 4391−4397

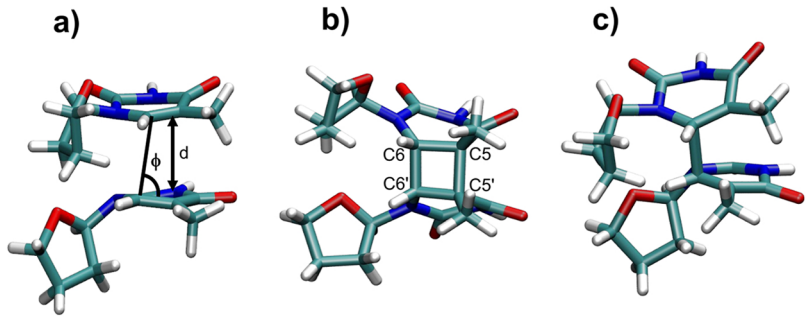


Figure 2. Representative atomic structures of the two thymine nucleosides. (a) Standard DNA; (b) CTD; (c) CIZ. The C atoms C5, C6 (and C5′, C6′ in the other thymine base) are indicated in (b). The reaction coordinates used in our study are shown in (a); d is the distance between the midpoints of the two C5−C6 bonds and ϕ is the angle for atoms C5′−C6′−C6.

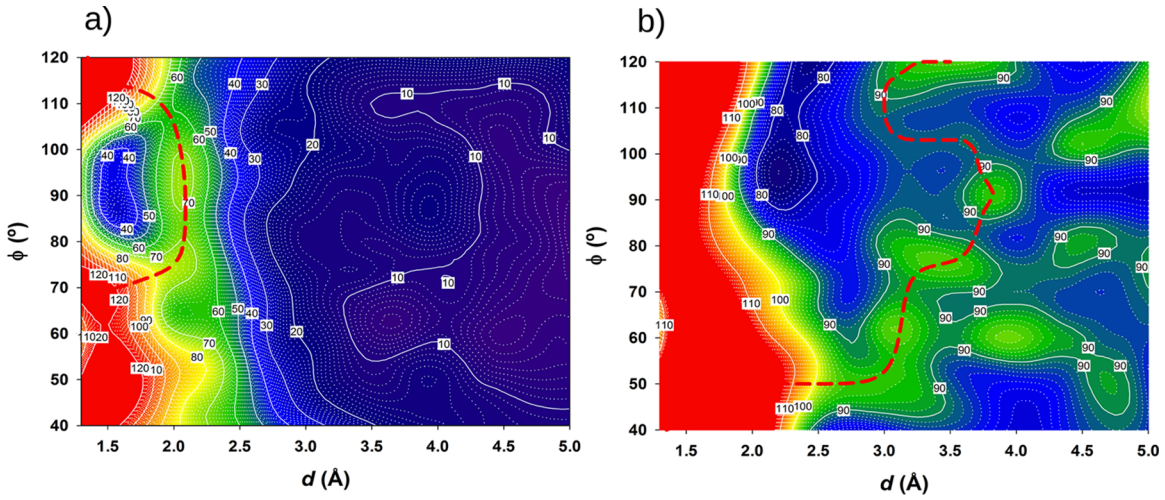


Figure 3. Free energy (kcal/mol) maps at RT as a function of the reaction coordinates d and ϕ for (a) the ground-state electronic state, S_0, and (b) first excited state, S_1. The red dashed line in (a) indicates the barrier between B-DNA and CTD configurations in the S_0 state, while the red dashed line in (b) marks the barrier for reactive conformations in the S_1 excited state. The origin of energies is the same in both figures.

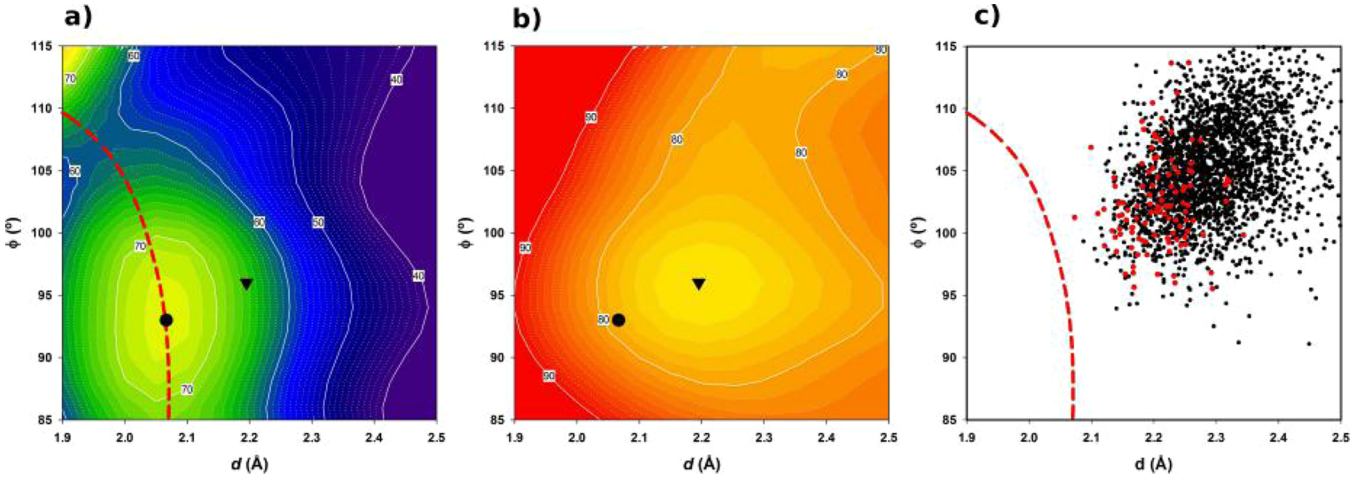


Figure 4. CIZ and $S_1 \rightarrow S_0$ transitions. (a,b) Zoom of the free energy maps (Figure 3) around the CIZ: (a) S_0 state; (b) S_1 state. The black triangle and black circle indicate the position of the S_1 minimum and S_0 maximum, respectively. (c) Spots corresponding to the positions in which the $S_1 \rightarrow S_0$ transition takes place in the different nonadiabatic QM/MM simulations; the red dots correspond to the trajectories that end up in CTD formation. The red dashed lines in (a) and (c) indicate the barrier between B-DNA and CTD configurations; see Figure 3a.

the region of reactive conformations; configurations above and at the left of this dashed line will move toward the CIZ and may result in CTD formation, while those below and at the right of this line move in the opposite direction, toward B-DNA conformations.

This minimum in the S_1 map is located close to a local maximum in the S_0 map; see Figure 4a,b. In the region between these two extremal points ($d \approx 2.1-2.2$ Å and $\phi \approx 92-98°$), the S_0 and S_1 free energy surfaces are very close. The characteristic atomic geometry for the two thymine nucleosides

DOI: 10.1021/acs.jpclett.6b02168
J. Phys. Chem. Lett. 2016, 7, 4391−4397

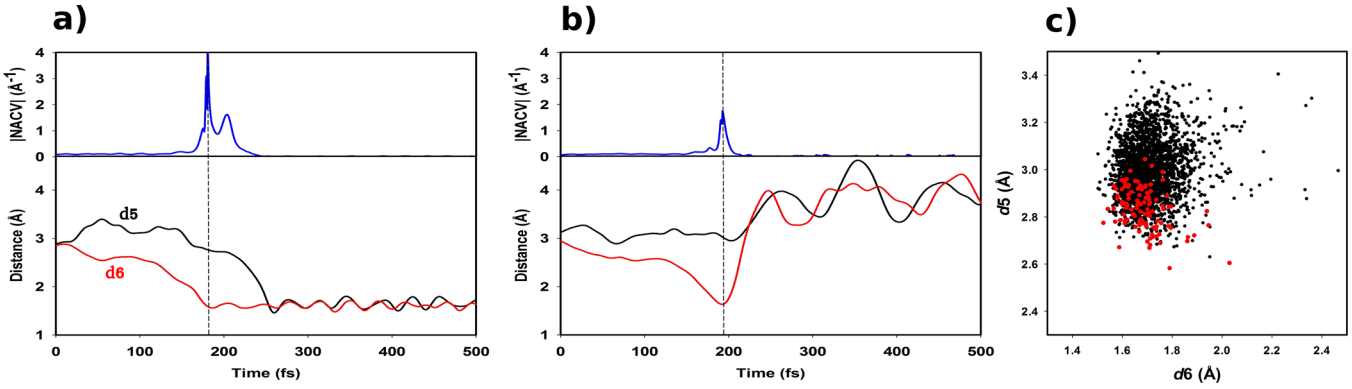


Figure 5. (a,b) Examples of nonadiabatic QM/MM trajectories: evolution of the S_1/S_0 nonadiabatic coupling vector (modulus) and of the C5−C5′ (*d*5; black line) and C6−C6′ (*d*6; red line) distances. In case (a), the trajectory ends up in CTD formation, while in (b), the system returns to B-DNA conformations. The dashed vertical lines indicate the instant of the $S_1 \rightarrow S_0$ transition. (c) Spots corresponding to the positions in which the $S_1 \rightarrow S_0$ transition takes place in the different nonadiabatic QM/MM simulations as a function of the *d*5 and *d*6 coordinates; only a small fraction (∼4%) of the trajectories end up in CTD formation (red spots).

in this region is shown in Figure 2c: a covalent bond between the C6 atoms of both thymine bases is forming, with a C6−C6′ distance, *d*6, of ∼1.7 Å, while the C5−C5′ distance, *d*5, is larger, ∼2.5 Å, reflecting the steric effect of the two methyl groups.

The free energy represented in Figures 3 and 4a,b is a statistical quantity, that is, each point in these maps is associated with an ensemble of accessible atomic structures at RT (see the Computational Methods). As shown below, in the region highlighted in Figure 4, there are many different atomic configurations for which the S_1 and S_0 PESs are very close and thus present a strong S_1/S_0 coupling and a large probability for $S_1 \rightarrow S_0$ transition.

In order to further investigate the atomic mechanisms for this photochemical reaction, we have performed nonadiabatic MD for an ensemble of 2560 trajectories using the fewest-switches surface-hopping method[38,39,41] as implemented in our QM/MM approach. In these simulations, we are interested in the fate of the excited-state (S_1) trajectories that reach the CIZ and thus may result in CTD formation. Therefore, the initial conditions for these trajectories were chosen so that at the moment of photoexcitation the system is already inside of the region of reactive conformations (see the Computational Methods).

Figure 5a,b show two representative examples of these trajectories. After the initial (t = 0) photoexcitation, the system very quickly (∼200 fs) moves to the CIZ. This ultrafast process is driven by the attraction between the C6 atoms in adjacent thymine bases; notice how the distance between the two C6 atoms, *d*6, is rapidly reduced to ∼1.6 Å, revealing the incipient formation of a covalent bond between the two bases. At the CIZ, the S_1/S_0 nonadiabatic coupling is large and the system quickly decays into the S_0 state (dashed vertical lines). Then, the C5 atoms may form a second covalent bond, leading to the CTD (e.g., Figure 5a), or the system may return to initial B-DNA configurations, breaking the newly formed C6−C6′ bond (e.g., Figure 5b).

Interestingly, out of the 2560 trajectories, only 110 (∼4%) end up in the photolesion, while in the other 2450 trajectories, the system goes back to standard DNA conformations. The values of the (d, ϕ) coordinates and (*d*5, *d*6) distances at the moment of the S_1/S_0 transition for the different trajectories are shown in Figures 4c and 5c, respectively; the red dots correspond to the cases that end up in the CTD. We observe (Figure 4c) that practically all of the transitions take place on the larger d values' side of the free energy barrier separating the CTD and B-DNA regions; this is related to the sloped topography[43] of the CIZ (see Figure 4a,b) as well as to the nonequilibrium nature of these ultrafast processes. Notice that the S_1/S_0 transitions occur in a relatively wide range of values, $d \approx 2.1-2.5$ Å and $\phi \approx 90-115°$, showing that there are many different DNA conformations that present strong S_1/S_0 coupling in this region. Figures 4c and 5c also show that CTD formation is more likely when the S_1/S_0 transition occurs for shorter d and *d*5 distances, closer to the barrier between CTD and B-DNA conformations. Finally, we mention that the statistical distribution of S_1/S_0 transition points shown in Figure 4c is displaced to larger d and ϕ values, as compared to the CIZ determined from the free energy maps, Figure 4a,b. This is a reflection of the "out of equilibrium" nature of the ultrafast decay after photoexcitation (Figure 5a,b).

The formation of photolesions in genomic DNA induced by UV radiation is a very complex process, and several factors should be considered in its analysis. These include the ground-state conformation at the moment of photoexcitation but also the role of collective excited states, the effects of the flanking bases, different pathways for excited-state quenching, long-lived excited states, and so forth.[7,16−19] Also, more detailed analysis of the interaction of UV radiation with DNA should be considered (e.g., change of the absorption probability with DNA conformation, effect of the energy spectrum of the UV source, etc.). Our nonadiabatic simulations show the important role of the excited-state dynamics. In particular, our results show that once a reactive conformation is photoexcited, the system reaches the CIZ in an ultrafast motion along the S_1 PES. Here, there is a large coupling between the S_1 and S_0 states, and consequently, there is a high probability that the system returns to the electronic ground state. At this point, the system may evolve in two different ways: (a) move back to initial DNA conformations, that is, internal conversion of the absorbed photon into heat, or (b) formation of the CTD photolesion. Due to the sloped topography of the CIZ and the ultrafast nature of the decay, case (a) is much more likely than case (b). The experimental quantum yield of CTD formation in genomic DNA[3] is a factor of 30−100 lower than for the case of oligonucleotides.[6,16,31] This large difference may be partially

DOI: 10.1021/acs.jpclett.6b02168
J. Phys. Chem. Lett. 2016, 7, 4391−4397

explained by the frequency of TT pairs and the UV absorption by other nucleobases;[7] our results show that the dynamical effects discussed here (free energy barrier for reactive conformations and ultrafast nonadiabatic excited-state dynamics) also play a very important role.

In conclusion, we have analyzed how the dynamical properties and biological environment affect the photostability of B-DNA. The barrier found in the S_1 free energy map defines the DNA conformations that can reach the CIZ in an ultrafast motion after photoexcitation. This free energy barrier is due to the constraints imposed by the Watson−Crick hydrogen bonds and DNA backbone; notice that in the cases of two thymine bases in the gas phase[22,23] or TpT in water[30,31] there is practically no energy barrier in the S_1 PES pathway to the conical intersection. In addition, our nonadiabatic simulations reveal that the probability of CTD formation from a photoexcited reactive conformation is low; this is due to the sloped topography of the CIZ and the nonequilibrium nature of the ultrafast nonradiative decay.

COMPUTATIONAL METHODS

In this study, we used a recently developed QM/MM method, Fireball/Amber,[37] based on the combination of techniques developed in the fields of molecular biology (Amber[44]) and condensed matter physics (Fireball[45,46]). In this technique, the QM region is described using a local-orbital DFT method[45,46] that presents an excellent balance between computational efficiency and accuracy. This allowed us to calculate, using QM/MM MD, the S_0 and S_1 free energy maps[34] as well as the ensemble of 2560 nonadiabatic trajectories.

In the MM part, we used the Amber[44] force field *parm99bsc0* for the DNA molecule and TIP3P for the water molecules. Regarding the DFT calculations, we used the BLYP exchange−correlation functional[47,48] and norm-conserving pseudopotentials.[49] We employed a basis set of optimized numerical atomic-like orbitals (NAOs)[50] with a 1s orbital for H and sp^3 orbitals for C, N, and O. In order to optimize this basis set, we analyzed different sets of organic molecules relevant for biological systems and considered intermolecular energies and distances as well as intramolecular distances. In particular, we used the S66 reference data set[51] for intermolecular energies and distances as well as the ionic hydrogen bond data set[52] for complexes featuring hydrogen bonds between ionic and neutral groups. Our optimized basis set yields a mean absolute deviation of 2.14 kcal/mol for the intermolecular energies in the S66 data set. Details will be published elsewhere.[53]

In the Fireball method, a practical tabulation/interpolation scheme is used to speed up the DFT simulations: all of the orbital integrals are calculated beforehand and stored in data tables that are loaded in memory in the initial step and used throughout the MD run. For this purpose, we used the multicenter weighted exchange−correlation density approximation[54] for calculation of the exchange−correlation terms.

The free energy maps shown in Figures 3 and 4 were obtained by means of long steered QM/MM MD simulations at RT along the S_0 or S_1 PES, as described in ref 34. In this way, $\sim 3 \times 10^6$ atomic configurations were generated for each map in order to properly explore the configuration space of the system for values of the reaction coordinates between 1.3 and 5 Å (d) and 40 and 120° (ϕ). The calculation of this huge number of QM/MM configurations is possible due to the computational efficiency of Fireball/Amber.[37] The maps were finally obtained by defining an (x,y) grid of ∼3000 points with 75 different d values (steps of Δd = 0.05 Å) and 41 different values for ϕ ($\Delta\phi$ = 2°). To each of these (x,y) grid points we associated ∼1000 different conformations, obtained from the steered QM/MM MD mentioned above, that were then used to calculate the free energy for that grid point; each point in the map represents an ensemble of accessible structures.[34] The profile lines for the energy barriers shown in Figure 3 were calculated using the program MEPSA.[55] For the calculation of the S_1 PES, we used a constrained DFT calculation, solving the Kohn−Sham equations self-consistently with constrained Kohn−Sham orbital occupations, in which one electron is promoted from the HOMO (highest occupied molecular orbital) to the LUMO (lowest unoccupied molecular orbital),[41] a method also known as ΔSCF.[56]

In the nonadiabatic simulations, the system was excited to the S_1 PES in the initial time step, and we let the system evolve freely for 1500 fs (3000 time steps for each trajectory), using the fewest-switches surface-hopping method[38,39,41] as implemented in our QM/MM approach. In these trajectories, the atoms moved initially in the adiabatic S_1 PES, but in each time step, there was a probability for the system to jump to the S_0 PES. The probability for these hops was determined from the nonadiabatic coupling[57] and time evolution of the electronic states, using the fewest-switches algorithm;[38] see ref 41 for all of the details. This probability is negligible except when the S_1/S_0 coupling is large, that is, in the CIZ.

The initial atomic positions and velocities for these trajectories were chosen from QM/MM steered MD simulations in the S_0 PES in the region $2.5 < d < 3.2$ Å and $75 < \phi < 100°$ (see Figure 3). This represents a small fraction of the RT-accessible ground-state conformations; however, we are interested here in trajectories that reach the CIZ in order to analyze the atomic dynamics associated with the $S_1 \rightarrow S_0$ transition and, in particular, to determine how many of these trajectories yield the CTD and how many undergo internal conversion back to nonreactive native B-DNA configurations.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpclett.6b02168.

Calculations for the thymine dimer in the gas phase, evolution of the frontier orbitals along a nonadiabatic trajectory, and reactive conformations from extensive MM MD simulations at RT(PDF)

AUTHOR INFORMATION

Corresponding Author

*E-mail: jose.ortega@uam.es.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

J.I.M.-M., D.G.T., J.M., and J.O. acknowledge financial support from the Spanish Ministerio de Economía y Competitividad (MINECO) under Project MAT2014-59966-R and through the "María de Maeztu" Programme for Units of Excellence in R&D (MDM-2014-0377). P.G-P. acknowledges financial support from the Spanish Ministerio de Economía y Competitividad (MINECO) under Project IPT2011-0964-900000. J.P.L. acknowledges support by the U.S. Department of Energy (DE-SC0004737).

DOI: 10.1021/acs.jpclett.6b02168
J. Phys. Chem. Lett. 2016, 7, 4391−4397

REFERENCES

(1) Schreier, W. J.; Gilch, P.; Zinth, W. Early events of DNA photodamage. *Annu. Rev. Phys. Chem.* **2015**, *66*, 497−519.
(2) Cadet, J.; Mouret, S.; Ravanat, J. L.; Douki, T. Photoinduced damage to cellular DNA: Direct and photosensitized reactions. *Photochem. Photobiol.* **2012**, *88*, 1048−1065.
(3) Douki, T. Effect of denaturation on the photochemistry of pyrimidine bases in isolated DNA. *J. Photochem. Photobiol., B* **2006**, *82*, 45−52.
(4) Beukers, R.; Eker, A. P. M.; Lohman, P. H. M. 50 Years Thymine Dimer. *DNA Repair* **2008**, *7*, 530−543.
(5) Crespo-Hernández, C. E.; Cohen, B.; Hare, P. M.; Kohler, B. Ultrafast excited-state dynamics in nucleic acids. *Chem. Rev.* **2004**, *104*, 1977−2019.
(6) Marguet, S.; Markovitsi, D. Time-resolved study of thymine dimer formation. *J. Am. Chem. Soc.* **2005**, *127*, 5780−5781.
(7) Schreier, W. J.; Schrader, T. E.; Koller, F. O.; Gilch, P.; Crespo-Hernandez, C. E.; Swaminathan, V. N.; Carell, T.; Zinth, W.; Kohler, B. Thymine dimerization in DNA is an ultrafast photoreaction. *Science* **2007**, *315*, 625.
(8) Schreier, W. J.; Kubon, J.; Regner, N.; Haiser, K.; Schrader, T. E.; Zinth, W.; Clivio, P.; Gilch, P. Thymine dimerization in DNA model systems: cyclobutane photolesion is predominantly formed via the singlet channel. *J. Am. Chem. Soc.* **2009**, *131*, 5038−9.
(9) Middleton, C. T.; de La Harpe, K.; Su, C.; Law, Y. K.; Crespo-Hernández, C. E.; Kohler, B. DNA Excited-State Dynamics: From Single Bases to the Double Helix. *Annu. Rev. Phys. Chem.* **2009**, *60*, 217−239.
(10) Gustavsson, T.; Improta, R.; Markovitsi, D. DNA/RNA: Building blocks of life under UV irradiation. *J. Phys. Chem. Lett.* **2010**, *1*, 2025−2030.
(11) Pilles, B. M.; Bucher, D. B.; Liu, L.; Clivio, P.; Gilch, P.; Zinth, W.; Schreier, W. J. Mechanism of the decay of thymine triplets in DNA single strands. *J. Phys. Chem. Lett.* **2014**, *5*, 1616−1622.
(12) Chen, J.; Zhang, Y.; Kohler, B. Excited states in DNA strands investigated by ultrafast laser spectroscopy. *Top. Curr. Chem.* **2014**, *356*, 39.
(13) Cadet, J.; Grand, A.; Douki, T. Solar UV Radiation-Induced DNA Bipyrimidine Photoproducts: Formation and Mechanistic Insights. *Top. Curr. Chem.* **2014**, *356*, 249−276.
(14) Law, Y. K.; Azadi, J.; Crespo-Hernández, C. E.; Olmon, E.; Kohler, B. Predicting thymine dimerization yields from molecular dynamics simulations. *Biophys. J.* **2008**, *94*, 3590−600.
(15) Johnson, A. T.; Wiest, O. Structure and dynamics of poly(T) single-strand DNA: Implications toward CPD formation. *J. Phys. Chem. B* **2007**, *111*, 14398−14404.
(16) McCullagh, M.; Hariharan, M.; Lewis, F. D.; Markovitsi, D.; et al. Conformational Control of TT Dimerization in DNA Conjugates. A Molecular Dynamics Study. *J. Phys. Chem. B* **2010**, *114*, 5215−5221.
(17) Markovitsi, D. UV-induced DNA Damage: The Role of Electronic Excited States. *Photochem. Photobiol.* **2016**, *92*, 45−51.
(18) Vayá, I.; Gustavsson, T.; Douki, T.; Berlin, Y.; Markovitsi, D. Electronic Excitation Energy Transfer between Nucleobases of Natural DNA. *J. Am. Chem. Soc.* **2012**, *134*, 11366−11368.
(19) Lee, W.; Matsika, S. QM/MM studies reveal pathways leading to the quenching of the formation of thymine dimer photoproduct by flanking bases. *Phys. Chem. Chem. Phys.* **2015**, *17*, 9927−9935.
(20) Lu, Y.; Lan, Z.; Thiel, W. Computational modeling of photoexcitation in DNA single and double strands. *Top. Curr. Chem.* **2014**, *356*, 89−122.
(21) Plasser, F.; Aquino, A. J. A.; Lischka, H.; Nachtigallova, D. Electronic excitation processes in single-strand and double-strand DNA: a computational approach. *Top. Curr. Chem.* **2014**, *356*, 1−38.
(22) Boggio-Pasqua, M.; Groenhof, G.; Schäfer, L. V.; Grubmüller, H.; Robb, M. A. Ultrafast deactivation channel for thymine dimerization. *J. Am. Chem. Soc.* **2007**, *129*, 10996−7.
(23) Blancafort, L.; Migani, A. Modeling Thymine Photodimerization DNA: Mech & Correlation Diagram. *J. Am. Chem. Soc.* **2007**, *129*, 14540−14541.
(24) Serrano-Pérez, J. J.; González-Ramírez, I.; Coto, P. B.; Merchán, M.; Serrano-Andrés, L. Theoretical insight into the intrinsic ultrafast formation of cyclobutane pyrimidine dimers in UV-irradiated DNA: Thymine versus cytosine. *J. Phys. Chem. B* **2008**, *112*, 14096−14098.
(25) Climent, T.; González-Ramírez, I.; González-Luque, R.; Merchán, M.; Serrano-Andrés, L. Cyclobutane Pyrimidine Photodimerization of DNA/RNA Nucleobases in the Triplet State. *J. Phys. Chem. Lett.* **2010**, *1*, 2072−2076.
(26) Barbatti, M. Computational Reference Data for the Photochemistry of Cyclobutane Pyrimidine Dimers. *ChemPhysChem* **2014**, *15*, 3342−3354.
(27) Durbeej, B.; Eriksson, L. A. Reaction mechanism of thymine dimer formation in DNA induced by UV light. *J. Photochem. Photobiol., A* **2002**, *152*, 95−101.
(28) Zhang, R. B.; Eriksson, L. A. A triplet mechanism for the formation of cyclobutane pyrimidine dimers in UV-irradiated DNA. *J. Phys. Chem. B* **2006**, *110*, 7556.
(29) Rössle, S.; Friedrichs, J.; Frank, I. The formation of DNA photodamage: The role of exciton localization. *ChemPhysChem* **2010**, *11*, 2011−2015.
(30) Improta, R. Photophysics and photochemistry of thymine deoxy-dinucleotide in water: a PCM/TD-DFT quantum mechanical study. *J. Phys. Chem. B* **2012**, *116*, 14261−14274.
(31) Banyasz, A.; Douki, T.; Improta, R.; Gustavsson, T.; Onidas, D.; Vayá, I.; Perron, M.; Markovitsi, D. Electronic Excited States Responsible for Dimer Formation upon UV Absorption Directly by Thymine Strands: Joint Experimental and Theoretical Study. *J. Am. Chem. Soc.* **2012**, *134*, 14834−14845.
(32) Yarkony, D. R. Diabolical conical intersections. *Rev. Mod. Phys.* **1996**, *68*, 985−1013.
(33) Matsika, S.; Krause, P. Nonadiabatic events and conical intersections. *Annu. Rev. Phys. Chem.* **2011**, *62*, 621−643.
(34) Mendieta-Moreno, J. I.; Marcos-Alcalde, I.; Trabada, D. G.; Gómez-Puertas, P.; Ortega, J.; Mendieta, J. A Practical Quantum Mechanics Molecular Mechanics Method for the Dynamical Study of Reactions in Biomolecules. *Adv. Protein Chem. Struct. Biol.* **2015**, *100*, 67−88.
(35) Improta, R.; Santoro, F.; Blancafort, L. Quantum Mechanical Studies on the Photophysics and the Photochemistry of Nucleic Acids and Nucleobases. *Chem. Rev.* **2016**, *116*, 3540−3593.
(36) Senn, H. M.; Thiel, W. QM/MM methods for biomolecular systems. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198−1229.
(37) Mendieta-Moreno, J. I.; Walker, R. C.; Lewis, J. P.; Gómez-Puertas, P.; Mendieta, J.; Ortega, J. FIREBALL/AMBER: An Efficient Local-Orbital DFT QM/MM Method for Biomolecular Systems. *J. Chem. Theory Comput.* **2014**, *10*, 2185−2193.
(38) Tully, J. C. Molecular dynamics with electronic transitions. *J. Chem. Phys.* **1990**, *93*, 1061.
(39) Craig, C. F.; Duncan, W. R.; Prezhdo, O. V. Trajectory surface hopping in the time-dependent Kohn-Sham approach for electron-nuclear dynamics. *Phys. Rev. Lett.* **2005**, *95*, 163001.
(40) Hammes-Schiffer, S. Current Theoretical Challenges in Proton-Coupled Electron Transfer: Electron-Proton Nonadiabaticy, Proton Relays, and Ultrafast Dynamics. *J. Phys. Chem. Lett.* **2011**, *2*, 1410−1416.
(41) Zobač, V.; Lewis, J. P.; Abad, E.; Mendieta-Moreno, J. I.; Hapala, P.; Jelínek, P.; Ortega, J. Photo-induced reactions from efficient molecular dynamics with electronic transitions using the FIREBALL local-orbital density functional theory formalism. *J. Phys.: Condens. Matter* **2015**, *27*, 175002.
(42) Park, H.; Zhang, K.; Ren, Y.; Nadji, S.; Sinha, N.; Taylor, J.-S.; Kang, C. Crystal structure of a DNA decamer containing a cis-syn thymine dimer. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 15965−70.
(43) Levine, B. G.; Martínez, T. J. Isomerization through conical intersections. *Annu. Rev. Phys. Chem.* **2007**, *58*, 613−634.

DOI: 10.1021/acs.jpclett.6b02168
J. Phys. Chem. Lett. 2016, 7, 4391−4397

(44) Salomon-Ferrer, R.; Case, D. A.; Walker, R. C. An Overview of the Amber biomolecular simulation packages. *WIREs Comput. Mol. Sci.* **2013**, *3*, 198−210.
(45) Lewis, J. P.; Glaesemann, K. R.; Voth, G. A.; Fritsch, J.; Demkov, A. A.; Ortega, J.; Sankey, O. F. Further developments in the local-orbital density-functional-theory tight-binding method. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2001**, *64*, 195103.
(46) Lewis, J. P.; Jelínek, P.; Ortega, J.; Demkov, A. A.; Trabada, D. G.; Haycock, B.; Wang, H.; Adams, G.; Tomfohr, J. K.; Abad, E.; et al. Advances and applications in the FIREBALL ab initio tight binding molecular dynamics formalism. *Phys. Status Solidi B* **2011**, *248*, 1989−2007.
(47) Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A: At., Mol., Opt. Phys.* **1988**, *38*, 3098.
(48) Lee, C.; Yang, W.; Parr, R. G. Results obtained with the correlation energy density functionals. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785.
(49) Fuchs, M.; Scheffler, M. Ab initio pseudopotentials for electronic structure calculations of poly-atomic systems using density-functional theory. *Comput. Phys. Commun.* **1999**, *119*, 67.
(50) Basanta, M. A.; Dappe, Y. J.; Jelínek, P.; Ortega, J. Optimized atomic-like orbitals for first-principles tight-binding molecular dynamics. *Comput. Mater. Sci.* **2007**, *39*, 759.
(51) Řezáč, J.; Riley, K. E.; Hobza, P. S66: A well-balanced database of benchmark interaction energies relevant to biomolecular structures. *J. Chem. Theory Comput.* **2011**, *7*, 2427.
(52) Řezáč, J.; Hobza, P. Advanced corrections of hydrogen bonding and dispersion for semiempirical quantum mechanical methods. *J. Chem. Theory Comput.* **2012**, *8*, 141.
(53) Mendieta-Moreno, J. I.; Trabada, D. G.; Ortega, J. To be published.
(54) Jelínek, P.; Wang, H.; Lewis, J. P.; Sankey, O. F.; Ortega, J. Multicenter approach to the exchange-correlation interactions in ab initio tight-binding methods. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2005**, *71*, 235101.
(55) Marcos-Alcalde, I.; Setoain, J.; Mendieta-Moreno, J. I.; Mendieta, J.; Gómez-Puertas, P. MEPSA: minimum energy pathway analysis for energy landscapes. *Bioinformatics* **2015**, *31*, 3853−3855.
(56) Maurer, R. J.; Reuter, K. Assessing computationally efficient isomerization dynamics: ΔSCF density-functional theory study of azobenzene molecular switching. *J. Chem. Phys.* **2011**, *135*, 224303.
(57) Abad, E.; Lewis, J. P.; Zobač, V.; Hapala, P.; Jelínek, P.; Ortega, J. Calculation of non-adiabatic coupling vectors in a local-orbital basis set. *J. Chem. Phys.* **2013**, *138*, 154106.

DOI: 10.1021/acs.jpclett.6b02168
J. Phys. Chem. Lett. 2016, 7, 4391−4397

Supporting information for: Quantum Mechanics / Molecular Mechanics Free Energy Maps and Nonadiabatic Simulations for a Photochemical Reaction in DNA: Cyclobutane Thymine Dimer

Jesús I. Mendieta-Moreno,[†,‡] Daniel G. Trabada,[†] Jesús Mendieta,[†,‡,§] James P. Lewis,[¶] Paulino Gómez-Puertas,[‡] and José Ortega[*,†]

†*Departamento de Física Teórica de la Materia Condensada and Condensed Matter Physics Center (IFIMAC), Universidad Autónoma de Madrid, ES-28049 Madrid, Spain.*

‡*Molecular Modelling Group, Center of Molecular Biology Severo Ochoa (CSIC-UAM), ES-28049 Madrid, Spain.*

¶*Department of Physics, West Virginia University, Morgantown, West Virginia 26506-6315, USA.*

§*Departamento de Biotecnología, Universidad Francisco de Vitoria, ctra. Pozuelo-Majadahonda, km 1,800, 28223 Pozuelo de Alarcón (Madrid), Spain.*

E-mail: jose.ortega@uam.es

Thymine Dimer in the Gas Phase

Figure S1 shows the potential energy profiles for the S_0 and S_1 electronic states as a function of d (distance between the midpoints of the C5-C6 bonds) along a set of energy-minimized

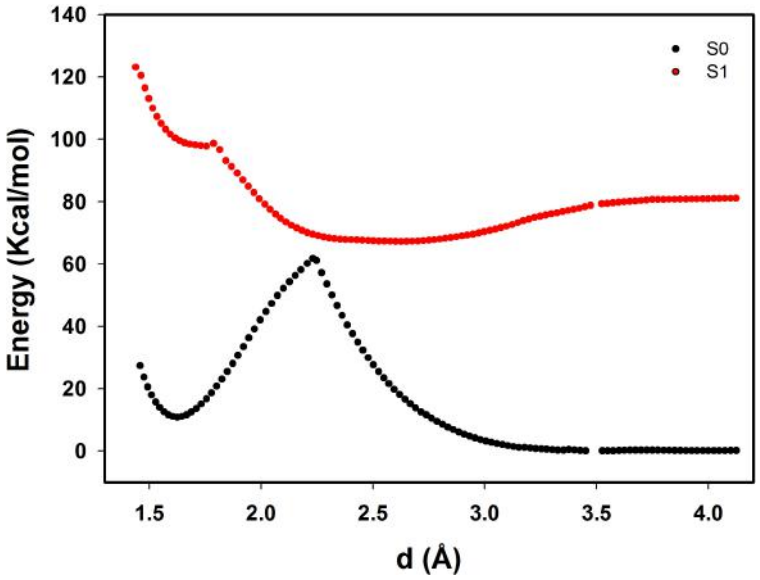


Figure S1: Potential energy profiles in the ground state (S_0) and first excited state (S_1) for the cycloaddition reaction of two stacked thymine bases leading to the formation of a cyclobutane thymine dimer, as a function of the distance d between the midpoints of the C5-C6 bonds.

configurations, calculated with the local-orbital DFT code Fireball with the same calculational parameters (BLYP functional, NAOs basis set, pseudopotentials, etc.) as used in the QM/MM simulations. The S_0 and S_1 energy curves have been calculated using optimized geometries corresponding to their own PES. The S_0 curve shows a large energy barrier, 62 kcal/mol, in agreement with previous results for the thymine dimer in the gas phase using quantum-chemistry methods, *e.g.* 63 kcal/mol,[S1] or DFT calculations, 80 - 88 kcal/mol.[S2,S3] On the other hand, the S_1 curve shows that in this excited electronic state there is no barrier between initial DNA conformations and the conical intersection configuration, also in agreement with previous results.[S1] In other words, in the case of the thymine dimer in the gas phase initial DNA conformations belong to the reactive conformations region. Previous calculations for the gas phase thymine dimer have yielded different geometries for the conical intersection, with values $d6 \sim 1.8$ - 2.3 Å and $d5 \sim 2.2$ - 2.5 Å;[S1–S6] in our calculations (Figure S1) we obtain $d6 = 2.1$ Å and $d5 = 2.3$ Å. Notice that this geometry is different from the one obtained in our QM/MM free energy calculations for the double helix B-DNA (Figure 2c), $d6 \sim 1.7$ Å and $d5 \sim 2.5$ Å, a result that reflects the important effect of the

double helix structure and biomolecular conditions in this photochemical reaction (see also Figure 5c).

Evolution of Frontier Orbitals

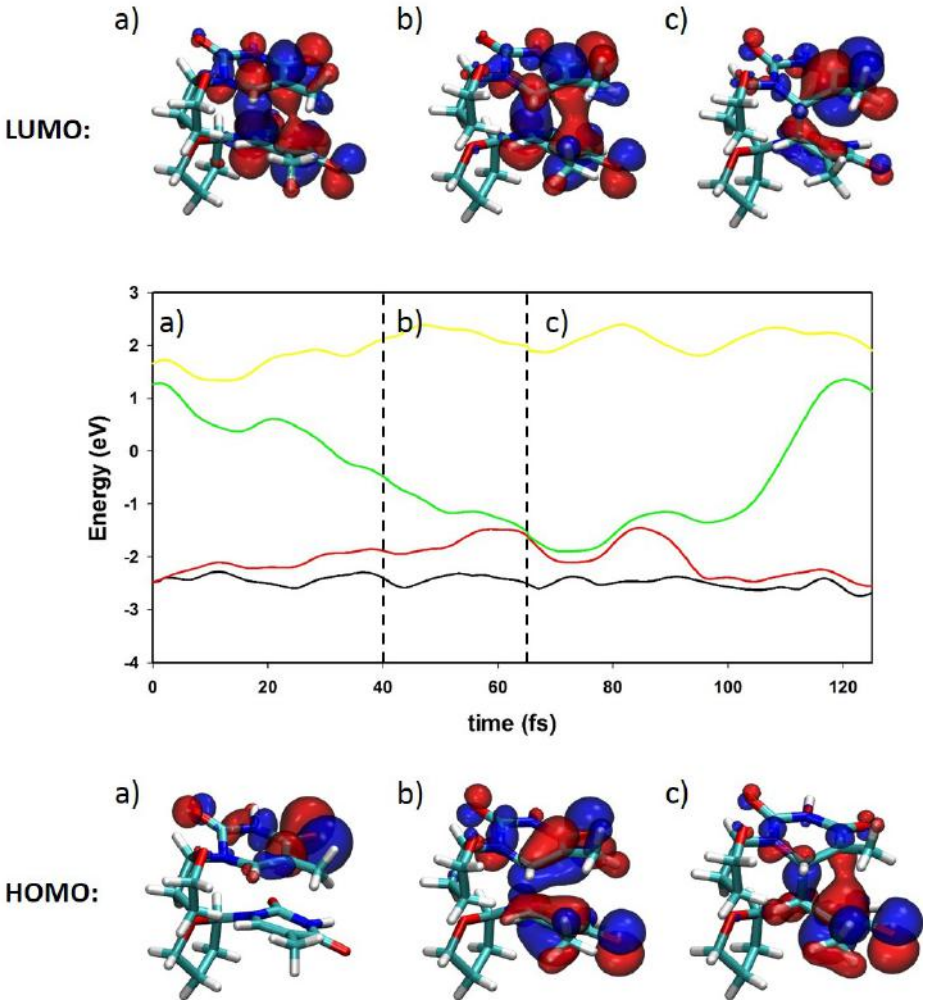


Figure S2: Example of the frontier orbitals evolution along a nonadiabatic trajectory. Middle panel: Kohn-Sham eigenvalues for HOMO-1, HOMO, LUMO and LUMO+1 along the trajectory. Up (Down): shape of the LUMO (HOMO) at selected instants (a-c) along the trajectory. The $S_1 \rightarrow S_0$ transition takes place at c).

Figure S2 shows the evolution of the frontier orbitals along a nonadiabatic trajectory. In general, these orbitals are delocalized between the two bases. The LUMO presents a bonding character between the two bases at a) and b). The occupation (after photo-excitation) of

this state induces the formation of C-C bonds between the two bases (mainly C6-C6'). At b) the HOMO presents a clear anti-bonding character. At c) (CIZ) the bonding orbital becomes the HOMO and the antibonding orbital becomes the LUMO. The $S_1 \rightarrow S_0$ transition takes place at c).

Reactive Conformations

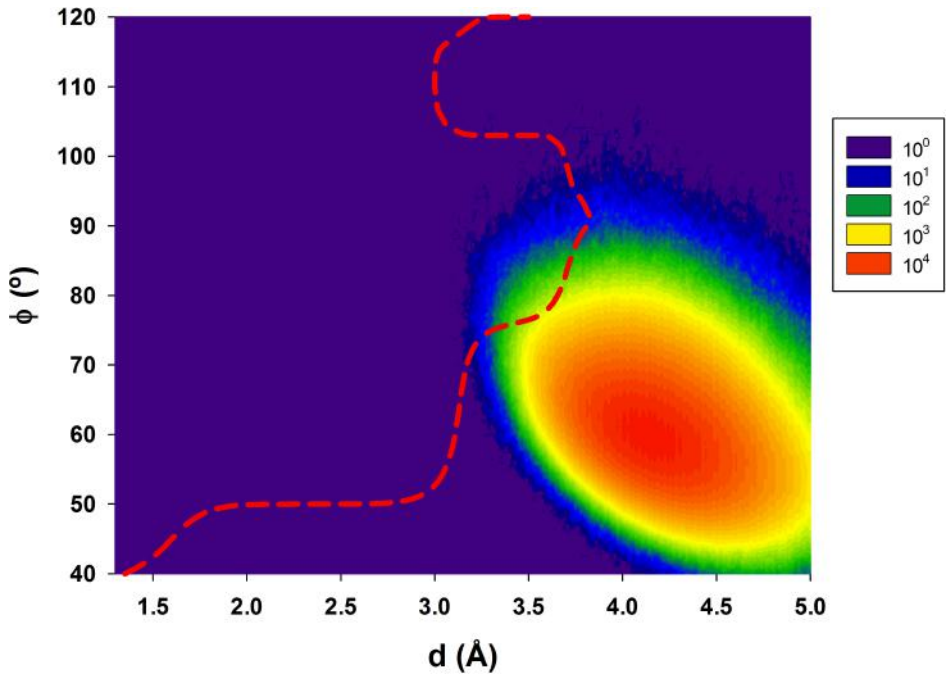


Figure S3: Distribution of DNA configurations as a function of the reaction coordinates from long (60 ns) MM MD in the ground state at RT. The different colors (notice the logarithmic scale) indicate the number of configurations per grid point; in this case we have used a finer grid of $\sim$ 48000 points with $\Delta d = 0.0125$ and $\Delta\phi = 0.5^o$. The red dashed line indicates the S_1 free energy barrier, see Figure 3b.

Figure S3 shows the distribution of DNA configurations as a function of the reaction coordinates d and ϕ found in a long (60 ns) MM MD simulation at RT in the electronic ground state S_0. In total, there are 3×10^7 different configurations ($\Delta t = 2$ fs). The reactive conformations are defined as those that are located on the left side of the red dashed line (S_1 free energy barrier, see Figure 3b). Although the MM force field is probably not very accurate for conformations close to this S_1 barrier, these simulations can be used to estimate the order

of magnitude of the fraction of time, f, that the system spends in reactive conformations. We find $\sim 1.1 \times 10^5$ reactive conformations; this yields a value of f of the order of 10^{-2} - 10^{-3}.

References

(S1) Boggio-Pasqua, M.; Groenhof, G.; Schäfer, L. V.; Grubmüller, H.; Robb, M. A. *J. Am. Chem. Soc.* **2007**, *129*, 10996–7.

(S2) Durbeej, B.; Eriksson, L. A. *J. Photochem. Photobiol. A* **2002**, *152*, 95–101.

(S3) Zhang, R. B.; Eriksson, L. A. *J. Phys. Chem. B* **2006**, *110*, 7556.

(S4) Blancafort, L.; Migani, A. *J. Am. Chem. Soc.* **2007**, *129*, 14540–14541.

(S5) Serrano-Pérez, J. J.; González-Ramírez, I.; Coto, P. B.; Merchán, M.; Serrano-Andrés, L. *J. Phys. Chem. B* **2008**, *112*, 14096–14098.

(S6) Barbatti, M. *ChemPhysChem* **2014**, *15*, 3342–3354.