

# Nuclear Magnetic Resonance Structural Studies and Molecular Modeling of Duplex DNA Containing Normal and 4'-Oxidized Abasic Sites<sup>†,‡</sup>

Jingyang Chen,<sup>§,||</sup> François-Yves Dupradeau,<sup>⊥,@</sup> David A. Case,<sup>⊥</sup> Christopher J. Turner,<sup>#</sup> and JoAnne Stubbe<sup>\*,§,+</sup>

Departments of Chemistry and Biology and Francis Bitter Magnet Laboratory, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, and Department of Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037

Received November 22, 2006; Revised Manuscript Received January 20, 2007

**ABSTRACT:** A 4'-oxidized abasic site (**X**) has been synthesized in a defined duplex DNA sequence, 5'-d(CCAAAG**X**ACCGGG)-3'/3'-d(GGTTTCATGGCCC)-5' (**1**). Its structure has been determined by two-dimensional NMR methods, molecular modeling, and molecular dynamics simulations. **1** is globally B-form with the base (A) opposite **X** intrahelical and well-stacked. Only the  $\alpha$  anomer of **X** is observed, and the abasic site deoxyribose is largely intrahelical. These results are compared with a normal abasic site (**Y**) in the same sequence context (**2**). **Y** is composed of a 60:40 mixture of  $\alpha$  and  $\beta$  anomers (**2 $\alpha$**  and **2 $\beta$** ). In both **2 $\alpha$**  and **2 $\beta$** , the base (A) opposite **Y** is intrahelical and well-stacked and the abasic site deoxyribose is predominantly extrahelical, consistent with the reported structures of the normal abasic site in a similar sequence context [Hoehn, S. T., Turner, C. J., and Stubbe, J. (2001) *Nucleic Acids Res.* 29, 3413–3423]. Molecular dynamics simulations reveal that the normal abasic site appears to be conformationally more flexible than the 4'-oxidized abasic site. The importance of the structure and flexibility of the abasic site in the recognition by the DNA repair enzyme Ape1 is discussed.

Bleomycins (BLMs)<sup>1</sup> are clinically important chemotherapeutic agents presently used for the treatment of a variety of cancers (1–3). Their toxicity is thought to cause the death of cancer cells by initial generation of both single-stranded (ss) and double-stranded (ds) DNA lesions (4–6). Cleavage of DNA requires the presence of both ferrous iron and O<sub>2</sub>, which generates the “activated BLM” that is responsible for the oxidative damage to the deoxyribose sugar of DNA (7–9). The sugar damage is initiated by abstraction of a 4'-H atom from a pyrimidine 3' to a guanine (10). The resulting

4'-radical intermediate can lead to two types of DNA damage products. O<sub>2</sub>-rich conditions lead to direct strand cleavage generating a gapped 3'-phosphoglycolate/5'-phosphate lesion (3'-PG/5'-P, Scheme 1) (11, 12) and a base propenal; O<sub>2</sub>-depleted conditions lead to an intact strand with a 4'-oxidized abasic site (Scheme 1) and a free nucleic acid base (13, 14). The knowledge of how these lesions are recognized and repaired in vivo is essential to understanding the cytotoxicity and drug resistance of BLM in cancer cells.

BLM-induced DNA damage belongs to a class of DNA lesions initiated by oxidative damage to the deoxyribose backbone. Pulse radiolysis studies on DNA in the presence of O<sub>2</sub> in the 1980s demonstrated that 80–90% of the DNA damage is associated with modified nucleic acid bases, while 10–20% is associated with modified deoxyribose moieties (15, 16). In the former case, excellent progress has been made in elucidation of the lesion structures in duplex DNAs and the mechanism of their repair (17, 18). In the latter case, products that result from initial hydrogen atom abstraction from each of the five carbons of the deoxyribose have been identified (4, 19). However, their structures within a DNA duplex and their mechanism of repair have remained less tractable. Only recently have analytical methods for analyzing the prevalence of each of these lesions in vivo under different growth conditions become available (20, 21).

Besides the 4'-oxidized abasic sites and the 3'-PG/5'-P lesion generated by the BLMs, enediynes (22, 23), hydroxyl radicals (15), and ionizing radiation, another physiologically prevalent lesion is the 2-deoxyribonolactone generated by oxidative damage to C-1' (Scheme 1) (15, 24–27). In all three cases, studies have shown that if these lesions are left unrepaired, they give rise to mutations and, ultimately, cell death (17, 28).

<sup>†</sup> This work is supported by NIH Grants GM 34454 to J.S. and GM 45811 to D.A.C., and the NMR facility is supported by NIH Grant RR-00995. F.-Y.D. was funded by the French Ministère de l'Éducation Nationale et de la Recherche.

<sup>‡</sup> Coordinates and NMR restraint files have been deposited in the Protein Data Bank (PDB entries 2HOU, 2HPX, 2HSK, 2HSL, 2HSR, and 2HSS).

\* To whom correspondence should be addressed. Telephone: (617) 253 1814. Fax: (617) 258-7247. E-mail: stubbe@mit.edu.

<sup>§</sup> Department of Chemistry, Massachusetts Institute of Technology.

<sup>||</sup> Present address: Department of Molecular Biology, Massachusetts General Hospital, 185 Cambridge St., Boston, MA 02114.

<sup>⊥</sup> The Scripps Research Institute.

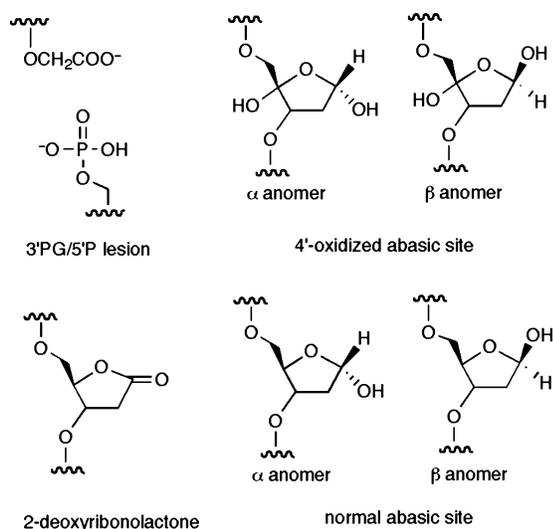
<sup>@</sup> Present address: DMAG, EA 3901, Faculté de Pharmacie, Université de Picardie Jules Verne, 1-3, rue des Louvels, 80037 Amiens Cedex 1, France.

<sup>#</sup> Francis Bitter Magnet Laboratory, Massachusetts Institute of Technology.

<sup>+</sup> Department of Biology, Massachusetts Institute of Technology.

<sup>1</sup> Abbreviations: BLM, bleomycin; ss, single-stranded; ds, double-stranded; 3'-PG/5'-P, 3'-phosphoglycolate/5'-phosphate; BER, base-excision repair; HPLC, high-performance liquid chromatography; 4'-N<sub>3</sub>-dU, 4'-azido-2'-deoxyuridine; UDG, uracil-DNA glycosylase; HIV-1 RT, HIV-1 reverse transcriptase; TSP, sodium 3-(trimethylsilyl)-1-propanesulfonate; NOE, nuclear Overhauser effect; RESP, restrained electrostatic potential; MEP, molecular electrostatic potential; THF, tetrahydrofuran; pol $\beta$ , DNA polymerase  $\beta$ .

Scheme 1



In humans, the 4'-oxidized abasic sites, the 3'-PG/5'-P lesions, and the 2-deoxyribonolactone abasic sites are thought to be recognized and repaired by Ape1, which hydrolyzes the phosphodiester bond 5' to the lesion and generates a 3'-hydroxyl group that is further processed by other enzymes in the base-excision repair (BER) pathway (29, 30). The most prevalent substrate for Ape1 *in vivo* is DNA containing abasic sites that arise from spontaneous nucleic acid base loss or excision of damaged bases by a battery of DNA glycosylases (17) (Scheme 1). Questions about how DNA repair enzymes like Ape1 specifically recognize DNA lesions such as normal or 4'-oxidized abasic sites remain.

Our group and others have used synthetic tools to generate these deoxyribose lesions in duplex DNA and to characterize their structures and interactions with repair enzymes *in vitro*. Among these lesions, limited information has been reported for 4'-oxidized abasic sites because of the synthetic challenges associated with their chemical instability. Recently, two synthetic methods for generating the 4'-oxidized abasic site in any sequence context have been developed (31, 32). The ability to site-specifically generate these lesions in duplex DNA offers the opportunity to study their structure, their repair, and the consequences of leaving the lesion unrepaired (33, 34).

In this study, two-dimensional (2D) NMR methods have been used to determine the first structure of a duplex oligonucleotide containing a single 4'-oxidized abasic site (**X**), d(CCAAAGXACCGGG)·d(CCCGGTACTTTGG) (**1**). This structure is compared with the structure of a normal abasic site (**Y**), d(CCAAAGYACCGGG)·d(CCCGGTACTTTGG) (**2**), in the same sequence context. The structural models that have resulted are the first step in understanding the mechanism(s) of recognition of these lesions by DNA repair enzymes.

## EXPERIMENTAL PROCEDURES

**Materials.** The oligonucleotides were synthesized on a 10  $\mu$ mol scale by Invitrogen Inc. and were further purified by anion-exchange high-performance liquid chromatography (HPLC) as described in the Supporting Information. *Escherichia coli* uracil-DNA glycosylase (UDG; 1 unit releases 60 pmol of uracil/min from ds uracil-containing DNA) was

purified from expression plasmid pET21a-UDG kindly supplied by J. Stivers (Johns Hopkins University, Baltimore, MD) as previously described (35). HIV-1 reverse transcriptase (HIV-1 RT; 1 unit incorporates 1 nmol of dNTP in 20 min at 37 °C and pH 8.3) was purified following the published protocol (36) from *E. coli* DH5 $\alpha$  containing the p66(His)/51 plasmid provided by P. Boyer and S. Hughes (National Cancer Institute, Bethesda, MD). 4'-N<sub>3</sub>-dUTP was prepared as previously reported (32). All other chemicals were purchased from Sigma-Aldrich Co.

**Synthesis of 1 and 2 for Analysis by 2D NMR Methods.** Compounds **1** and **2** were prepared by our published methods with modifications (32). The procedures and characterizations of **1** and **2** are described in detail in the Supporting Information.

**NMR Experiments.** The NMR experiments were performed on a custom-built 750 or 591 MHz spectrometer at the Francis Bitter Magnet Laboratory. The raw data were transferred to a Silicon Graphics workstation and were processed using Felix 2001 (Accelrys Inc.). Proton chemical shifts were referenced to an internal standard, sodium 3-(trimethylsilyl)-1-propanesulfonate (TSP) at 0.00 ppm. NOESY (100, 200, and 400 ms mixing times), TOCSY (60 and 110 ms mixing times), and E-COSY spectra were recorded at 25 °C in D<sub>2</sub>O at pH 6.5. For NOESY and TOCSY experiments, data sets of 4096 ( $t_2$ )  $\times$  512 ( $t_1$ ) complex points were acquired with a spectral width of 8012.4 Hz (750 MHz) or 7022.5 Hz (591 MHz) in both dimensions and 32 scans per  $t_1$  increment. For the E-COSY experiment, data sets of 4096 ( $t_2$ )  $\times$  1028 ( $t_1$ ) complex points were acquired with 48 scans per  $t_1$  increment. A presaturation pulse (2 s) was applied during the recycle delay period to suppress solvent signals in all NMR experiments carried out in D<sub>2</sub>O (37). Watergate-NOESY experiments (38) (200 and 400 ms mixing times) in a 90% H<sub>2</sub>O/10% D<sub>2</sub>O mixture were recorded at 4 °C. Data sets of 4096 ( $t_2$ )  $\times$  512 ( $t_1$ ) complex points were acquired with spectral widths of 13 020.8 Hz (591 MHz) in both dimensions and 32 scans per  $t_1$  increment. For all the data sets,  $t_1$  dimensions were linearly predicted to 4096 points and the data were processed with an exponential weighting function.

<sup>1</sup>H-<sup>13</sup>C GE-HSQC experiments (39) were conducted on the 591 MHz spectrometer at 25 °C with 1024  $\times$  512 complex points and a spectral width of 6009.6 Hz in the <sup>1</sup>H dimension and a spectral width of 16 025.64 Hz in the <sup>13</sup>C dimension. For each  $t_1$  increment, 64 transients were recorded. The <sup>13</sup>C dimension in the final data sets was linearly predicted to 1024 points. The spectrum was referenced indirectly through the transmitter frequency by external calibration on TSP.

<sup>1</sup>H-<sup>31</sup>P HSQC and H3'-selective PH-COSY experiments were performed on the 591 MHz spectrometer at 25 °C in D<sub>2</sub>O (40, 41). The <sup>31</sup>P chemical shifts were referenced to an external standard, trimethyl phosphate (0.00 ppm). <sup>1</sup>H-<sup>31</sup>P HSQC data were acquired with a spectral width of 6009.6 Hz in the <sup>1</sup>H dimension and a spectral width of 2000 Hz in the <sup>31</sup>P dimension (4096  $\times$  128 complex points). H3'-selective PH-COSY experiments (6009.6  $\times$  500 Hz spectral width, 4096  $\times$  256 complex points in each dimension) were performed to measure coupling constants of the DNA backbone (H3'-C3'-O-P). The H3' protons were specifically excited by using a band-selective 180° pulse with an

Table 1: Summary of Root-Mean-Square Deviations, Numbers of NMR-Derived Constraints, and Violations of Constraints of the Final 10 Structures

	1	2 $\alpha$	2 $\beta$
pairwise rmsd (Å)			
starting structures	4.80	4.54	4.55
final 10 structures	1.72	1.70	1.56
no. of NMR restraints			
NOE	410	482	482
dihedral angles	43	57	57
violations of NMR restraints			
NOE violations (Å)	0.056	0.14	0.11
dihedral angle violations (deg)	2.33	2.21	2.17

EBURP shape (41, 42). In the final data sets, the  $^{31}\text{P}$  dimension was linearly predicted to 1024 points. The Karplus equation ( $^3J_{\text{HCOP}} = 15.3 \cos^2 \phi - 6.1 \cos \phi + 1.6$ ) (43) was used to derive dihedral angle constraints for the H3'–C3'–O3'–P backbone from the experimentally determined coupling constants ( $J_{\text{H3'–P}}$ ).

**Calculation of NMR Restraints.** Distance restraints were derived from the volumes of the cross-peaks in NOESY spectra. The peak volumes were calculated with the peak picking protocol in Felix 2001. Volumes of cross-peaks between H2' and H2'' (1.772 Å) or between H5 and H6 (2.460 Å) of cytosine were used to calibrate the proton–proton distances. The cross-peak intensities were classified as weak, medium, or strong with distance restraints of 1.5–3.0, 2.0–5.0, or 3.0–6.0 Å, respectively. In addition, an iterative relaxation matrix analysis using MARDIGRAS was applied in estimating the distance restraints (44, 45) with excellent agreement. A total of 410 distance restraints were used for the subsequent structural calculation, including eight Watson–Crick H-bond interactions detected in the Watergate–NOESY experiments. Forty-three dihedral angle restraints (Table 1) were obtained by analysis of the coupling constants derived from E-COSY experiments. Both sets of restraints were used in the structural refinement of 1.

A similar analysis gave 482 distance restraints and 57 dihedral angle restraints for 2 $\alpha$  and 2 $\beta$  (Table 1) that were used in their structural refinement.

**NMR Structure Calculation and Refinement of 1, 2 $\alpha$ , and 2 $\beta$ .** All structural calculations were performed on an SGI Altix 3700 LINUX server with 128 Intel Itanium 1.3 GHz processors at The Scripps Research Institute using the SANDER module of AMBER 8.0 (46). The restrained electrostatic potential (RESP) atomic charges of the 4'-oxidized abasic site, the normal abasic site, the 2-deoxyribonolactone abasic site, and the 3'-PG/5'-P lesion were derived using the same strategy developed by the Kollman group (47, 48). Geometry optimizations and molecular electrostatic potential (MEP) calculations of each residue were performed at the HF/6-31G\*/HF/6-31G\* level of theory using the Gaussian 98 program (49) and the RED interface (50). The final AMBER 8.0 parameters of each abasic site residue have been deposited in the RESP ESP charge DData Base (R.E.DD.B.; <http://www.u-picardie.fr/labo/lbpd/REDDB/index.html>, Project Code F61–F64) and are summarized in Tables S1–S4 of the Supporting Information.

The starting model structures were built using the Biopolymer module of Insight 2000 (Accelrys Inc.). Three starting structures were constructed, including an A-form DNA, a

B-form DNA, and a third form with arbitrary helical parameters. The root-mean-square deviation (rmsd) for the three structures was  $\sim 4.5$  Å. A generalized Born implicit solvent model was applied to simulate the solvent dielectric constant. The initial model structures were energy-minimized by 500 steps of conjugated gradient minimization followed by 500 steps of steepest descent minimization to remove high-energy interactions introduced during the manual model building. The simulated annealing was carried out by heating the structure from 300 to 900 K in 5 ps followed by cooling to 300 K over the next 15 ps. During the cooling process, weak temperature coupling (slow cooling; TAUTP = 4.0) was applied for the first 13 ps followed by strong temperature coupling (rapid cooling; TAUTP = 0.1) for the last 2 ps (51). NMR restraints were enforced during the annealing process with an increasing force constant from 5 to 50 kcal/mol over the first 3 ps which was maintained at 50 kcal/mol for the rest of the calculation (51). The annealed structure was then subjected to 100 ps constant-temperature (300 K) molecular dynamics simulation with NMR restraints. Ten structures for each starting structure (A-form, B-form, and the third form) were calculated by sampling conformations during the last 10 ps of the molecular dynamics run at 1 ps intervals. The 30 resulting structures have an averaged rmsd of  $< 2.0$  Å, which indicates a structural convergence after refinement. Ten of the 30 structures were chosen on the basis of having the fewest NMR violations to represent the ensemble of the averaged structure. The final 10 structures and the averaged structures of 1, 2 $\alpha$ , and 2 $\beta$  have been deposited in the Nucleic Acid Database (entries 2HOU, 2HPX, 2HSK, 2HSL, 2HSR, and 2HSS). The statistical analysis of the 10 final structures is summarized in Table 1.

**Molecular Dynamics Simulations Using an Explicit Solvent Model.** The molecular dynamics simulations were performed at a constant temperature (300 K) using AMBER 8.0 following the protocol of Dupradeau et al. (52). The modified version of the Cornell et al. force field (53) by Cheatham et al. (54) was used in all simulations. The final averaged structures were first solvated using the TIP3P solvent model with approximately 6000 water molecules in an octahedral solvent box with a thickness of 10 Å. Twenty-four sodium ions were also included in the calculations to neutralize the system. The molecular dynamics calculations were carried out on a 3 ns time scale. The conformation was sampled every 2 ps and was analyzed by the PTRAJ module in Amber 8.0.

## RESULTS

**Proton Chemical Shift Assignments of 1 and 2.** Assignments of the vast majority of chemical shifts were based on a standard set of TOCSY (60 and 110 ms mixing times) and NOESY (100, 200, and 400 ms mixing times) experiments (55). The results are summarized in Tables S5 and S6 of the Supporting Information. The assignment of the protons associated with the 4'-oxidized abasic site is critical to efforts to understand its conformation within the oligomeric duplex and proved to be more challenging. Analysis of all of the H1' protons (Table S5) revealed one with an unusual chemical shift that is upfield from the others at 5.155 ppm. TOCSY experiments (Figure 1) revealed its interaction with H2' and H2'' protons (2.104 and 1.827 ppm, respectively), both of which interacted with H3' (4.559 ppm). As with H1',

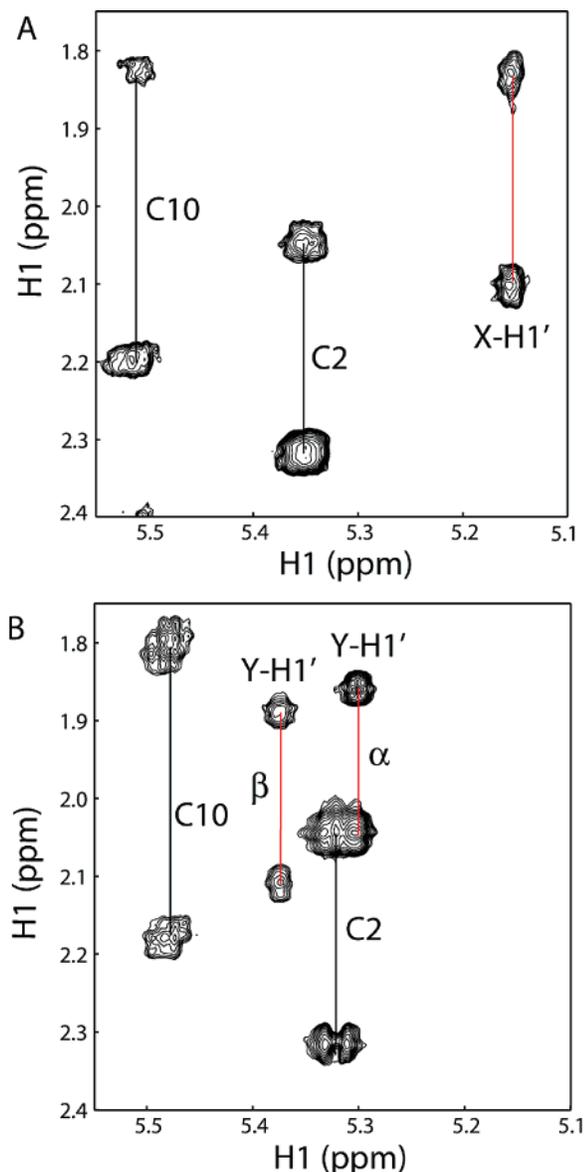


FIGURE 1: Assignments of the protons associated with the abasic sites using TOCSY spectra (60 ms mixing time) at 750 MHz. (A) Through-bond connectivities between H1' and H2'/H2'' protons (red line) of the 4'-oxidized abasic site (X) in **1**. (B) Through-bond connectivities between H1' and H2'/H2'' protons (red line) of the 4'-abasic site (Y) in **2**. The two sets of signals are associated with the  $\alpha$  and  $\beta$  anomers revealed by NOESY experiments. Cross-peaks associated with C2 and C10 in this region are also labeled.

H2'' and H3' stand apart from all other deoxyribose protons as they are all substantially upfield shifted relative to their "B-form" counterparts. These unusual chemical shifts and our inability to correlate them with assigned bases are the basis for the preliminary assignment of the deoxyribose moiety at the 4'-oxidized abasic site.

Additional experiments were carried out to substantiate this H1' assignment.  $^1\text{H}$ - $^{13}\text{C}$  GE-HSQC experiments were performed to establish the connectivity between H1' and its C1' relative to other deoxyriboses within the duplex (Figure 2A). The quality of natural-abundance  $^1\text{H}$ - $^{13}\text{C}$  GE-HSQC data is shown in the cross-peaks between base protons (H8 of purines and H6 of pyrimidine) and the attached carbon atoms (Figure 2B). H1' of the putative 4'-oxidized abasic site exhibits a cross-peak with a carbon signal at 98.41 ppm

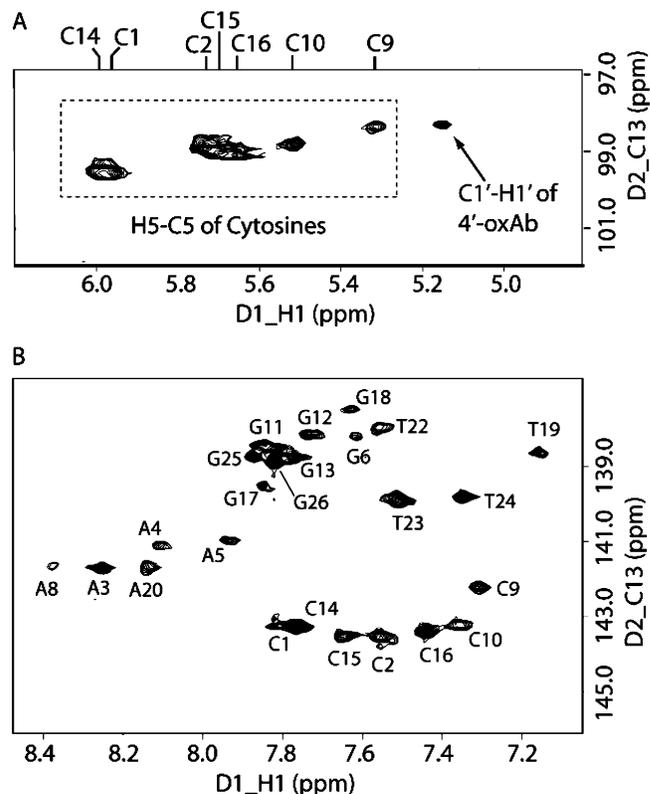


FIGURE 2:  $^1\text{H}$ - $^{13}\text{C}$  GE-HSQC spectrum of **1**. (A) Region showing cross-peaks between H5 and C5 of cytosines. The arrow indicates the cross-peak between H1' and C1' of the 4'-oxidized abasic site (4'-oxAb). (B) Regions showing cross-peaks between base protons (H6 of pyrimidines or H8 of purines) and their attached carbon atoms (C6 of pyrimidine or C8 of purine).

(Figure 2A). This value is substantially downfield shifted relative to all other C1' atoms (84–88 ppm, Table S7 of the Supporting Information) and very similar to the value of 101 ppm previously reported for C1' in a normal abasic site in duplex DNA (56) and 99.6 ppm for the anomeric carbon of 2,5-dihydroxy-2-furaldehyde dimethyl acetal (57). The  $^1\text{H}$  and  $^{13}\text{C}$  chemical shifts together are consistent with a 4'-oxidized abasic site.

The establishment of chemical shift assignments allowed determination of the anomeric configuration of the 4'-oxidized deoxyribose (Scheme 1) based on the relative NOE intensities of H2' and H2'' to H1' (55, 58). For the  $\alpha$  anomer, H1' is closer to H2' than to H2'' and should produce a NOE cross-peak stronger than the cross-peak to the latter. The reverse is true for the  $\beta$  anomer. The assignments of H2' and H2'' are based on the requirement for any sugar conformation that the cross-peak with H3' must be more intense to H2' than to H2''. Using this strategy, H2' and H2'' were assigned as shown in Figure 3A (red labels). The observation of a stronger cross-peak between H1' and H2' (B in Figure 3A) relative to that between H1' and H2'' (A in Figure 3A) requires an  $\alpha$  anomer. The assignment of a single conformation is also consistent with the NOE cross-peak patterns observed in Figure S1 of the Supporting Information, which show a single set of signals adjacent to the 4'-oxidized abasic site.

*Assignments of Abasic Site Protons in 2 $\alpha$  and 2 $\beta$ .* We have examined an abasic site (**2**) in a sequence identical to **1** containing the 4'-oxidized abasic site so that a direct

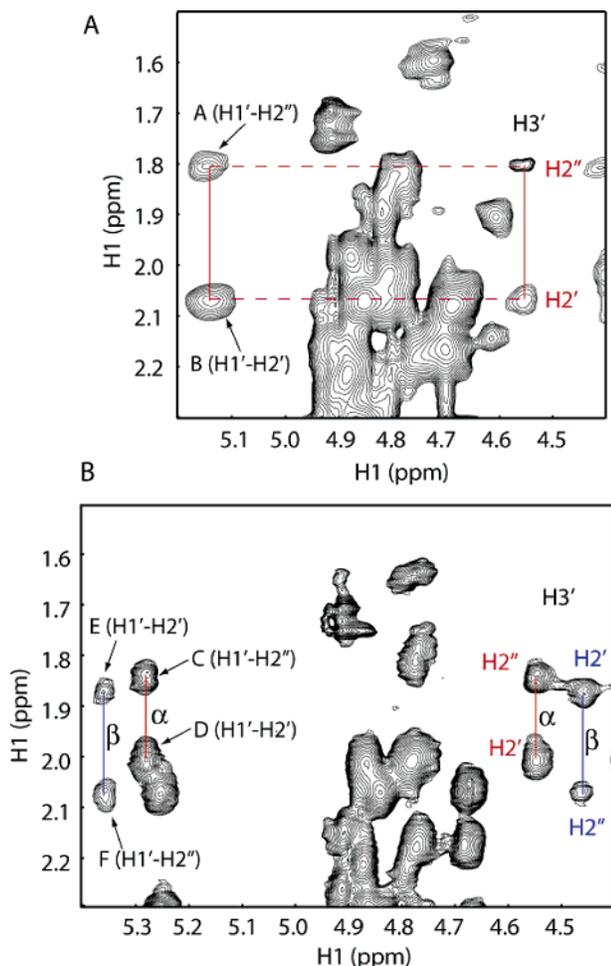


FIGURE 3: Determination of the anomeric configuration of **1** and **2** using the relative intensities of NOE cross-peaks. (A) For **1**, one set of signals are present and the connectivities are highlighted in red. (B) For **2**, two sets of signals are present (labeled in red and blue). The assignments of cross-peaks A–F are also listed.

comparison can be made. Duplex **2** is similar but not identical to a duplex containing an abasic site that we previously examined structurally (58). Two sets of signals in the 200 ms NOESY base to H1' region associated with T19 and T20 were readily apparent (Figure S2 of the Supporting Information) and suggested  $\alpha$  and  $\beta$  anomers (Scheme 1) (58). This conclusion is supported by the TOCSY spectra in which two H1' protons are observed to interact with two sets of H2' and H2'' (Y–H1' in Figure 1B). Their distinct chemical shifts compared with those in normal B-form DNA (Table S6), together with the inability to correlate them with any of the completely assigned bases, suggest that the protons must be associated with anomers of the abasic sites. Using the same strategy that was described for **1**, these two species were assigned anomeric configurations ( $2\alpha$  and  $2\beta$ ) (in Figure 3B, the  $\alpha$  anomer is denoted in red and the  $\beta$  anomer in blue) (56, 59). The chemical shift assignments of the abasic site protons of  $2\alpha$  and  $2\beta$  are very similar to those we previously reported (58).

**Intrahelicity of the Base Opposite from the Abasic Site Lesions.** The NOE connectivities between H1' and the base protons of **1** are presented in Figure S1. The results are indicative of B-form DNA. Disruption of the NOE connectivity between G6 and A8 indicated with a times sign

(dashed red line, Figure S1) is consistent with the presence of the 4'-oxidized abasic site between these residues. The observed connectivity among T19, A20, and C21 opposite from the 4'-oxidized abasic site (Figure S1, bottom, red lines) is indicative of an intrahelical conformation adopted by A20 opposite the 4'-oxidized abasic site.

The NOE connectivities between H1' and the base protons of **2** are shown in Figure S2. This connectivity, as with **1**, is disrupted by the abasic site between G6 and A8 indicated with a times sign (Figure S2, top, red dashed lines). Complete NOE connectivities are also observed for residues on the opposite strand (Figure S2, bottom, dashed red lines). In contrast to the observations with **1**, two distinct sets of NOE cross-peaks are observed for T19, A20, and C21 (Figure S2, bottom, red dashed lines). These observations are consistent with a mixture of two DNA conformations: one associated with the  $\alpha$  anomer and one with the  $\beta$  anomer of the abasic site ( $2\alpha$  and  $2\beta$ ) (58).

**Sugar Pucker and Phosphodiester Backbone Conformations of **1**,  $2\alpha$ , and  $2\beta$ .** E-COSY experiments at 750 MHz were performed to measure the coupling constants in an effort to assign the sugar pucker for the deoxyribose moieties in **1**,  $2\alpha$ , and  $2\beta$ . The H1'–H2' and H1'–H2'' coupling constants in **1** were measured for all residues except for the 4'-oxidized abasic site (Table S8A of the Supporting Information), the intensities of which were too low. For the residues whose coupling constants are available, the results are consistent with a Southern, C2'-endo conformation observed in B-form DNA (60).

For  $2\alpha$  and  $2\beta$ , these coupling constants were measured for all residues (Table S8B of the Supporting Information), and with the exception of the abasic site, they are consistent with a C2'-endo conformation (60) expected for B-form DNA. The coupling constants associated with abasic site protons are similar to those previously determined for a normal abasic site in a similar sequence context (58) (Figure S3 of the Supporting Information). Comparison of the experimentally determined coupling constants with the calculated values for different sugar conformations (60) suggests that the  $\alpha$  anomer adopts a near Northern conformation (*O1'-endo*) and that the  $\beta$  anomer adopts a near Southern conformation (*C1'-exo*).

In an effort to determine if the 4'-oxidized abasic site in **1** or the normal abasic site in  $2\alpha$  and  $2\beta$  gives rise to perturbed phosphodiester backbone conformations relative to B-form DNA,  $^1\text{H}$ – $^{31}\text{P}$  HSQC experiments were carried out. All of the  $^{31}\text{P}$  chemical shifts for **1**,  $2\alpha$ , and  $2\beta$  were within the normal envelope for B-form DNA (–4.2 to –5.3 ppm), consistent with limited distortions in the phosphodiester backbone conformation. Furthermore, the H3'-selective PH-COSY experiments gave H3'–C–O–P coupling constants in the range of 6–8 Hz, consistent with a B-form DNA conformation.

**Exchangeable Protons of **1**,  $2\alpha$ , and  $2\beta$ : Insight into the Conformation Adjacent to the Abasic Site.** Exchangeable base protons involved in the Watson–Crick base pair interactions were assigned by Watergate–NOESY experiments in a 90% H<sub>2</sub>O/10% D<sub>2</sub>O mixture at 4 °C. For **1**,  $2\alpha$ , and  $2\beta$ , eight imino protons were detected and assigned on the basis of their NOE interactions with their base-paired partners (Figure S4 of the Supporting Information). The chemical shifts and the NOE patterns were very similar for **1**,  $2\alpha$ , and  $2\beta$  (Figure

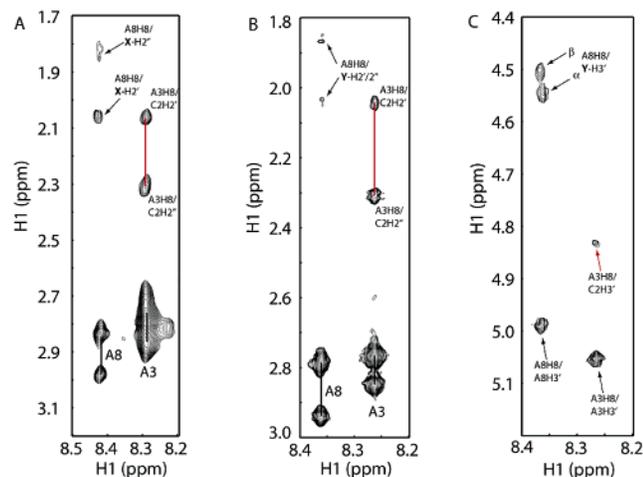


FIGURE 4: NOE interactions associated with the protons of the abasic site and the adjacent 3'-base proton (A8 H8) studied by NOESY (100 ms mixing time) at 750 MHz. (A) For **1**, A8 H8 exhibits a medium NOE signal with the H2' proton of the 4'-oxidized abasic site (X H2') and a weak NOE signal with the X H2'' proton (indicated by the arrows). (B) For both **2α** and **2β**, A8 H8 exhibits weak NOE interactions with the H2' and H2'' protons of the abasic site (Y H2'/2'') as indicated by the arrows. (C) For both **2α** and **2β**, A8 H8 exhibits medium NOE interactions with the H3' proton of the abasic site (Y H3') as indicated by the arrows. The red lines and arrow indicate NOE signals between C2 and A3 in a typical B-form DNA.

S4), which suggests that they adopt similar global structures. The imino protons at each end of the duplex (G13 and G26) and adjacent to the abasic site (G6 and T19) were not observed, presumably due to their fast exchange with solvent on the NMR time scale. These results suggest that the base pair interactions adjacent to the abasic site are weaker than those in regular B-form DNA.

**NOE Interactions Associated with Abasic Site Protons of 1, 2α, and 2β.** The assignments of the sugar protons associated with the deoxyribose lesions allowed analysis of NOESY data for interactions of these protons with adjacent residues. As revealed in the NOESY spectrum (Figure 4A), a medium NOE signal is observed between A8 H8 and X H2' and a weak one between A8 H8 and X H2'' in **1**. For comparison, a NOE of medium intensity is shown between A3 H8 and C2 H2'/H2'' (red line, Figure 4A), typical of generic B-form DNA. The unusually weak NOE interaction between A8 H8 and H2'' of the 4'-oxidized abasic sites suggests that this lesion adopts a conformation distinct from that of B-form DNA.

For both **2α** and **2β**, two weak A8 H8–Y H2'/H2'' NOE signals are observed (arrows in Figure 4B). Their intensity can be compared with cross-peaks between A3 H8 and C2 H2'/H2'' (red line in Figure 4B), typical of B-form DNA. In addition, an A8 H8–Y H3' NOE is observed for both **2α** and **2β** (Figure 4C) of medium intensity. In B-form DNA, the distance between H3' and the adjacent base proton on the 3' side is ~4.4 Å, which results in a weak NOE signal demonstrated by the NOE cross-peak between A3 H8 and C2 H3' (red arrow, Figure 4C). These three unusual NOE interaction patterns suggest that the conformation of the abasic site sugar in **2** is distinct from that of B-form DNA.

**Structural Models of 1, 2α, and 2β Derived from NMR Restraints and Molecular Dynamics Simulations.** NOE and dihedral restraints were used in conjunction with simulated

annealing followed by 100 ps molecular dynamics to obtain model structures of **1**, **2α**, and **2β** (Table 1). The two NOEs for **1** and three NOEs for **2α** and **2β** between the abasic site and the base proton on the 3'-side dictate the consequence of the modeling exercise. Ten final structures were obtained for **1**, **2α**, and **2β** without any NOE distance violations larger than 0.5 Å or dihedral angle violations larger than 5° (Table 1). The overlay of these 10 structures and the averaged structure is shown in panels A and B of Figure S5 of the Supporting Information. The overall structures in all three cases are consistent with regular B-form DNA without significant backbone distortion. In all structures, the base (A20) opposite from the lesion (colored blue, Figure S5A) is intrahelical and stacks between adjacent base pairs, similar to B-form DNA.

**Sugar Conformations at the Abasic Sites.** Despite the overall similarity of global conformation among **1**, **2α**, and **2β**, distinct sugar conformations at the abasic site are observed in the 10 calculated structures (colored red, Figure S5A). The subsequent description will focus on the conformation of each average structure. For **1**, the 4'-oxidized deoxyribose adopts predominantly an intrahelical conformation (Figure 5A). This conformation requires a relatively short (2.59 Å) H2'–A8 H8 distance and a long (4.22 Å) H2''–A8 H8 distance (Figure 6B) compared to those in B-form DNA (Figure 6A). This conformation is consistent with the intensities of NOE cross-peaks involving these protons (Figure 4A).

To evaluate the significance of the two observed NOE interactions in determining the final structure, NMR structural refinement was carried out without these two distance restraints. The results show that after sampling a variety of conformations during simulated annealing, the deoxyribose moiety of the 4'-oxidized abasic site adopts an intrahelical conformation that is similar to B-form DNA.

Our previous studies of an abasic site in a similar sequence context revealed that the abasic site adopted a partial extrahelical conformation (58). In this study, for both **2α** and **2β**, the deoxyribose of the abasic site is also partially extrahelical (Figure 5B,C). These results require the H2' and H2'' protons to move away from A8 H8 (red dashed lines in panels B and D of Figure 7) and the H3' proton to move closer to A8 H8 (blue dashed lines in panels B and D of Figure 7) in both **2α** and **2β** compared to that in B-form DNA (panels A and C of Figure 7). The extrahelical conformation is thus governed by the unusually weak NOE interactions between A8 H8 and X H2' and H2'' (Figure 4B) and the medium NOE signals between A8 H8 and H3' (Figure 4C). NMR structural refinements without these NOE restraints resulted in an intrahelical conformation for the abasic site.

**Molecular Dynamics Simulations.** To probe the conformational flexibility at each abasic site, the final average structure of **1**, **2α**, and **2β** was subjected to 3 ns molecular dynamics simulations. Water molecules and counterions were included in the calculation to better simulate the solution state of DNA molecules. To monitor the time-dependent extrahelical motion, the distances between C1' of the abasic site and the N1 atom of the opposite base (A20 N1) were monitored during the course of the simulation (Figure 8A–C). In B-form DNA, the C1'–N1 distance, ~6.2 Å, is restrained by Watson–Crick base pairing. As an example,

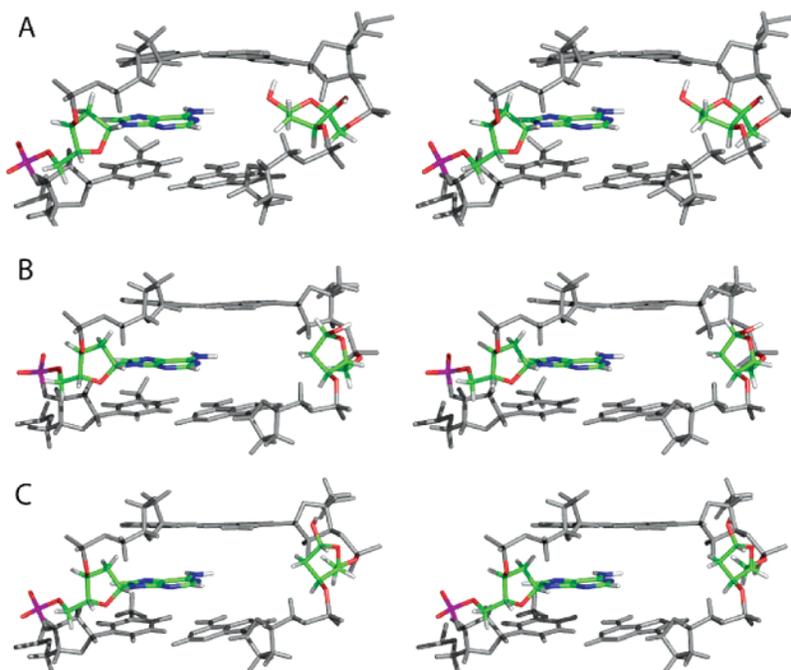


FIGURE 5: Stereoview of the abasic site and the neighboring base pairs in the final averaged structures: (A) **1**, (B) **2 $\alpha$** , and (C) **2 $\beta$** . The abasic site and the opposite base (A20) are colored in CPK mode.

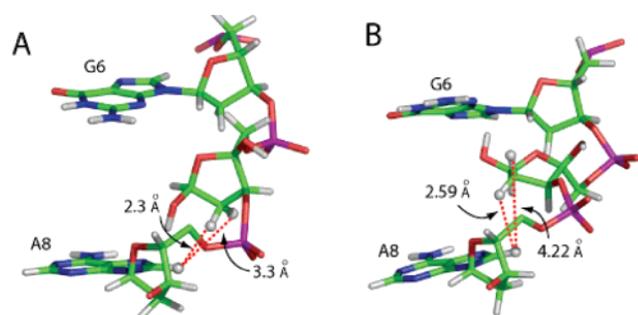


FIGURE 6: Conformation of the 4'-oxidized abasic site in **1** compared to that in a theoretical B-form DNA. (A) Computer-generated conformation of the 4'-oxidized abasic site in a B-form DNA. (B) Conformation of the 4'-oxidized abasic site in the final averaged structure.

the distance between T24 C1' and A3 N1 in **1**, **2 $\alpha$** , or **2 $\beta$**  (blue traces in Figure 8A–C) was monitored during the simulation and exhibited a very small fluctuation ( $<0.5$  Å). However, for the normal or 4'-oxidized abasic site, because of the lack of base pairing interactions, the distance between C1' of the lesion and N1 of A20 is sensitive to the motion of the abasic site (red traces in Figure 8A–C). Specifically, the extrahelical motion of the abasic site would cause an increase in this distance.

The analysis of the molecular dynamics simulation of **1** revealed that the 4'-oxidized abasic site is mainly intrahelical (red trace, Figure 8A). The average distance between C1' of the 4'-oxidized abasic site and N1 of A20 is  $\sim 6$  Å, similar to the value in normal B-form DNA. The extrahelical motion of the 4'-oxidized abasic site is also present during the simulation, which is indicated by the increase in the distance between C1' and A20 N1 (Figure 8A). The fluctuation of this distance is less than 2.5 Å. During most of the simulation, the 4'-OH group of the abasic site was shown to form H-bond interactions with the oxygen atom of the 5'-

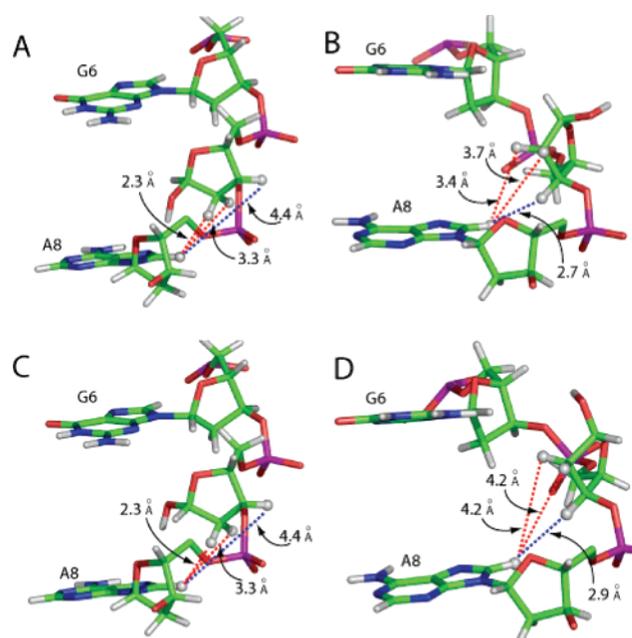


FIGURE 7: Conformation of the normal abasic site (Y) in **2 $\alpha$**  and **2 $\beta$**  compared to that in a theoretical B-form DNA. (A) Computer-generated conformation ( $\alpha$  anomer) in a B-form DNA. (B) Conformation in the final structure of **2 $\alpha$** . (C) Computer-generated conformation ( $\beta$  anomer) in a B-form DNA. (D) Conformation in the final averaged structure of **2 $\beta$** .

phosphate group, which might contribute to the conformational rigidity of the abasic site (Figure S6A of the Supporting Information, blue dashed line). Interestingly, a transient H-bond interaction between the 1'-OH group of **X** and N1 of A20 was observed, indicated by the decrease in the C1'–A20 N1 distance (arrow, Figure 8A). A stereoview of this transient H-bond interaction is presented in Figure 8D.

For both **2 $\alpha$**  and **2 $\beta$** , the analysis indicates that there is a larger extrahelical motion associated with the abasic site than

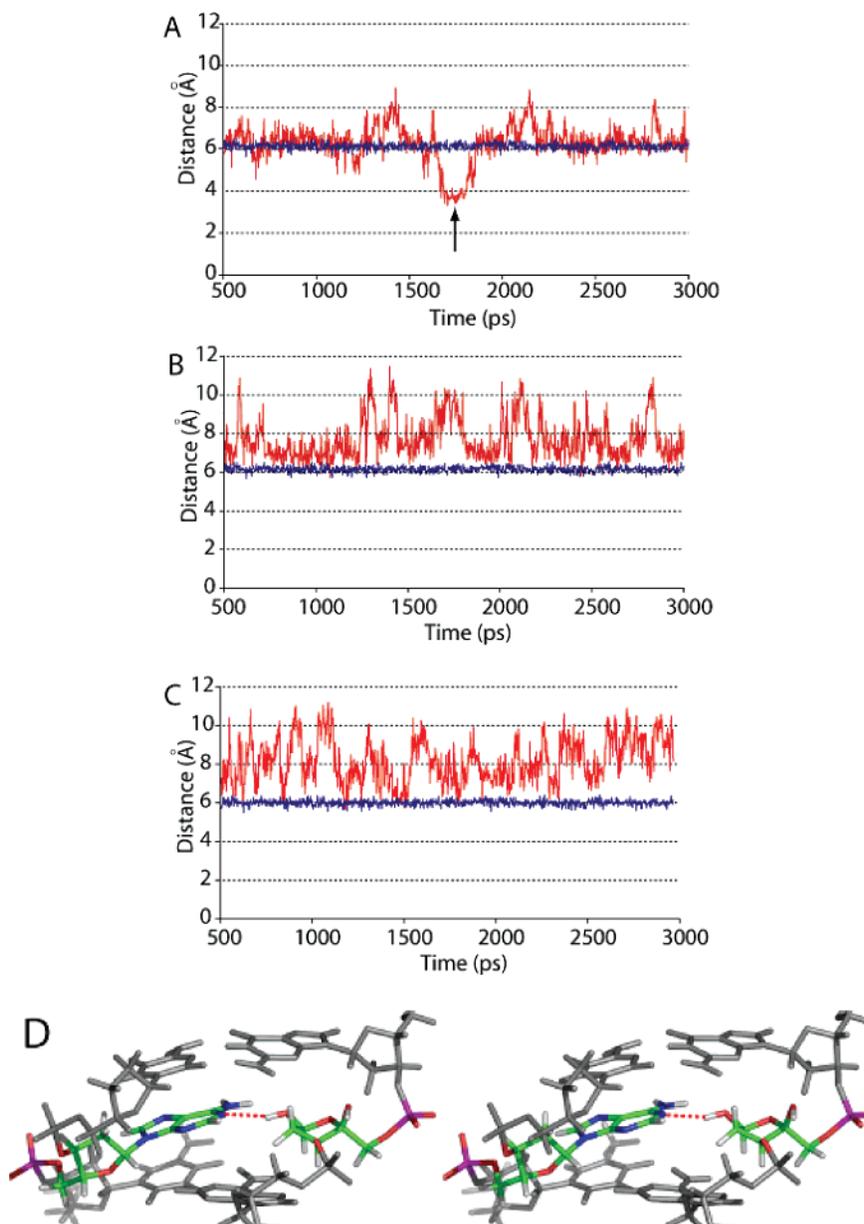


FIGURE 8: Time-dependent extrahelical motion of the abasic site in **1**, **2α**, and **2β**. The extrahelical motion was monitored by the change in the distance between C1' of the abasic site and the N1 atom of the opposite base (A20 N1) (red trace). The blue traces represent the distance between A3 N1 and T24 C1'. (A) For **1**, the arrow in panel A indicates the transient formation of a H-bond between the 1'-OH group and A20 N1. (B) For **2α**. (C) For **2β**. (D) Stereoview of the snapshot of **1** with a transient H-bond between the 1'-OH group and A20 N1 (indicated by the arrow in panel A).

with **1**. This motion is reflected by the large fluctuation of the distance between C1' of the abasic site and N1 of A20 ( $>5$  Å) (Figure 8B,C, red trace). The extrahelical motion occurs approximately every 50 ps, much more frequently than in **1**. Calculation also revealed subtle differences between **2α** and **2β**, with the abasic site spending more time in the extrahelical conformation in **2β** than in **2α** during the simulation.

*Molecular Dynamics Simulations on a 3'-PG/5'-P Lesion and a 2-Deoxyribonolactone Abasic Site.* We have previously synthesized and characterized by 2D NMR spectroscopic methods a 3'-PG/5'-P gapped lesion (5'-CCAAAG<sup>-PG</sup> P<sup>-</sup>-ACTGGG-3'/3'-GGTTTCATGACCC-5') (61). The structural models resulting from NMR restraints revealed that this lesioned DNA is B-form with the base opposite the gap well-stacked between adjacent base pairs. Modeling studies further

revealed that both the 3'-PG and 5'-P group adjacent to the gap are extrahelical. However, no NOE cross-peaks were observed between the methylene group of 3'-PG and the remainder of the DNA (61). We submitted this structure to molecular dynamics calculations. These simulations revealed that both the 3'-PG and 5'-P groups are conformationally flexible without significant perturbation of the DNA backbone (Figure S7 of the Supporting Information), supporting our previous model.

Molecular dynamics simulations were also performed on a 2-deoxyribonolactone (**L**) abasic site in the same sequence context as **1**, **2α**, and **2β** (5'-GLAC-3'/3'-CATG-5') starting with a computer-generated B-form structure. The structure of this lesion in the sequence of 5'-ACLCA-3'/3'-TGTGT-5' has previously been determined using 2D NMR methods (62). The results of their modeling based on two weak NOE

interactions between H2' and H2'' of **L** and the 3' base proton revealed that the overall DNA is B-form and that the 2-deoxyribonolactone ring adopts an intrahelical conformation. The base (T) opposite from **L** was intrahelical, as suggested by the unperturbed NOE interactions on the opposite strand. Interestingly, in contrast to **1**, **2 $\alpha$** , and **2 $\beta$** , stable base pair interactions were observed adjacent to **L** as revealed by the observed imino protons (62). Our molecular dynamics simulation suggests that the 2-deoxyribonolactone moiety adopts a predominantly intrahelical conformation with a much reduced flexibility compared with the normal abasic site (Figure S8A of the Supporting Information). Only two extrahelical motions were observed during the 3 ns simulation (Figure S8B,C). During the molecular dynamics calculation, no distance restraints were included for the 2-deoxyribonolactone residue, which suggests that the predominantly intrahelical conformation is probably intrinsic in this sequence context.

## DISCUSSION

Solution structural models of several abasic site lesions and their analogues in duplex DNA have been determined using 2D NMR methods, including normal abasic sites (58, 63–65), 2-deoxyribonolactone (62), the 3'-PG/5'-P gapped lesion (61), and the tetrahydrofuran (THF) (66, 67) and cyclopentane (68) abasic site analogues. Common structural features have been observed for duplex DNA containing a single abasic site lesion, including minimal overall backbone distortion and intrahelicity of the base opposite from the lesion. However, because these structural models are in different sequence contexts, no general conclusions can be drawn about the differences in the local conformation and the flexibility of these abasic sites and analogues, which might be key to understanding their recognition and subsequent repair by DNA repair enzymes (69–71).

In our study, the first solution structure of a 4'-oxidized abasic site has been determined using NMR-derived restraints and compared with the structure of a normal abasic site in the same sequence context. The 4'-oxidized abasic site in **1** exists as a ring-closed hemiacetal as previously indicated by MALDI-TOF mass spectrometry (31, 32) and by the H1' signal in TOCSY and GE-HSQC experiments in this study. Surprisingly, only the  $\alpha$  anomer is observed in **1** compared with both  $\alpha$  and  $\beta$  anomers observed with the normal abasic site in the same sequence context. The assignment of a single anomeric configuration in **1** is supported by the NOE cross-peak pattern, which shows a single set of signals adjacent to the 4'-oxidized abasic site. So far, it is not clear whether the preference for the  $\alpha$  anomer is sequence-dependent or inherent to the 4'-oxidized abasic site structure. Stereoelectronic effects or interactions resulting from the 4'-hydroxyl group could be responsible for the predominant  $\alpha$  configuration. Indeed, molecular simulations of **1** revealed a transient H-bond between the 1'-OH group and the oxygen atom of the 4'-OH group (Figure S6 A, red dashed line), which occurred four times and lasted for a total of 32 ps during the 3 ns simulation. This H-bonding interaction was further studied by quantum mechanics calculations performed at the HF/6-31G\* and B3LYP/6-31G\* levels of theory on a model compound of the 3',5'-diphosphate-4'-oxidized abasic site. Geometry optimization of this model compound confirmed the presence of the H-bonds observed during molecular

dynamics simulations (Figure S6B, red dashed line). Thus, the energetically favorable interaction between the 1'-OH group and the oxygen atom of the 4'-OH group, which could occur only in the  $\alpha$  anomer, might result in a predominant  $\alpha$  anomeric configuration for the 4'-oxidized abasic site. Unfortunately, in **1**, we have not been able to determine the sugar conformation of the 4'-oxidized abasic site due to the inability to assign coupling constants. Additional studies on small model oligonucleotides and/or  $^{13}\text{C}$ -enriched duplex samples are required to gain insight into the conformation of the 4'-oxidized abasic site and the influence on its surroundings.

Our long-range goal is to understand the molecular mechanism by which abasic site lesions generated by BLM and other DNA-damaging agents are recognized and repaired by DNA repair enzymes. In human cells, 4'-oxidized abasic sites, 3'-PG sites, and 2-deoxyribonolactone are all thought to be repaired via pathways initiated by Ape1 (72). Demple and co-workers by studying oxidized and normal abasic sites in a defined sequence context have determined that the 4'-oxidized abasic site is repaired by Ape1 at  $1/3$  of the rate of that observed for the normal abasic sites (29). Repair of a 3'-PG site by Ape1 occurred at  $1/400$  of the rate of the 4'-oxidized abasic site in the same sequence context (73, 74). Demple, Greenberg, and colleagues (30) have also demonstrated that Ape1 can incise the 2-deoxyribonolactone lesion, generating a 3'-hydroxyl end and 5'-phospho-2-deoxyribonolactone with a  $k_{\text{cat}}$  similar to that of the normal abasic site and a  $K_{\text{m}}$  elevated 5-fold relative to that of the normal abasic site. The basis for the observed rate differences and whether they are sequence-dependent are unclear as the rate-limiting step for Ape1 has not been established. Therefore, it becomes essential to examine interactions of these lesions with Ape1 structurally and to identify whether and how the conformation and its flexibility of each lesion dictate their recognition and repair.

The X-ray structure of Ape1 has been determined by the Tainer group in a complex with duplex DNA containing a THF abasic site analogue and provides a paradigm for thinking about the structural basis for lesion recognition by Ape1 (75). In this structure, Ape1 was shown to interact with the abasic site (THF) using residues from a flexible loop and a helix–turn–helix motif that binds to the DNA from both major and minor grooves. The duplex DNA is bent  $\sim 35^\circ$  at the site of the lesion. The most striking feature of the structure is that the THF moiety is flipped out of the phosphodiester backbone into a binding pocket formed by protein side chains that exclude intact bases. The extrahelical conformation has been observed in the structures of a number of abasic site–protein complexes, including Endo IV (76), uracil glycosylase (77), and 3-methyladenine glycosylase (78). However, whether the repair enzyme selectively recognizes the flipped-out abasic site or whether it actively flips the abasic site to the extrahelical position given the structural flexibility in the abasic site region remains unclear.

The flexibility of the abasic site has been proposed to be exploited by the repair enzymes for recognition (69–71). The study presented here suggests that there are differences in the structure and conformational flexibility among different abasic site lesions. However, the relevance of these differences in the recognition by Ape1 requires further study. Our results for the normal abasic site suggest that in solution,

the deoxyribose of the abasic site is flexible and adopts a wide range of extrahelical conformations. Therefore, it is likely that the extrahelicity of the abasic site could be recognized by DNA repair enzymes. For 4'-oxidized and 2-deoxyribonolactone abasic sites, a more active role of Ape1 in flipping the abasic site into its binding pocket is likely required due to their reduced flexibility. While the 3'-PG/5'-P lesion appears to be most flexible on the basis of the simulations, the structural features that govern their interaction with Ape1 remain unclear. The charge differences of the 3'-PG/5'-P lesion relative to the other lesions could play a key role in its slow turnover. The structural studies described here present a starting point for probing the conformational diversity of different abasic site lesions. Further studies on a fast time scale in the presteady state with Ape1 are required to gain information that can correlate the difference in structures of abasic sites (normal, 1'-oxidized, or 4'-oxidized) to the recognition and repair by Ape1.

### SUPPORTING INFORMATION AVAILABLE

Synthesis and characterization of compounds **1** and **2**, NOESY (200 ms mixing time) and E-COSY spectra at 750 MHz, Watergate-NOESY (200 ms mixing time) spectra at 591 MHz, a superimposition of 10 structure models and averaged structures of **1**, **2 $\alpha$** , and **2 $\beta$** , molecular dynamics simulations of 2-deoxyribonolactone and PG lesions, a summary of sequences of oligonucleotides **1-9**, RESP atomic charges for each lesion, <sup>1</sup>H chemical shift assignments and coupling constants for **1** and **2**, and <sup>13</sup>C chemical shift assignments for **1**. This material is available free of charge via the Internet at <http://pubs.acs.org>.

### REFERENCES

1. Sikic, B. I., Rozenweig, M., and Carter, S. K. (1985) *Bleomycin Chemotherapy*, Academic Press, Orlando, FL.
2. Levi, J. A., Raghavan, D., Harvey, V., Thompson, D., Sandeman, T., Gill, G., Stuart-Harris, R., Snyder, R., Byrne, M., Kerestes, Z., et al. (1993) The importance of bleomycin in combination chemotherapy for good-prognosis germ cell carcinoma. Australasian Germ Cell Trial Group, *J. Clin. Oncol.* **11**, 1300-1305.
3. Einhorn, L. H. (2002) Curing metastatic testicular cancer, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 4592-4595.
4. Stubbe, J., and Kozarich, J. W. (1987) Mechanisms of bleomycin-induced DNA degradation, *Chem. Rev.* **87**, 1107-1136.
5. Burger, R. M. (1998) Cleavage of nucleic acids by bleomycin, *Chem. Rev.* **98**, 1153-1169.
6. Boger, D. L., and Cai, H. (1999) Bleomycin: Synthetic and mechanistic studies, *Angew. Chem., Int. Ed.* **38**, 449-476.
7. Burger, R. M., Peisach, J., and Horwitz, S. B. (1981) Activated bleomycin: A transient complex of drug, iron, and oxygen that degrades DNA, *J. Biol. Chem.* **256**, 11636-11644.
8. Burger, R. M., Projan, S. J., Horwitz, S. B., and Peisach, J. (1986) The DNA cleavage mechanism of iron-bleomycin, *J. Biol. Chem.* **261**, 15955-15959.
9. Neese, F., Zaleski, J. M., Zaleski, K. L., and Solomon, E. I. (2000) Electronic structure of activated bleomycin: Oxygen intermediates in heme versus non-heme iron, *J. Am. Chem. Soc.* **122**, 11703-11724.
10. Wu, J. C., Kozarich, J. W., and Stubbe, J. (1985) Mechanism of bleomycin: Evidence for a rate-determining 4'-hydrogen abstraction from poly(dA-dU) associated with the formation of both free base and base propenal, *Biochemistry* **24**, 7562-7568.
11. Burger, R. M., Berkowitz, A. R., Peisach, J., and Horwitz, S. B. (1980) Origin of malondialdehyde from DNA degraded by Fe(II)-bleomycin, *J. Biol. Chem.* **255**, 11832-11838.
12. Giloni, L., Takeshita, M., Johnson, F., Iden, C., and Grollman, A. P. (1981) Bleomycin-induced strand-scission of DNA, *J. Biol. Chem.* **256**, 8608-8615.
13. Sugiyama, H., Xu, C., Murugesan, N., and Hecht, S. M. (1985) Structure of the alkali-labile product formed during iron(II)-bleomycin-mediated DNA strand scission, *J. Am. Chem. Soc.* **107**, 4101-4105.
14. Rabow, L. E., Stubbe, J., and Kozarich, J. W. (1990) Identification and quantitation of the lesion accompanying base release in bleomycin-mediated DNA degradation, *J. Am. Chem. Soc.* **112**, 3196-3203.
15. Deeble, D. J., Schulz, D., and von Sonntag, C. (1986) Reaction of  $\cdot$ OH radicals with poly(U) in deoxygenated solutions: Sites of OH attack and the kinetics of base release, *Int. J. Radiat. Biol.* **49**, 915-926.
16. von Sonntag, C. (1987) *The chemical basis of radiation biology*, Taylor and Francis, London.
17. Demple, B., and Harrison, L. (1994) Repair of oxidative damage to DNA: Enzymology and Biology, *Annu. Rev. Biochem.* **63**, 915-948.
18. Berti, P. J., and McCann, J. A. (2006) Toward a detailed understanding of base excision repair enzymes: Transition state and mechanistic analyses of N-glycoside hydrolysis and N-glycoside transfer, *Chem. Rev.* **106**, 506-555.
19. Pogozelski, W. K., and Tullius, T. D. (1998) Oxidative strand scission of nucleic acids: Routes initiated by hydrogen abstraction from the sugar moiety, *Chem. Rev.* **98**, 1089-1107.
20. Zhou, X., Liberman, R. G., Skipper, P. L., Margolin, Y., Tannenbaum, S. R., and Dedon, P. C. (2005) Quantification of DNA strand breaks and abasic sites by oxime derivatization and accelerator mass spectrometry: Application to gamma-radiation and peroxyxynitrite, *Anal. Biochem.* **343**, 84-92.
21. Zhou, X., Taghizadeh, K., and Dedon, P. C. (2005) Chemical and biological evidence for base propenals as the major source of the endogenous M1dG adduct in cellular DNA, *J. Biol. Chem.* **280**, 25377-25382.
22. Kappen, L. S., and Goldberg, I. H. (1992) Neocarzinostatin acts as a sensitive probe of DNA microheterogeneity: Switching of chemistry from C-1' to C-4' by a G-T mismatch 5' to the site of DNA damage, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 6706-6710.
23. Kappen, L. S., Goldberg, I. H., Frank, B. L., Worth, L., Jr., Christner, D. F., Kozarich, J. W., and Stubbe, J. (1991) Neocarzinostatin-induced hydrogen atom abstraction from C-4' and C-5' of the T residue at a d(GT) step in oligonucleotides: Shuttling between deoxyribose attack sites based on isotope selection effects, *Biochemistry* **30**, 2034-2042.
24. Kappen, L. S., and Goldberg, I. H. (1989) Identification of 2-deoxyribonolactone at the site of neocarzinostatin-induced cytosine release in the sequence d(AGC), *Biochemistry* **28**, 1027-1032.
25. Chen, C. B., Gorin, M. B., and Sigman, D. S. (1993) Sequence-specific scission of DNA by the chemical nuclease activity of 1,10-phenanthroline-copper(I) targeted by RNA, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 4206-4210.
26. Pope, L. M., Reich, K. A., Graham, D. R., and Sigman, D. S. (1982) Products of DNA cleavage by the 1,10-phenanthroline-copper complex. Inhibitors of *Escherichia coli* DNA polymerase I, *J. Biol. Chem.* **257**, 12121-12128.
27. Kuwabara, M., Yoon, C., Goyne, T., Thederahn, T., and Sigman, D. S. (1986) Nuclease activity of 1,10-phenanthroline-copper ion: Reaction with CGCGAATTCGCG and its complexes with netropsin and EcoRI, *Biochemistry* **25**, 7401-7408.
28. Povirk, L. F. (1996) DNA damage and mutagenesis by radiomimetic DNA-cleaving agents: Bleomycin, neocarzinostatin and other enediynes, *Mutat. Res.* **355**, 71-89.
29. Xu, Y. J., Kim, E. Y., and Demple, B. (1998) Excision of C-4'-oxidized deoxyribose lesions from double-stranded DNA by human apurinic/apyrimidinic endonuclease (Ape1 protein) and DNA polymerase  $\beta$ , *J. Biol. Chem.* **273**, 28837-28844.
30. Xu, Y. J., DeMott, M. S., Hwang, J. T., Greenberg, M. M., and Demple, B. (2003) Action of human apurinic endonuclease (Ape1) on C1'-oxidized deoxyribose damage in DNA, *DNA Repair* **2**, 175-185.
31. Kim, J., Gil, J. M., and Greenberg, M. M. (2003) Synthesis and characterization of oligonucleotides containing the C4'-oxidized abasic site produced by bleomycin and other DNA damaging agents, *Angew. Chem., Int. Ed.* **42**, 5882-5885.

32. Chen, J., and Stubbe, J. (2004) Synthesis and characterization of oligonucleotides containing a 4'-keto abasic site, *Biochemistry* 43, 5278–5286.
33. Greenberg, M. M., Weledji, Y. N., Kroeger, K. M., Kim, J., and Goodman, M. F. (2004) In vitro effects of a C4'-oxidized abasic site on DNA polymerases, *Biochemistry* 43, 2656–2663.
34. Kroeger, K. M., Kim, J., Goodman, M. F., and Greenberg, M. M. (2004) Effects of the C4'-oxidized abasic site on replication in *Escherichia coli*. An unusually large deletion is induced by a small lesion, *Biochemistry* 43, 13621–13627.
35. Drohat, A. C., Jagadeesh, J., Ferguson, E., and Stivers, J. T. (1999) Role of electrophilic and general base catalysis in the mechanism of *Escherichia coli* uracil DNA glycosylase, *Biochemistry* 38, 11866–11875.
36. Le Grice, S. F. J., Cameron, C. E., and Benkovic, S. J. (1995) Purification and characterization of human immunodeficiency virus Type 1 reverse transcriptase, *Methods Enzymol.* 262, 130–144.
37. Smallcombe, S. H., Patt, S. L., and Keifer, P. A. (1995) A WET solvent suppression and its applications to LC NMR and high-resolution NMR spectroscopy, *J. Magn. Reson., Ser. A* 117, 295–303.
38. Sklenar, V., Piotto, M., Leppik, R., and Saudek, V. (1993) Gradient-tailored water suppression for H-1-N-15 HSQC experiments optimized to retain full sensitivity, *J. Magn. Reson., Ser. A* 102, 241–245.
39. Sheng, S., and van Halbeek, H. (1998) Accurate and precise measurement of heteronuclear long-range couplings by a gradient-enhanced two-dimensional multiple-bond correlation experiment, *J. Magn. Reson.* 130, 296–299.
40. Schmieder, P., Ippel, J. H., van den Elst, H., van der Marel, G. A., van Boom, J. H., Altona, C., and Kessler, H. (1992) Heteronuclear NMR of DNA with the heteronucleus in natural abundance: Facilitated assignment and extraction of coupling constants, *Nucleic Acids Res.* 20, 4747–4751.
41. Sklenar, V., and BAX, A. (1987) Measurement of H-1-P-31 NMR Coupling-Constants in Double-Stranded DNA Fragments, *J. Am. Chem. Soc.* 109, 7525–7526.
42. Bax, A., and Lerner, L. M. (1988) Measurement of <sup>1</sup>H-<sup>1</sup>H Coupling-Constants in DNA Fragments by 2D NMR, *J. Magn. Reson.* 130, 296–299.
43. Lankhorst, P. P., Haasnoot, C. A., Erkelens, C., and Altona, C. (1984) In conformational analysis of nucleic acid fragments. 2. A reparametrization of the Karplus equation for vicinal NMR coupling constants in CCOP and HCOP fragments, *J. Biomol. Struct. Dyn.* 1, 1387–1405.
44. Borgias, B. A., and James, T. L. (1989) Two-dimensional nuclear Overhauser effect: Complete relaxation matrix analysis, *Methods Enzymol.* 176, 169–183.
45. James, T. L., Borgias, B. A., Bianucci, A. M., and Zhou, N. (1990) Determination of DNA and protein structures in solution via complete relaxation matrix analysis of 2D NOE spectra, *Basic Life Sci.* 56, 135–154.
46. Case, D. A., Darden, T. E., Cheatham, T. E., III, Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Merz, K. M., Wang, B., Pearlman, D. A., Crowley, M., Brozell, S., Tsui, V., Gohlke, H., Mongan, J., Hornak, V., Cui, G., Beroza, P., Schafmeister, C., Cadwell, J. W., Ross, W. S., and Kollman, P. A. (2004) *Amber 8*, University of California, San Francisco.
47. Bayly, C. I., Cieplak, P., Cornell, W. D., and Kollman, P. A. (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: The RESP model, *J. Phys. Chem.* 97, 10269–10280.
48. Cieplak, P., Cornell, W. D., Bayly, C. I., 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.
49. Frisch, M. J., et al. (2001) *Gaussian 98*, Gaussian Inc., Pittsburgh, PA.
50. Pigache, A., Cieplak, P., and Dupradeau, F.-Y. (2004) Automatic and highly reproducible RESP and ESP charge derivation: Application to the development of programs RED and X RED, 227th National Meeting of the American Chemical Society, Anaheim, CA, March 27–April 1.
51. Smith, J. A., GomezPaloma, L., Case, D. A., and Chazin, W. J. (1996) Molecular dynamics docking driven by NMR-derived restraints to determine the structure of the calicheamicin  $\gamma$ (I)(1) oligosaccharide domain complexed to duplex DNA, *Magn. Reson. Chem.* 34, S147–S155.
52. Dupradeau, F. Y., Case, D. A., Yu, C., Jimenez, R., and Romesberg, F. E. (2005) Differential solvation and tautomer stability of a model base pair within the minor and major grooves of DNA, *J. Am. Chem. Soc.* 127, 15612–15617.
53. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995) A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules, *J. Am. Chem. Soc.* 117, 5179–5197.
54. Cheatham, T. E., Cieplak, P., and Kollman, P. A. (1999) A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat, *J. Biomol. Struct. Dyn.* 16, 845–862.
55. Wijmenga, S. S., Mooren, M. W., and Hilbers, C. W. (1993) *NMR of Macromolecules: A Practical Approach* (Roberts, G. C. K., Ed.) pp 217–283, Oxford University Press, Oxford, U.K.
56. Manoharan, M., Ransom, S. C., Mazumder, A., Gerlt, J. A., Wilde, J. A., Withka, J. A., and Bolton, P. H. (1988) The characterization of abasic sites in DNA heteroduplexes by site specific labeling with <sup>13</sup>C, *J. Am. Chem. Soc.* 110, 1620–1622.
57. Ahmad, T., Andersson, R., Olsson, K., and Westerlund, E. (1993) On the formation of reductive acid from pentoses or hexuronic acids, *Carbohydr. Res.* 247, 217–222.
58. Hoehn, S. J., Turner, C. J., and Stubbe, J. (2001) Solution structure of an oligonucleotide containing an abasic site: Evidence for an unusual deoxyribose conformation, *Nucleic Acids Res.* 29, 3413–3423.
59. Wilde, J. A., Bolton, P. H., Mazumder, A., Manoharan, M., and Gerlt, J. A. (1989) Characterization of the equilibrating forms of the aldehydic abasic site in duplex DNA by <sup>17</sup>O NMR, *J. Am. Chem. Soc.* 111, 1894–1896.
60. van Wijk, J., Huckriede, B. D., Ippel, J. H., and Altona, C. (1992) Furanose sugar conformations in DNA from NMR coupling constants, *Methods Enzymol.* 211, 286–306.
61. Junker, H. D., Hoehn, S. T., Bunt, R. C., Marathius, V., Chen, J., Turner, C. J., and Stubbe, J. (2002) Synthesis, characterization and solution structure of tethered oligonucleotides containing an internal 3'-phosphoglycolate, 5'-phosphate gapped lesion, *Nucleic Acids Res.* 30, 5497–5508.
62. Jourdan, M., Garcia, J., Defrancq, E., Kotera, M., and Lhomme, J. (1999) 2'-Deoxyribonolactone lesion in DNA: Refined solution structure determined by nuclear magnetic resonance and molecular modeling, *Biochemistry* 38, 3985–3995.
63. Beger, R. D., and Bolton, P. H. (1998) Structures of apurinic and apyrimidinic sites in duplex DNAs, *J. Biol. Chem.* 273, 15565–15573.
64. Goljer, I., Kumar, S., and Bolton, P. H. (1995) Refined Solution Structure of a DNA Heteroduplex Containing an Aldehydic Abasic Site, *J. Biol. Chem.* 270, 22980–22987.
65. Wang, K. Y., Parker, S. A., Goljer, I., and Bolton, P. H. (1997) Solution structure of a duplex DNA with an abasic site in a dA tract, *Biochemistry* 36, 11629–11639.
66. Kalnik, M. W., Chang, C. N., Grollman, A. P., and Patel, D. J. (1988) NMR studies of abasic sites in DNA duplexes: Deoxy-adenosine stacks into the helix opposite the cyclic analogue of 2-deoxyribose, *Biochemistry* 27, 924–931.
67. Coppel, Y., Berthet, N., Coulombeau, C., Garcia, J., and Lhomme, J. (1997) Solution conformation of an abasic DNA undecamer duplex d(CGCACXCACGC)•d(GCGTGTGTGCG): The unpaired thymine stacks inside the helix, *Biochemistry* 36, 4817–4830.
68. de Los Santos, C., El-Khateeb, M., Rege, P., Tian, K., and Johnson, F. (2004) Impact of the C1' configuration of abasic sites on DNA duplex structure, *Biochemistry* 43, 15349–15357.
69. Rachofsky, E. L., Seibert, E., Stivers, J. T., Osman, R., and Ross, J. B. (2001) Conformation and dynamics of abasic sites in DNA investigated by time-resolved fluorescence of 2-aminopurine, *Biochemistry* 40, 957–967.
70. Stivers, J. T. (1998) 2-Aminopurine fluorescence studies of base stacking interactions at abasic sites in DNA: Metal-ion and base sequence effects, *Nucleic Acids Res.* 26, 3837–3844.
71. Brauns, E. B., Madaras, M. L., Coleman, R. S., Murphy, C. J., and Berg, M. A. (1999) Measurement of local DNA reorganization on the picosecond and nanosecond time scales, *J. Am. Chem. Soc.* 121, 11644–11649.
72. Demple, B., and DeMott, M. S. (2002) Dynamics and diversions in base excision DNA repair of oxidized abasic lesions, *Oncogene* 21, 8926–8934.
73. Chaudhry, M. A., Dedon, P. C., Wilson, D. M., III, Demple, B., and Weinfeld, M. (1999) Removal by human apurinic/aprimidinic

- endonuclease I (Ape 1) and *Escherichia coli* exonuclease III of 3'-phosphoglycolates from DNA treated with neocarzinostatin, calicheamicin, and gamma-radiation, *Biochem. Pharmacol.* 57, 531–538.
74. Suh, D., Wilson, D. M., III, and Povirk, L. F. (1997) 3'-Phosphodiesterase activity of human apurinic/aprimidinic endonuclease at DNA double-strand break ends, *Nucleic Acids Res.* 25, 2495–2500.
75. Mol, C. D., Izumi, T., Mitra, S., and Tainer, J. A. (2000) DNA-bound structures and mutants reveal abasic DNA binding by APE1 and DNA repair and coordination, *Nature* 403, 451–456.
76. Hosfield, D. J., Guan, Y., Haas, B. J., Cunningham, R. P., and Tainer, J. A. (1999) Structure of the DNA repair enzyme endonuclease IV and its DNA complex: Double-nucleotide flipping at abasic sites and three-metal-ion catalysis, *Cell* 98, 397–408.
77. Parikh, S. S., Mol, C. D., Slupphaug, G., Bharati, S., Krokan, H. E., and Tainer, J. A. (1998) Base excision repair initiation revealed by crystal structures and binding kinetics of human uracil-DNA glycosylase with DNA, *EMBO J.* 17, 5214–5226.
78. Lau, A. Y., Scharer, O. D., Samson, L., Verdine, G. L., and Ellenberger, T. (1998) Crystal structure of a human alkylbase-DNA repair enzyme complexed to DNA: Mechanisms for nucleotide flipping and base excision, *Cell* 95, 249–258.

BI6024269