BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

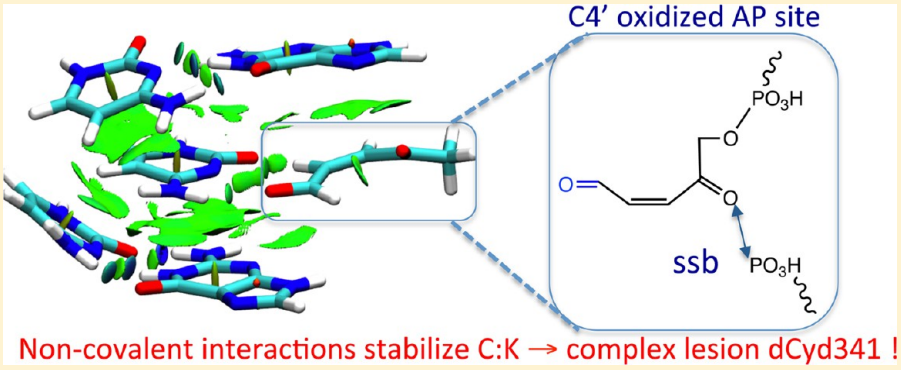
Structure, Dynamics, and Interactions of a C4′-Oxidized Abasic Site in DNA: A Concomitant Strand Scission Reverses Affinities

Chandan Patel,[†] Tomàš Dršata,[‡] Filip Lankaš,*[,‡] and Elise Dumont*[,†]

[†]Laboratoire de Chimie, UMR 5182 CNRS, École Normale Supérieure de Lyon, 46, allée d'Italie, 69364 Lyon Cedex 07, France

[‡]Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo nám. 2, 166 10 Praha 6, Czech Republic

Supporting Information



ABSTRACT: Apurinic/apyrimidinic (AP) sites constitute the most frequent form of DNA damage. They have proven to produce oxidative interstrand cross-links, but the structural mechanism of cross-link formation within a DNA duplex is poorly understood. In this work, we study three AP-containing d[GCGCGCXCGCGCG]·d[CGCGCGKGCGCGC] duplexes, where X = C, A, or G and K denotes an α,β-unsaturated ketoaldehyde derived from elimination of a C4′-oxidized AP site featuring a 3′ single-strand break. We use explicit solvent molecular dynamics simulations, complemented by quantum chemical density functional theory calculations on isolated X:K pairs. When X = C, the K moiety in the duplex flips around its glycosidic bond to form a stable C:K pair in a near-optimal geometry with two hydrogen bonds. The X = A duplex shows no stable interaction between K and A, which contrasts with AP sites lacking a strand scission that present a preferential affinity for adenine. Only one, transient G:K hydrogen bond is formed in the X = G duplex, although the isolated G:K pair is the most stable one. In the duplex, the stable C:K pair induces unwinding and sharp bending into the major groove at the lesion site, while the internal structure of the flanking DNA remains unperturbed. Our simulations also unravel transient hydrogen bonding between K and the cytosine 5′ to the orphan base X = A. Taken together, our results provide a mechanistic explanation for the experimentally proven high affinity of C:K sites in forming cross-links in DNA duplexes and support experimental hints that interstrand cross-links can be formed with a strand offset.

Apurinic/apyrimidinic (AP) sites are generated spontaneously or as a result of chemical or radiation damage to DNA. They constitute an important class of DNA damage whose repair has been extensively investigated.[1,2] If left unrepaired, they lead to replication errors according to the so-called A-rule.[3] A less understood aspect is the chemistry of oxidized abasic sites,[4] which are prone to undergoing *in situ* coupling reactions with other nitrogen-rich nucleobases. The resulting interstrand cross-links (ICLs) are among the most deleterious DNA lesions and can trigger genomic instability.[5]

In the past decade, ICLs have been identified on a regular basis.[6−10] Recent improvements in analytical techniques such as high-performance liquid analysis coupled to tandem mass spectrometry[11] allow for highly reliable detection in spite of the rarity of such defects (typically a few per 10^8 nucleotides). Recently, the focus has shifted to some of the more complex lesions. A particular lesion called dCyd341 has attracted much attention because of the time it was discovered in cellular DNA.[8] The lesion is initiated by abstraction of a hydrogen atom at the C4′ position of the deoxyribose ring, followed by fragmentation that leaves a 3′ single-strand break (ssb) and a ketoaldehyde (K). The structure of K is depicted in Figure 1. It is a bis-electrophile, prone to condense with a nucleobase X featuring a -N=C(NH_2)- motif, as demonstrated for cytosine in the experimental work of Ravanat and co-workers.[8]

The structural elucidation of an ICL lesion such as dCyd341 raises a series of questions, the first of which being that only cytosine was found experimentally to react with K. One would

Received: September 12, 2013
Revised: October 16, 2013
Published: October 17, 2013

   dx.doi.org/10.1021/bi401268q | *Biochemistry* 2013, 52, 8115−8125

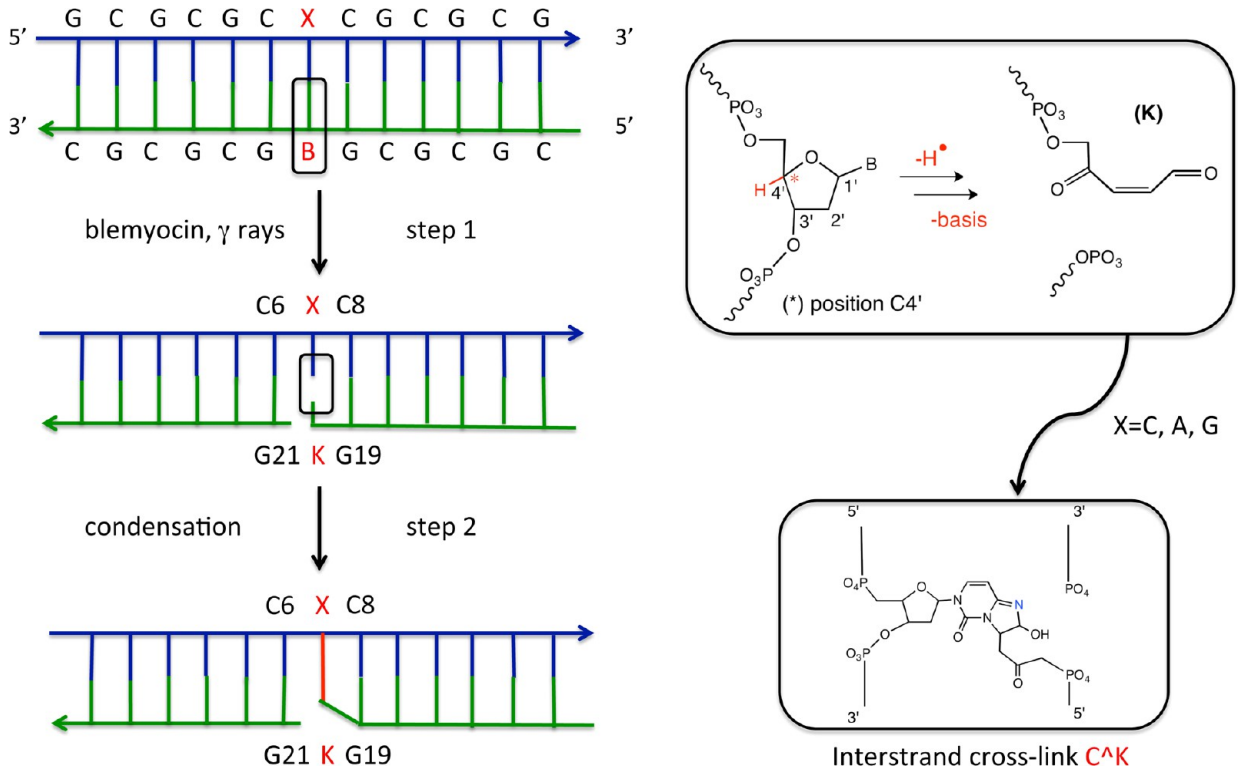


Figure 1. Scheme of a C4′-oxidized abasic site embedded in the middle of an alternate 13 bp poly(GC) sequence studied in this work. B denotes a generic base, which undergoes hydrogen abstraction on the deoxyribose ring (position C4′; the corresponding hydrogen atom is colored red). The outcome is an oxidized ketoaldehyde K. The damaged oligonucleotide can evolve toward an interstrand cross-link, represented here for C as found experimentally.

need to address the ease of formation of purine analogues, as K can *a priori* also react with adenine or guanine. The reactivity order is relatively easy to establish on isolated fragments, using experiments combined with theoretical investigations.[12] However, this provides only an intrinsic ordering that can be reversed once the two reactants are embedded in a B-DNA double helix.[13] Moreover, information about the location of the double nucleophilic attack is lost upon enzymatic digestion; thus, it is unclear whether K couples only with the opposite nucleobase or whether it can also react with other nucleobases. The feasibility of such an offset clearly depends on the AP duplex flexibility, probably exacerbated by the strand cleavage, and also on the interactions that the lesion can develop with proximal nucleobases. Greenberg and co-workers have suggested that the attack can proceed with a strand offset of either +1 or −1 [i.e., with base C6 or C8 in our case (see Figure 1)],[9] yet one lacks a more direct and quantitative proof that ICLs may indeed occur with a strand offset.

The relative positioning of the two ICL coupling partners, imposed by the B-DNA helix, is known to be a decisive factor for the ease of formation of a tandem lesion: the closer the reactive centers, the lower the free activation energy.[14] A 2 Å increase in the average approaching distance increases the barrier for the intrastrand peroxyl-nucleobase attack by a factor of 2. This provides a rationale for a significantly different ratio (1:10), a phenomenon sometimes called a sequence effect.[15,16] More recently, the reactivity of furan,[17] a photoactivated precursor of ketoaldehydes, was rationalized. Once the reaction mechanism has been firmly established,[18] it then becomes essential to characterize the structure of the reactants within B-DNA, i.e., the structure of the AP-containing duplexes. Molecular modeling provides a tool well-suited to elucidating the structural consequences of having an oxidized C4′-AP within a double-stranded oligonucleotide roughly 10−20 bp long.[19,20]

The case of AP sites arising from the spontaneous or enzymatic cleavage of the N-glycosylic bond without strand scission has been extensively studied.[21] These studies are based on experimental structures (either X-ray[10] or nuclear magnetic resonance[20,22,23]) and are often combined with molecular modeling.[24,25] A general feature is the potential of AP-containing duplexes to display considerable structural diversity, because they lack the Watson−Crick pairing and π-stacking that underpin DNA helicity. It has nevertheless been reported that duplexes with an adenine residue opposing the abasic site retain a right-handed, B-DNA helical structure in solution.[7] However, the existing studies of abasic sites within short oligonucleotides have so far always considered the situation where the integrity of the strand is preserved. Thus, information about the dynamics and binding of a C4′-AP site within B-DNA with the inclusion of a single-strand break is still missing. The single-strand break almost certainly implies additional flexibility, and it is not straightforward to predict how the DNA helix is perturbed and whether it can maintain a close X:K pairing, which would then favor ICL formation.

In this work, we set out to investigate the structure, dynamics, and interactions of ketoaldehyde K paired with A, G, or C nucleobases. We first characterize the intrinsic pairing between K and X (X = A, G, or C) based on quantum chemical density functional theory (DFT). We then study the effect of embedding the K:X pair in a B-DNA duplex with a 3′ single-strand break using atomic-resolution, explicit solvent molecular

dx.doi.org/10.1021/bi401268q | *Biochemistry* 2013, 52, 8115−8125

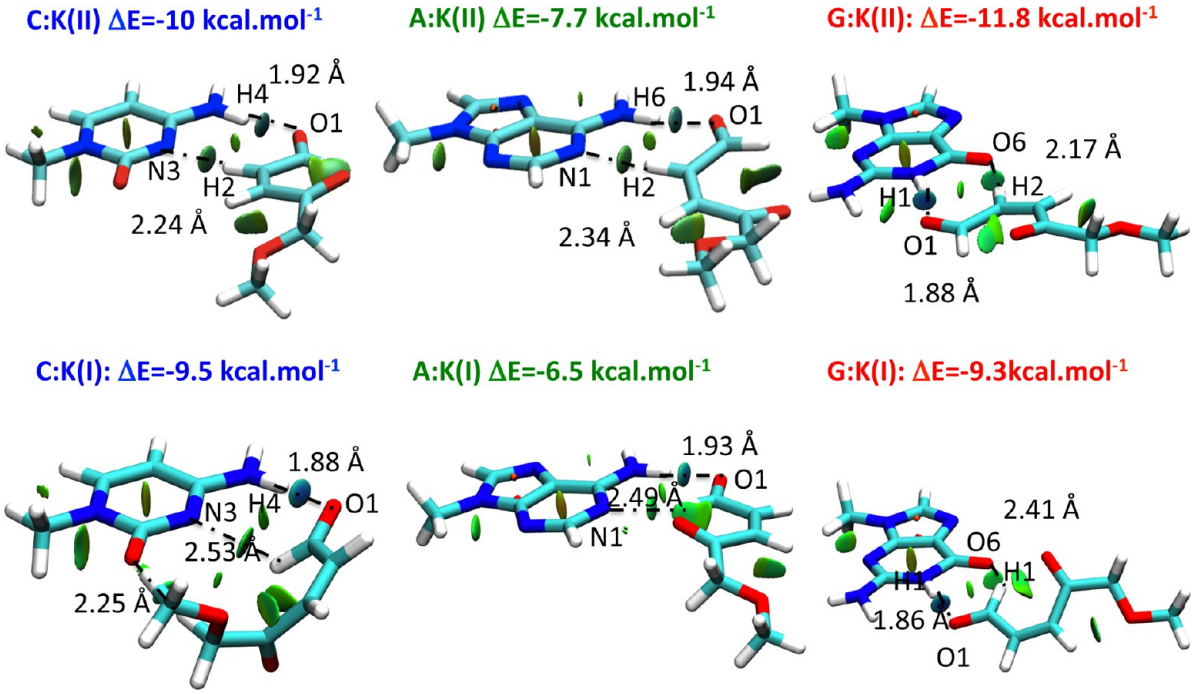


Figure 2. Characterization of the optimal interactions for isolated X:K pairs. The blue and green NCI isodensities (the isovalue is 0.5 au, for a density cutoff of 0.1 au) reflect the presence of noncovalent interactions. Two nearly equivalent orientations of the ketoaldehyde moiety (top and bottom rows) can be distinguished.

dynamics (MD) simulations. The structural changes induced by the lesion are described by a range of conformational parameters such as interatomic distances and angles, inter-base pair coordinates, and global bending characteristics.

METHODS

Here we briefly summarize the methods employed in our study. More details are provided in the Supporting Information. All classical MD simulations were performed using the Amber11 suite of programs.[26] Three 13 bp d[GCGCGCXCGCGCG]·d[CGCGCGKGCGCGC] double-stranded oligomers differing by the central base pair were generated *in silico* using the Nucleic Acid Builder module of AmberTools. The symbol X denotes either C, A, or G, whereas K is the α,β-unsaturated ketoaldehyde shown in Figure 1. For the nonstandard residue K, the bonded force field parameters were taken from the Generalized Amber Force Field (GAFF) and parm99 parameter sets. Atomic charges were assigned using the Restrained Electrostatic Potential (RESP) protocol[27] and are given in Figure S1 of the Supporting Information. In the absence of any experimentally resolved structure, an alternating poly(G-C) sequence was chosen to provide higher stability.

The three macromolecules were first neutralized using 23 potassium ions and then explicitly solvated using TIP3P water molecules with a buffer of 10 Å. Long range interactions were calculated with the particle mesh Ewald (PME) method. The SHAKE algorithm was used to constrain bonds involving hydrogen atoms. The temperature was controlled using Langevin dynamics with a collision frequency of 1 ps^{-1}. A time step of 1.5 fs was used, and snapshots were taken every 1.5 ps. A series of minimization and MD runs were performed to equilibrate the system (see the Supporting Information). A production run of 20 ns was then conducted for each duplex. Because our simulation lengths do not exceed 20 ns, the so-called Barcelona corrections (bsc0) were not applied. This simulation time was sufficient to identify key differences betweeen the conformational dynamics of the C-, A-, and G-containing duplexes.

The inter-base pair (or step) coordinates and the orthonormal frames attached to DNA bases were computed using 3DNA.[28] The normal vectors of the C and K planes were defined as follows: the normal of C was just the *z*-axis of its base-fixed frame as computed by 3DNA, the normal of K was the vector product $\mathbf{v}_1 \times \mathbf{v}_2$ with $\mathbf{v}_1 = \mathbf{r}(C2) - \mathbf{r}(C4)$ and $\mathbf{v}_2 = \mathbf{r}(C1) - \mathbf{r}(C4)$. The magnitude and direction of global bending were measured essentially as described previously.[29] Briefly, the two B-DNA helices flanking the lesion were each represented by one effective right-handed, orthonormal frame. The helix frame was obtained as the average of base-fixed frames of six bases chosen in the helix (bases in base pairs 2−4 in the first helix and base pairs 10−12 in the second). The bending magnitude was defined as the angle between the *z*-axes of the two helix frames, and the direction was measured with respect to a middle frame. The middle frame, in turn, was computed as the average of frames fixed to bases in the two Watson−Crick (WC) pairs surrounding the lesion (base pairs 6 and 8). The direction of 0° indicates bending into the major groove at the location of the lesion, and the direction of 180° means bending into the minor groove.

Auxiliary DFT calculations were performed without any constraints from the B-DNA environment to quantify the optimal geometry and strength of the loose, non-Watson−Crick hydrogen bonding between K and X (X = C, A, or G). DFT has proven its reliability in that respect: we follow the procedure calibrated by Bickelhaupt and co-workers, relying on the BP86 functional, combining Becke's 88 exchange with Perdew's 86 correlation functional.[30,31] The triple-ζ Dunning and cc-pVTZ basis sets were used alongside. At that level of theory, a noncovalent interaction (NCI) analysis[32] is conducted to visualize the interactions taking place between K and X. We

dx.doi.org/10.1021/bi401268q | *Biochemistry* 2013, 52, 8115−8125

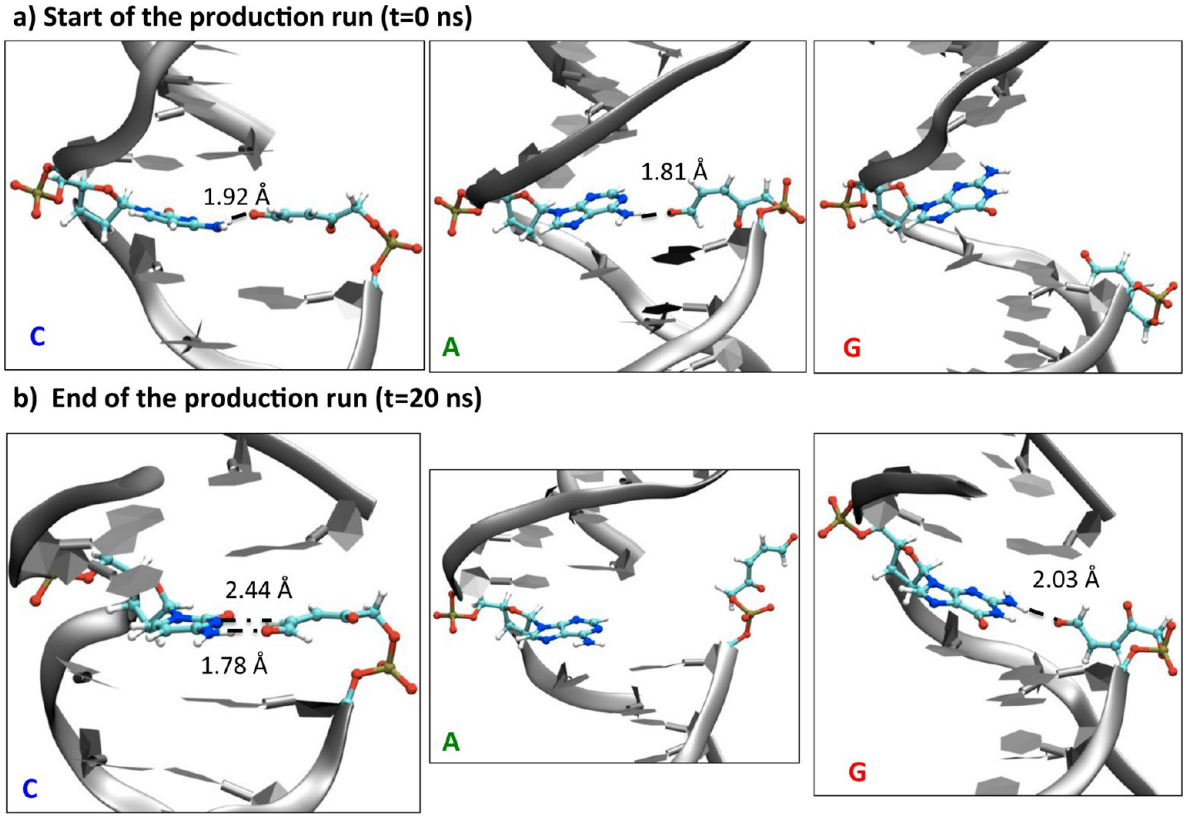


Figure 3. Cartoon representations of the X:K site and the surrounding bases in the DNA duplex at the beginning (a) and end (b) of the 20 ns MD trajectories.

recall that this method is based on the electronic density gradient and Hessian principal values and allows the pictorial identification of dispersive interactions and π-stacking (green), hydrogen bonds (blue core), or steric clash (red). The π-stacking was quantified using dispersion-corrected DFT with Grimme's D2 correction (see the Supporting Information).

RESULTS AND DISCUSSION

Intrinsic X:K Association and Attenuation of π-Stacking. The destabilization of double-stranded DNA (ds-DNA) upon C4′-AP formation arises from three main contributions: (i) the loss of π-stacking originally present between intact base B and the surrounding bases, (ii) a disruption of the WC pairing between B and X, and (iii) 3′ strand cleavage. The first two interactions are weakened yet not completely lost, as K will seek stabilization by hydrogen bonding with the opposite strand or by π-stacking with proximal nucleobases. Any preference in the interactions that K can develop with X (X = C, A, or G) is likely to be reflected in the trajectories of the damaged oligonucleotides.

The three fully optimized X:K geometries, obtained at the DFT level of theory, are depicted in the top row of Figure 2. They correspond to an optimal interaction of the two isolated moieties, in the absence of any constraints from the B-DNA environment. The strongest association with K accounts for −10.0, −7.7, and −11.8 kcal/mol for cytosine, adenine, and guanine, respectively (Table S1 of the Supporting Information provides the full energetic characterization). Each of the three lowest-energy structures features two hydrogen bonds (HBs). The HB distances are shorter for guanine, in line with the fact that isolated K interacts more strongly with G. The G:K structure has a HB between the terminal oxygen of K (O1) and the hydrogen of the imine group of G (H1), with a length of 1.88 Å, and a second interaction, a weaker (2.17 Å) but cooperative one, between the oxygen of guanine (O6) and the acidic hydrogen in α of the ketoaldehyde (H2). The NCI analysis allows us to highlight the interactions taking place (isodensities in Figure 2), with a central core for the O1···H1 HB (blue) that appears as the strongest hydrogen bond of the G:K entity. We note that one hydrogen of the amino group (H2) also contributes to the O1···H1 HB, which is thus reinforced.

The association of K with cytosine or adenine differs because of the absence of the imine hydrogen; instead, the HB to O1(K) takes place via amino hydrogen H4(C) or H6(A), leading to slightly longer distances: 1.92 and 1.94 Å for C and A, respectively. The same second HB between N3(C) or N1(A) and H2(K) is involved. Notice that the C:K hydrogen bonding takes place between H2(K) and N3(C), and not between H2(K) and O2(C).

The bottom row of Figure 2 depicts alternative structures obtained by rotating K. The interaction with H2 of K is disrupted, and an interaction with H1(K) takes place instead. This results in associations that are found to be 0.5−2.5 kcal/mol weaker in energy. Both interaction modes come into play during the structural dynamics, and this will guide our analysis in the rest of the paper. Hence, the affinity of K for nucleobases featuring a -N=C(NH_2)- motif can be explained in terms of HB partnering. It is worth noting that the interaction energies ($\Delta E_{X:K}$) are considerably weakened compared to that of the original WC pairing of −25 kcal/mol for the C:G base pair and −12 kcal/mol for the A:T base pair.[31] Hence, $\Delta E_{X:K}$ is destabilized roughly by a factor of 2 by the B → K mutation.

dx.doi.org/10.1021/bi401268q | *Biochemistry* 2013, 52, 8115−8125

The loss of the native π-stacking between B and the two adjacent guanines [G19 and G21 (Figure 1)] also contributes to the destabilization of the regular helical pattern. This structural deviation will be further analyzed with standard B-DNA angles below. We evaluate the energy penalty based on several representative orientations of the X:K systems: they were not performed arbitrarily but correspond to motions observed in the MD simulations (see supra). London forces between B and G19 and G21 versus those between K and G19 and G21 account for approximately −19 kcal/mol for GGG and approximately −12 kcal/mol for GTG and GCG (at the DFT-D2 level of theory). The geometries of the three nucleobases (G···B···G and G···K···G) were extracted from the initial structure of B-DNA and capped with hydrogens to saturate the glycosidic bond. The energies of the van der Waals interactions decrease to −9.3 kcal/mol for the GKG moiety, on a geometry common to the three AP duplexes. Of course, the weaker π-stacking arises from a net loss of atoms. This initial decrease is comparable to the loss of WC pairing and contributes to the destabilization of the X:K duplexes. We shall see hereafter that during the MD simulations, X can penetrate into the B-DNA helix and adopt a more favorable conformation that also partially restores stacking interactions.

We have identified two contributions that trigger an alteration of the initially regular B-DNA helix. The 3′ strand break along the phosphodiester linkage constitutes a third factor that can be expected to further compromise the stability and helicity of the AP duplex. The impact of a strand cleavage within an oligonucleotide has been investigated by MD studies[24] and found to be highly sequence-dependent. Furthermore, it is certainly amplified here by the loss of a nucleobase. We will now explore the structural consequences of the three destabilizing factors in the AP-containing DNA duplexes.

C4′-AP Site within B-DNA. Figure 3a shows structures of the C4′-AP site within the B-DNA duplex at the end of the equilibration procedure. Their inspection reveals that cytosine and adenine both adopt an orientation favoring a stabilizing interaction with K. During equilibration, a hydrogen bond between O1(K) and the amino hydrogen of C or A is formed with a distance that rapidly fluctuates around 1.92 Å for cytosine and 1.81 Å for adenine. The first value is equal to that inferred by DFT (Figure 2); the other one is even shorter. In contrast, G is a unique case as the B-DNA conformation cannot natively accommodate any HB with ketoaldehyde K, probably because of the initial repulsion between O1 and the oxygen of guanine. This explains why the G:K duplex already experiences a major strand opening at the end of the equilibration run.

These structures strongly suggest that the intrinsic stability order G > C > A for the X:K pairs (Figure 2) no longer holds and can be dramatically altered once K is embedded in a B-DNA environment. [Additional MD simulations performed on the K and C fragments immersed in a water box (not reported here) conserve the intrinsic affinity.] Yet K is flexible even within B-DNA, and a central point is the possibility of accommodating a more favorable X:K association after relaxation of the damaged duplex. This urges us to analyze the time evolution of the X:K-containing duplexes. They inevitably readapt, seeking intermolecular stabilization to fill the gap at the AP site. The mechanisms by which such AP sites readapt within B-DNA have been identified in the absence of a strand break:[19] purines and notably adenine were shown to present a higher affinity for oxidized C4′-AP sites. We will outline key differences between the structural consequences of the continuous versus broken strand, showing in turn how the strand scission impacts the X:K interaction.

A global measure of the stability of our AP-containing oligonucleotides with a strand break is the root-mean-square deviation (rmsd) computed for the backbone atoms with respect to the starting conformation, built as an ideal B-DNA structure so that the reference backbone is the same for the three systems. The rmsd time evolutions are plotted in Figure 4; they show large fluctuations for purines [X = A (green), and X = G (red)], but the most flexible structure corresponds to cytosine (blue). The rmsd values all lie above 2.0 Å. A significant distortion takes place, as a result of the three destabilizing factors we identified in the previous paragraph (DFT exploration). Clearly, the three X:K-containing duplexes deviate from the ideal B-DNA helical pattern. On the basis of the rmsd, our simulations show a much enhanced flexibility compared to analogous simulations of AP-containing duplexes for which no single-strand break is implied.[19]

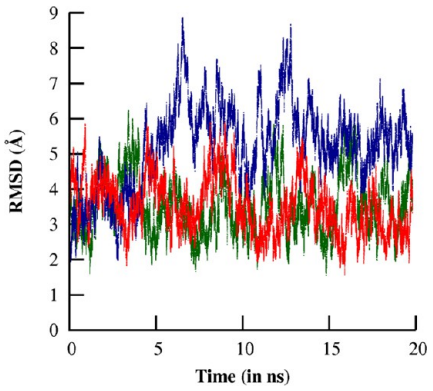


Figure 4. Time evolution of the backbone rmsd for the three X:K duplexes, computed with respect to the starting canonical B-DNA helix. The color code adopted for X is as follows: green for adenine, blue for cytosine, and red for guanine. Cytosine exhibits a higher rmsd.

It is clear that the B → K mutation is associated with major structural deviations. It then becomes essential to assess the locality of the structural perturbation caused by the strand-broken C4′-AP site. A strong hint pointing to the local character of the C4′-AP ds-DNA distortion is given by the time evolution of the rmsd computed for the two subhelices on both sides of the abasic site (Figure S2 of the Supporting Information), which remain small and stable; it is confirmed by monitoring the 36 Watson−Crick HBs that are extremely stable even for the C6:G21 and C8:G19 base pairs sandwiching the C:K moiety. These data indicate that the two $(GC)_6$ helices flanking the lesion are not affected by the fragmentation of base B.

The Flipping of K Stabilizes a Near-Attack C:K Complex in B-DNA. The structural inspection of the X:K-containing duplexes in our 20 ns trajectories indicates flipping fluctuations of K around the glycosidic bond that occur for all three opposite nucleobases. This motion is observed spontaneously, which suggests that the free energy of flipping for K is low compared to those of other lesions.[33] Obviously, the flipping of K tends to move it away from nucleobase X. In the rest of this study, the separating distance $d_{X\leftrightarrow K}$ will denote the distance between the nitrogen of the amino group of X (N4 for C, N6 for A, and N2 for G) and the C1 atom of ketoaldeyde K

dx.doi.org/10.1021/bi401268q | Biochemistry 2013, 52, 8115−8125

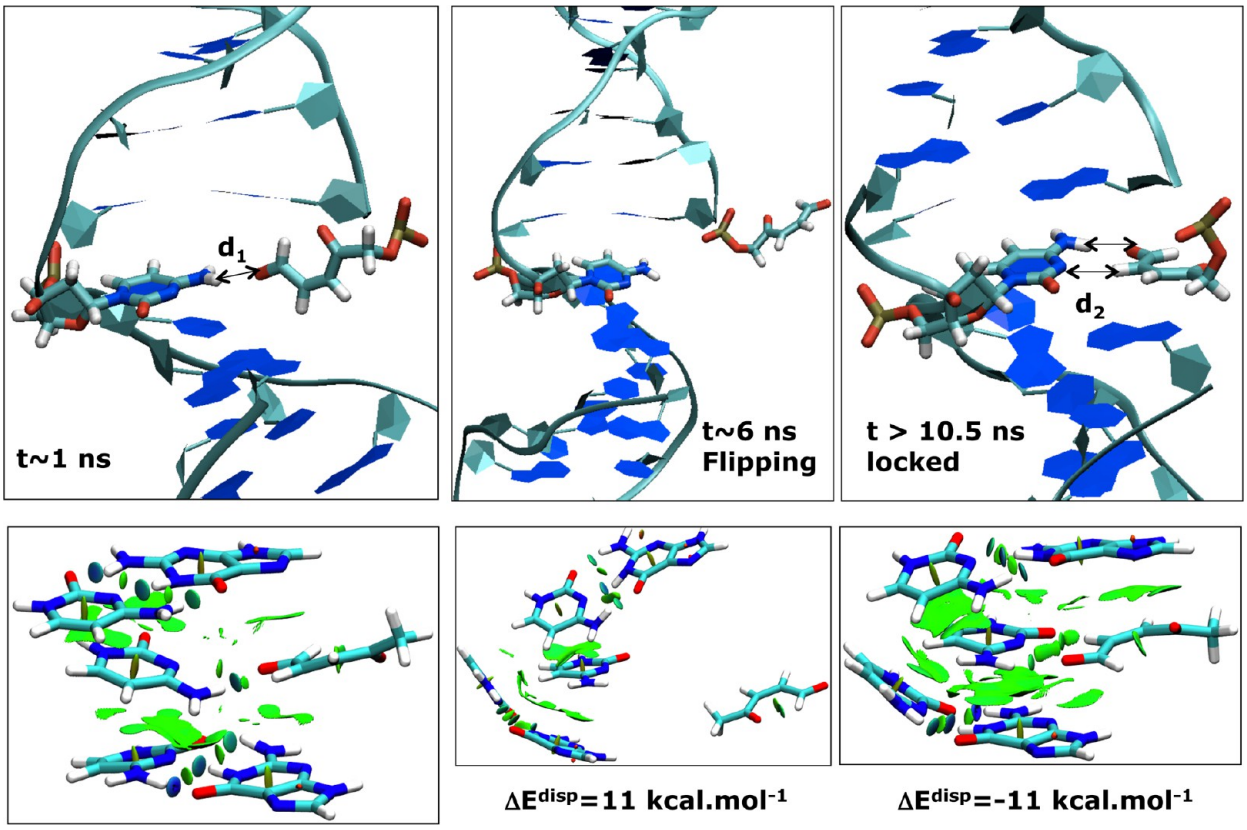


Figure 5. Cartoon representations for three snapshots (extracted at 1, 6, and 20 ns) of the C:K pairing along the MD trajectory (top). Dispersion energies, visualized in green and evaluated for the three structures (bottom).

(Figure S3 of the Supporting Information). This choice is the most relevant for gaining insight into the formation of a linear aminol, for which this distance is naturally the unidimensional reaction coordinate.

We first discuss the results obtained for cytosine (X = C). A lower bound for $d_{C\leftrightarrow K}$ is 3.41 Å, which corresponds to the optimal association featuring a 1.92 Å hydrogen bond between H4(C) and O1(K) (Figure 2). The time evolution of $d_{C\leftrightarrow K}$ (Figure S3 of the Supporting Information) shows that the distance is globally stable up to 4.3 ns, fluctuating around 4.0 Å. It then progressively increases, reaching values of >10 Å. This corresponds to a flipping of K around its glycosidic bond. Interestingly, the flipping is not accompanied by water intrusion, a common feature of other AP sites.[19] At 10.5 ns, the ketoaldehyde restores its pairing with cytosine, but now involving two hydrogen bonds. In addition to the HB between H4(C) and O1(K), a second HB forms betwen N3 of C and H2 of the aldehyde, just as in the optimal interaction mode (Figure 2). Representative structures shown in Figure 5 illustrate the initial arrangement (left), the transient flipping intermediate (middle), and the resulting stable structure (right). The time evolutions of the HB distances and angles shown in Figure 6 confirm not only stable HB lengths but also favorable angular positioning. Once the second hydrogen bond is formed, K and C remain almost coplanar, as demonstrated by the angle between the K and C planes plotted in Figure 6. The tight pairing between C and K is reflected by the probability distribution of $d_{C\leftrightarrow K}$ (Figure 7) that exhibits a sharp and neat peak centered at ∼4 Å. The B-DNA helix hence can accommodate a near-optimal pairing between C and K, very close to the intrinsic association inferred by DFT in the absence of any constraints.

The $d_{C7\leftrightarrow K}$ distance measures a horizontal separation, while one can also monitor another distance, $d_{G19\leftrightarrow K}$, to account for the strand cleavage [vertical separation, G19 is the 5′ neighbor of K (Figure 1)]. At the end of the trajectory, the latter distance fluctuates around 4.0 Å, the reference value for undamaged oligonucleotides. This demonstrates a stable stacked arrangement between K and G19. The stability of the C:K duplex is further enhanced by a transverse stacking among C6, C7, and G21 that takes place for >10.5 ns (Figure 5). These stacking interactions imply a cooperative stabilization of −11 kcal/mol, in addition to the two HBs that bring a similar increment (Figure 5). Moreover, the C7 base penetrates deeper toward the minor groove; this tends to further lock the structure by preventing the intrusion of water.

Taken together, there is no ambiguity about whether K and C can adopt a near-attack conformation[34] that then favors the multistep formation of the interstrand cross-link dCyd341. On the basis of our results, we propose that the strand scission brings enough flexibility to fill the gap between K and X, even for the less bulky cytosine, and establish a stable structure favorable for the reaction.

Effects of the AP Site on DNA Structure and Bending. The mean values of the inter-base pair coordinates tilt, roll, twist, shift, slide, and rise reveal local conformational changes upon flipping of K and creation of the stable C:K pair (Figure 8). The two WC base pairs surrounding the lesion (base pairs 6 and 8) are treated as parts of one big base pair step, the 6/8 step. At the beginning (black lines in Figure 8), the oligomer retains its B-DNA structure, with the lesion-containing 6/8 step

dx.doi.org/10.1021/bi401268q | Biochemistry 2013, 52, 8115−8125

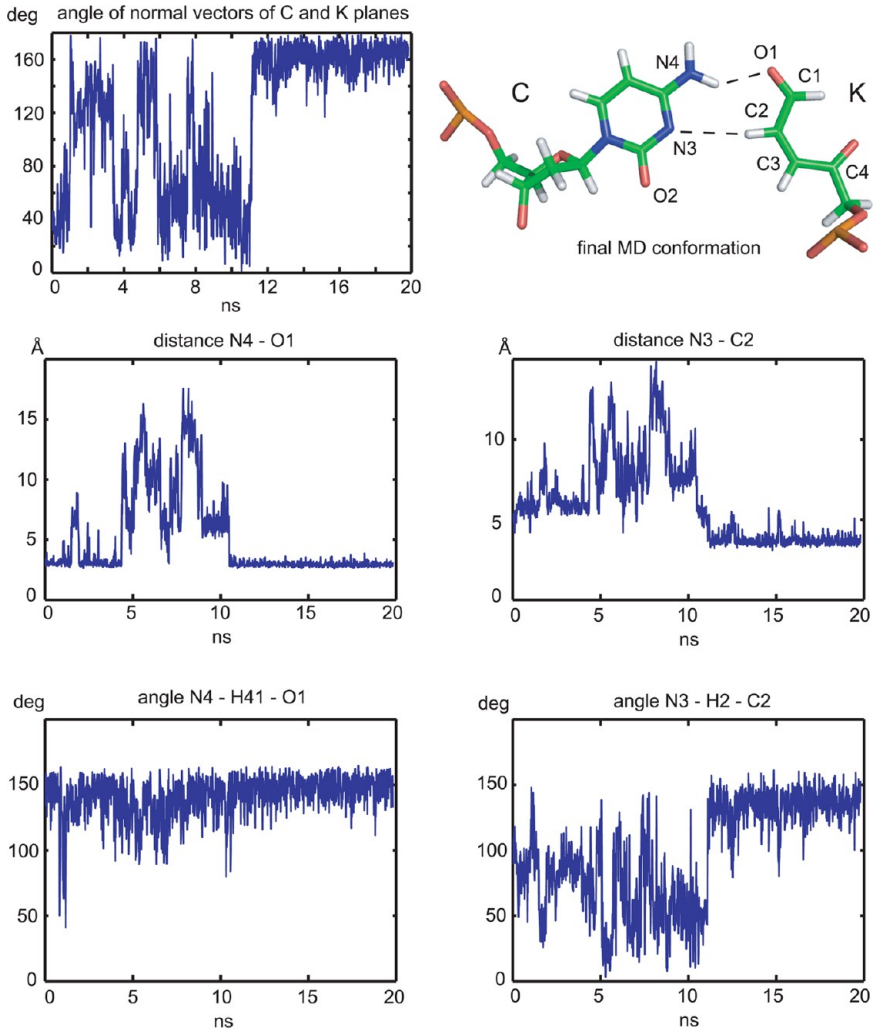


Figure 6. Time evolution of angles and distances characterizing the C:K pair. The data clearly indicate structural stabilization after 10.5 ns of the simulation by two hydrogen bonds with favorable distances and angles. Furthermore, the C and K moieties become almost coplanar.

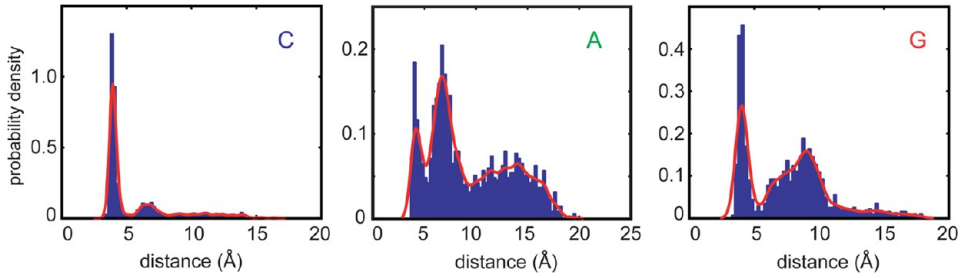


Figure 7. Probability distribution of distance $d_{X\leftrightarrow K}$. Red lines show Gaussian kernel density estimations. Probabilities for distance intervals are listed in Table S2 of the Supporting Information.

characterized by values of small tilt, shift, and slide, moderate positive roll, and values of twist (72°) and rise (7 Å) corresponding to two consecutive, regular base pair steps. After the flip and upon formation of the stable structure, however, the local conformation changes dramatically (green lines in Figure 8). In particular, the 6/8 step adopts high positive roll (28°) and negative tilt (−15°), indicating a sharp localized bend predominantly into the major groove at the point of the lesion. Furthermore, the 6/8 step is underwound (twist of 30°, roughly half of two regular base pair steps), and the two WC pairs are displaced laterally (shift of −1.5 Å and slide of 2.2 Å). On the other hand, rise does not change upon the lesion flipping and retains its value of two consecutive regular steps (7 Å), indicating stable stacking interactions around the lesion.

The local structural changes are reflected in global bending of the whole oligomer, measured as described in Materials and Methods. After the lesion has flipped, the bending fluctuates around the magnitude of 32° and the direction of 15°. This direction indicates that the oligomer is bent into the major groove at the lesion site.

dx.doi.org/10.1021/bi401268q | *Biochemistry* 2013, 52, 8115−8125

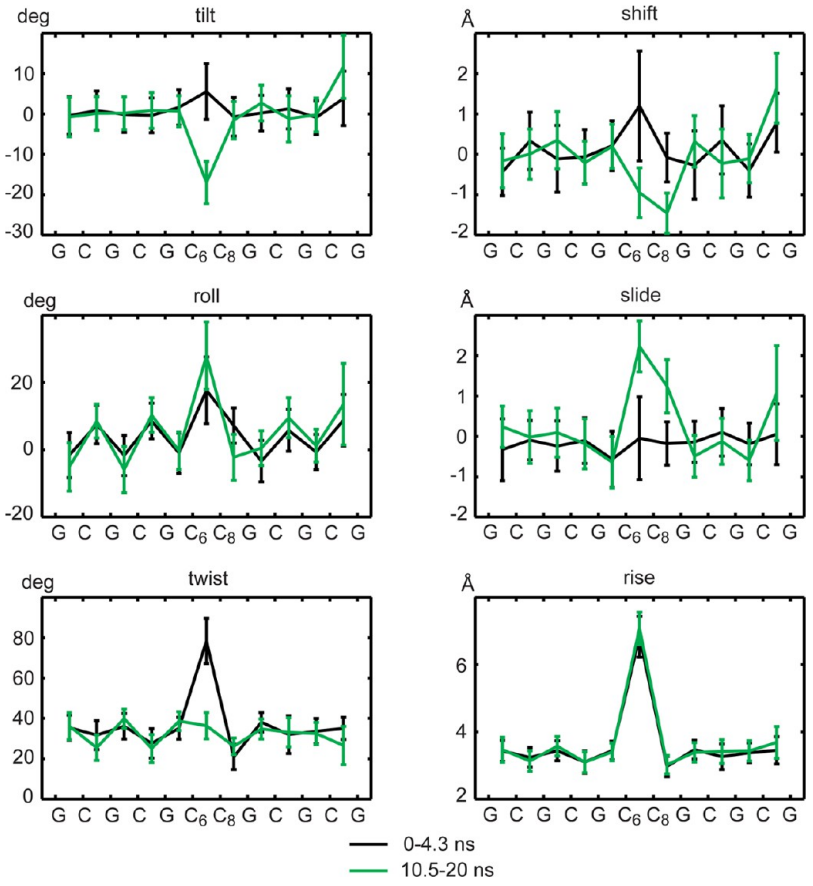


Figure 8. Mean inter-base pair (or step) coordinates from the C:K trajectory. Values for the initial portion (black) and for the final part after the lesion K has flipped and formed a stable C:K pair (green) are shown. Error bars are mean differences between the full trajectory and its halves.

Duplexes Containing Purine:K Sites. We now turn to the case of purine:K duplexes. For both X = A or G, rmsds (Figure 4) are lower, even though their association with K is less favored and significantly disrupted. The main reason is that the dramatic flip of K, which leads to a stable C:K pair, does not take place in A:K and G:K duplexes.

For adenine, it is clear from the probability distribution and time evolution of $d_{A\leftrightarrow K}$ (Figure 7 and Figure S3 of the Supporting Information) that no interaction is stabilizing the structure. The optimal interaction in the absence of B-DNA was found to be significantly weaker for A:K (7 kcal/mol) than for C:K [10 kcal/mol (Figure 2)]. The H6···O1 HB is initially present also for A:K embedded in B-DNA (Figure 3) but is rapidly disrupted and not sufficient to maintain a stable structure within the helix. Several times along the trajectory and notably toward the end, the duplex restores a close A:K pairing, but only transiently. This result for adenine is in constrast with what is known for AP sites that do not feature a strand scission: they present a higher affinity for purines.

Guanine is an interesting case. The intrinsic pairing with K is the strongest one, but it is lost in the B-DNA helix: already after the equilibration, the G:K duplex exhibits a marked vertical opening. It takes ∼6 ns for the system to decrease the $d_{G\leftrightarrow K}$ distance to the lower limit of ∼4 Å, hence restoring a more favorable G:K pairing (Figure S3 of the Supporting Information). The latter is driven by the formation of one intermolecular HB between O1(K) and H2(G), not present in the optimal structure. This HB stabilizes the G:K pair by approximately −8 to −9 kcal/mol and is maintained until 12 ns. In contrast, the HB between H2(K) and O6(G) identified in the optimal pairing is formed only marginally. As a consequence, the probability distribution of the distance $d_{G\leftrightarrow K}$ is more localized than for A:K but less than for C:K (Figure 7; see also Table S2 of the Supporting Information). The first sharp but weakly populated peak at ∼4 Å correponds to a 33% probability and is mostly due to the structures occurring between 6 and 12 ns. The majority of the population, however, corresponds to two larger peaks centered around 7 and 9 Å. A significant number of conformations (17%) show $d_{G\leftrightarrow K}$ distances of >10 Å and up to 19 Å (Table S2 of the Supporting Information). Such a separation reflects a total flipping of K out of the helix. At ∼19.5 ns, the system restores the tight G:K partnering, but it is maintained for <0.5 ns (see Figure 3 where a representative snapshot is shown in the bottom row). Thus, the B-DNA environment cannot accommodate the optimal G:K pairing. This is simply because of the different orientation of the amino group attached to C2 of G, whereas the amino group of C lies opposite the N-glycosidic bond. The alternative O1(K)···H2(G) hydrogen bond forms only transiently and cannot permanently stabilize the structure.

Overall, our results indicate that the B-DNA environment dictates the probability of positioning of the reactive conformers to form the oxidative interstrand cross-link, as already surmised for the case of intrastrand defects.[35] The cytosine:K interaction is the most stable one in the duplex and is capable of maintaining an intrahelical directionality. With purines, the disruption of pairing between X and K may still be compensated by the formation of a pairing with a strand

dx.doi.org/10.1021/bi401268q | *Biochemistry* 2013, 52, 8115−8125

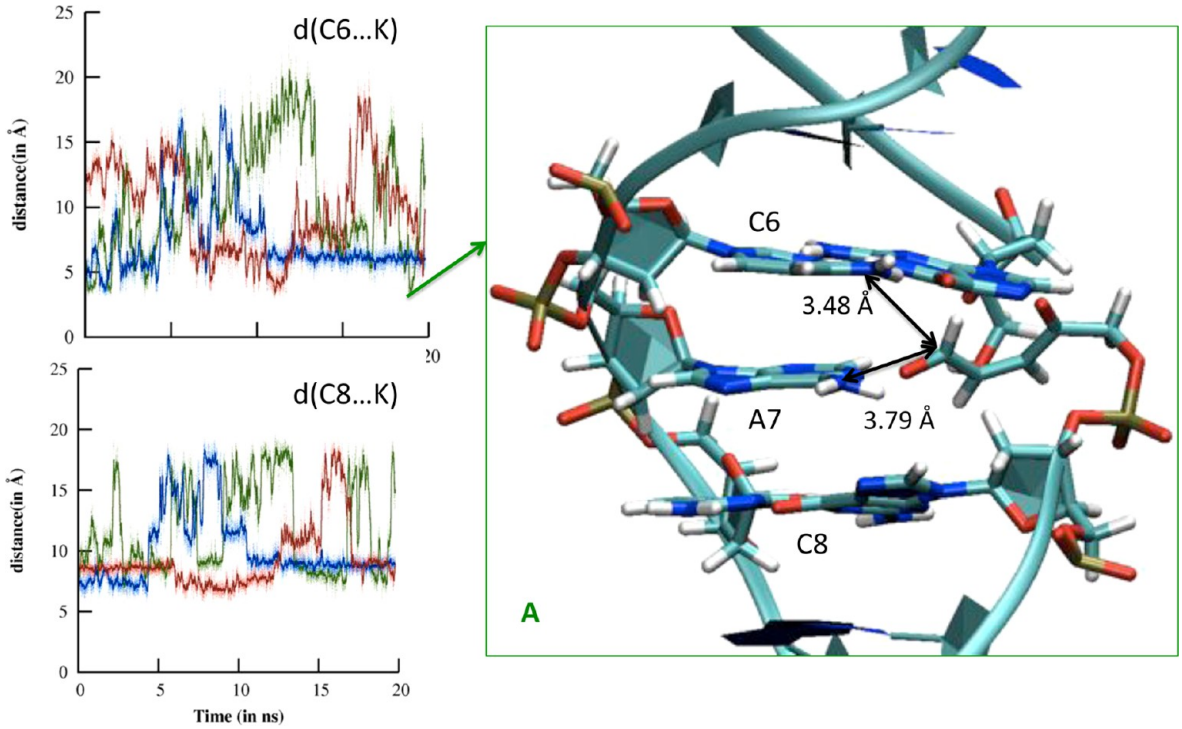


Figure 9. Time evolution of the $d_{C6\leftrightarrow K}$ and $d_{C8\leftrightarrow K}$ distances for duplexes containing C:K (blue), A:K (green), and G:K (red) base pairs. A cartoon representation of an MD snapshot taken around 19 ns is given to illustrate exceptions to the inequality (eq 1) observed in the A:K duplex.

offset, i.e., to the neighboring C6 or C8 bases. This might assist the formation of ICLs with a strand offset. We examine this possibility in the following section.

Possibility of ICL Formation with a Strand Offset. As the C4′-AP site is formed within the ds-DNA, it has been suggested experimentally that the attack (the first step being the formation of an aminol) may take place not with the orphan nucleobasis X but with nucleobases not initially paired with B.[9] This possibility depends in a crucial way on the conformation that the duplex, and hence the reactive partners, can adopt. We still lack more direct proof of the existence of these interactions (let us recall that such lesions are identified on the basis of tandem mass spectrometry, where information on the position of the ICL within B-DNA is lost upon enzymatic digestion). To obtain ICLs with an offset, a prerequisite is to have reactive conformers for the modified X:K base pair. To check this, we monitored distances $d_{C6\leftrightarrow K}$ and $d_{C8\leftrightarrow K}$ between K and its 5′ and 3′ neighbors in the opposite strand, C6 and C8 (Figure 9). Recall that the distances are measured between the C1 atom of K and the amino nitrogen of the nucleobases. For the starting structures in the canonical B-DNA conformation, the distances satisfy the inequality

$$d_{X\leftrightarrow K} < d_{C6\leftrightarrow K} < d_{C8\leftrightarrow K} \quad (1)$$

because of the inherent strand asymmetry. For instance, starting distances are 3.55, 4.05, and 5.61 Å for X = C. Transverse distances tend to be longer for purine:K duplexes: 3.55, 5.01, and 6.35 Å for X = A and 3.34, 6.45, and 6.60 Å for X = G. The C8···K interaction is *a priori* disfavored, but it is not clear whether the inequality is kept during MD, as K will seek to fill the gap within B-DNA. We found that for cytosine and guanine, the inequality holds over all the trajectories. Equivalently, no offset hydrogen bonds are formed. In contrast, the A:K duplex, which presents the largest structural deviations, shows $d_{C6\leftrightarrow K}$ distances that are slightly shorter (by 0.3 Å) than the $d_{A7\leftrightarrow K}$ distance between K and the opposite adenine, A7 (Figure 9).

The inversion of eq 1 may have a direct implication on reactivity: for two identical nucleobases X7 and X6, it is surmised that the ICL formed corresponds to the shortest approaching distance.[14] One can anticipate a close competition between C6 and A7, because (i) the free energy to develop an interaction between K and an offset nucleobase $X_{i\pm1}$ most likely can be assumed to be low and (ii) the cytosine is likely to be intrinsically more reactive than adenine.

Our simulations hence confirm the possible formation of transverse ICLs. To obtain a more quantititave understanding of this process, it would be necessary to test several sequences and to include an electronic description for K and the nucleobase acting as a bis-nucleophile.

■ CONCLUSIONS

In this work, we investigate structural properties of AP duplexes featuring a 3′ strand break using atomic-resolution, explicit solvent molecular dynamics (MD) simulations, complemented by quantum chemical DFT calculations on isolated X:K pairs (K denotes the lesion and X a nucleobase). Our results allow us to identify with no ambiguity a higher structural stability of the C:K pair embedded in the DNA duplex. The B-DNA helix can indeed accommodate the energetically most favorable C:K interaction with a pairing of two hydrogen bonds that lock the structure, which further benefits from transverse stacking interactions. This is in line with the experimental finding by Ravanat and co-workers that cytosine ICLs (dCyd341) are most preferentially formed.[8] In contrast, purines feature at most one hydrogen bond with K and flip constantly in and out.

Our analysis further shows that the stable C:K pair induces pronounced structural changes, namely a sharp bending into

dx.doi.org/10.1021/bi401268q | *Biochemistry* 2013, 52, 8115−8125

the major groove and a helix unwinding, which are localized to the site of the lesion. In contrast, the internal structure of the B-DNA helices flanking the lesion remains unperturbed by the damage.

In summary, our results strongly suggest that the behavior of an AP site within a DNA dublex is completely altered when accompanied by a single-strand break. MD simulations indicate the ease of formation of ICLs whose importance has been recently stressed. They provide the first insights into a complex reactivity that calls for reactive QM/MM−MD studies to obtain free energy profiles. Another question about ICL formation is the sequence dependence, presumably strong, which would deserve a dedicated study beyond the poly(GC) oligonucleotides studied here.

■ ASSOCIATED CONTENT

Supporting Information

Details of the computational procedure, Tables S1 and S2, and Figures S1−S3. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: filip.lankas@uochb.cas.cz. Phone: +420 220 410 319. Fax: +420 220 410 320.

*E-mail: elise.dumont@ens-lyon.fr. Phone: +33 (0)4 72 72 88 46. Fax: +33 (0) 4 72 72 88 60.

Author Contributions

C.P. and T.D. contributed equally to this work.

Funding

C.P. is grateful for a Ph.D. fellowship from the French Ministry of Research. This work was performed within the framework of the LABEX PRIMES (ANR-11-LABEX-0063) of Université de Lyon, within Program "Investissements d'Avenir" (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR). Calculations were performed using the local HPC resources of PSMN at ENS-Lyon. T.D. and F.L. were supported by Grant RVO61388963 provided by the Academy of Sciences of the Czech Republic.

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Dianov, G. L., Sleeth, K. M., Dianova, I. I., and Allinson, S. L. (2003) Repair of abasic sites in DNA. *Mutat. Res. 531*, 157−163.

(2) Lu, D., Silhan, J., MacDonald, J. T., Carpenter, E. P., Jensen, K., Tang, C. M., Baldwin, G. S., and Freemont, P. S. (2012) Structural basis for the recognition and cleavage of abasic DNA in Neisseria meningitidis. *Proc. Natl. Acad. Sci. U.S.A. 109*, 16852−16857.

(3) Obeid, S., Blatter, N., Kranaster, R., Schnur, A., Diederichs, K., Welte, W., and Marx, A. (2010) Replication through an abasic DNA lesion: Structural basis for adenine selectivity. *EMBO J. 29*, 1738−1747.

(4) Demple, B., and DeMott, M. S. (2012) Dynamics and diversions in base excision DNA repair of oxidized abasic lesions. *Oncogene 21*, 8926−8934.

(5) Zhang, P., Herbig, U., Coffman, F., and Lambert, M. W. (2013) Non-erythroid α spectrin prevents telomere dysfunction after DNA interstrand cross-link damage. *Nucleic Acids Res. 41*, 5321−5340.

(6) Kow, Y. W., Bao, G., Minesinger, B., Jinks-Robertson, S., Siede, W., Jiang, Y. L., and Greenberg, M. M. (2005) Mutagenic effects of abasic and oxidized abasic lesions in *Saccharomyces cerevisiae*. *Nucleic Acids Res. 33*, 6196−6202.

(7) Dutta, S., Chowdhury, G., and Gates, K. S. (2007) Interstrand Cross-Links Generated by Abasic Sites in Duplex DNA. *J. Am. Chem. Soc. 129*, 1852−1853.

(8) Regulus, P., Duroux, B., Bayle, P.-A., Favier, A., Cadet, J., and Ravanat, J.-L. (2007) Oxidation of the sugar moiety of DNA by ionizing radiation or bleomycin could induce the formation of a cluster DNA lesion. *Proc. Natl. Acad. Sci. U.S.A. 104*, 14032−14037.

(9) Sczepanski, J. T., Jacobs, A. C., Majumdar, A., and Greenberg, M. M. (2009) Scope and Mechanism of Interstrand Cross-Link Formation by the C4′-Oxidized Abasic Site. *J. Am. Chem. Soc. 131*, 11132−11139.

(10) Johnson, K. M., Price, N. E., Wang, J., Fekry, M. I., Dutta, S., Seiner, D. R., Wang, Y., and Gates, K. S. (2013) On the Formation and Properties of Interstrand DNA-DNA Cross-Links Forged by Reaction of an Abasic Site with the Opposing Guanine Residue of 5′-CAp Sequences in Duplex DNA. *J. Am. Chem. Soc. 135*, 1015−1025.

(11) Frelon, S., Douki, T., Ravanat, J.-L., Pouget, J.-P., Tornabene, C., and Cadet, J. (2000) High-Performance Liquid Chromatography-Tandem Mass Spectrometry Measurement of Radiation-Induced Base Damage to Isolated and Cellular DNA. *Chem. Res. Toxicol. 13*, 1002−1010.

(12) Dupont, C., Patel, C., Ravanat, J. L., and Dumont, E. (2013) Addressing the competitive formation of tandem DNA lesions by a nucleobase peroxyl radical: A DFT-D screening. *Org. Biomol. Chem. 11*, 3038−3045.

(13) Patel, C., Garrec, J., Dupont, C., and Dumont, E. (2013) What Singles Out the G[8−5]C Intrastrand DNA Cross-Link? Mechanistic and Structural Insights from Quantum Mechanics/Molecular Mechanics Simulations. *Biochemistry 52*, 425−431.

(14) Labet, V., Morell, C., Grand, A., Cadet, J., Cimino, P., and Barone, V. (2008) Formation of cross-linked adducts between guanine and thymine mediated by hydroxyl radical and one-electron oxidation: A theoretical study. *Org. Biomol. Chem. 6*, 3300−3305.

(15) Bellon, S., Ravanat, J.-L., Gasparutto, D., and Cadet, J. (2002) Cross-Linked Thymine-Purine Base Tandem Lesions: Synthesis, Characterization, and Measurement in γ-Irradiated Isolated DNA. *Chem. Res. Toxicol. 15*, 598-606,

(16) Hong, I. S., Carter, K. N., Sato, K., and Greenberg, M. M. (2007) Characterization and Mechanism of Formation of Tandem Lesions in DNA by a Nucleobase Peroxyl Radical. *J. Am. Chem. Soc. 129*, 4089−4098.

(17) Stevens, K., Claeys, D. D., Catak, S., Figaroli, S., Hocek, M., Tromp, J. M., Schurch, S., Van Speybroeck, V., and Madder, A. (2011) Furan-Oxidation-Triggered Inducible DNA Cross-Linking: Acyclic Versus Cyclic Furan-Containing Building Blocks—On the Benefit of Restoring the Cyclic Sugar Backbone. *Chem.—Eur. J. 17*, 6940−6953.

(18) Sviatenko, L., Gorb, L., Hovorun, D., and Leszczynski, J. (2012) Interaction of 2-Deoxyadenosine with cis-2-Butene-1,4-dial: Computational Approach to Analysis of Multistep Chemical Reactions. *J. Phys. Chem. A 116*, 2333−2342.

(19) Barsky, D., Foloppe, N., Ahmadia, S., Wilson, D. M., III, and MacKerell, A. D., Jr. (2000) New insights into the structure of abasic DNA from molecular dynamics simulations. *Nucleic Acids Res. 28*, 2613−2626.

(20) Chen, J., Dupradeau, F.-Y., Case, D. A., Turner, C. J., and Stubbe, J. (2008) DNA oligonucleotides with A, T, G or C opposite an abasic site: Structure and dynamics. *Nucleic Acids Res. 36*, 253−262.

(21) Pang, Y., Xu, Z., Sato, Y., Nishizawa, S., and Teramae, N. (2012) Base Pairing at the Abasic Site in DNA Duplexes and Its Application in Adenosine Aptasensors. *ChemBioChem. 13*, 436−442.

(22) Cuniasse, P., Sowers, L., Eritja, R., Kaplan, B., Goodman, M., Cognet, J., LeBret, M., Guschlbauer, W., and Fazakerley, G. (1987) An abasic site in DNA. Solution conformation determined by proton NMR and molecular mechanics calculations. *Nucleic Acids Res. 15*, 8003−8022.

(23) Hazel, R. D., and de los Santos, C. (2010) NMR Solution Structures of Clustered Abasic Site Lesions in DNA: Structural Differences between 3-Staggered and 5-Staggered Bistranded Lesions. *Biochemistry 49*, 8978−8987.

(24) Ayadi, L., Coulombeau, C., and Lavery, R. (1999) Abasic sites in duplex DNA: Molecular modeling of sequence-dependent effects on conformation. *Biophys. J. 77*, 3218−3226.
(25) Radzimanowski, J., Dehez, F., Round, A., Bidon-Chanal, A., McSweeney, S., and Timmins, J. (2013) An "open" structure of the RecOR complex supports ssDNA binding within the core of the complex. *Nucleic Acids Res. 41*, 7972−7986.
(26) Case, D. A., et al. *Amber 11*, University of California, San Francisco.
(27) Cieplak, P., Cornell, W. D., Bayly, C., and Kollman, P. A. (1995) Application of the multimolecule and multiconformational RESP methodology to biopolymers: Charge derivation for DNA, RNA, and proteins. *J. Comput. Chem. 16*, 1357−1377.
(28) Lu, X.-J., and Olson, W. K. (2008) 3DNA: A versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures. *Nat. Protoc. 3*, 1213−1227.
(29) Lankas, F., Spackova, N., Moakher, M., Enkhbayar, P., and Sponer, J. (2010) A measure of bending in nucleic acids structures applied to A-tract DNA. *Nucleic Acids Res. 38*, 3414−3422.
(30) Fonseca Guerra, C., and Bickelhaupt, F. M. (2002) Orbital Interactions in Strong and Weak Hydrogen Bonds are Essential for DNA Replication. *Angew. Chem., Int. Ed. 41*, 2092−2095.
(31) van der Wijst, T., Guerra, C. F., Swart, M., and Bickelhaupt, F. M. (2006) Performance of various density functionals for the hydrogen bonds in DNA base pairs. *Chem. Phys. Lett. 426*, 415−421.
(32) Johnson, E. R., Keinan, S., Mori-Sánchez, P., Contreras-García, J., Cohen, A. J., and Yang, W. (2010) Revealing Noncovalent Interactions. *J. Am. Chem. Soc. 132*, 6498−6506.
(33) Bianchi, C., and Zangi, R. (2013) Base-Flipping Propensities of Unmethylated, Hemimethylated, and Fully Methylated CpG Sites. *J. Phys. Chem. B 117*, 2348−2358.
(34) Bruice, T. C., and Lightstone, F. C. (1999) Ground State and Transition State Contributions to the Rates of Intramolecular and Enzymatic Reactions. *Acc. Chem. Res. 32*, 127−136.
(35) Bellon, S., Ravanat, J.-L., Gasparutto, D., and Cadet, J. (2002) Cross-Linked Thymine-Purine Base Tandem Lesions: Synthesis, Characterization, and Measurement in γ-Irradiated Isolated DNA. *Chem. Res. Toxicol. 15*, 598−606.