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

Insights into Intrastrand Cross-Link Lesions of DNA from QM/MM Molecular Dynamics Simulations

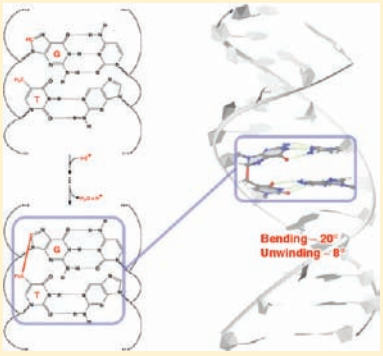
Julian Garrec,[†] Chandan Patel,[‡] Ursula Rothlisberger,[†] and Elise Dumont*[,‡]

[†]Laboratory of Computational Chemistry and Biochemistry, Institute of Chemical Sciences and Engineering, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland

[‡]Université de Lyon, Institut de Chimie de Lyon, CNRS, Ecole normale supérieure de Lyon, 46 allée d'Italie, 69364 Lyon Cedex 07, France

Supporting Information

ABSTRACT: DNA damages induced by oxidative intrastrand cross-links have been the subject of intense research during the past decade. Yet, the currently available experimental protocols used to isolate such lesions only allow to get structural information about linked dinucleotides. The detailed structure of the damaged DNA macromolecule has remained elusive. In this study we generated in silico the most frequent oxidative intrastrand cross-link adduct, G[8,5-Me]T, embedded in a solvated DNA dodecamer by means of quantum mechanics/molecular mechanics (QM/MM) Car−Parrinello simulations. The free energy of activation required to bring the reactant close together and to form the C−C covalent-bond is estimated to be ∼10 kcal/mol. We observe that the G[8,5-Me]T tandem lesion is accommodated with almost no perturbation of the Watson−Crick hydrogen-bond network and induces bend and unwinding angles of ∼20° and 8°, respectively. This rather small structural distortion of the DNA macromolecule compared to other well characterized intrastrand cross-links, such as cyclobutane pyrimidines dimers or cisplatin-DNA complex adduct, is a probable rationale for the known lack of efficient repair of oxidative damages.



1. INTRODUCTION

DNA intrastrand cross-links (ICL), in which two neighboring nucleobases are covalently tethered together, represent an important class of chemically induced DNA damage.[1] The best characterized lesions of these type are photoinduced,[2] such as cyclobutane pyrimidines dimers[3] and spore photoproduct.[4,5] They differ from oxidative ICLs,[6,7] whose ground-state formation is initiated by reactive oxygen species (ROS), generated during normal aerobic metabolism or through exogenous pathways.[8] Even a single radical (most frequently a hydroxyl HO$^{\bullet}$) can abstract a hydrogen atom from a nucleobase. Pyrimidine (Py) bases are more prone to this type of lesion, and among the several carbon or nitrogen atoms that can be attacked, thymine-methyl C−H bonds are the weakest linkage.[9] The resulting radical moiety (T−CH$_2^{\bullet}$) is prone to form covalent C−C or C−N linkages with adjacent nucleobases, preferentially purines (Pu), hence forming an intrastrand cross-link, denoted Pu^Py (and vice versa Py^Pu for the opposite 5′ → 3′ orientation).

Many aspects about ICLs formation and their structural impact on the regular B-DNA conformation are still unknown. Intensive research efforts over the past decade have mainly been directed toward the isolation and quantification of the repair of oxidative ICLs.[10−13] These studies show that these lesions are less abundant than other damages: 0.037 lesions per 10^9 nucleosides per Gy for G[8,5]C,[12] which is ∼300−2000 lower than the value reported for, e.g., 5-formyl-2′-deoxyuridine and 8-oxoguanine,[14] respectively. Oxidative ICLs nevertheless pose a potent threat to genome integrity, because of a mutagenic frequency reported to be close to 10%. This is due to the fact that recognition and nucleotide excision repair are significantly less effective than those for photolesions.[15,16] Their statistical rarity inevitably complicates the structural elucidation all the more, since tandem lesions imply a combinatorial chemistry with near-identical mass adducts. The rather large and increasing amount of experimental data collected on tandem lesions[10−13,15−17] spans most of the possible oxidative ICL. On-flow scan of damaged DNA by high-performance liquid chromatography coupled to tandem mass spectrometry HPLC-MS/MS[18] has lowered the detection threshold, hence leading over the last years to the identification of very exotic lesions. A first limitation is that, intrinsically, this technique cannot answer questions relative to the stereochemistry. For instance, for cluster DNA lesions with several chiral centers (e.g., ref 19), diastereoisomers can be obtained, but information on their relative ratio is scarce. A similar difficulty is encountered for intermediate peroxyl-bridged ICL.[20] Second, the currently available experimental protocols used to isolate ICL provide only limited structural informations about fragments of the system, typically restricted to linked dinucleotides. No NMR or X-ray structure of a DNA double

Received: September 6, 2011
Published: December 25, 2011

   dx.doi.org/10.1021/ja2084042 | J. Am. Chem.Soc. 2012, 134, 2111−2119

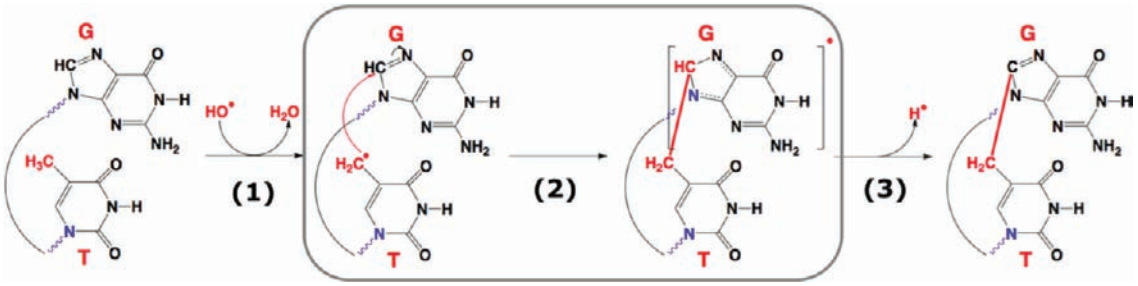


Figure 1. Three-step mechanism for the formation of the G[8,5-Me]T ICL adduct. The key radical attack from hydrogen-abstracted thymine to guanine leads to a covalently tethered adduct with a bridging $-CH_2-$ entity displayed with a bold red line. The round arc denotes the deoxyribo–phosphate–deoxyribo dpd surrounding. Step (2) is investigated in this study.

helix containing an oxidative ICL has been resolved so far. This is key missing information since the degree of structural distortion, usually quantified by means of bend and unwinding angles and the conservation or disruption of the Watson–Crick hydrogen-bond network, is intimately related to cytotoxity, lack of DNA repair, and ultimately cancer.[21] Molecular modeling offers an interesting alternative, acting as a "computational microscope"[22] which provides the atomic and even electronic level of detail required. This is particularly true for ICL as one can start from the known structure of solvated DNA and generate the lesion in silico. Computational approaches have played an increasingly important role in building up a sound understanding of DNA damage reactions, either relying on a classical force field description[23−25] or on density functional theory (DFT) calculations[26−29] or a combination of the two.[30]

In the present work, we investigate the formation of an oxidative ICL within a 12-base pair (12-bp) DNA duplex, using a hybrid quantum mechanics/molecular mechanics (QM/MM) framework in combination with Car–Parrinello (CP) molecular dynamics.[31,32] G[8,5-Me]T is taken as a case study, a legitimate choice in light of convergent experimental data and DFT studies on isolated dinucleotides[33,34] that identify this lesion as the most frequent oxidative ICL adduct; adenine, for instance, has proved to be less prone to radical additions.[35] G[8,5-Me]T formation involves a three-step ROS cascade represented in Figure 1. The initiating thymine-methyl H abstraction has been previously rationalized with the help of stochastic simulations,[36] in line with arguments based on solvent accessible surface area.[37] During the second step, the $T-CH_2^{\bullet}$ moiety attacks the C8 atom of an adjacent guanine and cyclization occurs. The third and last step, the G/H8 hydrogen departure is highly assisted by aqueous media; it restores a closed-shell species and is exergonic.

Our study focuses on the step (2) because cyclization requires a shortening of the distance between the two reactive bases, and the resulting local structural reorganization is expected to be the most important factor regarding the influence on the global structure of the DNA double strand. Steps (1) and (3) correspond to hydrogen atom transfers between the solvent and the macromolecule, and such chemical events have only minor effects in terms of electrostatic or steric interactions. (We verified this assumption by generating the product of step (3) and observing that its structural properties are not affected by the removal of the departing H atom.)

We show that the formation of the DNA lesion in step (2) is associated with a rather low activation free energy of ∼10 kcal/mol, in contrast to results from previous minimal gas-phase models (ca. 70 atoms, where the reactant exhibits a marked deviation from the π stacked mode and consequently activation energies of ∼25 kcal/mol for the defect formation).[34] In addition, the bend and unwinding angles we obtain are quite small compared to other ICL, showing that the induced DNA deformations are minor and thus hard to detect for DNA repair enzymes.

2. METHODOLOGY AND COMPUTATIONAL DETAILS

2.1. Structural Model and Equilibration at the Classical Level. Classical molecular dynamics (MD) simulations have been performed with the AMBER10[38] suite of programs. The initial configuration of an oligonucleotide dodecamer, with angles characterizing a canonical B-DNA, was generated in silico, using the nucgen module. Considering that experiments have suggested oxidative lesions are insensitive to sequence effects,[16] we have chosen the dodecameric sequence D[AGAGA **GT**AGAGT](·D-[TCTCTCATCTCA]). This choice has two main advantages: (i) the DNA double strand is long enough to maintain the B-helicity,[39] and (ii) the two nucleobases **G6** and **T7** engaged in the intrastrand lesion are placed at the center of the double helix, so that the lesion is as far as possible from the termini, which makes the model more representative of a longer double-stranded DNA sequence. Examples of previous QM/MM works on a DNA dodecamer with the reactive part located at the middle of the macromolecule can be found in refs 28 and 29.

A cartoon representation is given in Figure 2. The macromolecule was solvated explicitly with TIP3P water molecules[40] in a rectangular box with dimensions that were chosen to achieve a minimal distance of 28 Å between two periodic images of the solute, resulting in a cell of 49.2 × 49.7 × 72.7 Å^3. Twenty two potassium counterions were added to neutralize the double stranded DNA. The total system consists of ca. 17 000 atoms (including 5629 water molecules). The AMBER/parm99 force field[41] was used for the oligonucleotide and the potassium counterions. A hydrogen atom has been removed from the T7 methyl group (Me-C5), according to the numbering scheme in Figure 2. The residual charge of +0.07e was redistributed on the same T7/Me-C5 atom, while original van der Waals parameters were kept; a similar methodology was previously used for glycine radical.[42] Furthermore, this part of the T7 base is included in the QM part in the subsequent QM/MM simulations (see below). Long-range electrostatic interactions were computed using the smooth particle mesh Ewald (SPME) algorithm.[43] A cutoff of 10 Å was applied for the van der Waals interactions and the real space part of the electrostatic interactions. A series of classical MD runs (with a total length of 3 ns) was performed to reach an equilibrated state. All the bonds containing hydrogen atoms were

dx.doi.org/10.1021/ja2084042 | J. Am. Chem.Soc. 2012, 134, 2111−2119

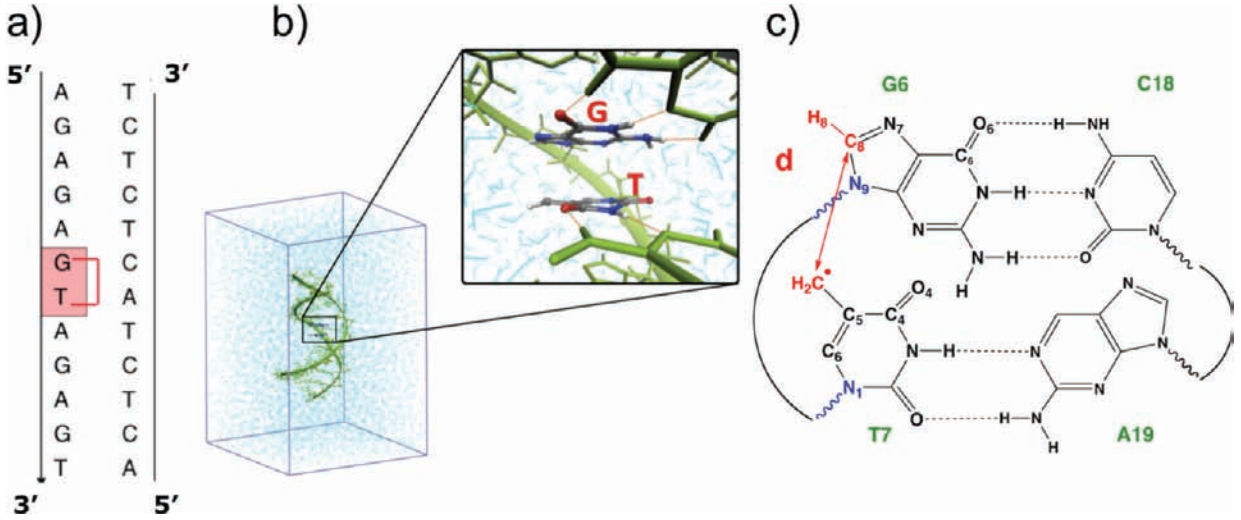


Figure 2. (a) Sequence of the 12-bp DNA with a centrally located G^T motif. (b) Cartoon representation in a water box (light blue lines), with a zoom on the two bases G6 and T7 constituting the reactive core. (c) Chemical structure and atom numbering scheme of the dinucleotide, paired with C18 and A19: the QM part encompasses the two nucleobase rings, with two nitrogen−carbon QM/MM boundaries indicated with blue wavy lines; the 22 QM atoms are represented in bold. The surroundings are treated with the AMBER/parm99 forcefield.

constrained with the SHAKE algorithm,[44,45] and a time step of 1.5 fs was used to integrate the equations of motion. Constant temperature was achieved using a Langevin thermostat[46] with a collision frequency of 5 and 1 ps^{-1} during equilibration and production phases, respectively. The Berendsen barostat[47] was used for NPT runs. The water box was relaxed around the solute by first conducting a NPT simulation (P = 1 atm and T = 300 K) with harmonic position restraints of 50 kcal/mol·$Å^2$ on the DNA atoms. Then the constraints were released, and the system underwent a new minimization. From this configuration, the system was heated up from 0 to 300 K using NVT stepwise simulations, for a total of 50 ps. Finally, a NPT MD production run of 5 ns was performed to provide starting configurations for subsequent QM/MM calculations. The root-mean-square deviation (rmsd) data, reported in Supporting Information, confirmed the stability of the unrestrained radical DNA structure.

2.2. QM Test Calculations. Two series of auxiliary QM calculations were performed with the Gaussian09 suite of programs.[48] The first one is based on a G-T dinucleotide in gas phase (70 atoms, corresponding to the two bases and the deoxyribo−phosphate−deoxyribo moiety). This model was used to compare different DFT functionals and assess the importance of dispersion corrections in our system. Indeed there is a key dispersion interaction between the bases of a same strand in DNA that standard DFT does not describe.[49] The corresponding results have been compiled in ref 50. Briefly, both BLYP and B3LYP yield reliable results, in terms of geometry and energetics, provided a dispersion correction is used.

Our second series of tests invoked a smaller model in order to make second-order Møller−Plesset (MP2) calculations tractable. It was verified on gas-phase reaction profiles (Supporting Information) that similar energy values were obtained when either BLYP-D ("-D" stands for "dispersion-corrected"), B3LYP-D, or MP2 are employed, the difference between relative energies being inferior to 1 kcal/mol.

An additional side result of our tests is that spin-contamination obtained with DFT lies in an acceptable range (inferior to 0.78 for an exact reference value of 0.75).

2.3. QM/MM MD Simulations. The QM/MM[51] simulations have been started from a structure of the solvated double strand DNA equilibrated at the classical level, in which the distance d between T7/C5-Me and G6/C8 atoms was ∼3.5 Å. The reason for this choice is that d is the reaction coordinate we have subsequently used to compute the free energy profile of the reaction (see below).

We used the approach developed by Rothlisberger and coworkers[31,32] which couples CPMD[52,53] and GROMOS.[54] This method was used in several previous studies dealing with similar DNA-based systems, with a QM region located at the center of the double helix.[28,29,55] Hence the choice we made for the simulation parameters of the present model relies partly on the tests performed in these studies.

The QM part is represented in Figure 2; it corresponds to the two reactive bases, i.e., the purine and pyrimidine moiety of G6 and T7, respectively (30 atoms in total). The dangling bonds were saturated with hydrogen atoms. Note that this particular treatment of the QM/MM frontier located at the C'_1−N_1 bond was tested extensively in ref 32. The DFT/BLYP[56,57] level of theory was used together with our dispersion-corrected atom-centered potentials (DCACP).[58,59] The Kohn−Sham orbitals were expanded in plane waves with a cutoff of 75 Ry and a quantum cell of dimensions 14.34 × 13.92 × 12.91 $Å^3$. The remaining solute atoms, the solvent, and the counterions were described with the AMBER/parm99 force field.[41]

All MD runs were performed with a fictitious electron mass of 500 au and a time step of 4 au (≈0.097 fs). The system underwent a first minimization using a simulated annealing-like procedure: Starting from a temperature of 20 K, atomic velocities were rescaled at each time step by a factor of 0.98. Then the system was heated from 0 up to 300 K in 6 ps. Subsequently, a 3 ps run of NVT simulation was performed, using a Nosé−Hoover chain thermostat[60−62] with 2000 cm^{-1} coupling frequency.

The free energy profile of step 2 of the reaction (Figure 1) was determined by thermodynamic integration, using the method of constraints.[63] The distance d between the C8 atom of G6 and the methyl carbon atom of T7 is an obvious choice

dx.doi.org/10.1021/ja2084042 | J. Am. Chem.Soc. 2012, 134, 2111−2119

for a simple reaction coordinate because the chemical step we are considering here corresponds to the formation of a single bond between these atoms. In order to perform the thermodynamic integration, d was decreased stepwise in 24 integration windows. For each window, the value of d was changed progressively from the previous calculation, with a growth speed of ∼0.8 Å/ps, then d was fixed for 3 ps and the Lagrange multiplier λ of the force acting on the constraint was collected and averaged from the last 2 ps. The error in the calculated free energy barrier was estimated as the sum of the statistical error on the Lagrange multiplier within each integration window and the error due to hysteresis. The former was estimated by computing the average λ from two segments of the window and then taking the difference between the two values, while the latter was obtained by repeating the thermodynamic integration for the backward reaction. At d = 1.5 Å, the constraint was removed in order to obtain the relaxed structure of the reaction intermediate (the third structure in Figure 1). A 3 ps unconstrained QM/MM MD run was performed to ensure the stability of this transient radical at a relatively short time scale.

Finally, we generated the final product of the reaction (the fourth structure in Figure 1) by removing manually the departing hydrogen atom and then running 25 ps of QM/MM MD simulation. We chose to adopt this simpler approach (with respect to the alternative, much more costly approach which would have consisted of performing a second thermodynamic integration) because the reaction mechanism is expected to be stepwise (see text) and the departure of a single hydrogen atom from the reacting subsystem is not expected to influence significantly the global structure of the DNA macromolecule. This latter assumption was validated a posteriori by observing that the reaction intermediate, and the products share very similar structural properties (see below).

The total QM/MM simulation time, including the time used to generate the lesion from the reactants by thermodynamic integration and the one used to relax the final product was 106 ps.

3. RESULTS AND DISCUSSION

We studied the reactivity of the 12-bp DNA (Figures 1 and 2) toward formation of the G[8,5-Me]T ICL adduct. Representative snapshots of the reaction states are represented in Figure 8. The free energy profile of step 2 of the reaction is given in Figure 4. The evolution of characteristic electronic and structural properties along the reaction path, such as the angle ϕ between the planes formed by G6 and T7, spin densities, and finally a set of helical angles are also plotted in Figures 5−7 respectively.

3.1. Structural Analysis of the Dodecamer after Hydrogen Abstraction From the T7/C5-Methyl Group. We start the analysis by considering the structure of the hydrogen-abstracted dodecamer (second structure in Figure 1), i.e., before we applied a constraint to enforce G[8,5-Me]T ICL formation. The rmsd along the classical equilibration and subsequent QM/MM−MD relaxation trajectories (Supporting Information) does not exhibit significant differences with respect to the "standard" nonradical dodecamer. The representative snapshot depicted in Figure 8a shows the Watson−Crick hydrogen-bond network around the reactive thymine (T7) is overall very well conserved. Note that the methylene radical of the hydrogen-abstracted T7 points toward the major groove that is in the opposite direction of the two Watson−Crick hydrogen bonds with A19 (see also Figure 2b). Like the original methyl group, the methylene radical is not involved in any contact with the rest of the macromolecule and is exposed to the solvent.

The distance between the C5 atom of the methylene radical of T7 and radical−philic centers of the purine bases can serve as a first possible indicator for the probability of forming an ICL, assuming that the longer the approaching distance, the higher the activation (free) energy. According to this criterion, one can relate the higher occurrence of Pu[8,5-Me]T vs T[5-Me,8]Pu adducts to the higher distance between the purine C8 atom and the T−$CH_2^{\bullet}$ radical center in the 3′ → 5′ orientation than in 5′ → 3′. This finding can be generalized to other tandem lesions, also by simply measuring distances on rigid canonical B-DNA structures. DNA is by essence a remarkably flexible system,[64] which can undergo a wide variety of structural deformations.[65] For the 12-bp sequence studied here, either G6 or A8 are possible partners for ICL formation with the hydrogen-abstracted T7; we denote the distance from the C5 atom of the methylene radical group of T7 (T7/C5) to the C8 atom of G6 (G6/C8) by d and the distance from T7/C5 to A8/C8 by d'. The time series of these two distances along our initial classical simulation in the reactant state are depicted in Figure 3. The $d < d'$ inequality is conserved during almost all

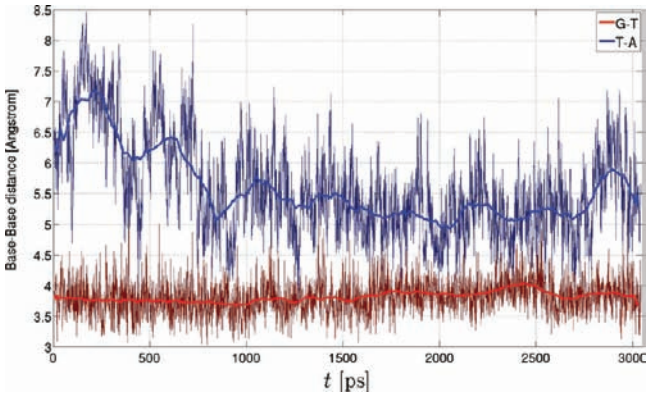


Figure 3. Comparison of the time evolution of the distances $d \equiv d[(T7/C5) - (G6/C8)]$ and $d' \equiv d[(T7/C5) - (A8/C8)]$ during our classical simulations. The G[8,5-Me]T lesion benefits from a shorter attack distance ($d(t) < d'(t)$ for most of the simulation time).

the simulation, despite a decrease of d' from ∼6.5 to 5.5 Å during the first nanosecond. In line with experimental data, this suggests that the formation of the T[5-Me,8]Pu lesion is less favorable than Pu[8,5-Me]T. In the next sections, we will focus on this latter type of lesion.

Although this simple geometrical criterion seems to rationalize the experimentally observed stereoelectronic preference, other more subtle factors are likely to come also into play. Thus we will study the evolution of energetic, electronic, and structural properties of the system upon formation of the covalent bond between the T7/C5-Me and G6/C8 atoms (step 2 in Figure 1).

3.2. Free Energy Profile of the Radical Attack during the Formation of G[8,5-Me]T. After an unconstrained 10 ps QM/MM MD simulation, the reaction coordinate d was kept fixed to 3.5 Å. It was then progressively decreased to 1.5 Å. The computed free energy profile (Figure 4) exhibits an activation barrier of $\Delta G^{\ddagger}$ of 9.7 ± 1.1 kcal/mol . The transition state is located at $d^{\ddagger}$ ∼2.05 Å ; it lies in the "usual" range of transition-state distances for C−C single bond formation in biological systems.[66]

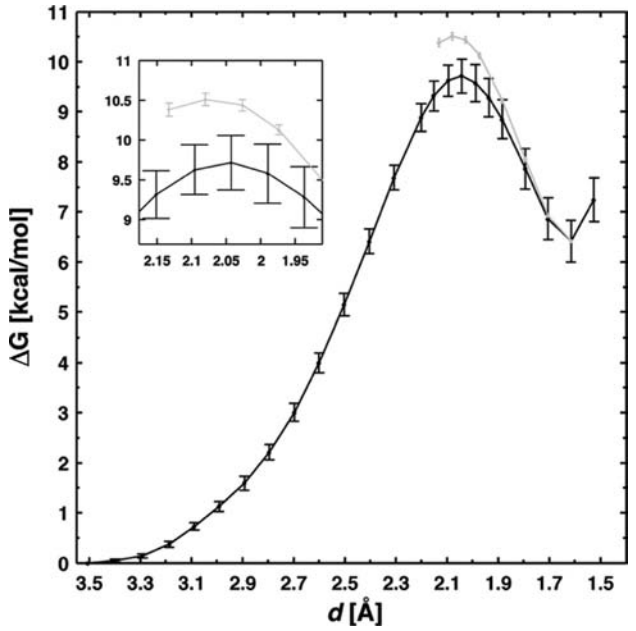


Figure 4. Free energy profile obtained by thermodynamic integration for the formation of the G[8,5-Me]T intrastrand cross-link embedded in a 12-bp oligonucleotide. The forward and backward reactions (computed to estimate the error due to hysteresis) are represented in black and gray, respectively. The error bars correspond to the statistical error on the Lagrange multiplier (see Section 2). The inset shows a zoom on the transition-state region.

Because steps 1 and 3 of the reaction mechanism (see Figure 1) only correspond to hydrogen atom transfers between the solvent and the DNA macromolecule, step 2 is expected to be the only one during which a global conformational transition, e.g., a bending of the B-DNA helix will occur. We will discuss these structural effects in Section 3.4. Considering the relatively low free energy barrier we obtained here, we can already say that, once a radical (most frequently a hydroxyl $HO^{\bullet}$ derived from bulk water) has diffused toward a solvent-exposed thymine methyl group and abstracted a hydrogen atom, the structural deformation required to bring the G and $T-CH_2^{\bullet}$ groups close together and form a covalent bond is a relatively facile process.

It is interesting to compare our results with previous calculations[34] that were conducted using a static model (i.e., based on geometry optimizations and not on MD simulations) of the GdpdT (d = deoxyribo, p = phosphate) dinucleotide in gas phase with a protonated phosphate. Our $d^{\ddagger}$ is slightly longer than the one obtained in ref 34 (2.13 Å). A more striking difference lies in $\Delta G^{\ddagger}$ which is as high as 24 kcal/mol with the static gas-phase model. We attribute this difference partly to the geometry of the reactants. In the gas-phase model of ref 34, the reactants geometry deviates substantially from a stacked arrangement, i.e., the angle ϕ between the planes formed by G6 and T7 is close to ∼50°, and the value of d is 6.3 Å. We also note that their optimized geometry exhibits extra intramolecular hydrogen bonds between nucleobases and desoxyribo rings, which do not occur in B-DNA. In our model instead, the stacked geometry is overall very well conserved, with average values of ϕ and d of ∼0 and 3.8 Å respectively. This illustrates the critical role of the supramolecular and aqueous environments for the tuning of tandem lesion formation: the available conformational landscape of the reactant within the solvated double strand DNA is limited by Watson−Crick hydrogen bonds and stacking interactions with neighboring bases, and the integrity of the macromolecule structure is itself strongly dependent on the presence of the solvent.

3.3. Spin Density Relocalization along the Reaction. The characteristic quantity to be monitored along a radical process is the unpaired spin density for the QM region that encompasses the two nucleobases (22 atoms). Figure 6 shows the unpaired spin distribution for five regularly spaced snapshots along d. The initial structure is a $T-CH_2^{\bullet}$ group within the dodecameric helix, with a reaction coordinate d = 3.5 Å and an ideally stacked structure. Its radical character is mainly localized on C5-Me (0.46), in line with the Lewis representation, also on the C6 atomic center, and to a lesser extent on the C4, N1 and O4 positions (respectively 0.41 and three ∼0.04 contributions). The reactivity of the thymine $CH_2^{\bullet}$ moiety is increased at d = 3 Å, as our calculations show a slight spin density transfer from C4, O4, and N1 to the C5-Me and C6 centers (values given in Supporting Information). The nearly alternating distribution seen in Figure 6 remains globally unaffected up to $d \sim 2.7$ Å, a distance at which the spin density begins to migrate progressively from T7 to G6. In Figure 6c, the unique role of the N7 nitrogen of the G6 five-membered ring is visualized, whereas the C8 position is free of any radical character. This thymine-to-guanine delocalization is, as expected, maximal near the transition state where the spin density is shared between T7 and G6, with a 0.4:0.6 ratio (case d). The thymine C5-Me and C6 positions and the guanine N7 center bear the largest contributions (∼0.4), but an increasing percentage of the unpaired spin density is delocalized in many other centers of the guanine bicyclic ring (i.e., N9, C4, C5, C6, O6, C2, and N2), as d further decreases toward the product.

This data provide complementary information to the incomplete Lewis picture (third structure in Figure 1). Purine nucleobases present more radical centers than pyrimidines (8 vs 5). This enhanced delocalization stabilizes both the transition-state structure and the product with respect to the reactant. The pyrimidine-to-purine transfer of the radical character is hence outlined as another electronic stabilization favoring thermodynamically the G[8,5-Me]T adduct; it is also likely to impede H8 departure, in agreement with a stepwise process as argued in the static description.[34] This inspection completes the analysis of the reaction process leading to the formation of a prototypical oxidative tandem lesion. Beyond the two bases directly involved in such lesions, the local ∼2 Å shrinkage of the guanine−thymine distance affects the global structur of the B-DNA helix. This is the scope of the next section.

3.4. Structural Distortion upon G[8,5-Me]T Formation. In relation with the nucleotide excision repair (NER) for this subfamily of DNA lesions, it is interesting to consider geometrical descriptors that characterize a damaged B-DNA. This may serve to further distinguish bulky tandem lesions that are refractory to repair by UvrABC endonucleases[16] and DNA glycosylases.[67]

A first local geometric signature of the DNA distortion is the continuous increase of the dihedral angle ϕ between the G6 and T7 base planes (Figure 5). In the reactant state ($d \sim 3.5$ Å in Figures 4 and 5), ϕ fluctuates around an average value of ∼0°, which corresponds to a stacked geometry (see also Figure 6a). Along the reaction path, ϕ increases progressively up to ∼25°. A closer inspection reveals that ϕ initially obeys a pendulum-like motion, with two sets of positive and negative values centered around the zero baseline. Slightly before the transition

dx.doi.org/10.1021/ja2084042 | J. Am. Chem.Soc. 2012, 134, 2111−2119

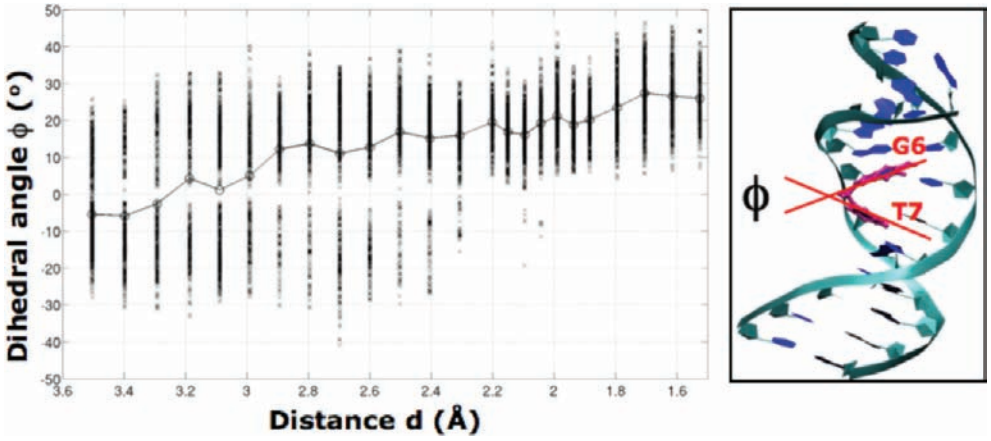


Figure 5. Evolution of the G6^T7 dihedral angle (ϕ) along the reaction coordinate d during our thermodynamic integration procedure. Starting from a near-zero averaged value reflecting an ideally stacked structure, ϕ progressively increases throughout lesion formation, to reach a final value of ~25°.

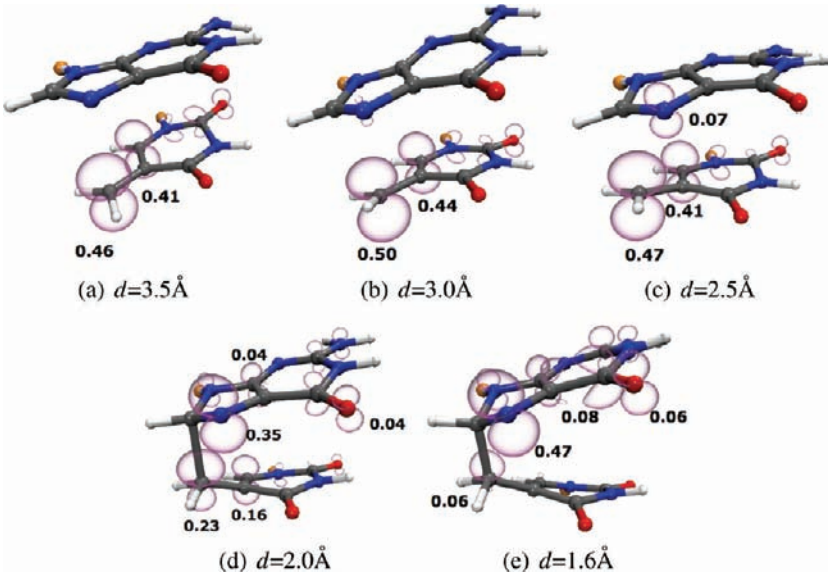


Figure 6. Spin density distributions between G6 and T7 for five distances d regularly spaced along the reaction profile. A progressive delocalization from thymine to guanine is observed. The rest of the DNA environment is omitted for the sake of clarity. Numerical data and the complete numbering are collected as Supporting Information.

state, at $d \sim 2.3$ Å, the C−C linkage becomes strong enough to lock the conformation in one mode, and ϕ only takes positive values. Hence, its fluctuations are notably diminished along the reaction coordinate, indicating a local increased rigidity. This dihedral pinch toward the major groove has a limited spatial extent, as time evolutions of the adjacent plane-to-plane angles (between A5 and G6 or T7 and A8) show a near-optimal planarity even after formation of the ICL.

Interestingly, this angular deformation does not have a significant impact on the G6:C18 and T7:A19 Watson−Crick hydrogen-bonds, as depicted in Figure 8b,c. Some alterations of the Watson−Crick network have been reported for photolesions, like, e.g., a weakening of the hydrogen-bond network for the cis−syn cyclobutane thymine dimer lesion[68] or its disruption upon formation of the (6−4) photolesion or intercalation of cisplatin.[26]

To assess more globally the structural impact of the G[8,5-Me]T adduct on the B-DNA conformation, the bend and unwinding angles are reported in Figure 7. These two angles, denoted α and γ hereafter, are crucial in the context of DNA repair. We first comment on the bend angle α. It initially presents a near-zero value, which progressively raises up to 40° halfway along the reaction path, i.e., for a reaction coordinate d = 2.5 Å. Afterward, the bend angle enters in a rather pronounced decreasing regime, to finally stabilize around 20 ± 5° in the G[8,5-Me]T product state. In Figure 7c, the bending dial does not reveal an orientation, whereas most of DNA lesions are known to bend toward the major groove.[68,69] Again, this pinpoints to the particular behavior of this oxidative ICL. One should keep in mind that bending is a soft deformation mode, with an energetic penalty most probably inferior to 5 kcal/mol, a value inferred for a 90° bend of a 15-bp sequence.[70]

Another key quantity that is often invoked to characterize the loss of B-DNA helicity is the unwinding (or more rarely overwinding) angle. It is defined as the difference between the mean twist angle for the product, 28° in our case, and the unperturbed, reference value of 36°. Our estimate of ~8° for the G[8,5-Me]T lesion, alongside the one for the bend angle,

dx.doi.org/10.1021/ja2084042 | J. Am. Chem.Soc. 2012, 134, 2111−2119

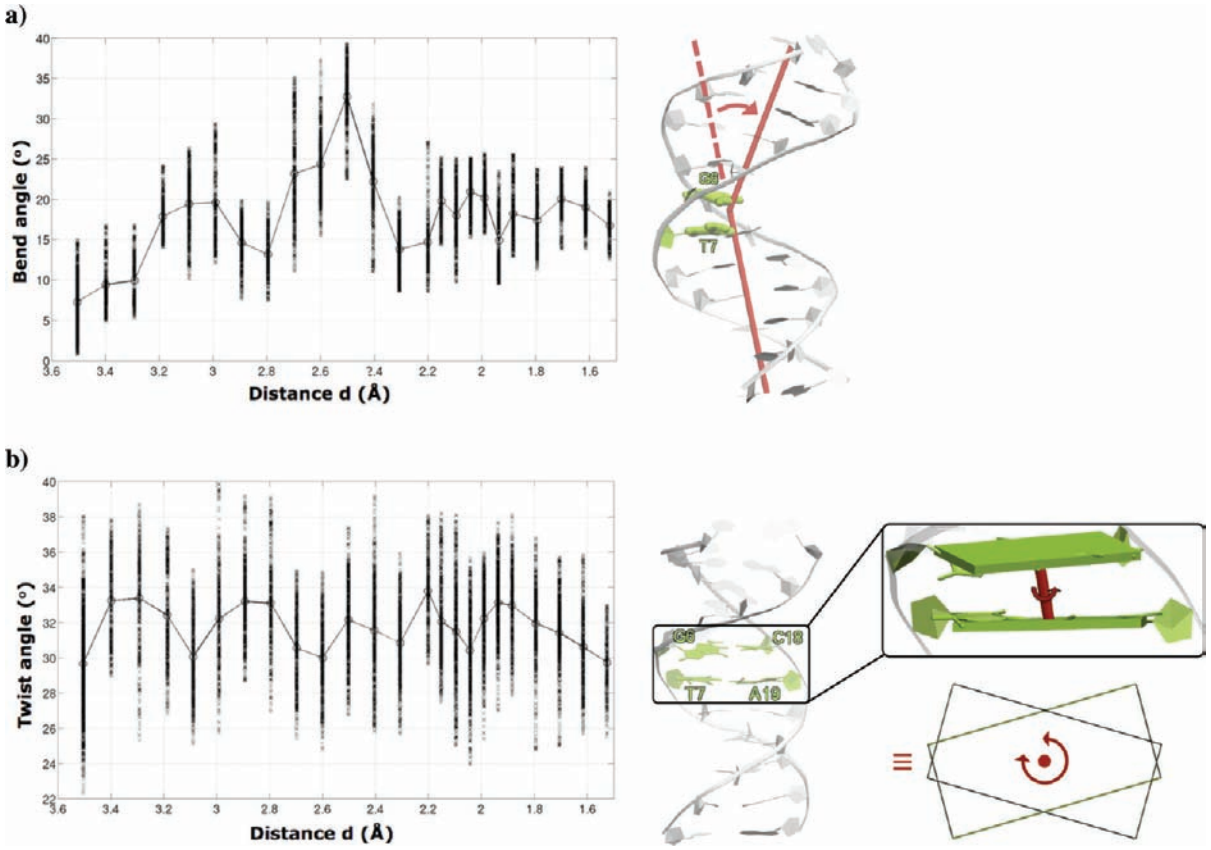


Figure 7. Evolutions of characteristic DNA angles along the reaction profile. The (a) bend and (b) twist angles. The definition of these parameters follows that of ref 74. The analysis was performed with the 3DNA software package.[74]

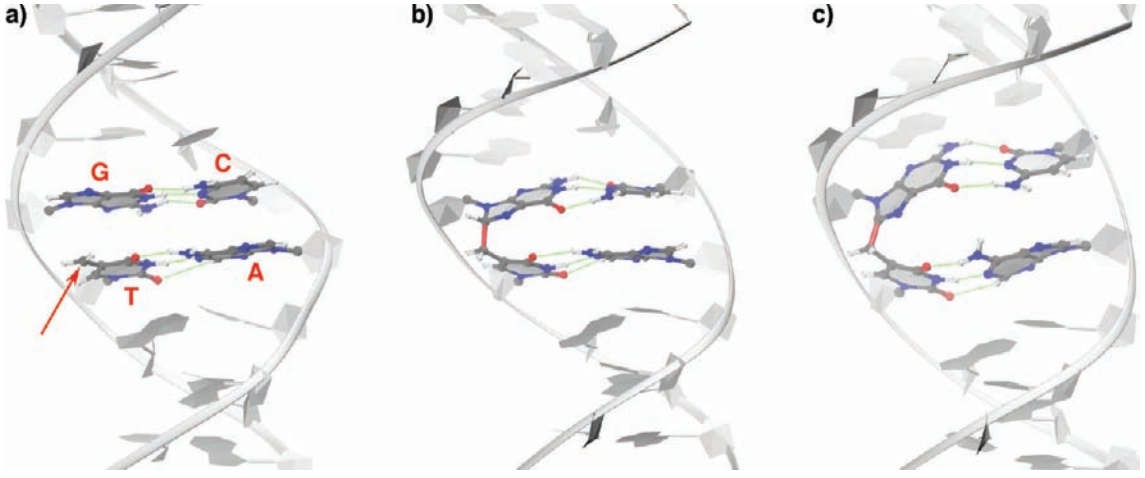


Figure 8. Representative snapshots of the reaction states we have modeled. (a) Reactant after hydrogen abstraction at $d \sim 3.5$ Å (second structure in Figure 1). The methylene radical is indicated with a red arrow. (b) Reaction intermediate at $d \sim 1.6$ Å, (third structure in Figure 1). The extra bond formed between the T and G nucleic bases is represented with a red stick. (c) Structure of the final G[8,5-Me]T lesion (fourth structure in Figure 1). The Watson−Crick hydrogen bonds, depicted with green lines, are preserved throughout the entire reaction mechanism.

anchors this prototypical oxidative ICL within the realm of different DNA lesions. In view of the absence of experimentally determined 3D structures for such embedded lesions and hampered by inherent difficulties to evaluate low bend angles[71] or unwinding angles, their estimation from first-principles simulations may palliate the lack of experimental data.

It could be argued that considering the bend and unwinding angles of the reaction intermediate instead of the structural parameters in the final product (respectively the third and fourth structures in Figure 1) is not relevant. In fact, we have conducted a QM/MM simulation of the product (see Section 2) and verified that the departure of one hydrogen atom from the reacting subsystem does not induce important modfications of the macromolecular global structure (Figure 8). In particular, the bend and unwinding angles remain essentially the same.

In Figure 9, we compare our values of the bend and unwinding angles to a set of available data for both intra- and interstrand cross-linked DNA lesions. Even though experimental measurements are

dx.doi.org/10.1021/ja2084042 | J. Am. Chem.Soc. 2012, 134, 2111−2119

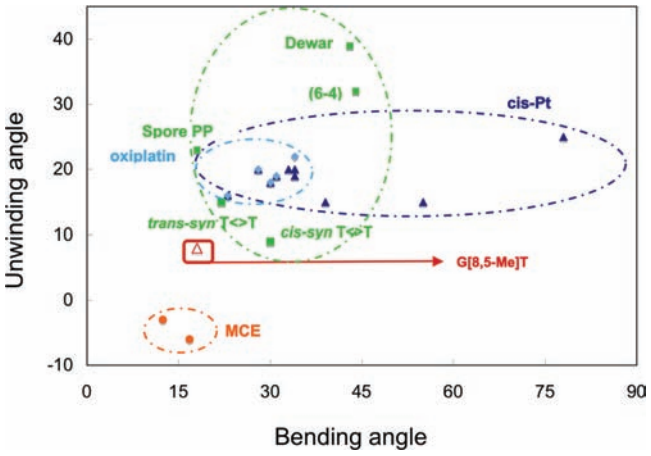


Figure 9. Bending vs unwinding angles for a collection of DNA lesions (induced by drugs or light). The complete list of references is given as Supporting Information. MCE stands for mechloroamine (orange). One can discern several families for each type of damage, our prototypical oxidative lesion being displayed in red at (20;8). These angles are inferior to the values reported for other intrastrand CL adducts (induced by cisplatin and oxiplatin in dark-blue triangles and light-blue diamonds; photolesions with green squares). By convention, negative angles denote an overwinding.

subject to important standard deviations depending on the underlying techniques (see the large range of values reported for cisplatin adducts, dark blue triangles), several families of well characterized lesions can indeed be identified. Chelating agents can induce either intra- or interstrand cross-links, as exemplified respectively by cisplatin adducts and related oxiplatin lesions (represented in Figure 9 with a blue ellipsoid) or by methylchloramine (interstrand CL, with an overwound). But perhaps the best systems to compare with are the five photolesions, depicted with green squares. It has been postulated that the contrasting structural distortions that exist among light-induced ICLs may explain their differences in repair efficiency.[72] Our data extend this idea to oxidative ICLs, alongside recent studies on G[8−5]U[73] that partly draw similar conclusions; the rather small structural deviations we observe (conservation of the Watson−Crick hydrogen-bonds network and low bend and unwinding angles) might hinder an efficient recognition, which would explain the lack of efficient repair for oxidative ICLs.

4. CONCLUDING REMARKS

We have simulated the formation of the G[8,5-Me]T oxidative ICL within a solvated DNA dodecamer using a QM/MM MD scheme. We show how the average distance between neighboring bases in the reactant state and the spin delocalization in the course of the reaction may favor a particular lesion (e.g., Pu^Py vs Py^Pu). More importantly, our calculations show that the critical step during which the G/C8 and T/C5-Me groups are brought close together to form the extra C−C bond is a facile process. The corresponding free energy barrier is estimated to be ∼10 kcal/mol. Despite a noticeable squeezing of the two G and T bases involved directly in the formation of the lesion, the effect of the ICL on the global structure of the DNA macromolecule is rather small. The computed structural indicators, such as Watson−Crick hydrogen bond pairing and bending and unwinding angles, all exhibit weaker deviations compared to other intrastrand cross-link adducts induced by chemotherapeutic drugs or photolesions. This investigation sheds new light on the lack of repair for such oxidative lesions, since the recognition may process less efficiently in absence of marked structural deviations.

■ ASSOCIATED CONTENT

Supporting Information

Overall and three local rmsds of the 12-bp T7/hydrogen-abstracted sequence, with respect to the starting frame for the classical equilibration. Energetic profiles computed with BLYP, B3LYP, and MP2 for the isolated QM system. Full list of references for Figure 9 and values for bend and unwinding angles reported for a series of intrastrand cross-link DNA lesions. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

elise.dumont@ens-lyon.fr

■ ACKNOWLEDGMENTS

The authors are grateful to Dr. Carine Michel and Dr. Paul Fleurat-Lessard for critical comments on the manuscript. E.D. acknowledges a financial support from the Région Rhône-Alpes (Explo'RA Pro fellowship). C.P. is grateful for a PhD fellowship from the French Ministery of research. Calculations were performed using the LCBC cluster and the local HPC resources of PSMN at ENS-Lyon. J.G. and U.R. acknowledge support by the Swiss National Science Foundation.

■ REFERENCES

(1) Kanvah, S.; Joseph, J.; Schuster, G. B.; Barnett, R. N.; Ceveland, C. L.; Landman, U. *Acc. Chem. Res.* **2010**, *43*, 280−287.
(2) Munzel, M.; Szeibert, C.; Glas, A. F.; Globisch, D.; Carell, T. *J. Am. Chem. Soc.* **2011**, *133*, 5186−5189.
(3) 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.
(4) Mantel, C.; Chandor, A.; Gasparutto, D.; Douki, T.; Atta, M.; Fontecave, M.; Bayle, P.-A.; Mouesca, J.-M.; Bardet, M. *J. Am. Chem. Soc.* **2008**, *130*, 16978−16984.
(5) Heil, K.; Kneuttinger, A. C.; Schneider, S.; Lischke, U.; Carell, T. *Chem.—Eur. J.* **2011**, *17*, 9651−9657.
(6) Cooke, M. S.; Evans, M. D.; Dizarogly, M.; Lunec, J. *FASEB J.* **2003**, *17*, 1195−1214.
(7) Cadet, J.; Douki, T.; Gasparutto, D.; Ravanat, J.-L. *Mutat. Res.* **2003**, *531*, 5−23.
(8) Waris, G.; Ahsan, H. *J. Carcinog.* **2006**, *5*, 14.
(9) Wang, W.; Razskazovskii, Y.; Sevilla, M. D. *Int. J. Radiat. Biol.* **1997**, *71*, 387−399.
(10) Zeng, Y.; Wang, Y. *Nucleic Acids Res.* **2006**, *34*, 6521−6529.
(11) Zeng, Y.; Wang, Y. *Biochemistry* **2007**, *46*, 8189−8195.
(12) Hong, H.; Cao, H.; Wang, Y. *Nucleic Acids Res.* **2007**, *35*, 7118−7127.
(13) Lin, G.; Zhang, J.; Zeng, Y.; Luo, H.; Wang, Y. *Biochemistry* **2010**, *49*, 2346−2350.
(14) Kriker, M. C.; Drake, J. W. *J. Bacteriol.* **1990**, *172*, 3037−3039.
(15) Yang, Z.; Collis, L. C.; Basu, A. K.; Zou, Y. *Chem. Res. Toxicol.* **2006**, *18*, 1339−1346.
(16) Gu, C.; Zhang, Q.; Yang, Z.; Wang, Y.; Zhou, Y.; Wang, Y. *Biochemistry* **2006**, *45*, 10739−10746.
(17) Gates, K. S. *Rev. React. Intermed. Chem.* **2007**, *8*, 333−378.
(18) Cao, H.; Wang, Y. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 611−617.

(19) Regulus, P.; Duroux, B.; Bayle, P.-A.; Favier, A.; Cadet, J.; Ravanat, J.-L. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 14032−14307.
(20) Hong, I. S.; Carter, K. N.; Sato, K.; Greenberg, M. M. *J. Am. Chem. Soc.* **2007**, *129*, 4089−4098.
(21) Helleday, T.; Petermann, E.; Lundin, C.; Hodgson, B.; Sharma, R. A. *Nat. Rev. Cancer* **2008**, *8*, 193−204.
(22) Lee, E. H.; Hsin, J.; Sotomayor, M.; Comellas, G.; Schulten, K. *Structure* **2009**, *17*, 1295−1306.
(23) Sczepanski, J. T.; Jacobs, A. C.; Majumdar, A.; Greenberg, M. M. *J. Am. Chem. Soc.* **2009**, *131*, 11132−11139.
(24) Cheng, X.; Kelso, C.; Hornak, V.; de los Santos, C.; Grollman, A. P.; Simmerling, C. *J. Am. Chem. Soc.* **2005**, *127*, 13906−13918.
(25) Song, K.; Hornak, V.; de los Santos, C.; Grollman, A.; Simmerling, C. *J. Comput. Chem.* **2008**, *29*, 17−23.
(26) Robertazzi, A.; Platts, J. A. *Chem.—Eur. J.* **2006**, *12*, 5747−5756.
(27) Vargiu, A. V.; Ruggerone, P.; Magistrato, A.; Carloni, P. *J. Phys. Chem. B* **2006**, *110*, 24687−24695.
(28) Gossens, C.; Tavernelli, I.; Rothlisberger, U. *J. Am. Chem. Soc.* **2008**, *130*, 10921−10928.
(29) Masson, F.; Laino, T.; Tavernello, I.; Rothlisberger, U.; Hutter, J. *J. Am. Chem. Soc.* **2008**, *130*, 3443−3450.
(30) Loos, P.-F.; Dumont, E.; Laurent, A. D.; Assfeld, X. *Chem. Phys. Lett.* **2009**, *475*, 120−123.
(31) Laio, A.; VandeVondele, J.; Rothlisberger, U. *J. Chem. Phys.* **2002**, *116*, 6941−6947.
(32) Laio, A.; Gervasio, F. L.; VandeVondele, J.; Sulpizi, M.; Rothlisberger, U. *J. Phys. Chem. B* **2004**, *108*, 7963−7968.
(33) Xerri, B.; Morell, C.; Grand, A.; Cadet, J.; Cimino, P.; Barone, V. *Org. Biomol. Mol.* **2006**, *4*, 3986−3992.
(34) Labet, V.; Morell, C.; Grand, A.; Cadet, J.; Cimino, P.; Barone, V. *Org. Biomol. Chem.* **2008**, *6*, 3300−3305.
(35) Bellon, S.; Ravanat, J.-L.; Gasparutto, D.; Cadet, J. *Chem. Res. Toxicol.* **2002**, *15*, 598−606.
(36) Begusova, M.; Spotheim-Maurizot, M.; Sy, D.; Michalik, V.; Charlier, M. *Proc. Conversation Biomol. Stereodyn., 11th* **2001**, *19*, 141.
(37) Balasubramanian, B.; Pogozelski, W. K.; Tullius, T. D. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *343*, 9738−9743.
(38) Case, D. et al. *AMBER 10*; University of California: San Francisco, CA, 2006.
(39) Ferentz, A. E.; Wiorkiewicz-Kuczera, J.; Karplus, M.; Verdine, G. L. *J. Am. Chem. Soc.* **1993**, *115*, 7569−7583.
(40) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(41) Wang, J. M.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049−1074.
(42) Barone, V.; Capecchi, G.; Brunel, Y.; Andries, M.-L. D.; Subra, R. *J. Comput. Chem.* **1997**, *18*, 1720−1728.
(43) Cheatham, T. E. III; Miller, J. L.; Fox, T.; Darden, T. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 4193−4194.
(44) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(45) Miyamoto, S.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 952−962.
(46) Izaguirre, J. A.; Catarello, D. P.; Wozniak, J. M.; Skeel, R. D. *J. Chem. Phys.* **2001**, *114*, 2090−2098.
(47) Berendsen, H. J. C.; Postma, J. P. M; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(48) Frisch, M. J.;et al. *Gaussian 09*, revision A.02; Gaussian, Inc.: Wallingford, CT, 2009.
(49) Lin, I.-C.; Rothlisberger, U. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2730−2734.
(50) Dupont, C.; Patel, C.; Dumont, E. *J. Phys. Chem. B* **2011**, *115*, 15138−15144.
(51) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227−249.
(52) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471.
(53) Parrinello, M.; Andreoni, W.; Curioni, A. *CPMD*; IBM Corporation and Max-Planck Institut: Armonk, New York and Stuttgart, Germany, 2000.
(54) Scott, W. R. P.; Hunenberger, P. H.; Tironi, I. G.; Mark, A. E.; Billeter, S. R.; Fennen, J.; Torda, A. E.; Huber, T.; Kruger, P.; van Gunsteren, W. F. *J. Phys. Chem. A* **1999**, *103*, 3596−3607.
(55) Spiegel, K.; Magistrato, A.; Maurer, P.; Ruggerone, P.; Rothlisberger, U.; Carloni, P.; Reedijk, J.; Klein, M. L. *J. Comput. Chem.* **2008**, *29*, 38−49.
(56) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098−3100.
(57) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789.
(58) von Lilienfeld, O. A.; Tavernelli, I.; Rothlisberger, U.; Sebastiani, D. *Phys. Rev. Lett.* **2004**, *93*, 153004.1−153004.4.
(59) von Lilienfeld, O. A.; Tavernelli, I.; Rothlisberger, U.; Sebastiani, D. *J. Chem. Phys.* **2005**, *122*, 014113.1−014113.6.
(60) Nosé, S. *J. Chem. Phys.* **1984**, *81*, 511−519.
(61) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695−1697.
(62) Martyna, G. J.; Klein, M. L.; Tuckerman, M. *J. Chem. Phys.* **1992**, *97*, 2635−2643.
(63) Sprik, M.; Ciccotti, G. *J. Chem. Phys.* **1998**, *109*, 7737−7744.
(64) Prevost, C.; Takahashi, M.; Lavery, R. *Chem. Phys. Chem* **2009**, *10*, 1399−1404.
(65) Noy, A.; Perez, A.; Lankas, F.; Luque, F. J.; Orozco, M. *J. Mol. Biol.* **2004**, *343*, 627−638.
(66) Himo, F. *Biochim. Biophys. Acta* **2005**, *1*, 24−33.
(67) Bergeron, F.; Auvre, F.; Radicella, J. P.; Ravanat, J.-L. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 5528−533.
(68) Cooney, M. G.; Miller, J. H. *Nucleic Acids Res.* **1997**, *25*, 1432−1436.
(69) Stehlikova, K.; Kostrhunova, H.; Kasparkova, J.; Brabec, V. *Nucleic Acids Res.* **2002**, *30*, 2894−2898.
(70) Curuksu, J.; Zacharias, M.; Lavery, R.; Zakrzewska, K. *Nucleic Acids Res.* **2009**, *37*, 3766−3773.
(71) Wozniak, A. K.; Schroder, G. F.; Grubmueller, H.; Seidel, C. A. M.; Oesterhelt, F. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 18337−18432.
(72) Kim, J. K.; Patel, D.; Choi, B. S. *Photochem. Photobiol.* **1995**, *62*, 44−50.
(73) Churchill, C. D. M.; Eriksson, L. A.; Wetmore, S. D. *Chem. Res. Toxicol.* **2011**, *24*, 2189−2199.
(74) Lu, X.-J.; Olson, W. K. *Nat. Protoc.* **2008**, *3*, 1213−1227.