Solution structure of duplex DNA containing an extrahelical abasic site analog determined by NMR spectroscopy and molecular dynamics

Zhen Lin, Kwong-Ngai Hung, Arthur P. Grollman and Carlos de los Santos*

Department of Pharmacological Sciences, State University of New York at Stony Brook, Stony Brook, NY 11794-8651, USA

Received January 20, 1998; Revised and Accepted April 2, 1998

ABSTRACT

Translesional DNA synthesis past abasic sites proceeds with the preferential incorporation of dAMP opposite the lesion and, depending on the sequence context, one or two base deletions. High-resolution NMR spectroscopy and molecular dynamics simulations were used to determine the three-dimensional structure of a DNA heteroduplex containing a synthetic abasic site (tetrahydrofuran) residue positioned in a sequence that promotes one base deletions. Analysis of NMR spectra indicates that the stem region of the duplex adopts a right-handed helical structure and the glycosidic torsion angle is in anti orientation for all residues. NOE interactions establish Watson–Crick alignments for all canonical base pairs of the duplex. Measurement of distance interactions at the lesion site shows the abasic residue excluded from the helix. Restrained molecular dynamics simulations generated three-dimensional models in excellent agreement with the spectroscopic data. These structures show a regular duplex region and a slight bend at the lesion site. The tetrahydrofuran residue extrudes from the helix and is highly flexible. The model reported here, in conjunction with a previous study performed on abasic sites, explains the structural bias of one-base deletion mutations.

INTRODUCTION

The loss of DNA bases by the spontaneous hydrolysis of N-glycosyl bonds results in the formation of abasic sites (apurinic/apyrimidinic sites, AP sites). This process, which in mammalian cells may produce up to 10,000 AP sites per day (1), leaves primarily an inter-converting mixture of 2′-deoxy-α-D-ribose, 2′-deoxy-β-D-ribose and a small percentage of the open form of the sugar. Base modifications that favor the hydrolytic process, and the repair of damaged bases by DNA glycosylases, accelerate the formation of AP sites (2). Naturally occurring abasic sites are somewhat unstable, because the aldehydic (open) form is subject to β-elimination and strand scission (3). This limitation can be overcome by using 3-hydroxy-2-(hydroxymethyl)-tetrahydrofuran (F), a closed-form analog of AP sites (4,5). DNA containing this modified tetrahydrofuran group is efficiently cleaved by endonuclease IV, exonuclease III and ApeI, the major AP endonucleases involved in DNA repair in Escherichia coli and mammalian cells, and serves as a template for AMV reverse transcriptase as well as prokaryotic and eukaryotic DNA polymerases (5–10).

The mutagenic properties of natural and synthetic AP sites have been investigated using randomly or site-specifically modified substrates. In vitro studies using bacterial DNA polymerases revealed the preferential incorporation of dAMP opposite AP sites (5,6,11). In E.coli, translesional synthesis past abasic sites is facilitated by induction of the SOS response, (12) and proceeds mainly by incorporation of dAMP opposite the lesions and a small percentage of deletion mutations (13,14). DNA sequences containing an abasic site flanked by a thymidine residue at the 5′ side have been identified as a ‘hot spot’ for Δ1 mutations (15). In contrast, a broader mutational spectrum is observed in eukaryotic cells. In simian kidney cells, replication of double-stranded vectors containing a single AP site leads to the preferential incorporation of dAMP (16). In the same cellular system, replication of a single-stranded vector carrying a unique abasic site showed random incorporation of bases opposite the lesion (17). In human lymphoblastoid cells, deletions and frameshift mutations accounted for 55% of the mutations and, among single base substitutions, incorporation of dGMP was prevalent (18). On the other hand, preferential incorporation of dCMP opposite a unique natural AP site has been reported in Saccharomyces cerevisiae (19).

Three-dimensional structures of DNA duplexes containing natural and synthetic abasic sites have been solved by NMR spectroscopy. In these studies, the duplex retained a B-form geometry and the AP site induced only local perturbations, limited to the lesion site and flanking base pairs (20–27). In duplexes containing purine (dG or dA) residues opposite the lesion, both the abasic site and opposing purine are stacked inside the helix (20,22–25,27). However, when pyrimidine (dC or dT) bases are opposite the AP site, the duplex structure shows increased conformational flexibility. Depending on the temperature, the abasic sugar residue, and sometimes the opposing pyrimidine base, are extrahelical (21,26). Less is known about the structure of DNA duplexes containing abasic site lesions in the context of frameshift mutations. Proton NMR spectroscopy and molecular mechanics have been used to determine the structure of a (GC)-rich oligodeoxynucleotide duplex containing an extrahelical abasic site residue (28). In the refined models, the abasic sugar adopted two conformations outside of the helix, associated with a different pucker.

*To whom correspondence should be addressed. Tel: +1 516 444 3649; Fax: +1 516 444 3218; Email: cds@pharm.sunysb.edu
on the 3′-flanking sugar. In the present study, we describe the solution structure of a d(G-A-A-C-T-F-G-C-T-A-C) d(G-T-A-G-C-A-G-T-T-C) heteroduplex (referred to throughout this paper as the AP-duplex), where F is a 3-hydroxy-2-(hydroxymethyl)tetrahydrofuran residue (Fig. 1), as determined by high resolution NMR spectroscopy and restrained molecular dynamics simulations. The numbering scheme for the duplex sequence employed in this paper is shown in Figure 1.

MATERIALS AND METHODS

Synthesis and purification

The synthesis of a modified tetrahydrofuran residue and incorporation into DNA sequences by phosphoramidite chemistry procedures has been described (5). Purification of products by reverse phase HPLC and isolation of the sodium salt follows methods previously reported (29). NMR samples consisted of 3.2 × 10⁻³ mol of AP-duplex dissolved in 0.6 ml of 10 mM phosphate buffer, pH 6.9, containing 50 mM NaCl and 1 mM EDTA in either 99.96% D₂O or 90% H₂O–10% D₂O.

NMR methods

NMR spectra were recorded on Varian spectrometers operating at 600 and 500 MHz. Proton spectra were referenced relative to the (2,2,3,3-d₄) sodium 3-trimethylsilyl-propionate signal at 0 p.p.m., and phosphorous spectra relative to TMP. Phase-sensitive (30) NOESY (50, 90, 140, 200, 270 and 300 ms mixing times), COSY, DQF-COSY and TOCSY (120 ms isotropic mixing time) spectra in D₂O buffer were collected at 25°C with a repetition delay of 1.5 s, during which the residual water signal was suppressed by saturation. A NOESY (220 ms mixing time) spectrum in H₂O buffer was recorded at 5°C using a jump-return reading pulse (31). Time domain data sets consisted of 2048 by 256 complex data points in the t₂ and t₁ dimensions, respectively. A [1H–3¹³P]-correlation spectrum was recorded at 25°C using indirect detection mode (32) with 2048 by 128 complex data points in t₂ and t₁ dimensions, respectively. NMR data were processed using Felix95 (Biosym) running on Silicon Graphics computers. Time domain data were multiplied by shifted sinebell window functions prior to Fourier transformation. No baseline correction was applied to the transformed spectra.

Computational methods

Molecular dynamics (MD) simulations were performed on Silicon Graphics and Sun computers using X-PLOR 3.1 (33). Visualization of the structures was done using Insight95 (Biosym) and Midas Plus (UCSF, Computer Graphics Laboratory) and the structural parameters obtained with Curves (34,35).

Simulations were run ‘in vacuum’ using an all atom force field derived from CHARMM (36). Partial atomic charges on phosphate groups were not reduced, resulting in deoxynucleotide residues with a net charge of –1. Cross-peak volumes measured from all NOESY spectra collected in D₂O buffer were input to the program MARDIGRAS (UCSF) which calculates interproton distances, taking into consideration spin diffusion effects (37). A single set of distance bounds was obtained by adding ± 0.4 Å to the averaged value determined by MARDIGRAS. A total of 300 experimental distances were enforced during the simulation using an empirical square-well energy potential function with a force constant of 30 kcal/(mol Å²). Watson–Crick alignments were enforced by distance restraints on all canonical base pairs of the duplex. Backbone dihedral angles outside the lesion site were restrained by an empirical square well potential function with a width encompassing a range for both A- and B-form DNA. The SHAKE algorithm (38) was employed to maintain the length of all covalent bonds involving hydrogen atoms. At the lesion site, the ε (C₄–C₅–O₅–P) torsion angles of T5 and F6 residues were determined based on their experimental J_HH–P coupling constant, using the Karplus relation J = 15.3 cos²⁹ε – 6.1 cos φ + 1.6 (39) and ε = (–φ – 120°). Of the four possible solutions, a cis conformational range (ε = 330–30°) was considered forbidden. Subsequently, potential energy functions with a double minimum between 100–115° and 210–270° were enforced during the simulation. Similarly, based on the large J_HH (5'–P coupling constant value, the β (P–O₅–C₅–C₆) torsion angle of F6 was constrained within the gauche⁺ (30–90°) or gauche⁻ (270–330°) conformational ranges by using a double-minimum potential energy function.

Sugar puckers and glycosidic torsion angles were not restrained during the simulations.

Five different initial models containing a tetrahydrofuran residue were built by replacing the aromatic ring of adenosine by a proton. Two of the initial models were 10mer duplexes, derived from B-form DNA, having an additional tetrahydrofuran residue in the major or minor groove of the helix. The other three models were derived from 11mer duplexes by deleting a T on one strand and modifying a dA residue, either complementary to the deleted T or staggered towards the 3′ or 5′ direction. The resulting structures were termed F-major, F-minor, F-opposite, F-3′-staggered and F-5′-staggered. Initial models were energy-minimized by the conjugate gradient method prior to MD. Simulations consisted of an equilibration step during which the temperature was increased from 100 to 500 K in 4 ps and the scale of the NOE distance restraint potential function was gradually augmented. Once the equilibration temperature was obtained, the simulation continued at 500 K for 20 ps, followed by cooling to 300 K in 10 ps, and 105 ps of restrained dynamics at constant temperature. Seven MD simulations were performed, starting from each initial model, using a different set of initial velocities. Coordinates of the last 5 ps of the simulation were averaged and subjected to 1000 steps of energy minimization, yielding distance-refined structures. Converging structures were defined as those having rms deviations in the helical region of the duplex smaller than 0.5 Å with respect to the average structure.
RESULTS

NMR spectra of non-exchangeable protons

The 1D proton NMR spectrum of the AP-duplex recorded in D₂O buffer, pH 6.9, 25°C is characterized by the presence of sharp signals and the absence of minor resonances (Supplementary Material, Fig. 1). Assignment of the non-exchangeable base and sugar protons followed the analysis of NOESY, COSY and TOCSY spectra using standard procedures (40,41). Figure 2 shows distance interactions detected on a NOESY (300 ms mixing time) spectrum between the base (purine H8/pyrimidine H6) and the sugar H1′ proton regions. Indicative of a right-handed helix, each base proton of the duplex exhibited NOE peaks to its own and 5′-flanking H1′ sugar protons. The presence of the tetrahydrofuran residue did not impede the observation of sequential distance interactions at the lesion site. On the lesion-containing strand, NOE peaks were observed between G7(H8) and T5(H1′, H3′) (Fig. 2, peaks A and E, respectively) as well as T5(H6, H2′, H2″) protons (data not shown). On the complementary strand, normal distance interactions connected A17(H8) and C16(H1′, H2′, H2″) protons. In addition, inter-strand connectivities were detected between A17(H2) and the H1′ sugar protons of T5 and G7 (Fig. 2, peaks B and D, respectively).

Assignment of the tetrahydrofuran protons started with the identification of COSY interactions between F6(H1′/H1′′) and F6(H2′/H2″) protons, which occurred in a region of the spectrum devoid of other signals (Supplementary Material, Fig. 2). Subsequent analysis of DQF-COSY, TOCSY (120 ms isotropic mixing time) and NOESY (100 ms mixing time) spectra completed their assignment. None of the tetrahydrofuran protons showed distance interactions to the base or H1′, H2′, H2″ and H3′ sugar protons of flanking T5 and G7 residues. The chemical shift of the non-exchangeable protons measured at 25°C is listed in Table 1.

Table 1. Proton and phosphorous chemical shifts

<table>
<thead>
<tr>
<th></th>
<th>H8/H6</th>
<th>H2/H5/H3</th>
<th>H1′</th>
<th>H2′</th>
<th>H2″</th>
<th>H3′</th>
<th>H4′</th>
<th>31P</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1</td>
<td>7.81</td>
<td>5.50</td>
<td>2.43</td>
<td>2.65</td>
<td>4.80</td>
<td>4.16</td>
<td>12.62</td>
<td>–</td>
</tr>
<tr>
<td>A2</td>
<td>8.17</td>
<td>7.46</td>
<td>5.91</td>
<td>2.74</td>
<td>2.88</td>
<td>5.04</td>
<td>4.39</td>
<td>–</td>
</tr>
<tr>
<td>A3</td>
<td>8.11</td>
<td>7.71</td>
<td>6.12</td>
<td>2.62</td>
<td>2.82</td>
<td>5.01</td>
<td>4.45</td>
<td>–</td>
</tr>
<tr>
<td>C4</td>
<td>7.20</td>
<td>5.09</td>
<td>5.77</td>
<td>1.83</td>
<td>2.39</td>
<td>4.68</td>
<td>4.19</td>
<td>–</td>
</tr>
<tr>
<td>T5</td>
<td>7.30</td>
<td>1.51</td>
<td>6.07</td>
<td>2.43</td>
<td>2.43</td>
<td>4.91</td>
<td>4.21</td>
<td>13.24</td>
</tr>
<tr>
<td>F6</td>
<td>–</td>
<td>–</td>
<td>3.99</td>
<td>2.25</td>
<td>2.12</td>
<td>4.63</td>
<td>4.13</td>
<td>–</td>
</tr>
<tr>
<td>G7</td>
<td>7.58</td>
<td>–</td>
<td>5.87</td>
<td>2.44</td>
<td>2.63</td>
<td>4.68</td>
<td>4.36</td>
<td>12.58</td>
</tr>
<tr>
<td>C8</td>
<td>7.47</td>
<td>5.21</td>
<td>5.81</td>
<td>2.09</td>
<td>2.44</td>
<td>4.70</td>
<td>4.19</td>
<td>–</td>
</tr>
<tr>
<td>T9</td>
<td>7.42</td>
<td>1.59</td>
<td>5.70</td>
<td>2.10</td>
<td>2.45</td>
<td>4.85</td>
<td>4.12</td>
<td>13.67</td>
</tr>
<tr>
<td>A10</td>
<td>8.25</td>
<td>7.57</td>
<td>6.23</td>
<td>2.64</td>
<td>2.83</td>
<td>4.99</td>
<td>4.38</td>
<td>–</td>
</tr>
<tr>
<td>C11</td>
<td>7.26</td>
<td>5.18</td>
<td>6.01</td>
<td>2.11</td>
<td>2.12</td>
<td>4.27</td>
<td>3.99</td>
<td>–</td>
</tr>
<tr>
<td>G12</td>
<td>7.90</td>
<td>–</td>
<td>5.94</td>
<td>2.56</td>
<td>2.76</td>
<td>4.79</td>
<td>4.20</td>
<td>12.64</td>
</tr>
<tr>
<td>T13</td>
<td>7.36</td>
<td>1.38</td>
<td>5.70</td>
<td>2.13</td>
<td>2.45</td>
<td>4.88</td>
<td>4.19</td>
<td>13.67</td>
</tr>
<tr>
<td>A14</td>
<td>8.17</td>
<td>3.79</td>
<td>6.05</td>
<td>2.70</td>
<td>2.88</td>
<td>5.03</td>
<td>4.39</td>
<td>–</td>
</tr>
<tr>
<td>G15</td>
<td>7.61</td>
<td>–</td>
<td>5.73</td>
<td>2.46</td>
<td>2.63</td>
<td>4.91</td>
<td>4.37</td>
<td>12.88</td>
</tr>
<tr>
<td>C16</td>
<td>7.17</td>
<td>5.12</td>
<td>5.61</td>
<td>1.86</td>
<td>2.30</td>
<td>4.72</td>
<td>4.14</td>
<td>–</td>
</tr>
<tr>
<td>A17</td>
<td>8.01</td>
<td>7.46</td>
<td>5.94</td>
<td>2.61</td>
<td>2.82</td>
<td>4.87</td>
<td>4.33</td>
<td>–</td>
</tr>
<tr>
<td>G18</td>
<td>7.63</td>
<td>–</td>
<td>5.85</td>
<td>2.50</td>
<td>2.70</td>
<td>4.91</td>
<td>4.37</td>
<td>12.74</td>
</tr>
<tr>
<td>T19</td>
<td>7.18</td>
<td>1.29</td>
<td>5.97</td>
<td>2.53</td>
<td>2.53</td>
<td>4.78</td>
<td>4.19</td>
<td>13.89</td>
</tr>
<tr>
<td>T20</td>
<td>7.39</td>
<td>1.60</td>
<td>6.14</td>
<td>2.51</td>
<td>2.51</td>
<td>4.86</td>
<td>4.13</td>
<td>14.04</td>
</tr>
<tr>
<td>C21</td>
<td>7.51</td>
<td>5.55</td>
<td>6.22</td>
<td>2.27</td>
<td>2.27</td>
<td>4.55</td>
<td>3.98</td>
<td>–</td>
</tr>
</tbody>
</table>

*Values in p.p.m. relative to TSP (1H) or TMP (31P) at 0 p.p.m.
*Measured at 29°C in D₂O phosphate buffer (10 nM), pH 6.9, 50 mM NaCl.
*Measured at 5°C in 10% D₂O phosphate buffer (10 mM), pH 6.9, 50 mM NaCl.
*Assignment of G1(N₁H) and G12(N₁H) imino protons could be reversed. nd, not detected.
Exchangeable protons

The 1D spectrum of the exchangeable protons recorded in 10% D₂O buffer, pH 6.8, 5°C showed seven partially resolved signals, between 12.2 and 14.2 ppm, which account for the 10 imino protons of the AP-duplex (Supplementary Material, Fig. 3). Their assignment followed the analysis of a NOE spectrum (Fig. 3). Labeled peaks are assigned as follows: (A) T20(N3H)–A2(H2); B, T19(N3H)–A3(H2); C, T13(N3H)–A10(H2); D, T9(N3H)–A14(H2); E, T5(N3H)–A17(H2); F, G18(N1H)–C4(N4H)hb; F’, G18(N1H)–C4(N4H)nhb; F”, G18(N1H)–C4(N4H)nhb; G, G15(N1H)–C8(N4H)hb; G’, G15(N1H)–C8(N4H)nhb; H, G7(N1H)–C16(N4H)hb; H’, G7(N1H)–C16(N4H)nhb; I, T20(N3H)–A3(H2); J, T19(N3H)–A2(H2); K, G15(N1H)–A14(H2); L, G18(N1H)–A17(H2); M, G18(N1H)–A3(H2) and N, G7(N1H)–A17(H2). (A) G18(N1H)–C4(N4H) nhb, and (N4H) nhb refer to the hydrogen-bonded and non hydrogen-bonded cytidine amino protons, respectively.

Phosphorous spectra

The 31P-NMR 1D spectrum of the AP-duplex at 25°C spans ~1.1 ppm, with two phosphates resonating slightly down-field from the normal region observed on DNA duplexes. The specific assignment of all phosphorous signals of the duplex followed the analysis of the 1H–31P correlation spectrum recorded in D₂O buffer, pH 6.8, 25°C, shown in Figure 4. Each phosphorous exhibited COSY cross-peaks to the H3′ proton of the 3′-flanking nucleotide. Thus, the two most down-field signals were assigned to T5–p-F₆ and F₆–p-G₇ flanking the tetrahydrofuran residue. Phosphorous chemical shifts measured at 25°C are listed in Table 1.

The ε torsion angles of the T₅ and F₆ residues were determined from the experimental J_H3′–p coupling constant values (Fig. 4, peaks A and B). In both cases, ε values were within the anti(+ac)/anti(−ac) ranges. Determination of J_H5′–p and J_H5−c coupling constants is generally impaired by the severe overlap of the H₅ and H₅′ proton signals and the lack of stereospecific assignments. Due to small J_H5′–p and J_H5−c cross-peaks are usually observed in DNA duplexes (Fig. 4). This characteristic is the result of a preferred β-trans conformation, which implies small values for both coupling constants simultaneously. On the AP-duplex, a strong COSY cross-peak was detected between the T₅–p-F₆ phosphorous and the overlapping F(H₅′, H₅′) protons (Fig. 4, peak C), excluding a trans conformation for the β torsion angle of the F residue. Since F(H₅′) and F(H₅−c) proton signals are isochronous, we could not differentiate between β-gauche− (J₅′–p > J₅−c–p) and β-gauche+ (J₅′–p > J₅−c–p) conformations.
Three-dimensional structures

MD simulations were performed with an average of 17 experimental restraints on unmodified nucleotides of the sequence, but only four on the tetrahydrofuran residue. Five refined structures were immediately rejected, because the final values of $\epsilon$ and $\beta$ torsion angles at the lesion site were outside the experimental bounds. The remaining structures exhibited rmsd in the atom position from 0.60 to 1.60 Å, with respect to their average, no violations of experimental dihedral restraints and no NOE-violation larger than 0.30 Å. Analysis of the structures showed that rmsd were unevenly distributed throughout the duplex. The helical part (excluding the tetrahydrofuran) was well defined, while the lesion site was quite flexible, having average rmsd of 0.60 and 1.40 Å, respectively. Thus, structures having rmsd $\leq$ 0.5 Å for the helical part (total of eight) were selected as converging models and used to describe the AP-duplex structure. Distribution of rmsd in the heavy atom positions is given in Table 2.

An overlap view of the converging structures, as seen from the major groove of the helix, is displayed in Figure 5. The structures belong to the B-form DNA family with sugar puckers on the C2'-endo/C1'-exo range and anti glycosidic torsion angles. The presence of the abasic residue induces a small bend on the helix in the direction of the major groove. However, Watson–Crick base-pair alignments remain intact throughout the helical region of the duplex. On the lesion-containing strand, the sugar–phosphate backbone is perturbed at the lesion site by the presence of the tetrahydrofuran residue, while the backbone of the partner strand is quite regular. Figure 6 shows a close view of the central 4 base pair segment of the duplex, as seen from the minor groove of the helix. Base pairs flanking the lesion are properly aligned and almost undisturbed. Only T5·A17, 5'-flanking the tetrahydrofuran residue, showed increased propeller twist and is slightly buckled. In contrast, normal values are seen for G7·C16. The helix is under twisted between the T5·A17 and G7·C16 base pairs, but this perturbation is immediately compensated for by a proportional over twist at the G7-C16–C8-G15 step.

The abasic lesion is extruded from the helix and, as suggested by the values of rmsd, its position is flexible. In some models, the tetrahydrofuran residue is inclined toward one of the grooves of the helix, while in others remains equidistant. Figure 7 depicts the position of the tetrahydrofuran residue, as seen from the helical axis looking in the direction of the sugar–phosphate backbone. The flexibility of the lesion is the result of the conformational freedom of the backbone angles. Of the torsion angles restrained during the simulation, only $\epsilon$ showed some propensity for the $+_{ac}$ range while $\beta$ exhibited no preference between the gauche$^+$ and gauche$^-$ ranges. Among the angles not restrained during the simulation, only $\delta$ ($C_5$–$C_4$–$C_3$–$O_3'$) showed some conformational preference for the $\pm_{sc}$ range. In contrast, the pucker of the

<table>
<thead>
<tr>
<th>Table 2. Rms-deviations for converging structures$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial models$^b$</td>
</tr>
<tr>
<td>---------------------</td>
</tr>
<tr>
<td>F-minor</td>
</tr>
<tr>
<td>F-minor</td>
</tr>
<tr>
<td>F-opposite</td>
</tr>
<tr>
<td>F-opposite</td>
</tr>
<tr>
<td>F-5′-staggered</td>
</tr>
<tr>
<td>F-5′-staggered</td>
</tr>
<tr>
<td>F-3′-staggered</td>
</tr>
<tr>
<td>F-3′-staggered</td>
</tr>
</tbody>
</table>

$^a$Values in Angstroms relative to an average structure calculated from 35 distance-refined models.

$^b$None of the refinements that started with the F residue on the major groove converged within 0.5 Å of the average. For each of the other initial structures, two refined models fulfilled the selection criteria.

$^c$Overall heteroduplex.

$^d$10mer helical region, only.

$^e$Abasic residue, only.

Figure 5. Overlapping view of the three-dimensional models obtained after restrained molecular dynamics as seen from the major groove. The picture shows the convergence of the T5·A17 (red) and G7·C16 (yellow) base pairs of the duplex contrasting with the conformational freedom exhibited by the tetrahydrofuran residue (blue).

Figure 6. Stereo view of the central d(C4-T5-F6-G7-C8)/d(G15-C16-A17-G18) segment as seen from the minor groove side. T5·A17 (red) and G7·C16 (yellow) exhibit proper Watson–Crick alignment while F6 (blue) is very flexible.
tetrahydrofuran residue was well defined, falling in the C2′-exo range (amplitude 12–34°) in the converging structures.

**DISCUSSION**

**NMR spectra**

The fact that we have detected sharp proton and phosphorous resonances, together with their assignment to one set of NMR signals, suggests that the AP-duplex adopts a single averaged structure in solution. The sequential distance connectivities observed in the NOESY spectra between the base (purine H8/pyrimidine H6) and the H1′, H2′, H2″, H3′ sugar protons of the 3′-attached nucleotide establishes a right-handed helical structure. Furthermore, weak intensities characterized the intra-residue base to H1′ NOE cross-peaks (Fig. 2), specifically at short mixing times, suggesting long distances among these protons of the AP-duplex. This observation is consistent with residues having glycosidic torsion angles on the anti conformation (40,41). Consistent with this orientation, Watson–Crick alignments were established by distance interactions detected in the NOESY spectrum recorded in 10% D2O buffer, between T(N3H)–A(H2) and G(N1H)–C(N4H1)/C(N4H2) protons (Fig. 3A, peaks A–E and F/F′–H/H′, respectively). The JH1′-H2′ and JH1′-H2″ cross-peak patterns detected in the phase-sensitive COSY spectrum together with the absence of strong NOE interactions between the base (purine H8/pyrimidine H6) and its own H3′ sugar proton, determined that sugar puckers are in the south conformation (42). Distance connectivities among the imino protons, and between imino and A(H2) protons of adjacent base pairs of the sequence (Fig. 3), indicated favorable base-pair stacking throughout the AP-duplex (41). The perturbation of the DNA backbone induced by the presence of a tetrahydrofuran residue is somewhat limited. The phosphorous spectra shows that only T5-p-F6 and F6-p-G7 resonate outside the region usually observed on B-form DNA duplexes (–4.5 to –5.0 p.p.m.) (Fig. 4). Thus, distortions of the sugar–phosphate backbone are limited to the lesion site on the modified strand.

The tetrahydrofuran residue

Several spectroscopic characteristics define structural features of the AP-duplex at the lesion site. NOE interactions encountered between G7(H8) and the H1′, H2′, H2″, H3′ sugar protons of T5 establish distances shorter than 4.5 Å between these protons of the duplex. Furthermore, in several regions of the NOESY spectra not hindered by signal overlap, we have failed to observe distance connectivities between tetrahydrofuran protons and the 3′- and 5′-flanking residues. This fact cannot be ascribed to broadening of the tetrahydrofuran signals due to molecular motion, since they were clearly identified in the phase-sensitive COSY, COSY45, DQF-COSY and TOCSY spectra without any apparent increase in the line width. Thus, we interpret the lack of distance interactions as confining the F residue to conformations outside the helix.

**AP-duplex structure**

Distance-restrained molecular dynamics was able to drive quite different initial models to structures having rmsd smaller than 0.7 Å with respect to their average. In evaluating the precision of the three-dimensional models, we must consider the disparity in the number of experimental restraints enforced during the simulation for different regions of the AP-duplex. The helical region of the duplex, with an average of 17 restraints/residue, showed excellent convergence, while the tetrahydrofuran, with only four experimental constraints, was highly flexible (Table 2).

The duplex structure is a fairly regular helix where the F residue has been forced out (Fig. 5). Stacking of the T5-A17 and G7-C16 base pairs is almost unaffected by the presence of the lesion. The tetrahydrofuran H1″ proton points toward the bulk solvent suggesting that, in natural abasic sites, the OH group at this...
position could be readily stabilized by hydrogen bonding interactions. Our three-dimensional model resembles the structure reported by Cuniasse et al., where an extrahelical F residue is flanked by G-C base pairs (28). Perhaps the most notable difference between the structures is the small bend in the helical axis observed here (Table 3). It is possible that the reason for this subtle difference is the presence of a less stable A·T base pair 5'-flanking the lesion.

As mentioned above, the limited number of experimental restraints impeded the precise determination of the tetrahedralfruran conformations in the AP-duplex structure. There was no apparent spectroscopic reason, other than distances >5 Å, that could account for the absence of distance interactions between protons of the abasic and flanking residues. Molecular dynamics simulations showed that a large number of conformations are accessible to the F residue, while maintaining long distances with the flanking bases. The lack of conformational preference exhibited by the backbone dihedral angles of the F residue makes possible the pronounced preference of DNA polymerases to incorporate dAMP opposite the AP sites, sequence context, is the result of the pronounced preference of thymidine residue can such a structure be stabilized by Watson–Crick hydrogen bonds, thereby facilitating translesion synthesis and resulting in the observed ‘hot spot’ for Δ−1 mutations induced by AP sites. The three-dimensional model presented here is essentially identical to the structures reported by Cuniasse et al., where an F residue was flanked by G-C base pairs (28). However, in vitro and in vivo mutagenic studies performed to date have failed to observe frameshift mutations in such a sequence context. Therefore, the lack of Δ−1 mutations, in a 5'-dG lesion-flanking sequence context, is the result of the pronounced preference of DNA polymerases to incorporate dAMP opposite the AP sites, and not the consequence of the three-dimensional structure of the mutagenic intermediate.

ACKNOWLEDGEMENTS

We thank Mr Robert Rieger and Ms Cecilia Torres for preparing and purifying modified oligodeoxynucleotides. Atomic coordinates of the NMR refined structures have been deposited in the Brookhaven Data Bank: This research is supported by NIH Grant CA74995. The NMR facility at SUNY Stony Brook is supported by NSF Grants CHE8911350 and CHE9413510 and NIH Grant 1S10RR554701.

REFERENCES