High-Field NMR and Restrained Molecular Modeling Studies on a DNA Heteroduplex Containing a Modified Apurinic Abasic Site in the Form of Covalently Linked 9-Aminoellipticine†

Malvinder P. Singh,‡ G. Craig Hill,§ Didier Péoc'h,∥ Bernard Rayner,∥ Jean-Louis Imbach,∥ and J. William Lown*,‡

Department of Chemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2G2, College of Pharmacy, Xavier University of Louisiana, New Orleans, Louisiana 70125, and Laboratorie de Chimie Bio-Organique, Universite des Sciences et Techniques du Languedoc UA 488 du CNRS, 34060 Montpellier, France

Received November 4, 1993; Revised Manuscript Received June 8, 1994®

ABSTRACT: Two-dimensional NMR methods were used to model the possible solution structure of an intercalative complex of 9-aminoellipticine (Aell), a polycyclic pyridocarbazolamine, covalently bound to an apurinic ring-opened deoxyribose site of a duplex DNA fragment in the reduced Schiff base form. The required oligonucleotide single strand containing covalently attached aminoellipticine was obtained by reductive amination in the presence of sodium cyanoborohydride. The combined NMR–energy minimization methods were employed to refine the model structures of two distinct forms, intrahelical and extrahelical, of a control 9-mer duplex DNA, d(CGTG·dr·GTGC)·d(GCACTCACG), which contains an apurinic site positioned opposite a dT residue on the complementary strand. The model structure of an aminoellipticine conjugate with the same DNA sequence, derivatized via the aforementioned covalent attachment, was also obtained by incorporating intermolecular drug–DNA and intra- and internucleotide NOE-derived proton–proton distance estimates as restraints in energy minimization routines. The indole ring system of aminoellipticine, which is inserted at the apurinic site, intercalates between and is parallel to flanking GC base pairs. The pyridinic ring of aminoellipticine, in protonated form, also stacks between cytidine and thymidine bases on the complementary strand, which is consistent with the observation that the normal sequential NOE connectivity at the 5′-C^{13}-T^{14} step is broken and indeed diverted through the ellipticine moiety, e.g., C^{13}-Aell-T^{14} connectivities through the Aell-H4/C5Me protons. Interestingly, the partial stacking of the pyridinic ring is observed only between the 5′-CT step vs an adjacent 5′-TC step, owing to inherently weak stacking interactions associated with the former. In the absence of any potential groups that can participate in electrostatic or hydrogen-bonding interactions with the nucleic acid, π–π stacking and hydrophobic contacts at the intercalation site appear to be the important factors in determining stability and conformation of the aminoellipticine–DNA conjugate. Stacking interactions in such a bistranded intercalative complexation of aminoellipticine apparently govern the formation of a single intrahelical form of a right-handed B-type DNA duplex. The overall structural features lead us to propose working models for an enzyme-like DNA cleavage activity of 9-aminoellipticine and the observed inhibition of the AP endonuclease-dependent DNA excision–repair pathway.

Several aspects of DNA processing in biological systems have been exploited as attractive targets for intervention by therapeutic agents, and a number of DNA-binding agents that presumably inhibit replication and transcription are in clinical use (Potier, 1992; Propst & Perun, 1992; Lown, 1993). Another potential and yet unexplored target is the DNA repair systems, which are largely responsible for the repair of DNA lesions induced spontaneously or through exogenous assault by ionizing radiation and chemically reactive alkylating agents, etc. (Lindahl, 1993). The repair of naturally or chemically damaged DNA in both prokaryotes and eukaryotes is usually controlled by specific enzymes which are able to recognize particular DNA lesions or a nucleic acid local structure containing such lesions. These repair systems use a diverse combination of key enzymes for various transformations (Lindahl, 1982; Sancar & Sancar, 1988; Wallace, 1988).

The DNA excision–repair pathway is an example of removal of a damaged nucleobase involving a glycosylase enzyme that detects an individual unnatural base and cleaves it from the deoxyribose sugar, initially leaving the sugar–phosphate backbone intact (Friedberg, 1985; Lindahl, 1982, 1993). The gap created by the loss of a nucleobase is called an AP site (for apurinic or apyrimidinic) which in turn is recognized by an AP endonuclease that cuts the backbone to afford a primer end from which DNA polymerase initiates synthesis to replace the missing nucleotide along with a few adjacent nucleotides. Knowledge of how AP endonucleases recognize these specific DNA lesions is central to an understanding of such efficient repair processes. It is also important to know the detailed three-dimensional structure of the abasic site-containing DNA sequences to understand the possible roles of precise conformational features. However, few reports are available concerning structural details of the nicked duplex structure (Cuinasse et al., 1990; Kalnik et al., 1988, 1989; Withka et al., 1991). Complementary information toward these goals can also be obtained through characterization of DNA–ligand

† Financial support for this work was provided by the Natural Science and Engineering Research Council of Canada (to J.W.L.), the Centers of Excellence Grant HHS-2-D3400006-07 (to G.C.H.), and grants (to J.L.I.) from L'Association pour la Recherche centre le Cancer (ARC) of France and Isis Pharmaceuticals at Carlsbad, CA.

* Author to whom correspondence should be addressed.
‡ University of Alberta.
§ Xavier University of Lousiana.
∥ CNRS.

® Abstract published in *Advance ACS Abstracts,* August 1, 1994.

0006-2960/94/0433-10271$04.50/0

complexes, particularly with agents that influence the repair processes by directly interacting with the abasic sites.

Ellipticines are a general class of coplanar annelated polycyclic compounds that interact with DNA in a primarily intercalative mode and possess in vitro as well as in vivo cytostatic activity (Potier, 1992). A number of synthetic ellipticine derivatives have been reported in an effort to rationally design potent biologically active analogues. Aminoellipticine (Aell), or 5,11-dimethyl-6*H*-pyrido[4,3-*b*]carbazol-9-amine, is one such agent that has been used as a mechanistic probe for enzymes involved in the base excision–repair pathways, wherein Schiff base formation with the apurinic sites leading to covalent adduct formation has been implicated as evidence for a selective inhibition of the AP endonuclease activity of *Escherichia coli* exonuclease III (Bertrand et al., 1989a; Lefrançois et al., 1990). The reaction of Aell, together with that of other primary amines and other reagents reactive toward the ring-opened aldehydic forms of the sugar at the abasic sites (Liuzzi et al., 1987; Livingstone, 1964; Vasseur et al., 1987, 1988), also serves as a model for the cleavage activity of AP endonuclease via a β-elimination of 5′-phosphonomonoesters of double-stranded DNA (Bertrand et al., 1989a).

The present study was undertaken in an attempt to understand the structural characteristics of a specific duplex DNA–aminoellipticine conjugate and its plausible role in the mechanisms of DNA cleavage at AP sites by aminoellipticine and of inhibition of the DNA repair process and why only minor modifications of this agent result in significant changes in activity (Malvy et al., 1986). We report on the model solution structure of a DNA oligomer duplex containing a covalently attached 9-aminoellipticine, in the form of a reduced Schiff base, at a specific abasic site positioned opposite a dT residue on the complementary strand. These are the first definitive experimental results which establish that the covalent attachment of aminoellipticine at such a gap formed by an AP site is characterized by insertion of the coplanar 4-ring ellipticine system into the DNA helix in an intercalative mode with additional evidence of partial stacking of the pyridinic ring between the dipyrimidinic 5′-CT step on the partner strand.

MATERIALS AND METHODS

Oligonucleotide Syntheses

In the procedures described below, all the reactions, handling, and storage of the derivatives containing the photosensitive *o*-nitrobenzyl group were carried out in the absence of light. Purifications by preparative high-performance liquid chromatography (HPLC) were performed using a Nucleosil C_{18} (5 μm) column (250 × 10 mm^2) obtained from SFCC-Shandon (France). The HPLC system (Waters, Millipore) employed a two-pump gradient controller interfaced with a NEC APC IV computer system. Analytical HPLC analysis employed the same system and a Beckman C_{18} XL-ODS (3 μm) column (70 × 4.6 mm^2). Solvents for chromatography were of HPLC grade and used without further purifications. The 2-cyanoethyl *N*,*N*-diisopropylphosphoramidite of *o*-nitrobenzyl 2-deoxy-D-ribofuranoside in its 5′-dimethoxytrityl-protected from (**1**) was prepared from 3,5-ditoluoyl-2-deoxyribofuranosyl chloride in four steps as described previously (Péoc'h et al., 1991). The desired product was isolated by silica gel chromatography and its purity confirmed by ^{1}H and ^{31}P NMR spectroscopy.

d(GCACTCACG). This nonamer deoxyribonucleotide was synthesized on an Applied Biosystems 381A synthesizer using β-cyanoethyl phosphoramidites and 10 μmol columns. The oligonucleotide was deprotected and purified by ion-exchange chromatography on a DEAE-Sephadex A25 (Pharmacia) column using triethylammonium hydrogen carbonate buffer (pH 7.5) as eluent (150 mL of a 0.5 M solution followed by a linear gradient from 0.5 to 1 M).

*d(CGTG·X·GTGC), where X Represents an o-Nitrobenzyl 2-Deoxy-*D*-ribofuranoside Residue.* This DNA sequence was prepared using a solid-phase phosphoramidite synthetic cycle starting with a DMT-protected dC attached to a DPG support on a 10 μmol column. After sequential couplings and the usual protocol for extending the chain with dG, dT, and dG, respectively, the phosphoramidite building block of *o*-nitrobenzyl 2-deoxy-D-ribofuranoside (**1**) was introduced in the fourth cycle, and further extension continued to afford efficient and site-specific incorporation of a protected deoxyribose residue (without the nucleobase) within the 9-mer oligonucleotide. After the application of the usual deprotection procedure, the oligonucleotide was passed through an ion-exchange DEAE-Sephadex A25 (HCO_3 form; Pharmacia) column (20 × 1.7 cm^2). The column was eluted with a linear gradient of triethylammonium bicarbonate buffer (0.5 → 1.0 M; pH 7.5) at a flow rate of 1 mL/min. Fractions corresponding to the main peak (at 254 nm) were collected, combined, and evaporated to dryness. The residue was dissolved in water (5 mL) and passed through a Dowex 50W (Na^+ form; Fluka) column. Elution with water afforded fractions with UV 254 nm absorbance which were combined and lyophilized. Purity (in excess of 97%) of the sodium form of the *o*-nitrobenzyl-protected abasic oligomer was estimated by HPLC analysis.

Abasic d(CGTG·dr·GTGC). A solution of the preceding oligomer (19.7 mg, 6.7 mmol) in 0.05 M aqueous HOAc (6 mL) was degassed, distributed into two thermostated (21–24 °C) quartz cuvettes, and irradiated with a high-pressure mercury lamp (HPK, 125; Phillips) through a pyrex filter (3 mm thickness) for 20 min. The solutions were combined and lyophilized to afford a residue which was gel filtered through a Sephadex G-25 column (70 × 2 cm^2) using 5 mM aqueous HOAc as eluent. The fractions containing the desired abasic nonamer were combined, passed through a 0.45 μm filter, and lyophilized to afford the product in 58% yield and 94% purity as determined by HPLC analysis.

Aminoellipticine–DNA Conjugate d(CGTG·Aell·GTGC). To a 0.7 mM aqueous solution (7 mL) of the preceding abasic oligomer was added a freshly prepared 0.25 M solution (2 mL) of $NaCNBH_3$ in 1 M aqueous NaOAc (pH 5) followed by a 8 mM solution (1 mL) of 9-aminoellipticine in 20 mM aqueous HCl. The resulting mixture was stirred for 1 h at 37 °C and centrifuged. The supernatant was further passed through a 0.45 μm filter, and the filtrate was concentrated to a final volume of 4 mL. This solution was loaded in five portions onto a preparative reverse-phase HPLC column and eluted with a linear gradient of acetonitrile (5–20%) in 0.05 M triethylammonium acetate (pH 7). Fractions corresponding to the main peak (R_t 15.9 min) were combined and lyophilized to afford a residue which was dissolved in water and passed through a Na^+ Dowex 50W column. The fractions indicating a HPLC purity in excess of 98% were combined and lyophilized to afford the desired modified oligomer in 90% yield. Owing to a slow degradation of this conjugate at pH 7, all further manipulations and handling of this compound in the solution state were done in pH 5.8 buffer.

NMR Spectroscopy

NMR samples were prepared by dissolving 12.4 mg of the 9-mer oligonucleotide d(GCACTCACG) in 600 μL of 10 mM phosphate buffer (pH 5.8) containing 10 mM sodium chloride and 0.1 mM Na_2EDTA. The solution was divided into two equal portions. Heteroduplex I containing the abasic site was prepared by annealing one of these portions with an equivalent amount of oligonucleotide **3** added in the form of a 0.3 mL solution in the same buffer as above. Heteroduplex II was obtained by similar treatment of the second portion with a 0.3 mL solution of the single-stranded aminoellipticine–DNA conjugate **4**. The annealing to the double-stranded forms was checked by comparing NMR spectra of the samples with those of the single-stranded forms. In each case, the sample was lyophilized to dryness and, for NMR experiments carried out in D_2O, further lyophilized twice from 99.9% D_2O and finally redissolved in 0.65 mL of 99.96% D_2O. For experiments in H_2O, the solid was redissolved in 0.65 mL of 9:1 H_2O/D_2O mixture.

All NMR experiments were performed on a Varian Unity 500 spectrometer, interfaced with a Sun workstation, at 499.2 MHz ^1H frequency. All experiments were temperature regulated by the spectrometer-interfaced computer and cooling units using a dry ice bath; the temperature was monitored via a thermocouple implant in the probe. The experiments performed in the 9:1 H_2O/D_2O mixture utilized a Sklenar–Bax water suppression pulse sequence (Sklenar & Bax, 1987) with a 90° pulse of 8.5 μs and a sweep width of 10 kHz. The delay time between the 90° pulses was optimized to 7.4 ms with a spin–lock pulse of 2 ms duration, in order to maximize the HOD suppression and the selective detection of the exchangeable imino proton signals. The data were Fourier transformed with a 3 Hz line broadening.

The two-dimensional NOESY spectra with mixing times of 50, 80, 150, 250, and 400 ms were acquired in the phase-sensitive mode using the hypercomplex method (States et al., 1982). The water (HOD) resonance was presaturated during the 2.2 s delay between scans. Sixty-four scans for each of the 512 t_1 values were collected with 2K data points. The spectral windows in F1 and F2 were adjusted to cover the entire spectral region that contained the resonance signals. In all experiments, the free induction decay, fid, along t_1 was zero filled to 2K data points prior to Fourier transformation to give a final 2K × 2K data matrix with digital resolution of 3.2 Hz/pt in each dimension. Data processing for each experiment was optimized using interactive sine-bell apodization and shifted sine-bell constants in t_1 and t_2 for resolution enhancement and a first-order polynomial drift correction in both F1 and F2 dimensions. The longer mixing time NOESY spectra were analyzed for the purposes of complete assignment of the nonexchangeable proton signals, while the interproton distances were estimated from the magnitude of cross peaks in NOESY spectra, acquired with a relatively long recycle time of 5 s and a relatively short mixing time of 150 ms, and using the isolated spin pair approximation (ISPA) relationship:

$$r_{ij} = (v_{ref}/v_{ij})^{1/6} r_{ref}$$

The cross peak volume integrals were determined for the individual cross peaks by evaluating the averaged volumes (peak height × area) for a series of equally sized rectangles circumscribing the peak which accounted for corrections (usually less than 10% of the peak volume) due to base line fluctuations. The cytosine H5—H6 interproton vector was taken to be a reference distance of 2.46 Å (Gronenborn & Clore, 1985), and the distance estimates were accurate compared with other fixed distances in the molecule, e.g., thymidine H6—Me and intraresidue H2′—H2″, etc.

Computations and Restrained Molecular Modeling

A normal B-DNA duplex, d(CGTGAGTGC)·d(GCACTCACG), was constructed using the assisted molecular building and energy refinement program AMBER 4.0 (Pearlman et al., 1991) and Arnott's coordinates (Arnott & Selsing, 1975) and further transformed into either heteroduplex I or II as follows.

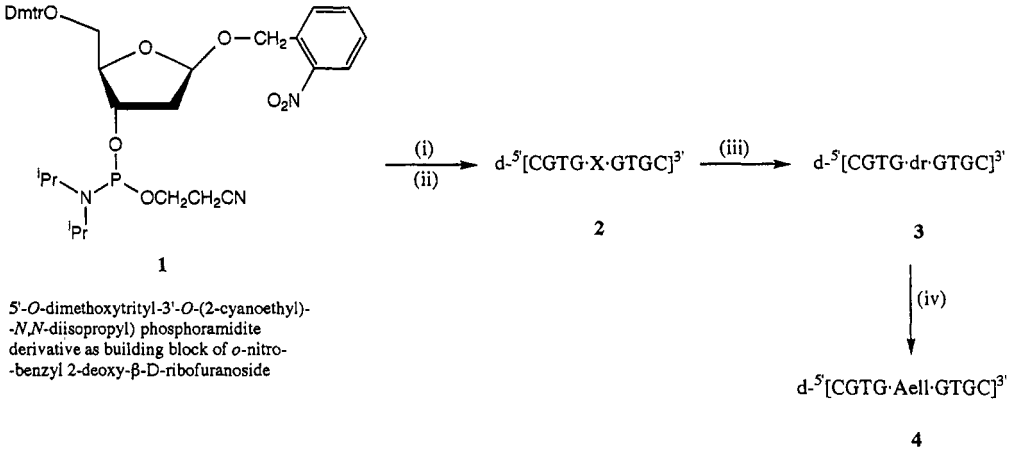
In order to obtain the coordinates for the abasic residue in heteroduplex I, those of the adenosine residue at position 5 in the first strand were imported from the AMBER 4.0 database into the program Macromodel V3.5X (Still, 1992) followed by deletion of the purine base atoms except for the N9 nitrogen atom of the glycosidic bond, which was replaced by an oxygen atom while retaining the β-configuration at the anomeric center. The ring-closed hemiacetal form of the abasic sugar residue was obtained by adding a hydrogen atom to this new oxygen. The resulting coordinates were minimized using the MM2 force field parameters (Still, 1992) and further optimized using the program MOPAC ESP 5.0 (Besler et al., 1990). The data on the final geometry and electrostatic potential charges of the abasic residue were used to generate an input file to be incorporated back for modeling the heteroduplex I in AMBER 4.0.

The aminoellipticine-derivatized sugar residue for heteroduplex II was constructed using Macromodel V3.5X and the preceding abasic residue. In this case, the C1′—O1′ bond of the deoxyribose sugar and the glycosidic C1′—N9 bonds were broken and the 9-Aell molecule was attached through its amino group to the C1′-position of the ring-opened sugar. Hydrogen atoms were added in compliance with the chemical structure shown in Figure 1c. Minimization of the Aell-derivatized sugar residue using MM2 parameters followed by optimization using MOPAC ESP 5.0 provided the final geometry and electrostatic potential charges that were employed to build the input file on the Aell residue of heteroduplex II for subsequent AMBER 4.0 simulations.

The model intra- and extrahelical conformations of the heteroduplex I were simulated and minimized using the AMBER 4.0 package. The initial intrahelical conformation was obtained by replacing the coordinates for the adenosine residue at position 5 in the normal B-DNA duplex with those obtained above for the abasic residue. The extrahelical conformation, where both dr^5 and the dT^{14} residue on the complementary strand are protruding out of the helix, was generated using MidasPlus (Ferrin et al., 1988), by rotating these two residues approximately 180° through the major groove around the helical axis. The transformed coordinates were then brought back in AMBER 4.0. All simulations utilized the all atom parameters (Weiner et al., 1984), except for angle OS—CT—OH, which was defined from the similar angle CT—CT—OH (Table S1, supplementary material).

Typically, minimizations were first done in vacuo with a distance-dependent dielectric constant until the rms deviation was less than 0.1 kcal/mol Å. The method of steepest descent was used for the first 100 cycles and for 10 cycles after each update of the nonbonded pair list; all subsequent cycles employed the conjugate gradient method. In order to attenuate the charge on the phosphate groups, 16 sodium counterions were then added to the model, each positioned at a distance of 3.0 Å from the phosphorus atoms within the phosphate O—P—O planes. The now charge-balanced system was

Scheme 1: Syntheses of the Single-Strand DNA Sequences Containing True Abasic Lesion Site **3** and Its Covalently Modified Aminoellipticine Conjugate **4**[a]



1

5'-*O*-dimethoxytrityl-3'-*O*-(2-cyanoethyl)--*N*,*N*-diisopropyl) phosphoramidite derivative as building block of *o*-nitro--benzyl 2-deoxy-β-D-ribofuranoside

2

3

4

[a] (i) Incorporation of **1** in the fourth cycle of an automated solid support DNA synthesizer cycle starting with 5'-DMT-dC; (ii) deprotection, purification by ion-exchange chromatography, and conversion to the Na^+ form (for details see Materials and Methods; X = *o*-nitrobenzyl 2-deoxy-β-D-ribofuranoside); (iii) *hν*, 20 min, 0.05 M acetic acid; dr = 2-deoxy-D-ribofuranoside; (iv) 9-aminoellipticine in the presence of $NaBH_3CN$, NaOAc, pH 5, 1 h, 37 °C; Aell represents the covalent adduct formed between the abasic site dr and aminoellipticine via its NH_2 group as shown in Figure 1.

reminimized in vacuo under the same conditions as above. The solvation of the resulting models was accomplished by immersion in a box of TIP3P water molecules with a minimum solvent shell thickness of 10 Å from DNA and counterions. The solvated models were then subjected to a "belly" minimization and molecular dynamics at constant volume followed by molecular dynamics at constant pressure and reminimization (Dr. William Ross, University of California at San Francisco, private communication). In the first solvated minimization, a "belly" is placed on the DNA and counterions so that they are "frozen" and only the surrounding water molecules are allowed to move. This was done for 200 cycles with a 9.0 Å cutoff (all solvated simulations use this value), switching from the method of steepest descent to that of conjugate gradient after 100 cycles. The next 1000 cycles (0.001 ps step size), intended to allow free movement of the water molecules around the DNA and counterions, involved raising the temperature from 10 to 300 K at 5 K intervals, under constant volume conditions. After such randomization of water, the velocities were discarded and the simulations continued, without the "belly" constraints, by raising the temperature over the first 10 ps as before and increasing the time step to 0.002 ps. This was followed by 40 ps of productive dynamics under the same conditions, while the temperature was kept regulated at 300 ± 5 K. The final complex without the "belly" was reminimized as above until the rms deviation was <0.1 kcal/mol Å.

Simulations and minimizations on the aminoellipticine-derivatized heteroduplex II, both with and without the experimental NOE distance constraints, were also performed using the AMBER 4.0 package and general procedures illustrated above. The positively charged form of the aminoellipticine moiety, in compliance with the NMR assignments (vide infra) that indicate protonation of the pyridinic ring at the N2 position, was used. Several new parameters were developed on the basis of those in the all atom force field (Table S1, supplementary material). Similar energy values were chosen from those already existing, and the angles were defined according to the geometry of the Aell-derivatized sugar residue obtained from MOPAC ESP 5.0, as discussed above. The first model of heteroduplex II was simulated without the NOE distance constraints. In the second case, these constraints were introduced gradually, beginning at the 5 ps mark, during equilibration dynamics and maintained at this level until the end of simulation. The models obtained were energy minimized again after removing the "belly" until the rms devication was less than 0.1 kcal/mol Å.

RESULTS

Synthesis of d(CGTG·dr·GTGC) and d(CGTG·Aell·GTGC)

The synthesis of two nonadexoyribonucleotide strands containing a true abasic site and a aminoellipticine attached covalently to the same abasic site, respectively, is outlined in Scheme 1. Oligonucleotide **2** was synthesized using the automated solid-phase phosphoramidite method with high yields for the coupling reaction at the stage of introducing the *o*-nitrobenzyl 2-deoxy-D-ribofuranoside phosphoramidite **1**. After photolytic cleavage of the *o*-nitrobenzyl group using a high-pressure mercury lamp, the desired abasic oligonucleotide **3** was purified by gel filtration.

Derivatization of the same aldehydic abasic oligomer with aminoellipticine via a covalent attachment was accomplished efficiently by carrying out their coupling reaction in the presence of sodium cyanoborohydride to selectively reduce the imine intermediate to obtain the aminoellipticine–DNA conjugate **4** (Bertrand et al., 1989a).

NMR Analyses

The two duplex DNA sequences chosen in this study are shown in Figure 1 together with the numbering schemes. The structures of the modified sites are also indicated. The particular sequence context with a high GC content was selected to maximize stability, and its non-self-complementary asymmetric nature precludes possible complications due to self-association of individual strands. It is important to note the use of a true abasic site with the sugar residue in the same oxidation state as it is presumed to occur in the damaged AP sites of native abasic DNA intermediates.

Before describing the results, it should be mentioned that the properties of DNA duplexes containing analogues of the abasic site have been studied by NMR in previous investigations (Cuniasse et al., 1987, 1990; Kalnik et al., 1988, 1989).

Heteroduplex I:

5'-	C^1	G^2	T^3	G^4	dr^5	G^6	T^7	G^8	C^9
3'-	G^{18}	C^{17}	A^{16}	C^{15}	T^{14}	C^{13}	A^{12}	C^{11}	G^{10}

Heteroduplex II:

5'-	C^1	G^2	T^3	G^4	$AeII^5$	G^6	T^7	G^8	C^9
3'-	G^{18}	C^{17}	A^{16}	C^{15}	T^{14}	C^{13}	A^{12}	C^{11}	G^{10}

Numbering and structures at individual nucleotide units are shown below for: (a) general case; (b) abasic residue dr^5; and (c) covalently modified aminoellipticine residue $AeII^5$

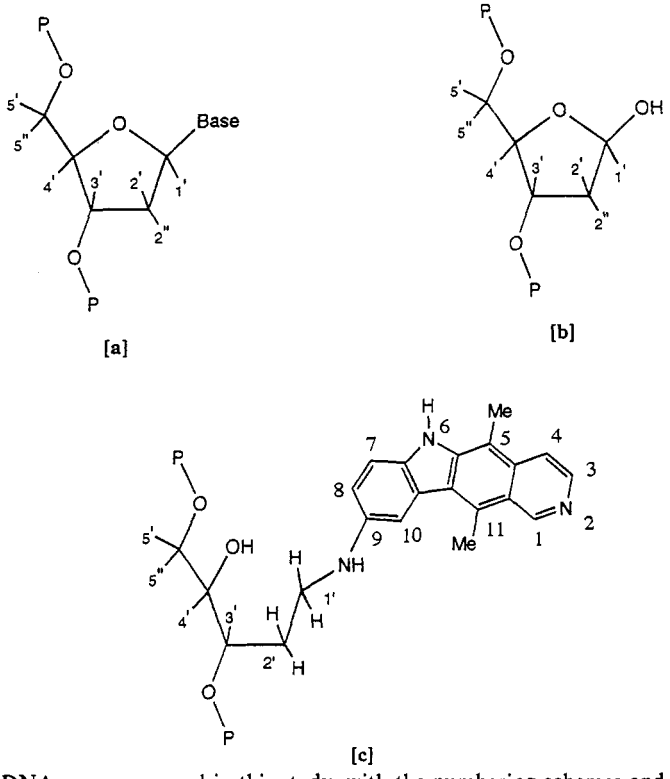


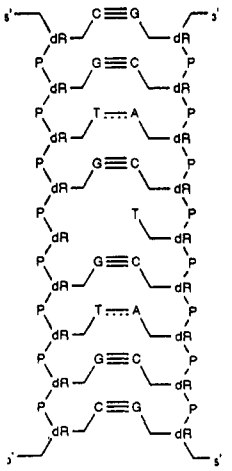
FIGURE 1: Heteroduplex DNA sequences used in this study, with the numbering schemes and the chemical structures of the abasic lesion site and the covalently modified 9-aminoellipticine adduct thereof. The degenerate sets of methylene protons are indicated in c as single labels 1′ and 2′.

We found the one by Cuniasse et al. (1990) to be very useful and most relevant for interpretations in the present work due to the same DNA sequence context as heteroduplex I (Figure 1), despite the difference in the nature of the abasic lesion site. The model sugar residue in that study, namely 1,4-anhydro-2-deoxy-D-ribitol or also termed the tetrahydrofuran analogue by some (Withka et al., 1991), did not feature the aldehydic oxidation state at the C1′-position. We focused our attention on the general structural features in that study that were proposed in relation to the abasic sites and the nature of the base residues on the complementary strand. Two such models for the systems with pyrimidine bases opposite the abasic site were considered (Figure 2). These two differ in the general configuration of the model sugar residues at the selected abasic lesion sites and illustrate a 'normal' DNA-like intrahelical conformation and a second distorted conformation where both the abasic residue and the residue positioned opposite it are displaced out of the helix.

In order to compare the possible structures for the true abasic site-containing sequence (heteroduplex I) and the aminoellipticine-derivatized duplex (heteroduplex II) under similar conditions, the NMR analyses were undertaken at pH 5.8, where the aminoellipticine–DNA conjugate was found to be stable toward degradation.

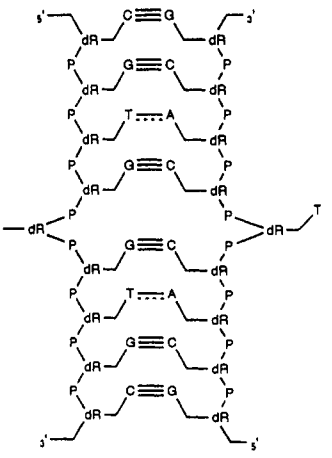
NMR Analysis of the Abasic Site-Containing Sequence

In making the assignments for the proton resonances of d(CGTG·dr·GTGC)·d(GCACTCACG) under the conditions of our NMR experiments (10 °C; pH 5.8 buffer), we were aided by the general trends and assignments reported previously for the same sequence with an analogue of the abasic site at 15 °C and pH 7 (Cuniasse et al., 1990). All nonexchangeable proton resonances (except for heavily overlapping signals due to 5′ and 5″ protons) were assigned in a sequential manner from 2D NOE, at mixing intervals of 100, 200, and 400 ms, and DQF-COSY spectra in D_2O following the well-established strategy for oligonucleotides [Feigon et al., 1983; Hare et al., 1983; Scheek et al., 1984; reviewed by Reid (1987)]. For example, the correlation cross peaks in the regions 7.0–8.5 and 5.0–6.5 ppm of 2D NOE spectra were used to assign aromatic and sugar C1′ proton resonances, and those in the 1.0–3.0 ppm region were used for the assignment of TMe and sugar C2′/2″ protons (Table S2, supplementary material). A predominantly B-form geometry for the duplex at 10 °C was confirmed from the relative intensities of the NOE cross peaks between base protons and their own and 5′-flanking sugar H1′, H2′, H2″, and H3′ protons, which provides the basis to distinguish between A- and B-type duplex helices. A complete map of B-DNA-type connectivities could be drawn out for the entire length of strand



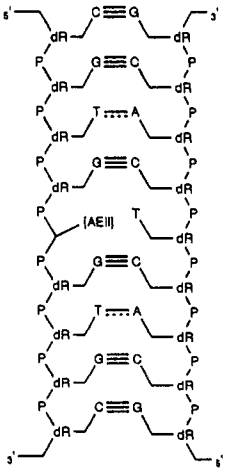
I_{intra}: abasic site and T^{14} intrahelical

[a]



I_{extra}: both abasic site and T^{14} extrahelical

[b]



Heteroduplex sequence II: Intrahelical incorporation of Aminoellipticine at the abasic site opposite the T^{14} residue

[c]

FIGURE 2: Two early models for the heteroduplex I, depicting the (a) intrahelical and (b) extrahelical arrangement of the abasic deoxyribose site and the oppositely positioned dT residue on the partner strand, on the basis of preliminary structures proposed by Cuniasse et al. (1990). Model c is a gross representation of the heteroduplex II used in this study and corresponds to the modified lesion site in the form of a covalently attached aminoellipticine as shown in Figure 1.

containing the unpaired T^{14} residue, consistent with T^{14} being a part of the helix.

The chemical shift value of 3.90 ppm for the glycoside H1′ proton at the abasic residue suggests that the sugar at the abasic residue exists predominantly in the cyclic deoxyribose (hemiacetal) form, since the H1′ proton in a ring-opened aldehydic form would be expected to show up at 8–10 ppm. Previous NMR work using specific ^{13}C-labeled derivatives (Manoharan et al., 1988) has also unequivocally established that there is minimal contribution (<1%) from the ring-opened aldehydic form of the sugar to the equilibrium mixture in solution. The ring-closed hemiacetal form can also equilibrate between the two anomeric forms due to α- or β-configuration of the OH group. However, a fast exchange of the OH group with solvent does not permit a firm distinction between the two configurations from ^1H NMR measurements.

The NOESY spectra recorded at temperatures higher than 10 °C required careful inspection of the NOE cross peaks, associated with the central trinucleotide segment containing the abasic site, for comparing the results with those in a previous NMR investigation of the same sequence with a chemically different abasic sugar moiety (Cuniasse et al., 1990). That study, employing the anhydro sugar analogue of the abasic site, had indicated an additional structurally distinct extrahelical form of the duplex predominating at 35 °C (in buffered solution containing 150 mM NaCl; pH 7) and where both the abasic sugar analogue and the oppositely positioned dT residue on the partner strand are located outside the regular helix. The overall conformation of this extrahelical form of the DNA heteroduplex was characterized by a doubling of resonances corresponding to the centrally located unpaired T^{14} residue. Such a feature was not observed under the conditions of our NMR experiments (pH 5.8; 10 mM NaCl), even upon warming the solution to 25 °C. Experiments performed above this temperature resulted in a degradation of heteroduplex I, possibly owing to a hydrolytic β-elimination reaction at the abasic lesion site (Bailly & Verly, 1987). Although the solution and experimental conditions employed in this study are also different from those in the previous study (Cuniasse et al., 1990), the results are qualitatively similar in terms of the gross conformational features associated with the extrahelical form. This is indicated in our interpretation in the form of a compression of the same strand that contains the abasic residue dr^5, which was characterized by NOEs between the flanking G^4 and G^6 residues on either side of the lesion, e.g., those corresponding to G^4H1'-G^6H8 and G^4H2''-G^6H8 at temperatures >20 °C, which may be explained by stacking of residues G^4 and G^6 to force the intervening dr^5 residue out of the helix. The minor differences between the results from the two studies may also arise from differences in the nature of the sugars at the abasic sites. As pointed out earlier, sugars in the aldehydic oxidation state were employed in this work, whereas the reduced 1,4-anhydro-2-deoxy-D-ribitol form was used previously (Cuniasse et al., 1990). This, and the use of different solution conditions, may account for the observed diversity in the nature of interconversion between intra- and extrahelical forms. Nevertheless, such marked conformational perturbation of duplex DNA at an abasic site could be responsible for the initial recognition and cleavage of such apurinic lesions by the DNA repair enzymes.

NMR Analysis of the 9-Aminoellipticine–DNA Adduct

A three-tier analysis of the experimental NMR data primarily in the form of NOE information was performed to obtain the most favored solution conformation of heteroduplex sequence II which contains a covalently attached aminoellipticine at the abasic site. The first two levels address the assignment of the resonances due to the AeII and DNA protons. These assignments were used to identify specific *intermolecular* NOE contacts between the ligand and DNA protons in what constitutes the last principal task in the analysis of both noncovalent and covalent ligand–DNA complexes in order to define the geometry and orientation of the ligand components.

Assignment of Aminoellipticine Proton Resonances in the Adduct. Two pairs of directly coupled aromatic spin systems, H3-H4 and H7-H8, were first identified via their respective

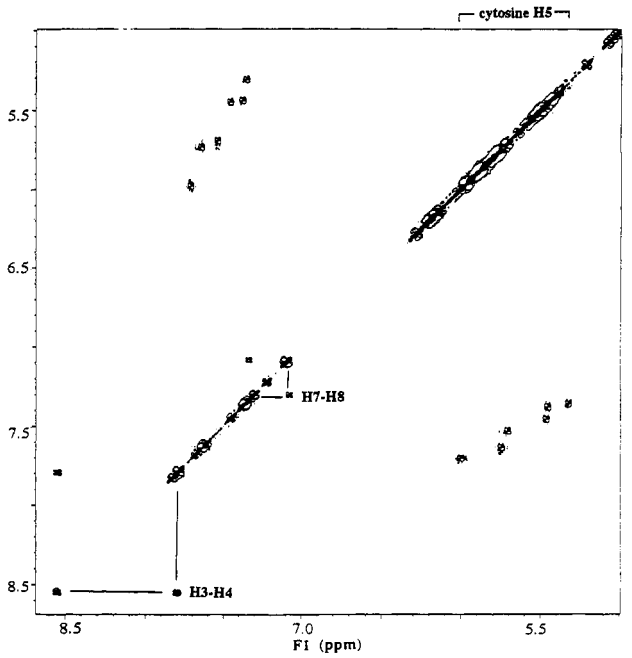


FIGURE 3: Selected region of a 500 MHz DQF-COSY spectrum of heteroduplex II, showing scalar-coupled cross peaks for aminoellipticine proton pairs in a region distinct from those containing the cytosine H6-H5 cross peaks.

cross peaks in a phase-sensitive COSY spectrum (Figure 3) in a region distinct from those containing the cross peaks from the cytosine H6-H5 spin systems of the DNA fragment. The cross peak corresponding to the H7-H8 pair was further assigned to Aell-H7 at 7.31 ppm and Aell-H8 at 7.04 ppm on the basis of chemical shift arguments and by analogy with our previous assignments on a trinucleotide segment containing covalently modified aminoellipticine (Bertrand et al., 1989a). The resonance signal at 8.55 ppm was assigned to Aell-H3 according to its coupling relationship with the Aell-H4 resonance at 7.79 ppm, which in turn also correlates (via NOE) with the Aell-C5Me resonance at 2.79 ppm in the 2D NOESY spectra. The pyridinic Aell-H1 proton appears most downfield, at 9.72 ppm, and could also be confirmed by its NOE cross peak with the Aell-C11Me signal at 2.38 ppm. The pyridinic ring of aminoellipticine (N2 position; Figure 4) is also believed to be in a protonated form under the slightly acidic conditions (pH 5.8) used for the experiments. This is evident from a downfield shift by ~0.5 ppm of the Aell-H1 and -H3 resonances compared with 9-aminoellipticine itself and is consistent with the experimentally determined pK_a values of 4.5 and 6.7 for protonation of the 9-NH_2 and pyridinic N2 groups, respectively, in aminoellipticine (Dr. C. Malvy, private communication). The labile amine and indole NH protons of aminoellipticine could not be assigned even from the experiments performed in 9:1 H_2O/D_2O, presumably because of fast exchange with solvent protons. The nonexchangeable proton resonances of the covalently attached aminoellipticine in the modified duplex are shown in Figure 4. A comparison of the chemical shift values obtained here in the case of aminoellipticine attachment, flanked on either side by G residues, with those from our previous analysis of aminoellipticine inserted between two T residues (Bertrand et al., 1989a) indicates a general upfield shift of all resonances associated with the indole ring system as a result of ring-current contributions from the flanking purine aromatics. This trend can be further attributed to an intercalative geometry for the aminoellipticine ring system between two GC base

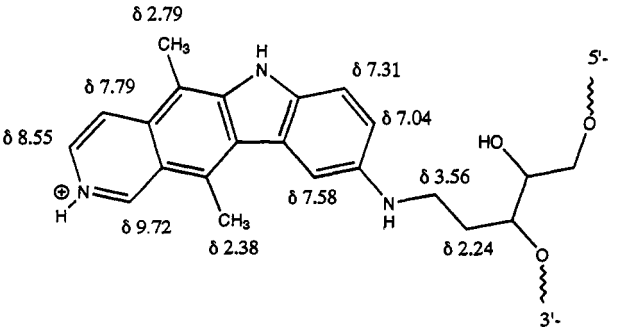


FIGURE 4: Proton resonance assignments for protonated aminoellipticine in its conjugated form in heteroduplex II.

pairs, as described below, from NOE observations and molecular modeling.

Nonexchangeable DNA Proton Resonances in the Modified Duplex. The well-established strategy of delineating sequential magnetization-transfer pathways in the 2D NOESY spectra of oligonucleotides [reviewed in Gronenborn and Clore (1985); Reid, 1987; van de Ven & Hilbers, 1988] was employed to obtain the assignments of nonexchangeable proton resonances in the aminoellipticine–DNA conjugate. This sequential mapping strategy for B-type DNA structures essentially relies on the relatively short (3–5 Å) distances between glycosidic protons (H1′, H2″, and H2′) and aromatic base protons (purine H8 and pyrimidine H6) on the same residue and to the ones on their 3′-neighbor residues. Additional cross checks on the assignments are provided by sequential connectivities involving thymine methyl and cytosine H5 via their proximity to protons on their respective 5′-neighbors.

The selected H8/H6-H1′ and H8/H6-H2″/H2′/TMe regions of a 250 ms mixing time NOESY spectrum are shown in Figure 5. For illustrative purposes, the NOE connectivity correlating base protons with the sugar H1′ protons is traced from G^{10} to G^{18} on the unmodified strand and from C^1 to C^9 on the modified strand (Figure 5A) containing the covalently attached aminoellipticine. Also shown in Figure 5B are the cross peaks for interactions with sugar H2′/H2″ and thymidine methyl groups in the upfield regions. The assignments for the nonexchangeable proton resonances obtained in this manner are provided in Table 1. Selected NOE enhancements between the protons on aminoellipticine and proximal DNA protons are labeled in the NOESY contour plots and discussed below.

The sequential NOE connectivity data acquired at 25 °C (shown in Figure 5) exhibits two important breaks (marked by X) reflecting structural perturbations associated with the incorporation of aminoellipticine at the abasic lesion site in the nonamer duplex. Thus, the NOE cross peak corresponding to internucleotide $C^{13}H1' \rightarrow T^{14}H6$ is lost, as is the one between G^4H1' and G^6H8. Note that the NOESY spectrum acquired at the same temperature for the heteroduplex I containing an abasic residue did show an NOE cross peak correlating G^4-H1′ and G^6H8, which was attributed to a collapsed helical DNA structure with an extrahelical protrusion of the abasic residue. In contrast, such a temperature-dependent conformational change is not observed for heteroduplex II containing the covalently attached aminoellipticine.

Intermolecular NOE Contacts and Distance Constraints for the Modified Lesion Site. The apparent broken connectivities in both the modified strand and its complement in the immediate vicinity of the aminoellipticine-containing lesion required further close inspection of the NOE relationships

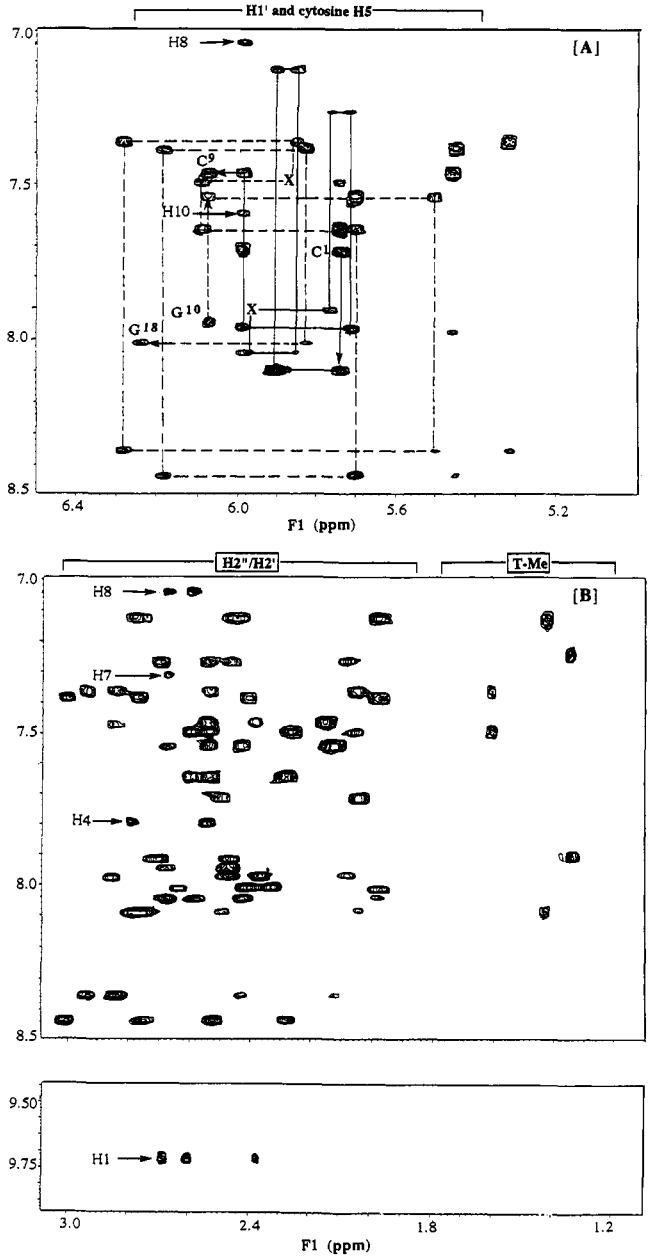


FIGURE 5: Expanded contour plots of the NOESY spectrum at 500 Hz of the heteroduplex II in D_2O, at 25 °C and pH 5.8. Part A corresponds to the expansion showing the NOE interactions between the aromatic base protons and the H1′/H5 protons. Sequential connectivities are traced from the residues C^1 to C^9 in strand A (solid lines) and from G^{10} to G^{18} in strand B (broken lines). Peaks labeled by X refer to breaks in the sequential NOE connectivity patterns, between G^4 and G^6 for the primary strand (along solid lines) and between C^{13} and T^{14} for the parner strand (along broken lines). Cross peaks labeled H8 and H10 refer to the NOE interactions between Aell-H8 ↔ G^4H1′ and Aell-H10 ↔ G^4H1′, respectively. Part B shows the H6/H8-H2′/H2″/Me region. Peaks along the labels H8, H7, H4, and H1 correspond to the NOE interactions between aminoellipticine and DNA. Specifically peaks are labeled as: along the label H8 (7.04 ppm), Aell-H8 → G4H2″ (2.67 ppm) and G4H2′ (2.60 ppm); along the label H7 (7.31 ppm), Aell-H7 → G4H2″ (2.67 ppm); along the label H4 (7.79 ppm), Aell-H4 → Aell-C5Me (2.79 ppm) and C13H2″ (2.52 ppm); and along the label H1 (9.72 ppm), Aell-H1 → G4-H2″ (2.67 ppm), G4H2′ (2.60 ppm), and Aell-C11Me (2.38 ppm).

between the aminoellipticine protons and the surrounding DNA protons. A number of such close *intermolecular* NOE contacts were identified on the basis of assignments obtained above for aminoellipticine and DNA resonances (Figure 6). Among these were specific contacts between the aminoellipticine ring system and the protons from the complementary DNA strand, e.g., the resonance due to the Aell-H4 proton at 7.79 ppm showed NOE cross peaks with C^{13}H2″ (2.54 ppm) and C^{13}H6 (7.36 ppm), as well as with T^{14}H6 (7.50 ppm). Similarly, the Aell-C5Me signal at 2.79 ppm correlates with both C^{13} (5.32 ppm for H5; 7.36 ppm for H6) and T^{14} (1.60 ppm for Me) residues. The aminoellipticine protons that lie close to the covalent linkage (H7, H8, and H10) also exhibit NOEs to individual DNA protons within the strand that contains the modification, e.g., with sugar H1′, H2′, and H2″ protons of the G^4 residue. These NOE interactions together with the enhancements observed between the methylene protons corresponding to the modified site itself, namely the reduced Schiff base, and the aromatic base proton of the G^6 residue on the 3′-side of aminoellipticine suggest an intrahelical conformation at the modified lesion site in such a way that the aminoellipticine ring system is stacked between G^4·C^{15} and G^6·C^{13} base pairs. Thus, one part of the heterocycle introduced into the duplex helix is interconnected with flanking G^4 and G^6 residues, while the other pyridinic part distal to the linkage site extends into the partner strand and lies between its C^{13} and T^{14} residues.

Table 1: Sequence-Specific Assignment of ^{1}H NMR Resonances (at 500 MHz) for the DNA Heteroduplex II, d(CGTG·Aell·GTGC)·d(GCACTCACG), at 25 °C in 10 mM Phosphate Buffer (pH 5.8) Solution Containing 10 mM NaCl and 0.1 mM Na_2EDTA

	H8/H6	H5/CH3	H1′	H2″	H2′
		Strand A			
C^1	7.72	5.99	5.73	2.49	2.04
G^2	8.10	–	5.91	2.77	2.74
T^3	7.12	1.42	5.85	2.43	1.98
G^4	8.05	–	5.98	2.67	2.60
Aell^5	–	–	3.56[a]	2.24[b]	2.24[b]
G^6	7.92	–	5.76	2.70	2.53
T^7	7.26	1.32	5.71	2.44	2.08
G^8	7.96	–	5.99	2.78	2.53
C^9	7.47	5.46	6.07	2.52	2.14
		Strand B			
G^{10}	7.95	–	6.07	2.67	2.53
C^{11}	7.54	5.70	5.50	2.43	2.12
A^{12}	8.36	–	6.28	2.94	2.84
C^{13}	7.36	5.32	5.86	2.52	2.02
T^{14}	7.50	1.60	6.09	2.59	2.26
C^{15}	7.65	5.74	5.70	2.51	2.27
A^{16}	8.44	–	6.18	2.97	2.75
C^{17}	7.39	5.45	5.82	2.41	1.96
G^{18}	8.02	–	6.24	2.62	2.35

[a,b] Degenerate sets of methylene protons.

It should be mentioned that it remains a notoriously challenging problem to accurately obtain the quantitative distance estimates from the NOESY experiments due to the uncertainties on spin diffusion effects at a given mixing time interval (Reid et al., 1989). In general, the spectra acquired with shorter mixing times yield cross peaks for only those nuclei with a high correlation rate (internuclear distance 2–3 Å), while cross peaks corresponding to lower cross relaxation rates, with interproton distances of up to 5 Å, are manifested at relatively longer mixing intervals. We have adopted the approach of performing the NOESY experiments at several mixing periods to determine 150 ms as the minimum length of the mixing delay required to observe all the *intermolecular* (Aell–DNA) NOE correlations listed in Table 2. The buildup curves for the corresponding cross peaks and the H5-H6 pairs of the C^{11} and C^{13} residues were observed to be nearly linear in the range of 50–150 ms for the mixing interval. Distance estimates for these close contacts were obtained from the

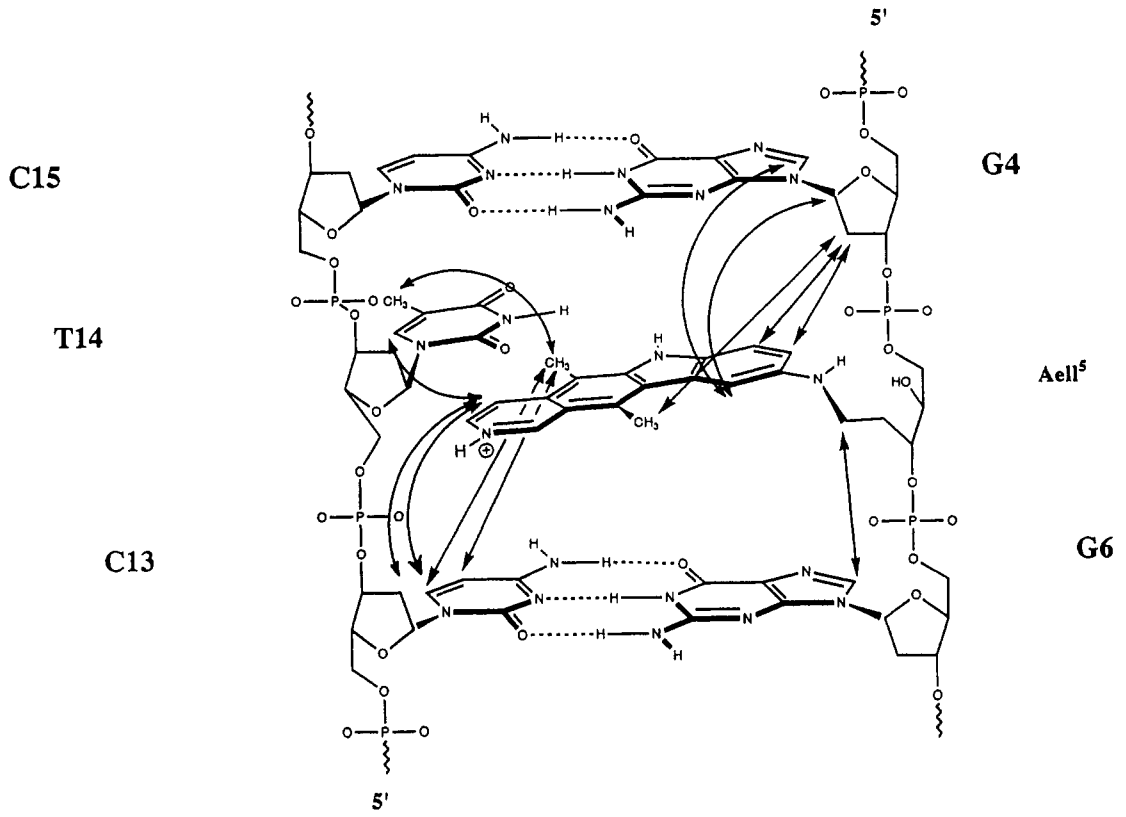


FIGURE 6: Experimentally observed *intermolecular* NOE interactions, indicated by arrows, between the covalently attached aminoellipticine heterocycle and the DNA protons in heteroduplex II.

Table 2: Experimentally Observed Proton–Proton NOE Contacts between the Covalently Attached Aminoellipticine Moiety and DNA Components from the 2D NOESY Spectra (Mixing time 150 ms; recycle time 5 s) of the Heteroduplex II[a]

strand A	strand B
AeII[b]-H1′ ↔ G^6H8(A)	AeII-H4 ↔ C^{13}H6 (B)
AeII-H1 ↔ G^4H2″ (B)	AeII-H4 ↔ T^{14}H6 (B)
AeII-H1 ↔ G^4H2′ (B)	AeII-H4 ↔ C^{13}H2″ (A)
AeII-H7 ↔ G^4H2″ (B)	AeII-C5Me ↔ C^{13}H5 (B)
AeII-H8 ↔ G^4H1′ (B)	AeII-C5Me ↔ C^{13}H6 (A)
AeII-H8 ↔ G^4H2″ (A)	AeII-C5Me ↔ T^{14}Me (B)
AeII-H8 ↔ G^4H2′ (B)	
AeII-H10 ↔ G^4H1′ (B)	
AeII-H10 ↔ G^4H8 (B)	
AeII-C11Me ↔ G^4H8 (B)	
AeII-C11Me ↔ G^4H1′ (B)	

[a] Distance constraints estimated for these contacts are classified as strong (A), medium (B), and weak (C) as given in parentheses, which correspond to the range of distances as 2.75–3.50, 3.50–4.50, and >4.50 Å, respectively. [b] Corresponds to two degenerate CH_2 protons.

measured volume integrals of the NOE cross peaks in a 150 ms NOESY spectrum acquired using a recycle delay of 5 s between consecutive increments, using isolated spin pair approximation and the fixed-distance cytosine H5—H6 vector as a reference (Gronenborn & Clore, 1985). Only those distance constraints that relate to the protons of aminoellipticine were used in the restrained molecular modeling refinements described below. The calculated values were arbitrarily classified (Table 2) by ranking the relative intensities of the cross peaks in three distance ranges, A (strong NOEs, 2.75–3.50 Å), B (moderate NOEs, 3.50–4.50 Å), and C (weak NOEs, >4.50 Å). In view of the fact that such a classification assumes the validity of isolated spin pairs and does not explicity account for spin diffusion effects that may be prevalent with the use of moderately long 150 ms mixing intervals for NOESY experiments, we must point out that the final model structures obtained with the use of distance information therefrom may represent certain limitations in terms of the "exactness" of the structures. Further improvements in the quality of the starting model structures derived from NOE distances would require extensive computations and corrections for the spin diffusion contributions, which were not attempted.

RESTRAINED MOLECULAR MODELING REFINEMENTS

Our approach to model the possible most favored solution structure of the aminoellipticine–DNA conjugate starts with the selection of two different models for molecular dynamics and energy minimization. The difference between these two is in terms of using the experimental NOE distance constraints. In the first case, the intermolecular NOE contacts were employed only to constrain the aminoellipticine moiety in approximately the right position but were not used to drive the positions of the atoms during simulation. In the second model, the constraints were enforced in a time-averaged manner to actually drive the simulation to the final model for heteroduplex II. The results on energy calculations from these experiments are given in Table 3. For comparison, two additional models of the unmodified abasic site duplex DNA were also considered. These two relate to d(CGTG·dr·GTGC)·d(GCACTCACG) containing the ring-closed hemiacetal form of the abasic sugar (dr^5) residue and differ in terms of the orientation of the lesion site and its complementary base on the partner strand, namely, T^{14}. The intrahelical model contains the dr residue at position 5 in an intrahelical orientation (similar to normal DNA), while in the extrahelical model, the residues dr^5 and T^{14} are both

Table 3: Results from Restrained Molecular Mechanics Calculations Showing Energies (kcal mol^{-1}) of the Heteroduplex I and II Models

model	total energy	vdw[a]	elstat[b]	net total[c]
	Heteroduplex I			
intrahelical	−50 856.5	115.3	−695.1	−579.8
extrahelical	−51 751.8	154.9	−666.6	−511.7
	Heteroduplex II			
without constraints	−48 164.3	124.7	−792.4	−667.7
with constraints	−46 130.1	130.5	−756.6	−626.1

[a] van der Waals or steric contributions. [b] Electrostatic contributions. [c] The net energy is the addition of contributions from footnotes a and b.

projected out of the helix into the solvent and away from the rest of the DNA nucleotides.

Although the exact numerical values of Table 3 can not be directly compared, certain trends are nevertheless evident. Thus, the intrahelical form is preferred for the abasic site heteroduplex I over the extrahelical form by 68.1 kcal/mol. In the case of heteroduplex II containing aminoellipticine, the unconstrained model is preferred over the constrained system by 41.6 kcal/mol. Such comparisons were made for the final model structures obtained by simulations in the presence of counterions and solvating water molecules, and the quoted values correspond to the contributions from van der Waals and electrostatic terms (i.e., net total term in Table 3). The total energy values provided in Table 3 can not be compared directly due to varying degrees of solvation of the intra- and extrahelical forms of heteroduplex I and of the T^{14} residue that is displaced out of the plane for heteroduplex II. Examination of the hydrogen-bonding networks of the lesion sites (dr and Aell, respectively) and the rest of the DNA residues (Table S3, supplementary material) reveals significant changes that have occurred during the molecular dynamics. All the models were obtained from the same initial coordinates corresponding to a normal DNA with the adenine residue at position 5. The residues dr and Aell were then substituted for this adenine, and as the molecular dynamics runs proceeded, the geometries evolved to form new interactions.

The intra- and extrahelical forms of the abasic heteroduplex have been characterized here and in a previous study (Cuinasse et al., 1990). The purpose of modeling is to postulate the most favored solution structure for this interesting DNA oligomer. In the intrahelical model (Figure 7), the overall conformation of DNA is maintained as the B-type with the sugar residues in C2′-*endo* configurations, which is consistent with the experimental NMR work outlined above. In the extrahelical form (Figure 8), the structure of DNA compensates for the exclusion of dr^5 and T^{14} residues from the helix by contiguous $\pi-\pi$ stacking interactions via contraction of the space between the base pairs on either side of these residues. The intrahelical form also exhibits better van der Waals (Δ39.6 kcal/mol) and electrostatic interactions (Δ28.5 kcal/mol). As pointed out above in the comparison of the combined energy terms, a clear preference is shown for the intrahelical form. In the final refined model structure for the intrahelical form, a single hydrogen bond is observed between the anomeric OH of the dr^5 residue in the hemiacetal form and the N3 nitrogen of the preceding G^4 residue on the same strand.

The slight difference between the two models for the modified aminoellipticine–DNA conjugate in terms of the use of experimental NOE distance constraints is also important in terms of reinforcing the predictive effects of the molecular modeling procedures for this complex, e.g., a theoretical simulation of the unconstructed model vs the constrained model structure guided by experimental evidence. The aminoellipticine, in principle, can be envisioned to act as an adenine substitute and form a *normal* base pairing with T^{14} by hydrogen bonding through its nucleophilic pyridinic nitrogen. However, after preliminary modeling efforts to investigate such interactions, it was not feasible to simply position Aell in an adenine-like configuration due to the lack of space for this extended tetracyclic heterocycle. The constrained model (Figure 9) was furnished by incorporating experimental NMR evidence on the partially intercalative geometry of the aminoellipticine system between G^4 and G^6 residues of the same strand and between C^{13} and T^{14} residues on the complementary strand, as indicated by break down and diversion of the normal internucleotide NOE connectivities discussed above. In both the models, an overall B-type conformation of DNA could be maintained with accommodation of aminoellipticine's steric bulk in an intercalative configuration. In the absence of any hydrogen-bonding interactions, this intercalative insertion at the lesion site is evidently governed by hydrophobic $\pi-\pi$ stacking forces and sterically forces T^{14} to a new position in which it is capable for forming a hydrogen bond with G^4 across the strand.

The energy differences between the two models of the aminoellipticine–DNA conjugate highlight the small changes due to the inclusion of constraints. Addition of penalties for the differences in the distances between the ideal and time-averaged values imparts a significant amount of energy. Much less favorable interactions are seen in both the van der Waals and electrostatic components for the constrained structure, which can be attributed, in large part, to the conformation of the T^{14} residue. The steric bulk of the Aell-C5Me group displaces the T^{14} residue into the major groove in both models with slight differences in the degree of displacement. In the unconstrained model, the T^{14} base is displaced approximately 45° from the helix axis, while in the constrained model, this angle is closer to 90° (Figure 9). Consequently, a transmitted displacement of the nearest 3′-neighbor C^{15} residue is also observed. A significant decrease in the van der Waals energy component, evidently due to a loss of stacking interactions between Aell and T^{14}, contributes to an overall difference of 5.8 kcal/mol between the two models. In addition, a consequence of bending T^{14} relative to the helix axis is that new hydrogen bonds are possible in the unconstrained model but not in the constrained model where normal base pairing on the adjacent residues, G^4-C^{15}, is maintained. In terms of the contribution from the electrostatic term to the energy differentials, the constrained model is destabilized by 35.8 kcal/mol, presumably due to a 90° rotation of T^{14} from the helical axis, that precludes the formation of any new hydrogen bonds (Table 3). The structural differences between the two models had to be considered since it was intended to model the most favored structure by matching the experimental evidence and since the theoretical modeling could not accurately predict the structure for the aminoellipticine–DNA conjugate.

DISCUSSION

Mechanistic diversity in the biological DNA repair systems is an important problem of current research focus. Several types of damage are known to cause structural alterations in the DNA double helix, e.g., bulky groups attached to nucleic acid bases and spontaneous or chemically induced cleavage of the nucleobase–sugar glycosidic bond, etc. (Friedberg, 1985;

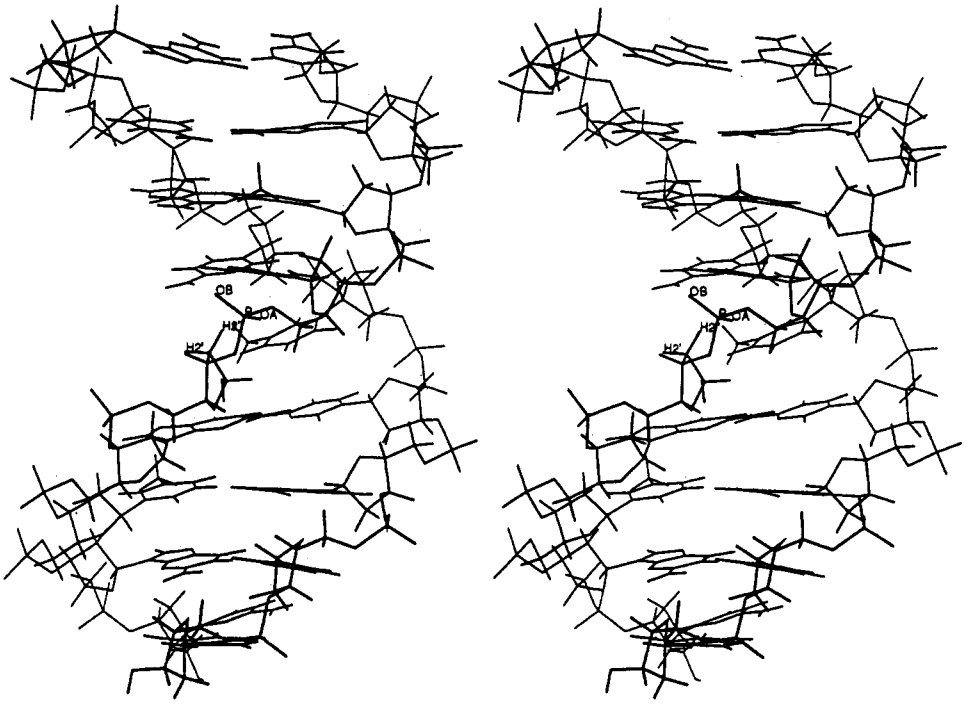


FIGURE 7: A steroview along the long axis, into the minor groove, of the intrahelical heteroduplex I model, obtained from restrained molecular modeling.

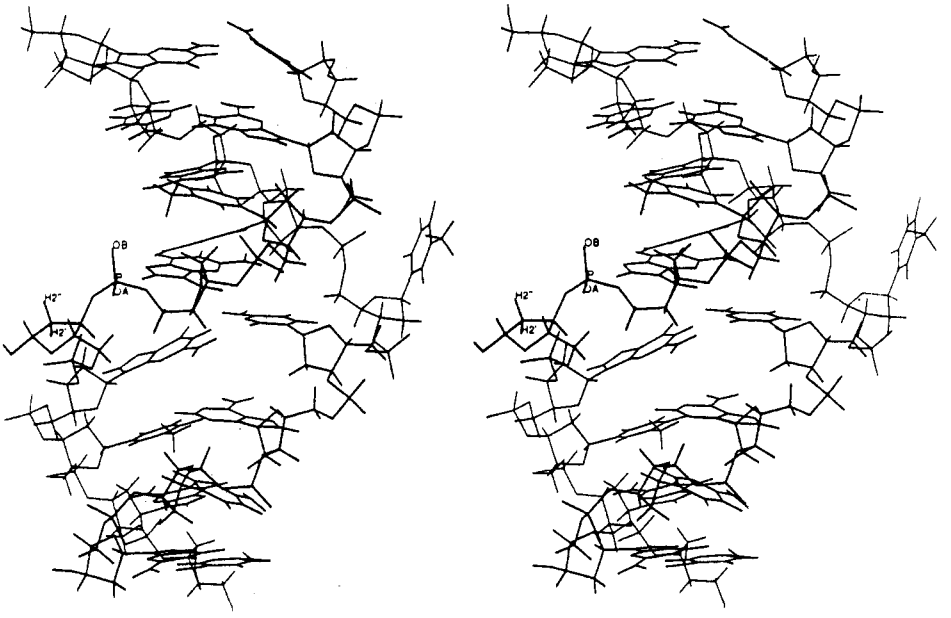


FIGURE 8: A stereoview of the model extrahelical conformation of heteroduplex I. Note the out of helix displacement of the abasic site and the thymidine base on the complementary strand.

Lindahl, 1982, 1993). One of the efficient repair mechanisms is the prereplicative excision–repair pathway which induces an initial introduction of chain breaks and a subsequent removal of the DNA segment containing the damaged sites. A group of enzymes, such as DNA glycosylases and AP endonucleases, are used for specific reactions involving the lesions (Lindahl, 1982; Sancar & Sancar, 1988; Wallace, 1988). In all these cases, the substrate specificity appears to arise from the recognition of an apparent deformation of the altered region in the DNA double helix. Thus it is important to understand the structural features of the altered DNA substrates, containing such damaged sites or derivatives thereof, as a first step toward delineating enzymatic mechanisms of damaged-site recognition and repair.

In this study, we have a chosen a duplex DNA sequence containing a true abasic site on one strand and positioned opposite a deoxyribothymidine (dT) residue on the complementary strand. The same sequence containing an analogue of the abasic sugar residue was also employed in a previous NMR study (Cuniasse et al., 1990). The authors reported on a comparative analysis of DNA oligomers containing lesions in the form of model abasic sites to show preliminary but distinct structural features due to the location of the model anhydroribitol form of the sugar at the abasic site opposite all four (dA, dG, dT, and dC) residues in a 9-mer duplex DNA. The proposed structure were characterized by a regular right-handed B-type DNA geometry for the oligomers containing the abasic site opposite purine residues A and G,

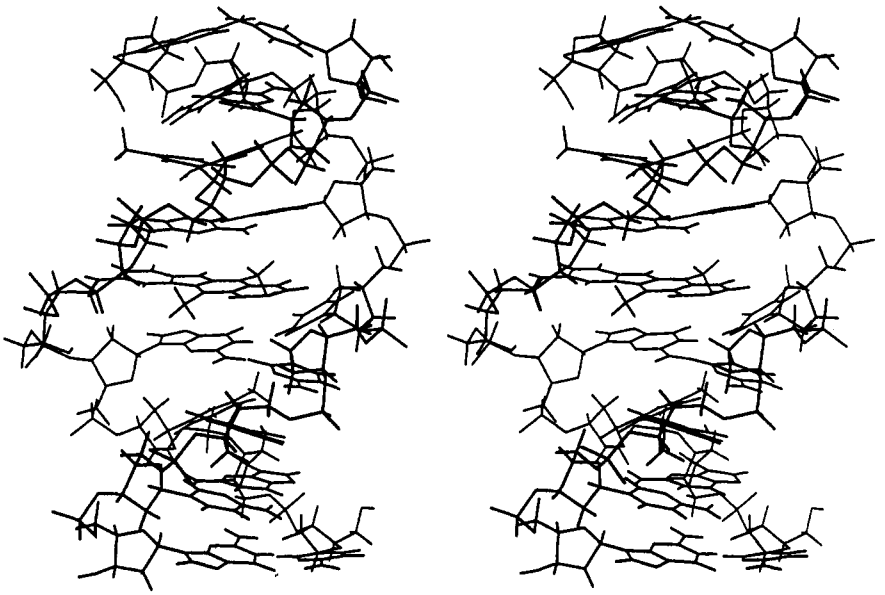


FIGURE 9: The model energy-minimized structure, in stereo, of the aminoellipticine–DNA conjugate (heteroduplex II) that was derived by incorporating distance constraints obtained from the experimental NOE data.

while those against pyrimidinic residues T and C induce an additional form of a collapsed duplex DNA helix which is characterized by an extrahelical arrangement of both the abasic site and the opposite residues. Further, the system comprising the abasic residue positioned against the dT residue is shown to be an equilibrium mixture of the extrahelical form (Figure 8) and the evidently predominant intrahelical form (Figure 7) with characteristics of a B-DNA conformation. Our analysis of the heteroduplex d(CGTG·dr·GTGC)·d-(GCACTCACG) and the model structures derived for the intra- and extrahelical forms emphasizes those earlier observations and shows that the displacement of the abasic residue out of the helix is apparently driven by the stacking of G^4 with G^6. The significance of the discrepancy on whether the displacement of T^{14} positioned opposite the abasic site accompanies the conformational change in the primary strand is not clear at this time, and presumably arises from the differences in the nature of the abasic sugar residues and the experimental conditions used for NMR.

The synthesis of phosphoramidite derivatives which can be directly used in the solid-phase synthetic methods is central to our strategy (Scheme 1) for preparing oligonucleotides containing a true abasic site and to covalently attach aminoellipticine at such sites via a reductive amination reaction (Bertrand et al., 1989a; Péoc'h et al., 1991). This approach obviates the need for elaborate procedures of solution-phase oligonucleotide synthesis and minimizes the degradation at aldehydic abasic sites due to β-elimination of the 5′-phosphonomonoester of 3′-residues. Additionally, the availability of *o*-nitrobenzyl 2-deoxyribofuranoside phosphoramidites, their efficient and convenient incorporation in the solid-phase synthetic protocol, and a selective photolytic cleavage of the *o*-nitrobenzyl group in essentially quantitative yields provide an extremely useful strategy without any restrictions on the length or sequence context of the desired abasic site-containing oligonucleotide chain. Preparation of the aminoellipticine–DNA conjugate is also facile for such abasic oligomer via a coupling reaction in the presence of $NaCNBH_3$ for a selective reduction of the imine intermediate to afford covalent attachment of aminoellipticine at the C1′ position of abasic sugar residue.

The present study is primarily directed at probing the effect of covalently incorporating aminoellipticine on the structural features of a double-stranded DNA and detecting specific perturbations in the structure due to the presence of aminoellipticine in the interior of a duplex DNA sequence containing a "hole" in the form of an abasic site. Thus, we have focused on the comparison of the abasic site-containing duplex with an aminoellipticine-derivatized counterpart. The derivatization of the abasic site by aminoellipticine has implications for an understanding of key enzymes involved in DNA excision–repair pathways in terms of the structural and enzyme recognition factors.

The evidence for Schiff base intermediates in the reaction of the aldehydic form of sugar at the abasic sites with aminoellipticine and related agents has been presented previously (Bertrand et al., 1989a,b; Livingston, 1964; Vasseur et al., 1987, 1988, 1989). Due to the chemical instability of imine species, we have focused on the covalently modified aminoellipiticine–DNA conjugate which was shown by theoretical computations to be a good model system (Letellier et al., 1991). The structural and conformational properties of the DNA–ellipticine conjugate may provide the basis for understanding how 9-aminoellipticine acts to specifically accelerate the cleavage of abasic sites via a β-elimination mechanism (Malvy et al., 1986; Vasseur et al., 1989) and also how the aminoellipticine-modified DNA duplex could afford very good protection against hydrolytic AP endonuclease activity (Figure S4, supplementary material) and in turn inhibit the excision–repair pathways (Lefrançois et al., 1990).

Well-resolved cross peaks in the two-dimensional NOESY spectra of the modified heteroduplex II have permitted assignment of the nonexchangeable DNA base and sugar protons and aminoellipticine protons (Table 1). In contrast with the parent abasic heteroduplex I, where two structurally distinct intra- and extrahelical forms exist, the aminoellipticine-linked heteroduplex II appears to adopt an intrahelical conformation exclusively, which was characterized on the basis of sequential NOE patterns. NOE interactions associated with the central (G^4·Aell·G^6)·(C^{13}·T^{14}·C^{15}) trinucleotide segment and specific aminoellipticine–nucleotide NOEs provide strong evidence for an occupancy of the abasic gap by the indole ring system between dG^4·dC^{15} and dG^6·dC^{13}

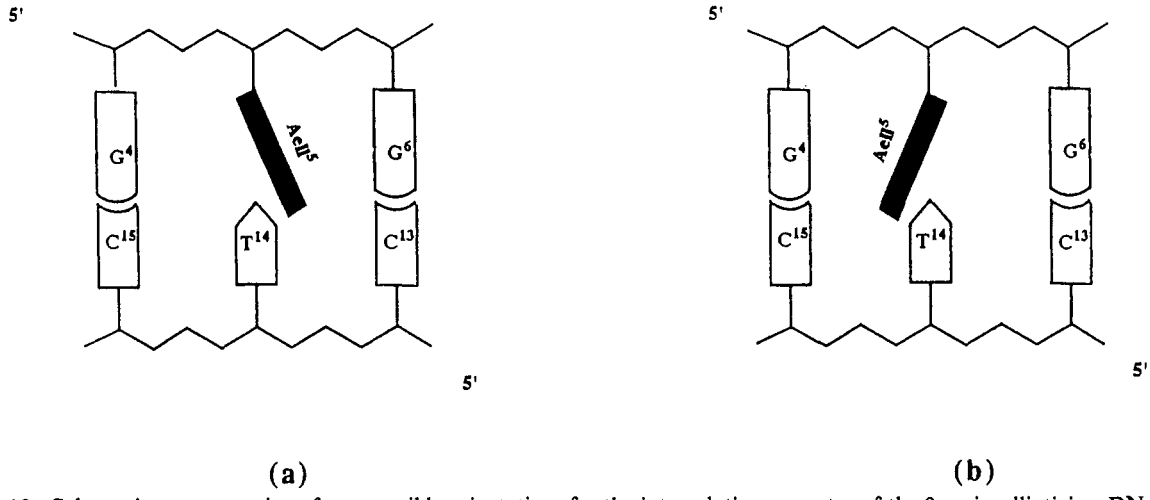


FIGURE 10: Schematic representation of two possible orientations for the intercalative geometry of the 9-aminoellipticine–DNA conjugate. The experimentally observed complex corresponds to a with the pyridine ring of the aminoellipticine ring system stacked between the 5'-CT step.

base pairs. A key feature of such intercalative geometry for the tetracyclic ellipticine is that the pyridinic ring of ellipticine further extends across into the complementary strand and lies between the dC^{13} and dT^{14} residues (Figure 6). This alignment accounts for the observed break in the sequential NOE connectivity at the C^{13}-T^{14} step and a reestablished connectivity pathway through AeII-H4 and -C5Me protons in the new C^{13}·AeII·T^{14} stacked arrangement. It is important to note that two dipyrimidinic steps, dC^{13}-dT^{14} and dT^{14}-dC^{15}, should be equally accessible for an across the strand stacking of the pyridinic ring of ellipticine. Nevertheless, the observed preference (Figure 10) for the former suggests the inherently weaker stacking interactions for a 5'-CT step compared to a 5'-TC step as also shown by previous theoretical calculations (Kollman et al., 1981).

Proximities between the AeII-H10/H7/H8 protons of the indole ring system and the G^4/G^6 residues within the modified DNA strand and those between AeII-H4/C5Me and the C^{13}/T^{14} residues on the complementary strand are observed in the form of intermolecular NOE contacts (Table 2). Restrained molecular modeling calculations by incorporating these NOE constraints provide the most favored model structure which shows an overall B-type DNA conformation and intercalative geometry for the ellipticine ring spanning across both the strands. A significant structural change can be seen in the T^{14} residue positioned opposite the aminoellipticine modification (Figure 9). The thymidine base is displaced into the major groove from coplanarity with the adjacent C^{13}/C^{15} nucleobases in both simulations. However, the magnitude of this change is closer to 90° out of the normal plane of base pairs in the constrained structure while the same angle is only 45° out of the plane in the unconstrained model. Thermal melting investigations (Rayner and Imbach, unpublished results) on analogous aminoellipticine-modified sequences indicate a significant increase in duplex stability relative to the unmodified abasic site sequences (Bertrand et al., 1989b), as determined by an increase in the transition temperature by 20–25 °C. To our knowledge, there is only one other example of the displacement of DNA bases at the expense of accommodating a covalently linked polycyclic aromatic system intercalatively, and this has been reported recently for an adduct of a benzo[*a*]pyrene epoxide with a deoxyguanosine residue in DNA (Cosman et al., 1993).

The base-displaced intercalation-type structure for the aminoellipticine–DNA conjugate also offers attractive possibilities for further functionalization of an aminoellipticine pharmacophore by appending sequence-selective DNA-binding agents. The orientation of aminoellipticine shows that AeII-H10/C11Me/H1 protons that lie on one edge toward the minor groove (Figure 6) are potential positions where a sequence-selective minor groove-binding agent, e.g., netropsin and/or distamycin with AT base pair specificities [for review see Bailly and Hénichart (1991), Dervan (1986), Lown (1990), and Zimmer and Wahnert (1986)], could be attached. Similarly, elaborations at AeII-H3/H4/C5Me/NH6/H7/H8 positions on the opposite edge pointing toward the major groove with triplex-forming oligonucleotides would afford new hybrid agents (Hélène, 1991; Thuong & Hélène, 1993). This will provide excellent examples whereby the selective abasic site-directed reactivity of aminoellipticine is reinforced by binding/recognition selectivity of the groove-binding agents.

Implications for the Observed Inhibition of DNA Excision–Repair Pathways by Aminoellipticine

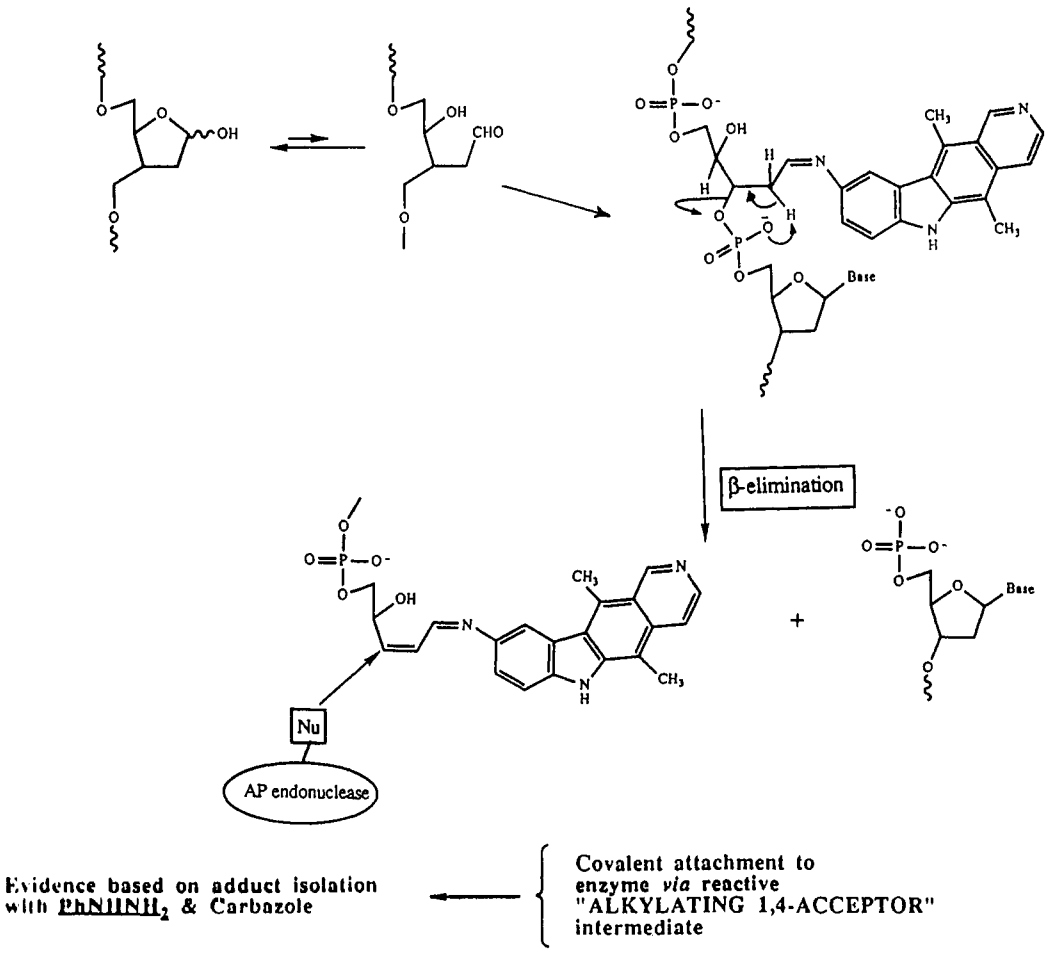
It was envisaged that the structural characteristics of such conjugates would provide information on the essential features of the observed inhibitory activity of aminoellipticine toward the AP endonuclease component of the excision–repair pathway in *E. coli* (Lefrançois et al., 1990). On the basis of our model structural studies, several hypotheses can be put forth to rationalize the structural and functional role of the high abasic site-directed reactivity of aminoellipticine.

The first step in the DNA excision–repair pathway is proposed to be the recognition of an altered region or a defect in the double-helical structure of DNA (Lindahl, 1982; Sancar & Sancar, 1988; Wallace, 1988). An abasic site could serve as the recognition signal in the form of a conformational distortion of the double helix as a result of the created nucleobase-free gap. One possible form of this change is an extrahelical arrangement of the abasic residue, possibly accompanied by the protrusion of the opposite residue on the partner strand, in the helix as shown in this NMR study, as well as in a previous NMR investigation of the same nonamer duplex sequence containing an analogue of the basic site (Cuniasse et al., 1990). Both intra- and extrahelical conformations have been observed for this sequence only when an apurinic site is positioned against pyrimidinic residues dT and dC.

The modification of the AP site in the form of a covalently attached aminoellipticine gives rise to only the intrahelical

Scheme 2: Proposed Chemical Basis for the Observed Inhibition of AP Endonuclease and DNA Excision–Repair Pathway by 9-Aminoellipticine

9-Aminoellipticine Mediated Production of Reactive Electrophile That May Alkylate AP Endonuclease



form of DNA where the ligand is able to stack inside the helix. Aminoellipticine is known to inhibit the repair process associated with AP endonuclease by forming metastable Schiff base intermediates in its reaction with the aldehydic abasic sites (Bertrand et al., 1989a; Lefrançoise et al., 1990). Such imine species, possibly in the cyclic hemiaminal forms, is assumed, on the basis of a theoretical study (Letellier, 1991), to be structurally similar to the covalently modified aminoellipticine–DNA adduct studied here. The observed occupancy of the abasic gap by the ellipticine pharmacophore, its unusual mimicry of a normal AT base pair, at least in terms of shape, and the attachment in the form of a covalent deoxyribose C1′—N bond thus provide a strong rationale in terms of a masked local structure of the modified site toward initial recognition by the enzyme(s).

A second step in the mechanism of damage-AP site recognition has also been proposed which implicates an aromatic amino acid residue at the active site of AP endonuclease in its catalytic action of the basis of enzyme-like catalytic activities of the oligopeptides Lys-Trp-Lys and Lys-Tyr-Lys (Hélène, 1985; Toulmé & Saison-Behmoaras, 1985) and a similar site-directed reactivity of aminoellipticine (Malvy et al., 1986). The covalent modification of abasic sites by aminoellipticine and the consequent stability of an intercalative complex with a duplex DNA may also inhibit AP endonuclease, presumably by blocking the entry of the critical aromatic peptide residue required for the repair activity.

Alternatively, a chemical basis for the inhibition of DNA excision–repair pathways by aminoellipticine is conceivable (Scheme 2). Current evidence suggests that the nucleobase-free deoxyribose residues exist in an equilibrium between the major cyclic hemiacetal form and about 1% of the chemically reactive acyclic aldehyde form (Manoharan et al., 1988; Withka et al., 1991). The interaction with aldehyde-specific reagents, e.g., phenylhydrazine, aminocarbazole, and aminoellipticine (Vasseur et al., 1987, 1988; Bertrand et al., 1989a), leads to the formation of Schiff base intermediates which may have a direct role in inhibition of AP enonuclease. An α,β-unsaturated imine arising from a β-elimination reaction could act as a potent electrophilic alkylating agent for critical residues on the enzyme (Bailly & Verly, 1987; Bertrand et al., 1989a; Manoharan et al., 1988; Mazumder et al., 1989). Precedents for such mechanism-based inactivation and the reactive nature of the potential 1,4-Michael acceptor, as shown in Scheme 2, exist in the form of stable adducts characterized from model reactions of phenylhydrazine and aminocarbazole and the proposals of similar imine intermediates therein (Vasseur et al., 1987, 1988).

In conclusion, this contribution outlines the combined NMR–energy minimization studies on aminoellipticine-bound heteroduplex DNA which identify the wedge-shaped stacking of aminoellipticine across both the strands in a conformation that could block the action of DNA repair enzymes by chemical modification of an apurinic site so that (i) the recognition

and/or binding to the nucleobase-free site is inhibited and/or (ii) the critical catalytic active residues of the enzymes are not able to recognize/displace the exogenous ellipticine moiety. At present, we favor the structure-based hypothesis that the accessibility of an abasic lesion to degradation by AP endonuclease is impaired due to the presence of intercalated aminoellipticine in its reaction products with the target DNA.

ACKNOWLEDGMENT

We thank Dr. C. Malvy for kindly providing us with a sample of 9-aminoellipticine. Computations were performed at the facilities of the College of Pharmacy at New Orleans.

SUPPLEMENTARY MATERIAL AVAILABLE

Tables and schemes containing additional data on heteroduplexes I and II (8 pages). Ordering information is given on any current masthead page.

REFERENCES

Bailly, C., & Hénichart, J.-P. (1991) *Bioconjugate Chem. 2*, 379–393.
Bailly, V., & Verly, W. G. (1987) *Biochem. J. 242*, 565–572.
Bertrand, J.-R., Malvy, C., & Paoletti, C. (1987) *Biochem. Biophys. Res. Commun. 143*, 768–774.
Bertrand, J.-R., Vasseur, J.-J., Gouyette, A., Rayner, B., Imbach, J.-L., Paoletti, C., & Malvy, C. (1989a) *J. Biol. Chem. 264*, 14172–14178.
Bertrand, J.-R., Vasseur, J.-J., Rayner, B., Imbach, J.-L., Paoletti, C., & Malvy, C. (1989b) *Nucleic Acids Res. 17*, 10307–10319.
Besler, B. H., Merz, K. M., & Kollman, P. A. (1990) *J. Comput. Chem. 11*, 431–439.
Cosman, M., de los Santos, C., Fiala, R., Hingerty, B. E., Ibanez, V., Luna, E., Harvey, R., Geacintov, N. E., Broyde, S., & Patel, D. J. (1993) *Biochemistry 32*, 4145–4155.
Cuniasse, Ph., Sowers, L. C., Eritja, R., Kaplan, B., Goodman, M. F., Cognet, J. A. H., LeBret, M., Guschlbauer, W., & Fazakerley, G. V. (1987) *Nucleic Acids Res. 15*, 8003–8022.
Cuniasse, Ph., Fazakerley, G. V., Guschlbauer, W., Kaplan, B., & Sowers, L. C. (1990) *J. Mol. Biol. 213*, 303–314.
Dervan, P. B. (1986) *Science (Washington, DC) 232*, 464–471.
Feigon, J., Leupin, W., Denny, W. A., & Kearns, D. R. (1983) *Biochemistry 22*, 5943–5951.
Ferrin, T. E., Huang, C. C., Jarvis, L. E., & Langridge, R. (1988) *J. Mol. Graphics 6*, 13–27.
Freidberg, E. C. (1985) *DNA Repair*, W. H. Freeman, New York.
Gronenborn, A. M., & Clore, G. M. (1985) *Prog. Nucl. Magn. Reson. Spectrosc. 17*, 1–32.
Hare, D. R., Wemmer, D. E., Chou, S. H., Drobny, G., & Reid, B. R. (1983) *J. Mol. Biol. 171*, 319–336.
Hélène, C. (1985) *Adv. Biophys. 20*, 177–186.
Hélène, C. (1991) *Anti-Cancer Drug Des. 6*, 569–584.
Kalnik, M. W., Chang, C. N., Grollman, A. P., & Patel, D. J. (1988) *Biochemistry 27*, 924–931.
Kalnik, M. W., Chang, C. N., Johnson, F., Grollman, A. P., & Patel, D. J. (1989) *Biochemistry 28*, 3373–3383.
Kollman, P. A., Weiner, P. K., & Dearing, A. (1981) *Biopolymers 20*, 2583–2621.
Lefrançois, M., Bertrand, J.-R., & Malvy, C. (1990) *Mutat. Res. 236*, 9–17.
Letellier, R., Taillandier, E., Bertrand, J.-R., & Malvy, C. (1991) *J. Biomol. Struct. Dyn. 9*, 579–597.
Lindahl, T. (1982) *Annu. Rev. Biochem. 51*, 61–87.
Lindahl, T. (1993) *Nature (London) 362*, 709–715.
Liuzzi, M., Weinfeld, M., & Paterson, M. C. (1987) *Biochemistry 26*, 3315–3321.
Livingston, D. C. (1964) *Biochim. Biophys. Acta 87*, 538–540.
Lown, J. W. (1990) in *Molecular Basis of Specificity in Nucleic Acid-Drug Interactions* (Pullman, B., Jortner, J., Eds.) pp 103–122, Kluwer Academic Publishers, The Netherlands.
Lown, J. W. (1993) *Chem. Soc. Rev.* 165–176.
Malvy, C., Prevost, P., Gansser, C., Viel, C., & Paoletti, C. (1986) *Chem.-Biol. Interact. 57*, 41–53.
Malvy, C., Safraoui, H., Block, E., & Bertrand, J.-R. (1988) *Anti-Cancer Drug Des. 2*, 361–370.
Manoharan, M., Ransom, S. C., Mazumder, A., & Gerlt, J. A. (1988a) *J. Am. Chem. Soc. 110*, 1620–1622.
Manoharan, M., Mazumder, A., Ransom, S. C., Gerlt, J. A., & Bolton, P. H. (1988b) *J. Am. Chem. Soc. 110*, 2690–2691.
Mazumder, A., Gerlt, J. A., Absalon, M. J., Stubbe, J., Cunningham, R. P., Withka, J., & Bolton, P. H (1986) *Biochemistry 30*, 1119–1126.
Mazumder, A., Gerlt, J. A., Rabow, L., Absalon, M. J., Stubbe, J., & Bolton, P. H. (1989) *J. Am. Chem. Soc. 111*, 8029–8030.
Pearlman, D. A., Case, D. A., Caldwell, J. C., Seibel, G. I., Singh, U. C., Weiner, P. K., & Kollman, P. A. (1991) *AMBER 4.0*, University of California, San Francisco.
Péoc'h, D., Meyer, A., Imbach, J.-L., & Rayner, B. (1991) *Tetrahedron Lett. 32*, 207–210.
Potier, P. (1992) *Chem. Soc. Rev. 21*, 113–119.
Propst, C. L., & Perun, T. L. (1992) *Nucleic Acid Targeted Drug Design*, Marcel Dekker, New York.
Reid, B. R. (1987) *Q. Rev. Biophys. 20*, 1–34.
Reid, B. R., Banks, K., Flynn, P., & Nerdall, W. (1989) *Biochemistry 28*, 10001–10007.
Sancar, A., & Sancar, G. B. (1988) *Annu. Rev. Biochem. 57*, 29–67.
Scheek, R. M., Boelens, R., Russo, N., van Boom, J. H., & Kaptein, R. (1984) *Biochemistry 23*, 1371–1376.
Sklenar, V., & Bax, A. (1987) *J. Magn. Reson. 75*, 378–383.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286–292.
Still, W. C. (1992) *Macromodel V3.5X Interactive Molecular Modeling System*, Columbia University, New York.
Thuong, N. T., & Hélène, C. (1993) *Angew. Chem., Int. Ed. Engl. 32*, 666–690.
Toulmé, J.-J., & Saison-Behmoaras, T. (1985) *Biochimie 67*, 301–307.
van de Ven, F. J., & Hilbers, C. W. (1988) *Eur. J. Biochem. 178*, 1–38.
Vasseur, J.-J., Rayner, B., Imbach, J.-L., Verma, S., McCloskey, J. A., Lee, M., Chang, D.-K., & Lown, J. W. (1987) *J. Org. Chem. 52*, 4994–4998.
Vasseur, J.-J., Rayner, B., & Imbach, J.-L. (1988) *J. Heterocycl. Chem. 25*, 389–392.
Vasseur, J.-J., Rayner, B., Imbach, J.-L., Bertrand, J.-R., Malvy, C., & Paoletti, C. (1989) *Nucleosides Nucleotides 8*, 863–866.
Wallace, S. (1988) *Environ. Mol. Mutagen. 12*, 431–477.
Weiner, S. J., Kollamn, P. A., Case, D., Singh, U. C., Ghio, C., Alagona, G., Prafeta, S., & Weiner, P. K. (1984) *J. Am. Chem. Soc. 106*, 765–784.
Wilman, D. E. V. (1990) *The Chemistry of Antitumor Agents*, Blackie Inc., Glasgow, London.
Withka, J. M., Wilde, J. A., Bolton, P. H., Mazumder, A., & Gerlt, J. A. (1991) *Biochemistry 30*, 9931–9940.
Zimmer, C., & Wahnert, U. (1986) *Prog. Biophys. Mol. Biol. 47*, 31–112.