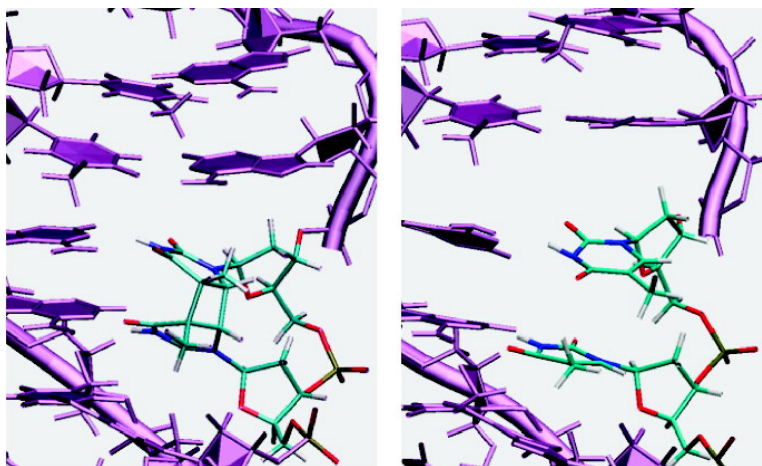Article

Computational Study of Thymine Dimer Radical Anion Splitting in the Self-Repair Process of Duplex DNA

Fanny Masson, Teodoro Laino, Ivano Tavernelli, Ursula Rothlisberger, and Jrg Hutter

More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

Computational Study of Thymine Dimer Radical Anion Splitting in the Self-Repair Process of Duplex DNA

Fanny Masson,*,[†] Teodoro Laino,[†] Ivano Tavernelli,[‡] Ursula Rothlisberger,[‡] and Jürg Hutter[†]

Physikalisch Chemisches Institut, Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland, and Laboratory of Computational Chemistry and Biochemistry, Federal Institute of Technology, EPFL, CH-1015 Lausanne, Switzerland

Received August 16, 2007; E-mail: masson@pci.uzh.ch

Abstract: Formation of the thymine dimer is one of the most important types of photochemical damage in DNA, responsible for several biological pathologies. Though specifically designed proteins (photolyases) can efficiently repair this type of damage in living cells, an autocatalytic activity of the DNA itself was recently discovered, allowing for a self-repair mechanism. In this paper, we provide the first molecular dynamics study of the splitting of thymine dimer radical anions, using a quantum mechanical/molecular mechanics (QM/MM) approach based on density functional theory (DFT) to describe the quantum region. A set of seven statistically representative molecular dynamics trajectories is analyzed. Our calculations predict an asynchronously concerted process in which C5−C5′ bond breaking is barrierless while C6−C6′ bond breaking is characterized by a small free energy barrier. An upper bound of 2.5 kcal/mol for this barrier is estimated. Moreover, the molecular dynamics study and the low free energy barrier involved in C6−C6′ bond breaking characterize the full process as being an ultrafast reaction.

1. Introduction

Cyclobutane pyrimidine dimer is the most abundant type of damage caused to DNA by ultraviolet (UV) light.[1] It is formed between two adjacent pyrimidine nucleobases, mainly thymine, via a [2 + 2] photocycloaddition. This lesion can lead to miscoding during DNA replication, which may cause mutations resulting in the development of skin cancer.[2] Recent studies directly connected the photochemical yield of thymine dimer (T⟨ ⟩T) formation to the deoxynucleotide (dN) sequence,[3] DNA conformation,[4,5] and binding of sequence-specific proteins to DNA.[6] The most ingenious strategy of a cell to repair this lesion is the light-driven photoreactivation catalyzed by DNA photolyase.[7−9] The commonly accepted model for this reaction proposes that blue or near-UV light energy (300−500 nm) is absorbed by an antenna pigment, which transfers the excitation energy to the reduced flavin coenzyme (FADH⁻) of the photolyase. The excited FADH⁻ then transfers an electron to the thymine dimer, leading to destabilization of the C5−C5′ and C6−C6′ bonds and thus boosting reversion to base monomers (see Figure 1).

However, this very elegant way of directly repairing damaged DNA, used by many organisms, does not explain how evolution may have been possible in a primordial RNA world, where the self-replicating nucleic acid population was highly vulnerable to mutations via UV light-mediated pyrimidine dimer formation. A possible and surprising solution was recently proposed along with the discovery of autocatalytic properties of nucleic acids in photoreactivation reactions: a deoxyribozyme containing a guanine quadruplex can efficiently repair the thymine dimer via a mechanism reminiscent of DNA photolyase.[10] It was observed that, under constant illumination, an equilibrium is established between the dimer and the normal base forms. The guanine quadruplex was implicated in serving as a light-harvesting antenna, with photoreactivation of the thymine dimer possibly proceeding via electron donation from an excited guanine base. Nucleic acids themselves could thus have played a role in preserving the integrity of such an RNA world, autocatalyzing the splitting of the pyrimidine lesion sites. More recently, an experimental study[11] shed light on the self-repair mechanism of T⟨ ⟩T, emphasizing the importance of the nucleobases adjacent to dipyrimidine sites in controlling the levels of T⟨ ⟩T. Lower levels of T⟨ ⟩T were found in sites surrounded by guanine versus adenine, consistent with the fact that guanine is the

[†] Universität Zürich.

[‡] Federal Institute of Technology, EPFL.

(1) Cadet, J.; Vigny, P. In *Bioorganic Photochemistry: Photochemistry and the Nucleic Acids*; Morrison, H., Ed.; John Wiley: New York, 1990; pp 1−272.

(2) Lima-Bessa, K. M.; Menck, C. F. M. *Curr. Biol.* **2005**, *15*, R58−R61.

(3) Kundu, L. M.; Linne, U.; Marahiel, M.; Carell, T. *Chemistry* **2004**, *10*, 5697−5705.

(4) Lyamichev, V. I.; Frank-Kamenetskii, M. D.; Soyfer, V. N. *Nature* **1990**, *344*, 568−570.

(5) Schreier, W. J.; Schrader, T. E.; Koller, F. O.; Gilch, P.; Crespo-Hernandez, C. E.; Swaminathan, V. N.; Carell, T.; Zinth, W.; Kohler, B. *Science* **2007**, *315*, 625−629.

(6) Schieferstein, U.; Thoma, F. *Biochemistry* **1996**, *35*, 7705−7714.

(7) Rupert, C. S.; Goodgal, S. H.; Herriott, R. M. *J. Gen. Physiol.* **1958**, *41*, 451−471.

(8) Sancar, A. *Chem. Rev.* **2003**, *103*, 2203−2237.

(9) Essen, L. O.; Klar, T. *Cell. Mol. Life Sci.* **2006**, *63*, 1266−1277.

(10) Chinnapen, D. J. F.; Sen, D. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 65−69.

(11) Holman, M. R.; Ito, T.; Rokita, S. E. *J. Am. Chem. Soc.* **2007**, *129*, 6−7.
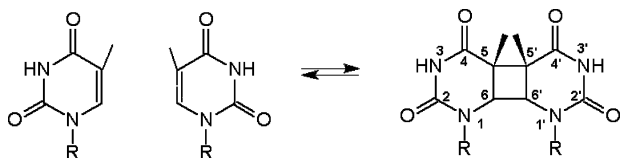
Figure 1. Formation and cycloreversion of the cis—syn thymine dimer.

nucleobase with the lowest oxidation potential, capable of acting as transient electron donor to promote repair of the T⟨ ⟩T lesion.

Many theoretical calculations regarding both the photochemical ring formation[12,13] and the thymine dimer repair mechanism[14-18] have been performed in order to elucidate the underlying molecular mechanism. Gas-phase quantum chemical studies predict a stepwise mechanism in which the initial ring-opening process occurs spontaneously upon electron uptake.[14,16] However, it was recently pointed out[18] that the cluster model underlying these gas-phase calculations is not sufficient to provide a realistic representation of the splitting process in condensed phase. Hydrogen bonding with surrounding water molecules or the complementary bases may stabilize the valence-bound state of the C4/C4′ carbonyl groups where the electron is mainly localized and make it energetically more favorable than the dipole-bound state in the gas phase.[19,20] Calculations including three water molecules showed that the hydrogen bond to the C4/C4′ carbonyl groups has a dramatic effect on the reaction mechanism.[17] On the basis of these considerations, Saettel and Wiest[17] suggested a quasi-concerted mechanism with an activation energy of 1.1 kcal/mol for the cleavage of the C5—C5′ bond and a subsequent barrierless breaking of the C6—C6′ bond. We note that the expression "quasi-concerted mechanism", presumably meaning an asynchronous concerted mechanism, was not defined by Saettel and Wiest.

The model of Saettel and Wiest[17] suffers from major drawbacks: the surrounding environment is made of only three water molecules, which are free to move and to sample a configurational space significantly larger than the one sampled by water molecules near the thymine dimer in duplex DNA. Though the model was considered as a great improvement with respect to the previous gas-phase theoretical studies, the lack of steric and electrostatic interactions due to the DNA backbone still make for a highly approximate representation of the real system. Until now, none of the published theoretical investigations could be considered conclusive in addressing the concerted or sequential[21] nature of the splitting mechanism of the thymine dimer radical anion.

Nonetheless, a very interesting aspect is the kinetics of the splitting process. A very recent experimental study[22] clarified several kinetic aspects of thymine dimer radical anion splitting

(12) Durbeej, B.; Eriksson, L. A. *Photochem. Photobiol.* **2003**, *78*, 159—167.
(13) Zhang, R. B.; Eriksson, L. A. *J. Phys. Chem. B* **2006**, *110*, 7556—7562.
(14) Durbeej, B.; Eriksson, L. A. *J. Am. Chem. Soc.* **2000**, *122*, 10126—10132.
(15) Voityuk, A. A.; Michel-Beyerle, M. E.; Rösch, N. *J. Am. Chem. Soc.* **1996**, *118*, 9750—9758.
(16) Voityuk, A. A.; Rösch, N. *J. Phys. Chem. A* **1997**, *101*, 8335—8338.
(17) Saettel, N. J.; Wiest, O. *J. Am. Chem. Soc.* **2001**, *123*, 2693—2694.
(18) Harrison, C. B.; O'Neil, L. L.; Wiest, O. *J. Phys. Chem. A* **2005**, *109*, 7001—7012.
(19) Richardson, N. A.; Wesolowski, S. S.; Schaefer, H. F. *J. Am. Chem. Soc.* **2002**, *124*, 10163—10170.
(20) Mantz, Y.; Gervasio, F. L.; Laino, T.; Parrinello, M. *Phys. Rev. Lett.* **2007**, *99*, 058104.
(21) Kao, Y.-T.; Saxena, C.; Wang, L.; Sancar, A.; Zhong, D. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 16128—16132.
(22) Chatgilialoglu, C.; Guerra, M.; Kaloudis, P.; Houée-Lévin, C.; Marignier, J.-L.; Swaminathan, V. N.; Carell, T. *Chem.—Eur. J.* **2007**, *13*, 8979—8984.

in water. Using pulse radiolysis and combining the experimental results with a computational model, Chatgilialoglu et al.[22] were able to show that the thymine dimer radical anion in water undergoes a fast splitting of the cyclobutane ring with a half-life of few picoseconds. In particular, they characterized the splitting process in water as a two successive carbon—carbon bond cleavages, estimating a half-life of 1—2 ps for the C5—C5′ bond breaking and 15—16 ps for the C6—C6′ bond cleavage. Thus, their conclusions contradict the results of the model of Saettel and Wiest,[17] confirming that the microsolvation study is an oversimplified model, unable to capture the essential mechanistic aspects of thymine dimer splitting both in water and in duplex DNA.

Interestingly, if further investigations on thymine dimer splitting in DNA confirm the estimates of the experimental study in water,[22] it will be proved that the efficiency of the ring cleavage is nearly 100%, since the charge recombination prior to thymine dimer splitting would require a much longer time (>560 ps in the photolyase active site[21]). This would imply that the reason for the low photochemical yield of the overall self-repair process in DNA (5%)[10] should be searched for in the ability of the DNA to transiently reduce the thymine dimer.

The description of the overall photoactivated self-repair process can be schematically divided into three main steps: (1) the ability of photoexcited DNA to transfer an electron to a thymine dimer, (2) the splitting dynamics of the thymine dimer radical anion, and (3) the fate of the unstable radical species created after the splitting of the thymine dimer. All of these steps are of relevant importance for a global understanding of the photoactivated self-repair process. The first point, the ability of DNA to transiently reduce a thymine dimer, has already been well-established experimentally.[11,48] The second and third points are technically more difficult to investigate experimentally due to the short lifetime of the radical species.

Until now, no theoretical or experimental studies have been performed to analyze the kinetics of bond breaking in the thymine dimer splitting process in DNA, and to the best of our knowledge, the present work represents the first attempt to clarify its mechanism.

We focus our analysis on the second step of the self-repair process, the dynamics of the transiently reduced thymine dimer, using a mixed quantum mechanical/molecular mechanics (QM/MM) method based on Born—Oppenheimer (BO) molecular dynamics, providing the possibility to compare the recently published experimental data[22] for splitting in water with our estimates for the same reaction in duplex DNA.

We perform the first finite temperature dynamics of the reaction as opposed to the static theoretical pictures published so far, which were based on the minimum energy path studies along chosen reaction coordinates. Our approach exploits a density functional description for the thymine dimer along with a classical mechanics treatment of the molecular frame and the solvent, explicitly taking into account the steric and electrostatic effects of the environment. A set of seven trajectories, initiated from an equilibrated run at constant pressure and temperature of the neutral T⟨ ⟩T structure, was used to sample the dynamics of the T⟨ ⟩T repair reaction. Statistical analysis of the performed trajectories shows that the mechanism behind T⟨ ⟩T repair can be described as an asynchronous concerted mechanism, in which the breaking of the C5—C5′ bond is spontaneous upon electron

uptake and is subsequently followed by C6−C6′ cleavage, with an upper bound of 2.5 kcal/mol for the activation energy of the latter process.

The C5−C5′ cleavage reaction in DNA is completed in a time range of 10−50 fs after electron uptake, instead of the 1−2 ps estimated experimentally for the same bond cleavage in water.[22] Regarding the C6−C6′ bond breaking reaction, we observe that four out of seven splitting processes are fully completed in less than 1 ps, thus providing an upper bound to the free energy barrier of the breaking process of 2.5 kcal/mol for the three trajectories locally trapped. This free energy barrier estimate is consistent with a characteristic reaction time of a few picoseconds. Our estimate of the free energy barrier for the C6−C6′ bond cleavage can be compared with the estimate of 3.2 kcal/mol from Chatgilialoglu et al.[22] Our kinetic results for dimer splitting in DNA are roughly 1 order of magnitude smaller than the experimental estimates for the same reaction in water, characterizing the thymine dimer radical anion splitting in DNA as an ultrafast reaction.

2. Methods

2.1. Structural Model. The initial configuration was taken from the X-ray structure of a DNA decamer containing a cis−syn thymine dimer, d[GCTTAATTCG]·d[CGAAT⟨⟩TAAGC] (PDB entry code 1N4E, 2.5 Å resolution).[23] The system was solvated with TIP3P water[24] in a rectangular box of $50 \times 51 \times 68$ Å^3. Eighteen potassium counterions were added to neutralize its charge. The total system contained ≈12 700 atoms (≈4000 water molecules). The AMBER-parm99 force field[25] was adopted for the oligonucleotide and the potassium counterions. For the thymine dimer (T15, T16) the atom types at the C5/C5′ and C6/C6′ positions (see Figure 1) were changed to "CT" atom type and at the H6/H6′ positions to "H1" atom type. The modified charges for the lesion sites were computed with the RESP module of AMBER8[26] for the isolated N-methyl derivatives of cis−syn T⟨ ⟩T following the same protocol as described in a previous paper.[27] A charge of 0.126 (the charge of the DNA backbone at the lesion sites) was imposed on the methyl groups and a RESP charge optimization was performed at the HF/6-31G* level of theory to yield a neutral model system. The RESP charges are provided as Supporting Information, together with the atom types and the additional force field parameters. Cross-links between T15/CT5−T16/CT5 and T15/CT6−T16/CT6 of the cis−syn dimer were also specified within the xleap module of AMBER8.

2.2. Classical MD Simulation. The system was treated within full periodic boundary conditions. Electrostatic interactions were computed with the smooth Particle-Mesh Ewald (SPME) algorithm,[28] with a cutoff of 10 Å for the real space part of the electrostatic interactions and the van der Waals term. A preliminary step involved geometry optimization of the full structure by use of a conjugate gradient algorithm. Energy minimization was carried out in two steps: first with harmonic restraints on the DNA decamer and then without any restraint.

Subsequently, MD simulations were conducted with restrained DNA at constant volume (NVT ensemble), increasing the temperature from 0 to 300 K in five steps, for a total of 50 ps. The last restrained
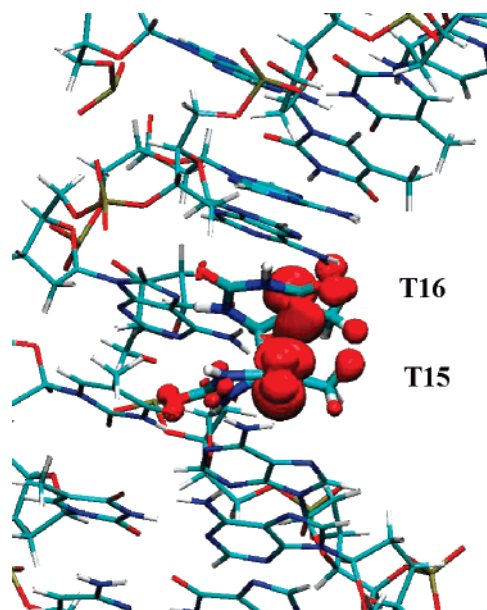

Figure 2. Spin density distribution of the optimized radical anion thymine dimer by use of a TZV2P basis set. The excess electron is fully localized on the QM part of the system.

configuration at 300 K was then used to start the unrestrained constant pressure−constant temperature (NPT ensemble) equilibration for 100 ps at $T = 300$ K. Finally, a MD production run of ≈5 ns at constant pressure (1 atm) and temperature (300 K) was performed to provide starting configurations for subsequent QM/MM calculations. All of the classical simulations used an integration time step of 1.5 fs. The room-temperature simulations were achieved by coupling the system to a Berendsen thermostat.[29] The rmsd data, reported in the Supporting Information, confirm the stability of the unrestrained DNA structure after full equilibration was achieved. All classical simulations have been performed with the AMBER suite of programs.[26]

2.3. QM/MM MD Simulation. The QM/MM driver is based on the quantum mechanical program QUICKSTEP[30,31] and the molecular mechanics driver FIST, which are both parts of the freely available CP2K package.[32] The general QM/MM scheme recently developed for large-scale molecular dynamics simulations[33,34] is based on a multigrid technique for computing the electrostatic potential due to the MM atoms. The description of the quantum region is treated at the density functional theory (DFT) level. The QM region is made up of two thymine bases, named T15 and T16 according to the sequence of the PDB structure 1N4E, resulting in a total number of 30 QM atoms. The thymines were cut at N1 and the valence of the terminal nitrogen atoms was saturated by the addition of capping hydrogen atoms. The remaining part of the system, including water molecules and counterions, was modeled at the classical mechanics level with the AMBER-parm99 force field. A triple-ζ valence basis set with two sets of polarization functions, TZV2P,[35] and an auxiliary plane-wave basis set with a density cutoff of 300 Ry were used to describe the wavefunction and the electronic density. It has previously been shown that this kind of high-level basis set is necessary for accurate geometries as well as energetics.[14] Dual-

(23) Park, H.; Zhang, K.; Ren, Y.; Nadji, S.; Sinha, N.; Taylor, J.-S.; Kang, C. *Proc. Natl Acad. Sci. U.S.A.* **2002**, *99*, 15965−15970.
(24) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(25) Wang, J. M.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049−1074.
(26) Case, D. A.; et al. *AMBER8*; University of California, San Francisco, CA, 2004.
(27) Spector, T. I.; Cheatham, T. E., III; Kollman, P. A. *J. Am. Chem. Soc.* **1997**, *119*, 7095−7104.
(28) Cheatham, T. E., III; Miller, J. L.; Fox, T.; Darden, T. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 4193−4194.
(29) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(30) Lippert, G.; Hutter, J.; Parrinello, M. *Theor. Chem. Acc.* **1999**, *103*, 124−140.
(31) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. *Comput. Phys. Commun.* **2005**, *167*, 103−128.
(32) CP2K developers group; freely available at the URL http://cp2k.berlios.de, released under GPL license 2007.
(33) Laino, T.; Mohamed, F.; Laio, A.; Parrinello, M. *J. Chem. Theory Comput.* **2005**, *1*, 1176−1184.
(34) Laino, T.; Mohamed, F.; Laio, A.; Parrinello, M. *J. Chem. Theory Comput.* **2006**, *2*, 1370−1378.
(35) Schaefer, A.; Huber, C.; Ahlrichs, R. *J. Chem. Phys.* **1994**, *100*, 5829−5835.

**Table 1.** RMSD Values between Selected Snapshots along the Classical Trajectory and the X-ray Structure[a]

	CPD1	CPD2	CPD3	CPD4	CPD5	CPD6	CPD7
X-ray	0.69, 2.62	0.57, 2.84	0.62, 2.65	0.78, 2.81	0.76, 2.94	0.73, 2.84	0.74, 2.90

[a] The first value refers to the cis−syn thymine dimer as shown in Figure 1, and the second value refers to the entire double helix. Values are given in angstroms.

Table 2. QM/MM Averaged Lengths of Hydrogen Bonds for the A5−T16 and A6−T15 Base Pairs of the Neutral cis−syn Thymine Dimer[a]

		CPD1	CPD2	CPD3	CPD4	CPD5	CPD6	CPD7	X-ray
A5 H61	T16 O4	2.26	2.27	2.37	2.57	2.27	2.29	2.22	1.89
		(0.15)	(0.15)	(0.17)	(0.36)	(0.20)	(0.19)	(0.17)	(0.03)
A5 N1	T16 H3	1.90	1.96	1.95	2.04	1.95	1.96	1.98	1.85
		(0.12)	(0.13)	(0.16)	(0.26)	(0.16)	(0.12)	(0.13)	(0.03)
A6 H61	T15 O4	2.50	2.29	2.45	2.84	2.68	2.57	2.57	2.49
		(0.26)	(0.15)	(0.29)	(0.29)	(0.45)	(0.26)	(0.23)	(0.03)
A6 N1	T15 H3	1.98	2.12	2.02	2.00	1.99	2.06	2.05	2.20
		(0.12)	(0.31)	(0.14)	(0.17)	(0.12)	(0.21)	(0.18)	(0.03)

[a] Values are given in angstroms; standard deviations are shown in parentheses.
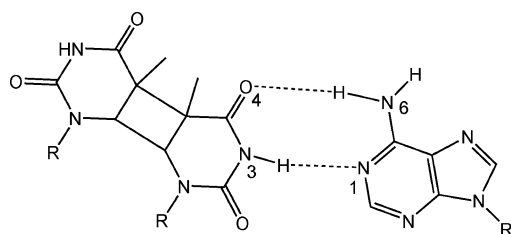
Figure 3. Hydrogen bonding between the thymine dimer and a complementary adenine in DNA.

space pseudopotentials[36,37] were used for describing core electrons and nuclei. We used the gradient corrected Becke exchange[38] and the Lee−Yang−Parr correlation functionals (BLYP),[39] and when an electron was added to the thymine dimer, the DFT calculations were performed within the local spin-density approximation (LSD). As DFT encounters difficulties in describing van der Waals interactions, dispersion-corrected atom-centered potentials (DCACPs)[40,41] have been used as a correction to the BLYP functional for all atoms, namely, C, O, N, and H. This correction allows one to account for the van der Waals interactions between the two stacked thymines after splitting of the ring.[42] Energies were tested for convergence with respect to the wavefunction gradient $(5 \times 10^{-7}$ hartree) and cell size, which was required to be no smaller than 14.8 Å to achieve a correct decoupling between the periodic images.[43]

A set of seven snapshots was randomly chosen from the 5 ns of the classical NPT run, and the geometries were used as starting points for QM/MM simulations. First, the QM region of each of the QM/MM structures was relaxed by performing a geometry optimization, while the MM part was kept frozen. Then a molecular dynamics run of the whole system was performed at 300 K with an NVT ensemble and collection of ≈2 ps of unconstrained QM/MM dynamics. Ab initio molecular dynamics (MD) simulations within the Born−Oppenheimer approximation were performed in the canonical (NVT) ensemble with an integration time step of 0.5 fs. The temperature was kept constant by use of a Nosé-Hoover thermostat[44,45] with a time constant of 1 ps.

(36) Goedecker, S.; Teter, M.; Hutter, J. *Phys. Rev. B* **1996**, *54*, 1703−1710.
(37) Hartwigsen, C.; Goedecker, S.; Hutter, J. *Phys. Rev. B* **1998**, *58*, 3641−3662.
(38) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098−3100.
(39) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789.
(40) von Lilienfeld, O. A.; Tavernelli, I.; Rothlisberger, U.; Sebastiani, D. *Phys. Rev. Lett.* **2004**, *93*, 153004.
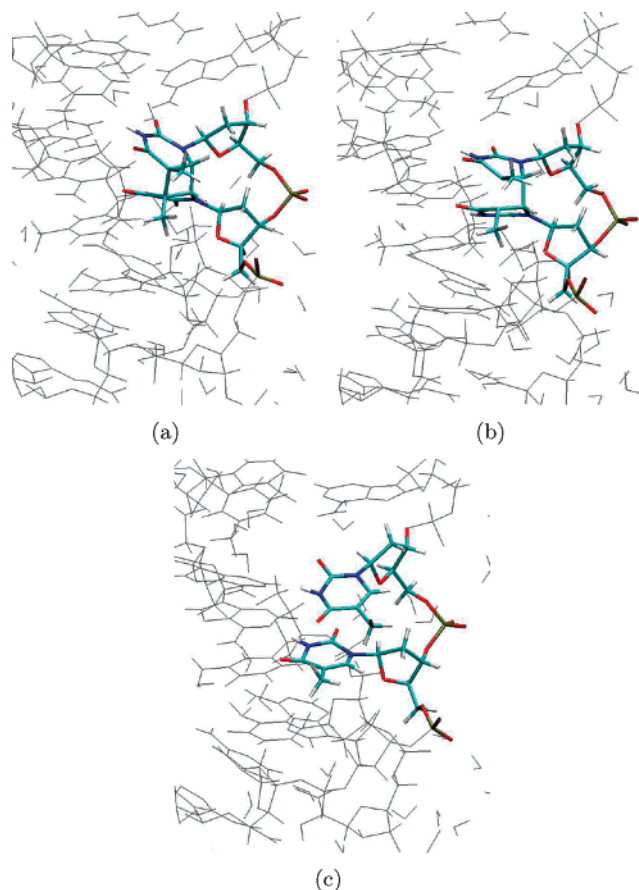(41) von Lilienfeld, O. A.; Tavernelli, I.; Rothlisberger, U.; Sebastiani, D. *J. Chem. Phys.* **2005**, *122*, 14113.
(42) Lin, I.-C.; von Lilienfeld, O. A.; Coutinho-Neto, M. C.; Tavernelli, I.; Rothlisberger, U. Manuscript in preparation.
(43) Bloechl, P. E. *J. Chem. Phys.* **1995**, *103*, 7422−7428.
(44) Nose, S. J. *J. Chem. Phys.* **1984**, *81*, 511−519.
(45) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695−1697.

Figure 4. Representation of the thymine dimer splitting reaction during the simulation. The starting configuration is the thymine dimer (a). Upon electron uptake, the C5−C5′ bond breaks spontaneously (b), followed by C6−C6′ bond cleavage (c) according to an asynchronous concerted mechanism. The structure subsequently relaxes to the canonical B-DNA conformation.

After completion of the QM/MM equilibration run, an electron was added to each system and the simulation was continued for ≈2 ps, unless the breaking of both bonds was observed before.

QM/MM simulations are particularly sensitive to unphysical delocalization of the electronic density from the QM to the MM region, the so-called spill-out effect. An appropriately modified short-range Coulomb potential for the interaction between the QM and the MM parts was used to ensure that no spill-out occurs.[46] The spin density distribution of the radical anion thymine dimer indicates that the excess
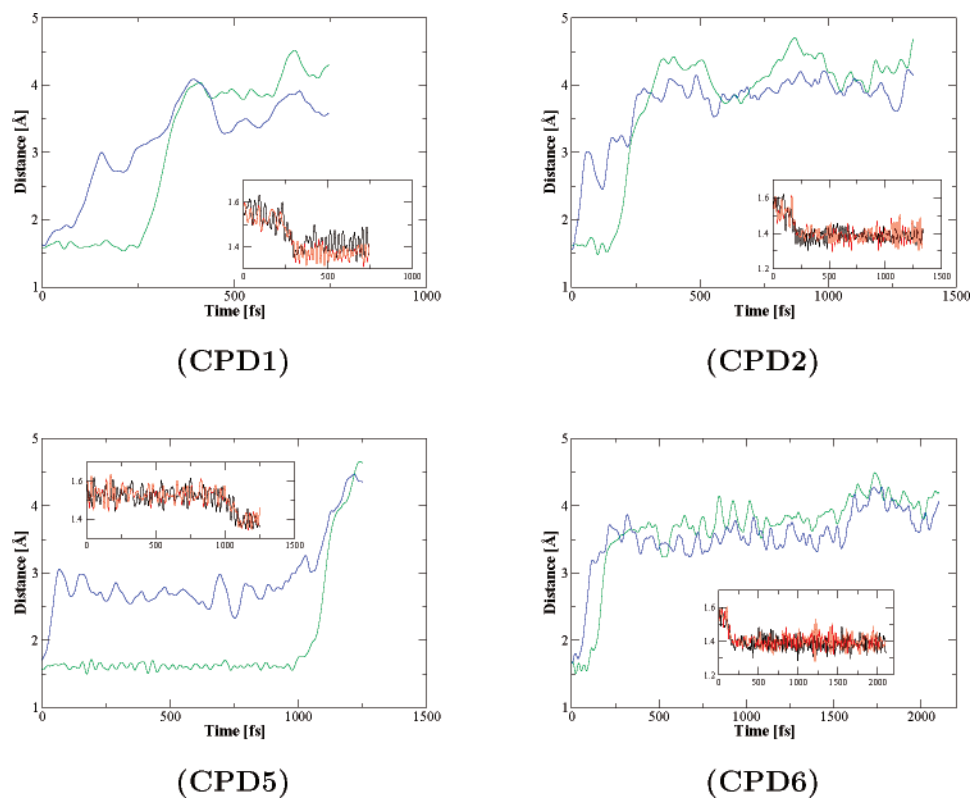
(CPD1)

(CPD2)

(CPD5)

(CPD6)

Figure 5. Time evolution of the C5−C5′ (blue) and C6−C6′ (green) bond lengths during the T⟨ ⟩T lesion reversion process along the four reactive trajectories. (Inset) C5−C6 (black) and C5′−C6′ (red) bonds become double bonds when the C6−C6′ bond is breaking.

electron is fully localized on the QM part while being delocalized over both thymines (see Figure 2).

Furthermore, since we are treating an odd-electron system, we assessed the influence of the self-interaction error (SIE) on the electronic structure by performing reference calculations with methods that are not suffering from SIE. The spin density was computed at the restricted open-shell Hartree−Fock (ROHF) and unrestricted MP2 (UMP2) levels. The spin density was also fully distributed on both thymines at these different levels of theory (data not shown) in line with our DFT results, indicating that the SIE might be less important. Finally, the basis set superposition error is assumed to be negligible due to the extended basis set used.

The choice of including only the two thymines in the QM region relies on a recent study showing that proton transfer from a complementary adenine is an endothermic process and therefore unlikely in DNA.[47] Both complementary bases and water molecules around the two thymines have interactions with the T⟨ ⟩T site only through hydrogen bonds, which renders their description by a classical force field sufficient. A possible quantum description of the complementary bases or of the water molecules surrounding the two thymines would result only in a small numerical discrepancy on the energy barriers but it would not change the global picture of this work.

Nevertheless, the proton transfer is expected to take place in the active site of DNA photolyase[9] and may increase the repair quantum efficiency by stabilizing the radical anion so that nonproductive back-electron transfer is avoided and bond cleavage is favored. In the DNA environment, the stabilization is not so crucial due to the small probability that the electron transfers back to an adjacent base. In fact, it was shown that electrons hop through DNA using pyrimidine bases as stepping stones.[48]

3. Results and Discussion

Seven different frames of the thymine dimer, namely, CPD1−CPD7, were selected from the classical MD trajectory as starting points for seven independent QM/MM MD simulations. The quality of the reference trajectory is assessed in Table 1, where we report the rmsd values of the sampled structures compared to the X-ray structure. The main geometrical differences between the lesion site of the different conformations and the X-ray structure are the C5−C6 and C5′−C6′ bonds. In the crystallographic structure, they remain close to typical values for double bonds ($\approx$1.32 Å) rather than to values for single bonds ($\approx$1.53 Å), as one may expect upon cycloaddition. The low resolution (2.5 Å) of the X-ray data may be an explanation for this discrepancy. In Table 2 we show the hydrogen-bond distances between the lesion site and the complementary bases (see Figure 3), where comparison with respect to the X-ray data shows a lengthening of hydrogen bonds in our simulations.

Consistent with the crystallographic structure that finds the thymine T15 hydrogen bonds broken, the thymine T15 hydrogen bonds are weaker than those of thymine T16. Inspection of the trajectories of the neutral thymine dimer shows that the average bridging C5−C5′ bond is longer than the C6−C6′ bond (1.65 $\pm$ 0.06 versus 1.61 $\pm$ 0.05 Å, respectively), which is attributed to the repulsion between the two methyl groups. The cyclobutane puckering angle −C5−C6−C6′−C5′ fluctuates around 14° $\pm$ 4°.

Electron Uptake. Upon electron uptake by the thymine dimer, the C5−C5′ bond cleavage occurs spontaneously (see
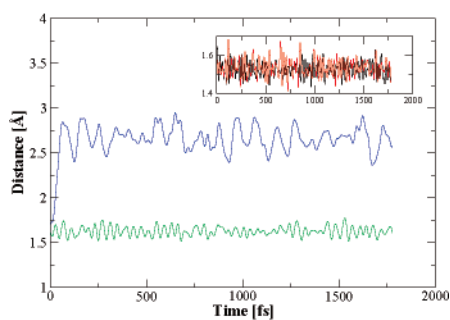
(46) Laio, A.; VandeVondele, J.; Rothlisberger, U. *J. Phys. Chem. B* **2002**, *106*, 7300−7307.
(47) Rak, J.; Voityuk, A. A.; Michel-Beyerle, M. E.; Rösch, N. *J. Phys. Chem. A* **1999**, *103*, 3569−3574.
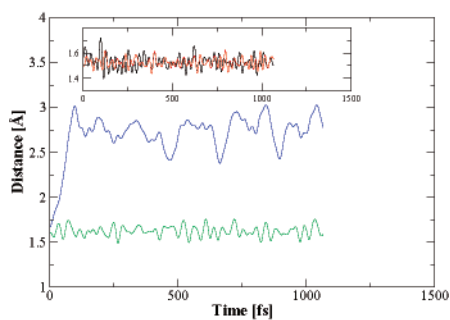(48) Behrens, C.; Cichon, M. K.; Grolle, F.; Hennecke, U.; Carell, T. In *Long-Range Charge Transfer in DNA I*; Schuster, G. B., Ed.; Springer: Berlin, 2004; pp 187−204.
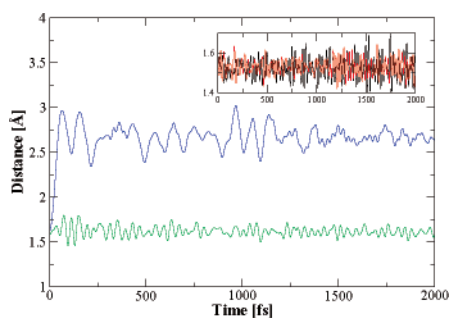
(CPD3)

(CPD4)

(CPD7)

Figure 6. Time evolution of the C5−C5′ (blue) and C6−C6′ (green) bond lengths along the three nonreactive trajectories. (Inset) C5−C6 (black) and C5′−C6′ (red) bonds.

Figures 5 and 6). This initial ring-opening process can be rationalized as a delocalization of the singly occupied C4-(C4′)−O π^* orbital (SOMO) into the C5−C5′ σ^* orbital.[49] The extra strain imposed on the puckered cyclobutane ring by the helical structure of DNA may enhance the rate of the C5−C5′ cleavage by making the interaction of the SOMO with the cyclobutyl ring orbital larger.[50] Averaging the results of the seven QM/MM simulations, we find that the C5−C5′ distance fluctuates around 2.70 ± 0.14 Å and shows large fluctuations from 2.3 to 3.0 Å. The C5−C6 and C5′−C6′ bonds still retain σ-character as they display an average value of 1.54 ± 0.06 Å. The puckering angle rotates toward more positive values, with fluctuations up to 55° and an average of 35° ± 4°.

Spontaneous C6−C6′ bond cleavage was observed in four out of seven simulations. This bond cleaves within 500 fs after the C5−C5′ bond breaking in all of the trajectories but one,

where the C6−C6′ bond breaks 1000 fs afterward (see Figure 5). The C6−C6′ distance increases strongly, fluctuating around 4.18 ± 0.47 Å, while the C5−C5′ distance rises at the same time and fluctuates around 3.88 ± 0.30 Å. These distances have to be compared with the values obtained by a 3 ns classical simulation on an undamaged decamer with the same base sequence, where C5−C5′ and C6−C6′ distances display an average value of 4.1 ± 0.3 and 4.5 ± 0.3, respectively. This indicates that the time scale of our QM/MM simulation (2 ps) does not allow us to recover a fully undamaged DNA conformation.

Spin Density Distribution along the Reaction. The splitting of the cyclobutane ring can be monitored by spin density variations on relevant atoms as reported in Figure 7. Upon electron uptake, the electron is delocalized over both fragments. During the breaking process, a slight decrease of the spin densities on C6 is observed (up to −0.06 e⁻). In the same region, the spin densities are close to their maximum value on O4 and at their maximum value on C5 (0.57 e⁻). However, on the other thymine, the spin densities are close to zero on O4′ and drop significantly on C5′. Thus, the electron is mostly localized on one thymine (T15) when the cleavage occurs with a maximum on C5 and a decrease of the spin densities on C6. This favorable interaction between C5 and C6 allows for the formation of a double bond between them. After the C6−C6′ bond dissociation, the spin densities on O4 and C5 decrease significantly and the electron is mostly localized on C6 and C6′.

Hydrogen Bonding. Gas-phase quantum chemical studies at the MP2[16] and B3LYP[14] levels of theory predicted stepwise splitting in which the C5−C5′ bond breaks spontaneously upon electron uptake. Durbeej and Eriksson[14] have calculated an activation energy of 2.3 kcal/mol for the C6−C6′ bond cleavage, which is the rate-determining step of the splitting process. It has been pointed out that such gas-phase calculations do not give an adequate representation of the reaction[17] since the valence-bound state of pyrimidine, which hosts the extra electron, can be stabilized by hydrogen bonding with water[51] or between base pairs.[52] Saettel and Wiest[17] suggested a quasi-concerted mechanism after including hydrogen-bonding solvent molecules in their model system. Through the introduction of three water molecules and single-point B3LYP/6-311++G** calculations, the activation energy of the C5−C5′ bond breaking is found to be 1.1 kcal/mol and the breaking of the C6−C6′ bond occurs without a barrier. These considerations lead us to investigate the hydrogen bonds of each species with the complementary adenines (A) and water, focusing our attention on the A H61−T O4 H-bonds, which stabilize the developing negative charge on the C4 carbonyl group. In particular, in Figures 8 and 9, we show, for both reactive and nonreactive trajectories, the hydrogen bonds involving one of the two thymines. In the case of CPD3, the A5 H61−T16 O4 H-bond is lost after ≈1850 fs and O4 forms a hydrogen bond with an approaching water molecule, while the A5 N1−T16 H3 H-bond is still maintained. The A5 H61−T16 O4 H-bond is also lost in the case of CPD6, which is reflected by the constant increase of the dihedral −A5 N1−A5 N6−T16 O4−T16 C4 (from −30° to 28°) during the simulation, while the A5 N1−T16 H3 H-bond

(49) Hartman, R. F.; Van Camp, J. R.; Rose, S. D. *J. Org. Chem.* **1987**, *52*, 2684−2689.
(50) Heelis, P. F. *J. Mol. Model.* **1995**, *1*, 18−21.
(51) Dolgounitcheva, O.; Zakrzewski, V. G.; Ortiz, J. V. *J. Phys. Chem. A* **1999**, *103*, 7912−7917.
(52) Al-Jihad, I.; Smets, J.; Adamowicz, L. *J. Phys. Chem. A* **2000**, *104*, 2994−2998.
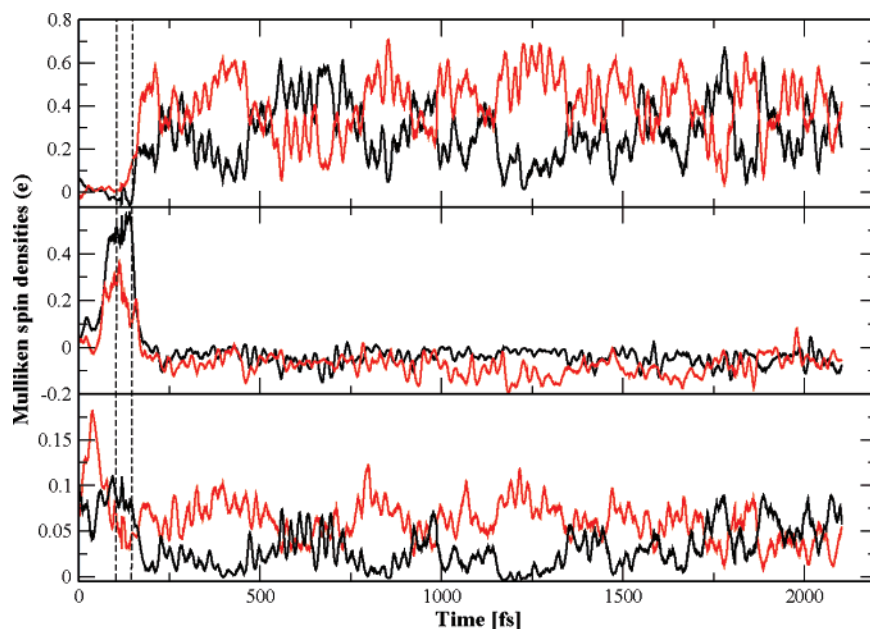
Figure 7. Time evolution of the Mulliken spin densities (e⁻) of selected atoms shown only for CPD6 species. The breaking process of the C6−C6′ bond occurs between the two dashed lines. Atoms of T15 are shown in black and atoms from T16 are shown in red. Upper panel: C6 (black) and C6′ (red) spin densities. Middle panel: C5 (black) and C5′ (red) spin densities. Lower panel: O4 (black) and O4′ (red) spin densities. The three panels are not on the same scale; the total lengths of the y-axis are 0.9, 0.8, and 0.2, respectively.
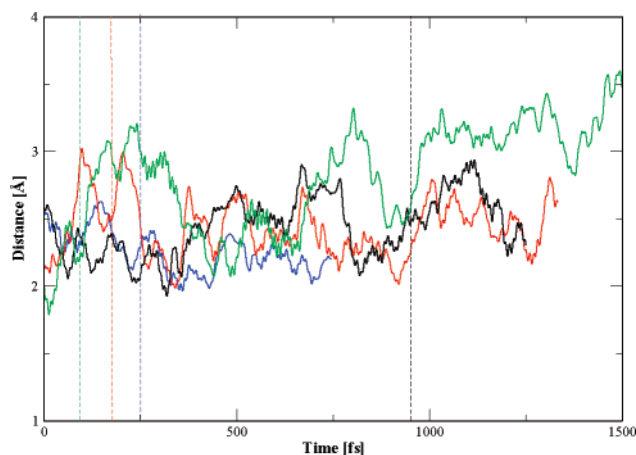

Figure 8. Variation with time of the A5 H61−T16 O4 hydrogen-bond length for the species CPD1 (blue), CPD2 (red), CPD5 (black), and CPD6 (green), where the breaking of the C6−C6′ bond is observed. The dashed lines indicate when the breaking of the C6−C6′ bond occurs.

Figure 9. Variation with time of the A5 H61−T16 O4 hydrogen-bond length for the species CPD3 (blue), CPD4 (red), and CPD7 (black), where the breaking of the C6−C6′ bond cannot be observed.

is preserved as well. It is evident that the hydrogen bond alone does not represent a driving variable for the reaction. On average, the hydrogen bond involving the reactive species (see Figure 8) is slightly shorter than the one from nonreactive species (see Figure 9). This contradicts the results of a previous theoretical study[17] where the water molecules were considered of fundamental importance for the dimer splitting mechanism. This inconsistency can be explained by analyzing in detail the model used by Saettel and Wiest. They studied the thymine dimer in vacuo (i.e., without any steric or electrostatic contributions due to the DNA backbone) with the presence of only three water molecules surrounding the thymine dimer. The water molecules were free to move, in contrast to what really happens in duplex DNA, where base stacking dramatically reduces the available range of configurations for solvating water molecules. Therefore the computational findings of Saettel and Wiest are valid only for their system, that is, a microsolvation of a free
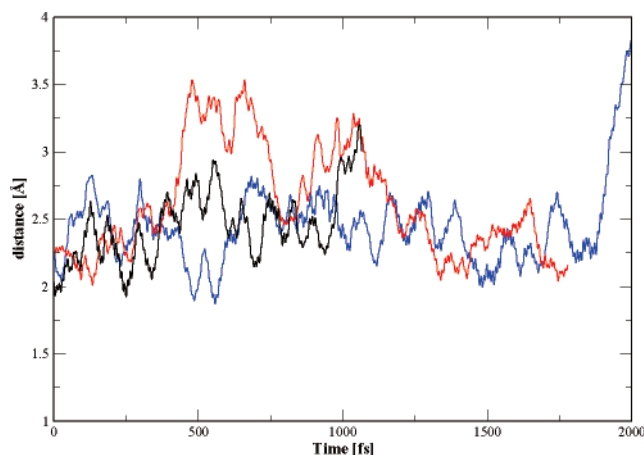
thymine dimer. On the other side, our model takes into account both electrostatic and steric effects from the full DNA backbone, water molecules, and ions. In the real environment, the importance of hydrogen bonding with water molecules is reduced since the global electrostatic and steric effects of the surrounding base pairs as well as the solvating water molecules drive the splitting reaction.

O4 was also found to interact with a water molecule in the case of CPD1, CPD2, and CPD3, but no significant water interactions with the thymine dimer could be detected for the other species.

4. Bridging the Time Scale: Free Energy Simulations

The limited time scale accessible to our QM/MM MD simulations is in part responsible for the missing observation of the lesion repair in a few trajectories. In other words, despite the minimal statistics, we have shown the evidence that a barrier is present in the self-repair mechanism and the lack of observing
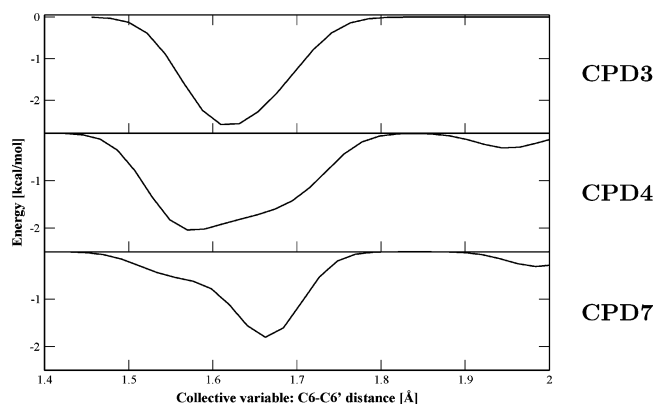
Figure 10. Free energy profiles of nonreactive trajectories (CPD3, CPD4, and CPD7) with respect to the bond length C6−C6′.

a spontaneous repair reaction in three out of seven simulations is due to the problem of having trajectories trapped in local free energy minima.

To provide an estimate of the free energy barrier, we have performed metadynamics runs.[53,54] It is worth noting that the choice of the collective variables (CV) in metadynamics is crucial for its successful application. In this case a natural choice was the C6−C6′ bond length. Other collective variables would be necessary to fully characterize the real transition state in the free energy landscape. Since the aim of the present analysis is simply to provide an upper bound to the barrier, we stress that the choice of the reaction coordinate is not only the most obvious but also the most reasonable in describing the repair process. The metadynamics runs were performed by use of Gaussian-shaped potential hills with a height of 0.3 kcal/mol and a width of 0.04 Å. The hills were spawned at intervals of 20 fs of QM/MM MD.

The free energy profiles of the three nonreactive trajectories are plotted as a function of the CV in Figure 10, and we can provide a value of 2.5 kcal/mol as a reasonable upper bound to the free energy barrier of the ring-splitting process. Splitting of the thymine dimer was also observed at 100 K for one reactive structure by performing a QM/MM simulation, confirming that the estimate of 2.5 kcal/mol is definitely an upper limit to the real free energy barrier.

5. Conclusion

We have studied the thymine dimer radical anion splitting of the photoactivated self-repair process in DNA using a QM/MM scheme based on DFT for treating the quantum region. This work represents the first dynamical study of the repair process of the thymine dimer lesion in DNA and in general one of the few QM/MM studies that, instead of relying on one

simple MD trajectory, performs several different statistically sampled runs.

Our computations, which take explicit account of the hydrogen-bonding network through the QM/MM interaction term, reveal an asynchronously concerted mechanism, providing an upper bound to the free energy barrier of 2.5 kcal/mol. The C5−C5′ bond of the thymine dimer breaks spontaneously upon electron uptake within 10−100 fs, and subsequent C6−C6′ cleavage occurs with a low barrier easily overcome with the thermal energy at room temperature, within a simulation time on the order of picoseconds. The exact determination of the energy barrier is beyond the goal of the present investigation and can be the subject for future studies, using novel methodologies.[55,56]

The results characterize the thymine dimer splitting process in DNA as an asynchronous concerted reaction, in line with the recent experimental findings[22] for thymine dimer splitting in water. The theoretical estimates of the C5−C5′ and C6−C6′ bond breaking rates in DNA are 1 order of magnitude smaller than the experimental rates[22] describing the same processes in water. Therefore, our study identifies thymine dimer splitting in DNA as an ultrafast reaction. The ultrafast nature of the splitting process implies that, in the overall self-repair mechanism, the back electron transfer is not a competing process, since, experimentally, it is observed to occur on a time scale 2 orders of magnitude larger than the splitting process.[21] Thus, the efficiency of the ring cleavage can be considered nearly 100%.

Special relevance should be given to the study of the unstable radical anion species formed after the thymine dimer radical anion splitting. This will be the subject of future studies.

The present work, combined with the available experimental data,[10,11] which show the ability of DNA to transiently reduce the thymine dimer, explains the observed photochemical stability of native DNA. It confirms the picture where, in a primordial world, life may have developed due to the intrinsic great resistance of the DNA/RNA molecules against UV light.

Supporting Information Available: Detailed analysis of the seven MD trajectories, complete ref 26, modified classical force field for the thymine dimer, atomic coordinates for the three representative structures of Figure 4, and full rmsd data for the equilibration processes. This material is available free of charge via the Internet at http://pubs.acs.org.

JA076081H

(53) Laio, A.; Parrinello, M. *Proc. Natl Acad. Sci. U.S.A.* **2002**, *99*, 12562−12566.
(54) Micheletti, C.; Laio, A.; Parrinello, M. *Phys. Rev. Lett.* **2004**, *92*, 170601.
(55) Maragliano, L.; Fischer, A.; Vanden-Eijnden, E.; Ciccotti, G. *J. Chem. Phys.* **2006**, *125*, 24106.
(56) Branduardi, D.; Gervasio, F. L.; Parrinello, M. *J. Chem. Phys.* **2007**, *126*, 054103.