Arch Pharm Res Vol 24, No 5, 455-465, 2001

Archives of Pharmacal Research
http://apr.psk.or.kr

DNA Structural Perturbation Induced by the CPI-Derived DNA Interstrand Cross-linker : Molecular Mechanisms for the Sequence Specific Recognition

Hyun-Ju Park

College of Pharmacy Sungkyunkwan, University, Suwon 440-746, Korea

(Received July 15, 2001)

The highly potent cytotoxic DNA-DNA cross-linker consists of two cyclopropa[c]pyrrolo[3, 4-3]indol-4(*5H*)-ones indoles [(+)-CPI-I] joined by a bisamido pyrrole (abbreviated to "Pyrrole"). The Pyrrole is a synthetic analog of Bizelesin, which is currently in phase II clinical trials due to its excellent *in vivo* antitumor activity. The Pyrrole has 10 times more potent cytotoxicity than Bizelesin and mostly form DNA-DNA interstrand cross-links through the N3 of adenines spaced 7 bp apart. The Pyrrole requires a centrally positioned GC base pair for high cross-linking reactivity (i.e., 5'-T̲(A/T)$_2$G(A/T)$_2$A*-3'), while Bizelesin prefers purely AT-rich sequences (e.g., 5'-T̲(A/T)$_{4\ or\ 5}$A*-3', where T̲ represents the cross-strand adenine alkylation and A* represents an adenine alkylation) (Park *et al.*, 1996). In this study, the high-field ^{1}H-NMR and rMD studies are conducted on the 11-mer DNA duplex adduct of the Pyrrole where the 5'-T̲TAGTTA*-3' sequence is cross-linked by the drug. A severe structural perturbation is observed in the intervening sequences of cross-linking site, while a normal B-DNA structure is maintained in the region next to the drug-modified adenines. Based upon these observations, we propose that the interplay between the bisamido pyrrole unit of the drug and central G/C base pair (hydrogen-bonding interactions) is involved in the process of cross-linking reaction, and sequence specificity is the outcome of those interactions. This study suggests a mechanism for the sequence specific cross-linking reaction of the Pyrrole, and provides a further insight to develop new DNA sequence selective and distortive cross-linking agents.

Key words: CPI-dimers, DNA interstrand cross-linker, High-field ^{1}H-NMR, 2D-NOESY, Restrained Molecular Dynamics

INTRODUCTION

One of the Bizelesin analogs that is a highly potent cytotoxic DNA-DNA cross-linker consists of two cyclopropa [c]pyrrolo[3,4-3]indol-4(*5H*)-ones indoles [(+)-CPI-I] joined by a bisamido pyrrole linker (abbreviated to "Pyrrole"). As shown in Fig. 1, the compound can form a DNA-DNA interstrand cross-link through N3 of adenines spaced 7 bp apart. It was determined that the [(+)-CPI-I]$_2$-bisamido-pyrrole (Pyrrole) requires a centrally positioned GC base pair for high cross-linking reactivity (i.e., 5'-T̲(A/T)$_2$G(A/T)$_2$A*-3') (Park *et al.*, 1996). The cross-linking reactivity of the Pyrrole was compared with that of the analog having an N-methylated pyrrole in the linker using 21-mer duplex oligomers containing either GC or IC base pairs. The results revealed the DNA sequence and linker requirements for high cross-linking reactivity. In the DNA duplex, to attain highest reactivity, a centrally placed GC base pair and the exocyclic 2-amino group were required, while for the linker in the Pyrrole, an unsubstituted amine in the pyrrole ring was necessary. On the basis of these results, we suggested a rationale for the structural requirements for the specific sequence cross-linked with high reactivity by the Pyrrole which involves in two hydrogen-bonding donoracceptor pairs.

To obtain further evidence for this suggestion, the structural aspects of the DNA cross-linking by the Pyrrole were investigated using high-field ^{1}H-NMR analysis on the 11-mer duplex adduct of the Pyrrole where the 5'-

Correspondence to: Hyun-Ju Park, College of Pharmacy, Sung kyunkwan University, Suwon 440-746, Korea
E-mail: hyunju@skku.ac.kr

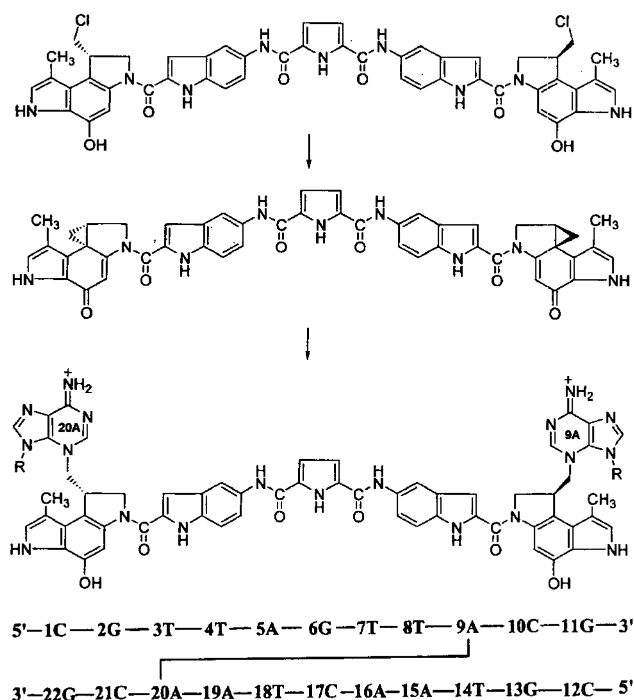

Fig. 1. A scheme for the reaction of the Pyrrole compound with adenines 9A and 20A on opposite strands, to form the interstrand cross-linked duplex adduct

TTAGTTA*-3′ sequence is cross-linked by the drug (Fig. 1). The analysis of the NMR data sets reveals a prominent structural perturbation in the intervening AGT segment, especially in the central 6G and 17C base pair. The results also show that a normal B-DNA behavior is retained in the region next to the drug-modified bases (9A and 20A). The study not only supports the suggested molecular mechanism for the cross-linking of DNA by CPI-dimers, but also reveal interesting features of the DNA structural distortion induced by the Pyrrole which relate to its unique GC preference in the cross-linking of DNA.

MATERIALS AND METHODS

Chemicals and enzymes

Bizelesin and the Pyrrole were a gift from The Pharmacia Upjohn Company and reagents used for the NMR and HPLC buffer, sodium phosphate (99.99%), sodium chloride (99.99%), and EDTA (99.99%), were purchased from Aldrich. HPLC water, methanol, and acetonitrile were from Fisher. Electrophoretic reagents (acrylamide, ammonium persulfate, bisacrylamide, and *N,N,N′,N′*-tetramethylethylene diamine) were purchased from Biorad. T4 polynucleotide kinase and T4 ligase were from United States Biochemical Co. (γ-^{32}P)-ATP and X-ray film were from Amersham, and intensifying screens developing and chemicals were from Kodak.

Oligonucleotide synthesis and purification for NMR studies

The 11-mer duplex d(CGTTAGTTACG) d(CGTAACTA ACG) was synthesized on an Applied Biosystems Model 381A automated DNA synthesizer using solid phase phosphoramidite chemistry. The crude oligonucleotides were deprotected with saturated ammonium hydroxide at 55°C overnight and then purified by reverse phase HPLC on a C18 column. The purified DNA was dialyzed using Spectra/Por CE membrane (MWCO: 1000) against 2 × 2000 ml of HPLC water and evaporated to dryness. The DNA was dissolved in NMR buffer (10 mM NaH_2PO_4, 100 mM NaCl (pH 6.85), and 0.5 mM EDTA) and then examined by ^{1}H-NMR.

Preparation of purification of the pyrrole-11-mer duplex adduct

To the 10 mg of the Pyrrole in 0.2 ml of DMF-d^7 solution, 0.03 ml of aqueous sodium bicarbonate was added, stirred for 20 min, and then 20 mg of purified 11-mer duplex in 0.8 ml of buffer (10 mM NaH_2PO_4, 100 mM NaCl, pH 6.95) was added. The reaction mixture was stirred for five days at room temperature in the dark, centrifuged to spin down unreacted drug molecules, and the supernatant was separated and lyophilized several times to dryness. The resulting residue (drug-DNA adduct) was purified by reverse-phase HPLC using Rainin C18 and C8 Dynamax-300A preparation columns (21.4 mm × 250 mm). The solvent gradient progressed from a 100 mM ammonium acetate buffer toward a 100% acetonitrile with a flow rate of 8 ml/min. The percentage of acetonitrile increased following a course of 0.0% (0 min), 5% (35 min), 10% (60 min), 40% (90 min), and 100% (120 min).

Proton NMR experiments

One- and two-dimensional NMR experiments were recorded on a Bruker AMX 500 FT NMR spectrometer. Phase-sensitive two-dimensional NOESY spectra in D_2O were obtained (TPPI) at 27°C for three mixing times, 100, 250, and 350 ms, and each spectrum was collected using 1024 data points in t_2 and 512 points in t_1, with a spectral width of 5000 Hz. Two-dimensional ROESY, TOCSY, and DQF-COSY experiments were also conducted on the sample in D_2O. For the assignments of exchangeable protons, two-dimensional NOESY spectra (150, 250, and 350 ms mixing time) in H_2O buffer solution were recorded at 27 °C using a 1-1 echo pulse sequence with 2.5 s pulse repetition time, a sweep width of 12000 Hz, and a 90° pulse width of 12 μs. Data were apodized and Fourier transformed using Felix 2.1 NMR data processing software.

Restrained molecular dynamics

Interproton distances were determined from the measured 2D NOE intensities in D_2O-NOESY spectrum using MARDIGRAS program. CORMA and MARDIGRAS were performed for the single mixing time of 150 ms and energy minimized starting structure to produce interproton distance restraint. The restraints were applied as the form of a flat well with a flat, parabolic, and linear part. The flat region was defined by the lower and upper bounds (r_2 and r_3), calculated with MARDIGRAS. The resulting restraint set consisted of 151 interproton restraints including hydrogen bonding restraints for in G/C and A/T base pairs through 1C/22G-4T/19A and 8T/15A-11G/12C.

The initial coordinates and charges for the Pyrrole was calculated by using MOPAC 5.0 implanted in Sybyl 6.2 (Tripos Associates). The starting coordinates for the Pyrrole adduct ("flipped-out" model) was generated by using Sybyl 6.2 and bad contacts in those initial models were removed by conjugate gradient minimization with the Amber force field incorporated in Sybyl 6.2. The initial models were refined by adding the counterions and energy minimization with Amber 4.1 and then used for solvated rMD calculation.

The solvated restrained molecular dynamics were performed with the SANDER (Simulated Annealing with NMR Driven Energy Restraints) module of Amber 4.1 on an SGI Indigo2 XZ workstation. For solvated rMD, the counterions and 64 boxes of 216 Monte Carlo waters were placed around the energy-minimized B-form adduct by using AMBER EDIT module. The structure was energy minimized until the energy gradient converged to 0.1 kcal/mol $Å^{-1}$. Belly dynamics for water only (0-300 K; 10 ps) was followed by solvated rMD. The rMD simulation was run for 100,000 × 1 fs steps (i.e., 100 ps) with periodic boundary condition, constant pressure, and SHAKE on all bonds. The system was slowly heated up to 700 K over a period of 65 ps and equilibrated at this temperature for 10 ps. The temperature was then lowered to 300 K over the next 5 ps and maintained at 300 K to the end of the rMD simulation (20 ps). Following the temperature protocol, force constants were increased 40 kcal/mol· $Å^2$ (distance restraints) and 20 kcal/mol $Å^2$ (hydrogen bonding restraints) at 700 K, subsequently reduced by half during cooling, and kept constant until the end of the run. After stabilization at 300 K for 20 ps, the final coordinate set was minimized without restraints to obtain the reported structure. The trajectories over the last 20 ps were used for hydrogen bonding analysis (HBOND) and measurement of the distance between H2 of adenine and H1′ of the 3′-neighbor of the complementary residue by using of CARNAL module (Amber 4.1). The HBOND option determines the hydrogen bonding donor-acceptor pairs located in the distance <4.0 Å and the angle $<60°$.

For the final stabilization step of the second rMD, the temperature (at 300 K) and restraint weight (10 kcal/mol $Å^2$ interproton distance restraints) kept constant for 30 ps. From the trajectories over the last 30 ps, the coordinate sets were extracted every 2 ps, averaged, and minimized to obtain the reported structure. The hydrogen bonding donor-acceptor pairs that located in the distance <4.0Å and the angle $<60°$ were determined, by using the HBOND option of CARNAL module (Amber 4.1).

RESULTS

Covalent bonding sites are 9A and 20A

The 7 bp cross-linked adduct of the Pyrrole at the 5′-TTAGTTA* site was prepared using an 11-mer duplex (sequence shown in Fig. 1). The complete formation of a single interstrand cross-linked adduct was confirmed by a strand breakage assay (Reynolds *et al.*, 1985) on 5′-single-end labeled 11-mer duplexes treated with the Pyrrole. The cross-linked species was purified by 20% denaturing polyacryamide gel electrophoresis (PAGE), thermally treated, and loaded onto the sequencing gel to determine the drug modified bases, 9A and 20A (unpublished results). The proton resonance assignments were determined by comparison of the two-dimensional NOESY, ROESY, DQF-COSY, and TOCSY spectra of the free DNA and the drug-DNA duplex adduct in D_2O and H_2O. In the NOESY spectra, not all of the expected NOEs are unambiguously assigned, and a complete assignment of the aromatic protons from the drug and DNA bases is not possible due to severe line broadening (Rinkel *et al.*, 1991; Kennedy *et al.*, 1993) in the region of 7.1-7.4 vs 7.7-7.8 ppm and 7.6-7.8 vs 8.0-8.3 ppm. Although a three-dimensional homonuclear TOCSY -NOESY experiment (Radhakrishnan *et al.*, 1992) was also conducted on the duplex adduct, the data does not improve the resonance assignment significantly.

As shown in Fig. 2, the sites for the covalent linkages of the Pyrrole are defined by the drug-to-DNA NOE connectivities obtained from the data of NOESY in D_2O and H_2O. The strong NOEs between adenine H2 protons and H4A and H4B protons, and those between sugar protons of flanking bases (10C, 11G, 21C, and 22G) and the drug protons from CPI-units, confirm the location of the drug modification site (Lin *et al.*, 1990; Seaman *et al.*, 1993). The NOE connectivities between pyrrole vinylic proton $H4''_R$ and the 6GH2″ proton are shown in the D_2O-NOESY spectrum (the strong NOEs for pyrrole $H3''_L/H4''_R$ and the 6GH2′/2″ protons are observed in the H_2O-NOESY). The strong NOEs between the indole $H7'_L$ and the 5AH5′/5″ are also shown in the D_2O-NOESY spectrum. These connectivities (Fig. 2) indicate the approximate positioning of the linker unit of the Pyrrole in the duplex.

Base-pairs 5A to 7T are not hydrogen-bonded

The assignments of the exchangeable protons for the duplex and the drug adduct are demonstrated in the one-dimensional proton NMR spectra in Fig. 3. The typically sharp hydrogen-bonded imino proton resonance signals (Wüthrich *et al.*, 1986) are absent through 5A/18T-7T/16A base pairs, but are detected for 3T/20A-4T/19A, 8T/15A-9A/14T, and the flanking GC base pairs (2G/21C and 10C/13G). A disruption of the hydrogen bonding in the central AGT segment indicates that the DNA conformational perturbation is induced by cross-linking. However, some broad resonance signals are observed in the range of 12.7-13.4 ppm, which is within the range of chemical shifts for imino protons involved in Watson-Crick base pairing. A line broadening of these signals may be related to a conformational exchange behavior of the imino protons (Cho *et al.*, 1994; Ippel *et al.*, 1995). This is in striking contrast to the 7 bp cross-linked adduct of Bizelesin on the same oligomer sequence, in which all the Watson-Crick base pairings remain intact except for the terminal frayed bases (Thompson *et al.*, 1995).

Nonexchangeable proton spectra: evidence for the structural perturbation of the intervening segment (4T/19A-8T/15A) in the cross-linked site

The nonexchangeable protons for the 11-mer cross-linked duplex adduct were assigned by analyzing NOESY, ROESY, DQF-COSY, TOCSY, and ^{1}H-^{31}P COSY data sets. The chemical shifts of the nonexchangeable protons and a comparison of the ^{1}H-NMR chemical shifts between the free DNA duplex and the drug-DNA duplex adduct are summarized in Table I. The intra- and inter-residual NOE connectivities of DNA in the duplex adduct are also summarized in Fig. 4.

The analysis of the ^{1}H-^{31}P COSY spectrum reveals that the chemical shifts of the ^{31}P resonances of the duplex adduct were dispersed within a range of 1.14 ppm (from -0.44 to -1.58 ppm relative to external 85% H_3PO_4, unpublished results). For a normal B-DNA helix, the ^{31}P resonances appear within a region of about 1 ppm in width (Patel *et al.*, 1982; Holak *et al.*, 1984). In comparison with the chemical shift dispersion of 0.66 ppm for free DNA duplex (-0.28 to -0.94 ppm), the structure of the cross-linked adduct does not significantly deviate from the B-DNA phosphodiester backbone structure (or the conformational feature of this adduct does not result

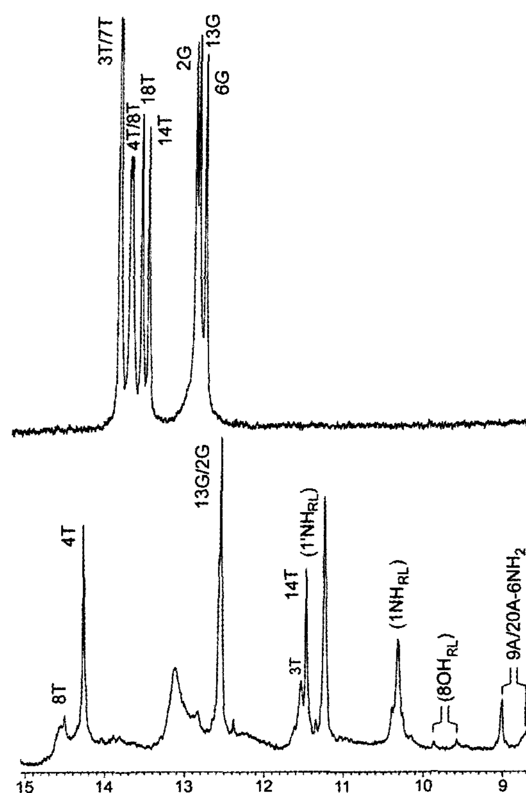

Fig. 3. One-dimensional proton NMR spectra (9-15 ppm) of the 11-mer duplex (top) and the cross-linked Pyrrole adduct in H_2O, pH 6.80, at 27°C. The imino proton assignments are shown above the spectra, and the imino and phenolic-OH protons listed in parentheses are from the Pyrrole.

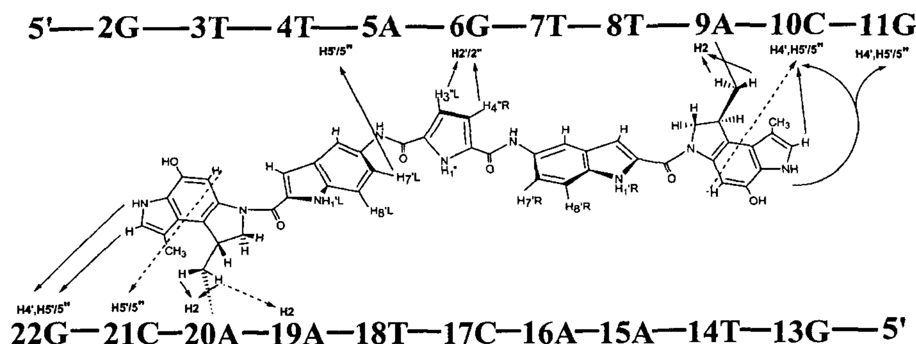

Fig. 2. Summary of intermolecular drug-to-DNA NOE connectivities between the Pyrrole compound and DNA protons. Solid arrows indicate strong NOE intensity, and dotted arrows indicate medium or weak.

Fig. 4. Summary of intra- and inter-residual DNA NOE connectivities in the Pyrrole-11-mer cross-linked adduct. The dotted line between the DNA bases indicates an assigned hydrogen bond.

in unusual phosphorus chemical shifts).

Selected H8/H6-H1′ and H8/H6-H2′2″/TCH_3 regions of a 350 ms mixing time NOESY spectrum obtained in D_2O buffer at 27°C are shown in Fig. 5 and 6, respectively. The sequential assignments of the base to H1′ connectivities (Hare *et al.*, 1983; Feigon *et al.*, 1983) for each strand are shown separately in Fig. 5A and B. A number of features of the spectra indicate that there is a significant perturbation of the structure in the cross-linked site (through 3T/20A-9A/14T). The H8/6-H1′ inter-residue NOE connectivities are broken at the 5A-6G step on the top strand, and through 14T-18T on the bottom strand, suggesting that these bases may not be completely stacked. Examination of the intra-residue NOE connectivities of each nucleotide shows that the NOE cross-peak for 5A is strong and broad, 6G and 14T are very weak, 15A and 16A appear very weak and broad, and 17C is absent. In Fig. 6A, the base to H2′/2″, and methyl region of the same NOESY experiment are shown. As was the case in the base to H1′ region, there is a disrup-tion of the inter-residue connectivities through 5A-6G, 8T-9A, and 16A-17C base steps. If the cross-linked site is bisected by the central 6G/17C base pair, the right-hand side is through 7T/16A-9A/14T and the left-hand side is through 18T/5A-20A/3T. The analyses of the H8/H6-H1′ and H8/H6-H2′2″/TCH_3 regions reveal that the structural perturbation is prominent on the right-hand side of the bottom strand (see Fig. 4), while the left-hand side is relatively well organized.

For right-handed helices, a relatively strong NOE between a base proton and the H5 (or 5-CH_3) proton of the subsequent base on the 3′-side is generally observed (Hare *et al.*, 1983; Feigon *et al.*, 1983). The strong NOEs between a base proton and the H5 (or CH_3) proton of the 3′-side proton are observed for most of dinucleotide step in the duplex except for the 5A/18T-7T/16A region. The corresponding NOE for 17CH6-18TCH_3 has a weak intensity even at 350 ms, a relatively high mixing time,

Fig. 5. Two-dimensional D_2O-NOESY spectrum (350 ms mixing time), showing connectivities for aromatic PuH8/PyH6 to H1′ proton and to H1′ proton of 5′-neighbor. (A) is for the top strand of the duplex (from 1C to 22G). (B) is for the bottom strand of the duplex (from 13G to 22G), the broad NOE cross-peaks for 17CH5-17CH6 connectivity are shown in the box.

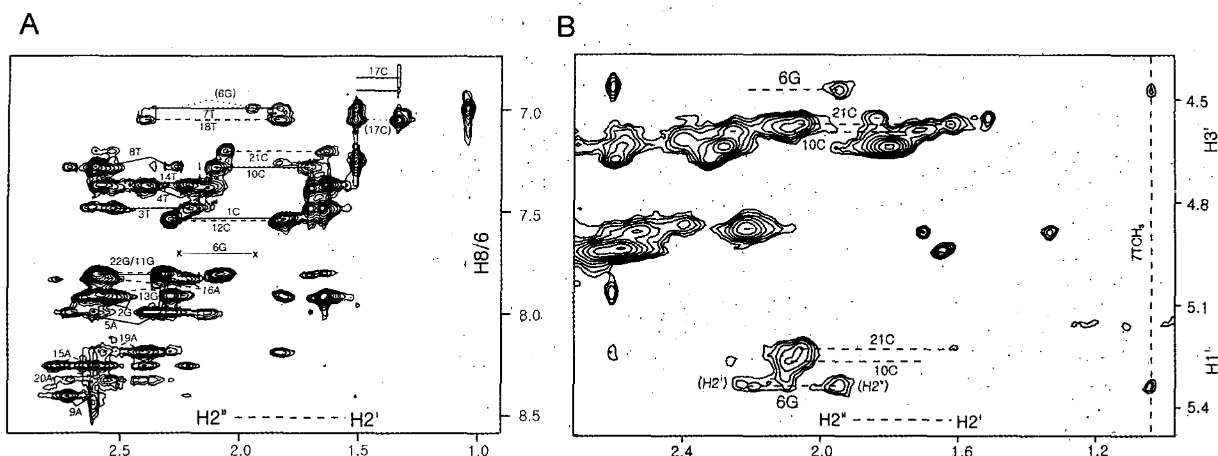

Fig. 6. (A) The PuH8/PyH6 to H2'/2" proton region of the NOESY spectrum (350 ms mixing time). Missing intra-residual connectivities for 6GH8-H2'/2" are marked as 'x's. (B) The H2'/2" to H1' and H3' proton region of the NOESY spectrum (150 ms mixing time) showing the chemical shift reversal of 6GH2' and 6GH2".

while the NOEs for 6GH8-7TCH$_3$ and 16AH8-17CH5 are absent. This pattern indicates that a certain amount of stacking between 17C and 18T residues is present, which may not be a conventional B-DNA right-handed helical stacking, while the 6G-7T and 16A-17C step are either totally unstacked or only slightly stacked. In summary, the results obtained from the ^{1}H-^{31}P COSY and D_2O-NOESY data suggest that some residues may have unusual linkages, structural perturbations, and severe motion in the cross-linked site of the drug adduct, while the global DNA conformation retains a right-handed helical structure.

Conformational perturbation of 6G and 17C

It was demonstrated that a centrally positioned G/C base pair is important for high cross-linking reactivity of the Pyrrole previously (Park *et al.*, 1996). The NMR data for the 11-mer duplex adduct obviously reveals a significant structural disturbance in the central AGT segment (through 5A/18T-7T/16A). As shown in Fig. 6A, an intra-residual NOE for 6GH2′/H2″-6GH8 is absent, but a very weak NOE to the 3′-side base 7TH6 is present, suggesting an unusual glycosidic linkage for 6G and the existence of relatively low stacking between residues 6G and 7T. The NOE intensity for H1′-H2″ is always stronger than that for H1′-H2′, and the H2″ proton resonance is downfield from the H2′ proton resonance in normal B-form DNA (Reid *et al.*, 1987; Patel *et al.*, 1987). As shown in D_2O-NOESY at 150 ms (Fig. 6B), the chemical shifts of H2′ and H2″ are reversed for 6G, while the corresponding H2″ proton resonance is downfield from the H2′ proton resonance for the unmodified 11-mer duplex. The corresponding couplings are not observed in the ROESY (100 ms) data set, and the through bond coupling for H1′-H2′ is very weak in the DQF-COSY data. The results indicate the conformational perturbation in sugar pucker geometry at 6G (Weiss *et al.*, 1984a; Wemmer *et al.*, 1984; Saenger *et al.*, 1984; Kintanar *et al.*, 1987).

Previous NMR studies on 6 or 7 bp cross-linked adduct of Bizelesin (Thompson *et al.*, 1995a; Thompson *et al.*, 1995b) have shown that the H8/H6 proton resonances for the centrally positioned DNA bases (2 or 3 base pairs)

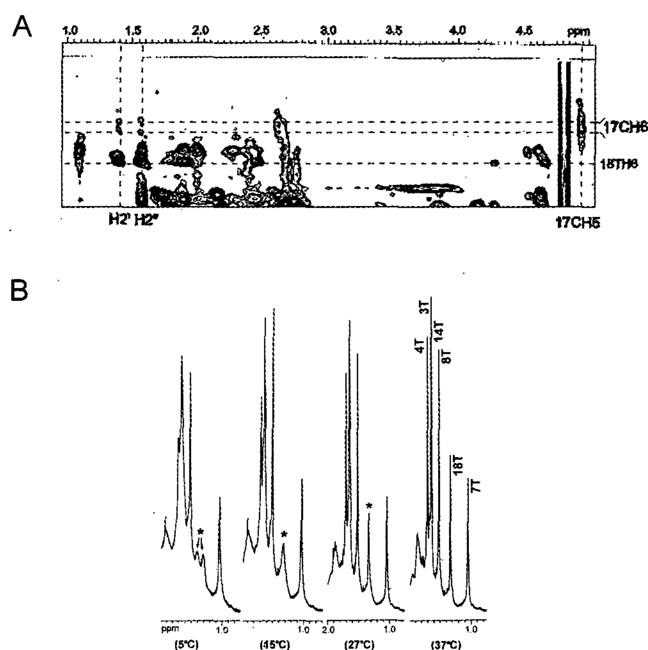

Fig. 7. (A) Expanded H_2O-NOESY contour plot (300 ms mixing time) of the Pyrrole-11-mer cross-linked adduct at 27 °C. Two sets of NOE cross-peaks for 17CH6-17CH2'/2" are assigned. (B) One-dimensional proton NMR spectra (1.0-2.0 ppm) in H_2O, pH 6.80 between 5°C and 37°C. The assignment of thymine 5-CH$_3$ protons are shown above spectra. The temperature-dependent change of line-shape for 18T5 CH$_3$ is denoted by asterisks.

Table I. Chemical shifts (ppm) of nonexchangeable protons in the [d(CGTTAGTTACG)·d(GCTAACTAACG)] Pyrrole compound cross-linked adduct at 27°C

	H8	H6	H2	H5	H1'	H2'	H2"	H3'	H4'	H5'	H5"
1C		7.54(-0.13)[a]		5.76(-0.16)	5.65(-0.15)	1.83(-0.23)	2.30(-0.16)	4.68(-0.04)	4.04(-0.53)	3.67(-0.41)	3.74(-0.01)
2G	7.93(-0.07)				5.84(-0.21)	2.50(-0.22)	2.67(-0.18)	4.96(-0.04)	4.32(-0.08)	4.12(+0.08)	3.67(-0.45)
3T		7.48(+0.17)		1.67(+0.23)	6.17(+0.10)	2.22(+0.09)	2.53(-0.05)	4.90(+0.03)	4.34(-0.06)	4.23(+0.05)	4.14(-0.12)
4T		7.38(+0.01)		1.72(+0.04)	5.78(+0.03)	2.14(+0.01)	2.41(-0.09)	4.88(-0.02)	4.30(-0.14)	4.09(-0.04)	4.10(-0.07)
5A	8.00(-0.23)		NA[b]		5.94 (-0.12)	2.34(-0.39)	2.72(-0.20)	4.64(-0.41)	4.65	3.85	3.48
6G	7.73(+0.18)				5.35(-0.48)	1.96(-0.48)	2.25(-0.46)	4.49(-0.38)	NA	NA	NA
7T		7.00(-0.19)		1.05(-0.16)	NA	1.86(-0.20)	2.40(-0.14)	4.56(-0.27)	NA	NA	NA
8T		7.27(-0.10)		1.53(-0.11)	5.51(-0.27)	2.29(+0.22)	2.54(+0.05)	4.60(-0.30)	NA	NA	NA
9A	8.40(+0.08)		8.18(+0.63)		5.95(-0.24)	2.74(+0.45)	NA	4.85(-0.19)	NA	NA	NA
10C		7.30(+0.02)		5.50(+0.15)	5.29(-0.37)	1.72(-0.13)	2.12(-0.17)	4.60(-0.18)	4.23	3.99(-0.13)	3.80(-0.36)
11G	7.81(-0.09)				6.06(-0.08)	2.61(+0.25)	2.35(-0.23)	4.69(+0.04)	4.22(+0.06)	4.13(+0.06)	4.13(+0.06)
12C		7.55(-0.10)		5.79(-0.13)	5.65(-0.13)	1.84(-0.18)	2.32(-0.12)	4.68(-0.02)	4.04(-0.33)	3.69(-0.38)	3.69(-0.03)
13G	7.91(-0.07)				5.84(-0.12)	2.26(-0.40)	2.54(-0.24)	4.96(-0.02)	4.31(-0.05)	4.05(+0.03)	3.93(-0.16)
14T		7.37(+0.13)		1.65(+0.14)	NA	2.23(+0.21)	2.42(+0.06)	NA	4.23(+0.04)	4.09(+0.03)	4.13(-0.03)
15A	8.26(+0.02)		7.73(+0.69)		NA	2.64(-0.11)	2.78(-0.14)	NA	4.24	4.14	4.12
16A	7.84(-0.26)				NA	2.25(-0.38)	NA	4.90(-0.09)	NA	NA	NA
17C		6.80[d](-0.39)		4.90(-0.21)	NA	1.35(-0.51)	1.53(-0.87)	NA	4.23	NA	NA
18T		7.06(-0.26)		1.34(-0.21)	5.74(+0.11)	1.85(-0.19)	2.43(+0.03)	4.63(-0.24)	NA	NA	NA
19A	8.19(-0.05)		7.74(+0.80)		5.75(-0.13)	2.42(-0.30)	2.57(-0.32)	4.78(-0.28)	NA	NA	NA
20A	8.33(+0.22)		8.43(+0.77)		5.76(-0.31)	2.63(+0.07)	2.73(-0.07)	NA	NA	NA	NA
21C		7.21(+0.01)		5.51(+0.27)	5.25(-0.38)	1.65(-0.17)	2.07(-0.19)	4.58(-0.29)	NA	NA	3.47(-0.59)
22G	7.80(-0.07)				6.06(-0.08)	2.61(+0.26)	2.35(-0.22)	4.68(+0.04)	4.21(+0.06)	4.11(+0.05)	3.81(-0.25)

[a]Chemical shift differences, δ(DNA-drug adduct)-δ(DNA), are shown in parentheses, and the chemical shift differences greater than 0.25 are underlined.
[b]Not assigned.
[c]Difference is not determined because the chemical shifts in duplex are unassigned.
[d]average chemical shift of overlapped broad signals.

in the cross-linked sites were uniformly upfield-shifted due to the Bizelesin cross-linking. A comparison between the chemical shift changes of the H8 and H6 protons (see Table I) of the central AGT/ACT bases shows that the chemical shifts are significantly up-field shifted except for the 6G, which is 0.18 ppm down-field shifted. It is known that an extrahelically positioned base proton is deshielded and largely down-field shifted (Morden *et al.*, 1993; Cosman *et al.*, 1994; Cosman *et al.*, 1996). The H8/H6 chemical shifts for the central AGT/ACT bases do not provide strong evidence for any of the bases being out of helix. However, a unique down-field shift of 6GH8 is noticeable, which may be related to different geometry and low stacking of 6G in the duplex adduct.

The H5 and H6 protons of 17C are assigned through a combined analysis of the NOESY, ROESY, and TOCSY spectra recorded in D_2O buffer. The NOE cross-peak for the H5-H6 of 17C is very broad, which is apparently a mixture of two cross-peak species at 27°C (see Fig. 5B). In addition, the NOEs between 17CH6 and 17CH2'/H2" appear as two separate sets in H_2O-NOESY, as shown in Fig. 7A. These results indicate that 17C may be in a slow conformational equilibrium state at 27°C (Cho *et al.*, 1994). Not unexpectedly, the through bond coupling for the H5-H6 of 17C is not observed in the DQF-COSY data, which is due to the extreme line broadening of the corresponding resonances (Weiss *et al.*, 1984b; Petros *et al.*, 1990). The temperature dependency of the thymine-$5CH_3$ protons (1.0-2.0 ppm) in the duplex DNA adduct between 5 and 37°C (Fig. 7B) reveals that the lineshape broadening is observed for $18TCH_3$ upon lowering the temperature. At 5°C, the resonance signal for $18TCH_3$ is detected as two separate peaks. The 5'-CT step is a structurally labile site with the low stacking interaction and has the propensity to form a kink in the DNA structure (Sun *et al.*, 1993; Singh *et al.*, 1994; el Hassan *et al.*, 1996; Dlakic *et al.*, 1996; Subirana *et al.*, 1997). The inherent conformational flexibility of the 17C-18T step may be entrapped by the Pyrrole cross-linking, and then revealed as a slow conformational equilibrium state in the final Pyrrole cross-linked adduct structure. The conformational equilibrium of 17C-18T while maintaining a certain amount of stacking interaction is a pivotal point for the more severe structural perturbation of the right-hand side, in contrast to a relatively organized left-hand side in the cross-linked site.

A model derived from restrained molecular dynamics (rMD) analysis reveals that 6G is displaced toward the major groove

Several different initial models of the cross-linked adduct were generated by using various drug conformers and a standard B-form DNA. Initially, the rMD calculation did not work properly on the B-form DNA models. Thus a "flipped-out" model was generated by breaking the backbones next to 17C and manually inserting a flipped-out cytosine into the gap. A "flipped-out" model is designed by taking into account the absence of hydrogen bonding around this base pair and the inherently weak stacking interaction in the 5'-CT step. By using the "flipped-out" model (Fig. 8A) as a starting structure, the rMD calculations were conducted in two continuous steps. For the first rMD, the DNA-to-DNA and the DNA-to-CPIs restraints were given, and a resulting coordinate set was used as a starting coordinate for the second rMD, where the DNA-to-(indole)-pyrrole unit restraints were added (see Fig. 2). The coordinate sets (every 2 ps) were extracted from the final 30 ps stabilization step of the MD trajectories, averaged, and minimized to produce a final

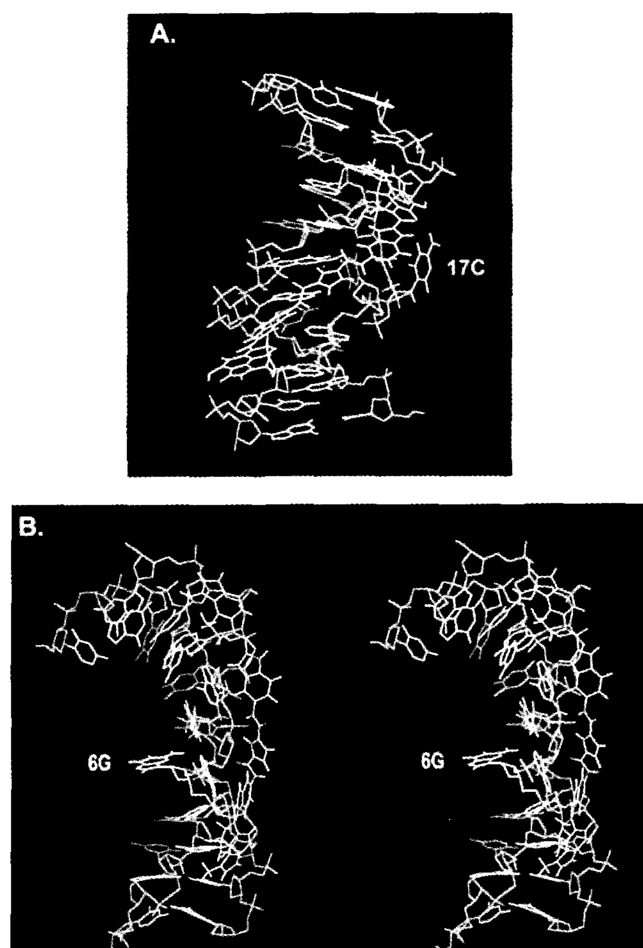

Fig. 8. (A) [d(CGTTAGTTACG)d(GCTAACTAACG)]-Pyrrole cross-linked adduct model. The rMD starting structure, a view looking into the minor groove. Extrahelical positioning of 17 C is indicated by the arrow. Colors are green (adenine), red (thymine), cyan (guanine), orange (cytosine), magenta (backbone), and yellow (drug molecule). (B) Side stereoview of the averaged rMD structure (30 ps, 300K) with the minor groove on the right and the major groove on the left. Note the displacement of the 6G (yellow) toward the major groove.

rMD structure shown in Figure 8B. The behavior of 6G and 17C observed in the rMD simulations agrees with the characteristics of those bases shown in the NMR analysis.

Examination of the structural changes during MD reveals that the flipped-out 17C becomes intrahelical and the 6G is displaced toward the major groove after the second stage of the rMD, as shown in Fig. 8B (the side view). A displacement of the 6G is due to the NOE restraints between the indole-pyrrole subunit of the drug and the 5A-6G step of the DNA. A steric repulsion may displace the 6G toward the major groove, and in the process cytosine can change its direction of motion into the helix.

An analysis of the second MD trajectories reveals that the hydrogen bonding between 6G-2NH_2 and the amide carbonyl oxygen (O10) occurs in an early stage of MD (similar to the 7 bp cross-link of Bizelesin on the same sequence), and the hydrogen bonding between 17C-4NH_2 and the amide carbonyl oxygen (O11) exists with a 100% occurrence in the final stabilization stage.

DISCUSSION

Previous ^{1}H-NMR analysis and rMD study on the 7 bp cross-linked adduct of Bizelesin on the 5′-TTAGTTA* site suggested that the hydrogen bonding between the carbonyl of the ureylene and the 2NH_2 group of 6G is important for the formation of the 7 bp cross-linked adduct rather than 6 bp. The cross-linking efficiency of the Bizelesin on the 5′-TTAGTTA* site is about 40%, and that for the Pyrrole is about 95% indicating a high preference for this site. Bizelesin is somewhat short for spanning 7 bp, while the Pyrrole is long enough to cover 7 bp. To compare the sizes of the most energetically stable conformers of Bizelesin and the Pyrrole, the difference is about 4 Å, which is close to the rise per residue for B-form DNA (3.4 Å). Why doesn't the shorter compound (Bizelesin) disturb the DNA structure, while the compound with a more suitable size (the Pyrrole) disrupts the DNA structure significantly? This phenomenon suggests a more active role of the linker unit rather than a simple size fit for the formation of the DNA cross-link.

For the Pyrrole, the length of the linker is extended in comparison with Bizelesin, but the linker unit may still be rigid (or more rigid) due to the stability of the pyrrolodicarbonyl unit, which consists of conjugated double bonds having delocalized electrons to generate resonance structures. In addition, intramolecular hydrogen bonding between pyrrole-NH and carbonyl oxygen may also be possible, and therefore those factors should restrict the geometry of the linker unit of the Pyrrole (i.e., pyrrole-NH and carbonyl oxygens are oriented toward the same side). As suggested previously, the favored geometry of the linker unit may be involved in hydrogen bonding donor-acceptor pairings with the 6G, which may be related with the preferential recognition of centrally positioned G/C base pair for high activity. To achieve those interactions, relatively hydrophilic moieties of the linker unit (pyrrole-NH and carbonyl oxygens) should contact the N3 and the 2-NH_2 group of the 6G in the hydrophobic DNA minor groove pocket. This may result in the disruption of the hydrogen bonding between base pairs of AGT segment and the displacement of the 6G toward the major groove.

Although the NMR data provide strong evidence of conformational heterogeneity in the intervening sequences of the Pyrrole adduct, a weak B-DNA-like characteristic is still observed in that region. A normal B-DNA behavior is retained in the region of drug-modified bases and flanking base pairs. Overall, the duplex adduct of the Pyrrole has more propensities for a B-form-like DNA structure rather than a highly distorted, looped-out structure, despite the absence of Watson-Crick base pairings in the three central base pairs. This structural consequence of the Pyrrole adduct also supports the previous suggestions (Seaman *et al.*, 1996) about the mechanism for the formation of a DNA-DNA interstrand cross-link. The first monoalkylation of the CPI unit generally induces a bent DNA structure, and the DNA should be reorganized to scatter bending distortion and restore a B-form like conformation at the second alkylation site to form a cross-link. The structural consequences of the Pyrrole cross-linked adduct at an asymmetric 5′-TTAGTTA*-3′ sequence implicate the pathway of cross-linking reaction. The previous study using the monoalkylative CPI drug revealed that, on the 5′-TTAGTTA-3′ sequence, the monoalkylation is preferred at the 5′-TTA* site rather than 5′-TAA* (Thompson *et al.*, 1995b). Definitely, the first monoalkylation by the Pyrrole should occur at the right-hand side 5′-TTA* (A* is 9A in the 11-mer duplex), and then the second alkylation occurs at the left-hand side 5′-TAA* (A* is 20A) to form a cross-link. To accomplish the second alkylation, the left-hand side 5′-TAA segment (18T/5A-20A/3T) should be reorganized to B-form like conformation. In this process, the ordered conformation of the prealkylated right-hand side 5′-TTA* segment (7T/16A-9 A*/14T) could be sacrificed to dissipate the monoalkylation-induced structural distortion that propagated toward the 5′-side. This explains the conspicuous structural perturbation on the right-hand side (especially through 14T-17C) and a relatively organized left-hand side (through 18T-20A).

REFERENCES

Cho, B. P., Beland, F. A., and Matilde Marques, M. NMR structural studies of a 15-mer DNA duplex from a ras protooncogene modified with the carcinogen 2-amino-

fluorene: conformational heterogeneity. *Biochemistry*, 33, 1373-1384 (1994).

Cosman, M., Fiala, R., Hingerty, B. E., Amin, S., Geacintov, N. E., Broyde, S., and Patel, D. J. Solution conformation of the (+)-trans-anti-[BP]dG adduct opposite a deletion site in a DNA duplex: intercalation of the covalently attached benzo[a]pyrene into the helix with base displacement of the modified deoxyguanosine into the major groove. *Biochemistry*, 33, 11507-11517 (1994).

Cosman, M., Hingerty, B. E., Luneva, N., Amin, S., Geacintov, N. E., Broyde, S., and Patel, D. J. Solution conformation of the (-)-cis-anti-benzo[a]pyrenyl-dG adduct opposite dC in a DNA duplex: intercalation of the covalently attached BP ring into the helix with base displacement of the modified deoxyguanosine into the major groove. *Biochemistry*, 35, 9850-9863 (1996).

Dlakic, M. and Harrington, R. E. The effects of sequence context on DNA curvature. *Proc. Natl. Acad. Sci. U.S.A.*, 93, 3847-3852 (1996).

el Hassan, M. A. and Calladine, C. R. Propeller-twisting of base-pairs and the conformational mobility of dinucleotide steps in DNA. *J. Mol. Biol.*, 259, 95-103 (1996).

Feigon, J., Leupin, W., Denny, W. A., and Kearns, D. R. Two-dimensional proton nuclear magnetic resonance investigation of the synthetic deoxyribonucleic acid decamer $d(ATATCGATAT)_2$. *Biochemistry*, 22, 5943-5951 (1983).

Hare, D. R., Wemmer, D. E., Chou, S. -H., Drobny, G., and Reid, B. R. Assignment of the non-exchangeable proton resonances of d(C-G-C-G-A-A-T-T-C-G-C-G) using two-dimensional nuclear magnetic resonance methods. *J. Mol. Biol.*, 171, 319-336 (1983).

Holak, T. A., Borer, P. N., Levy, G. C., van Boom, J. H., and Wang, A. H. J. ^{31}P-NMR analysis of the B to Z transition in double-stranded $(dC\text{-}dG)_3$ and $(dC\text{-}dG)_4$ in high salt solution. *Nucleic. Acids Res.*, 12, 4625-4635 (1984).

Ippel, J. H., Lanzotti, V., Galeone, A., Mayol, L., van den Boogaart, J. E., Pikkemaat, J. A., and Altona C. Slow conformational exchange in DNA minihairpin loops: a conformational study of the circular dumbbell d. *Biopolymers*, 36, 681-694 (1995).

Kennedy, M. A., Nuutero, S. T., Davis, J. T., Drobny, G. P., and Reid, B. R. Mobility at the TpA cleavage site in the T3A3-containing AhaIII and PmeI restriction sequences. *Biochemistry*, 32, 8022-8035 (1993).

Kintanar, A., Klevit, R. E., and Reid, B. R. Two-dimensional NMR investigation of a bent DNA fragment: assignment of the proton resonances and preliminary structure analysis. *Nucleic. Acids Res.*, 15, 5845-5862 (1987).

Lin, C. H. and Hurley, L. H. Determination of the major tauto-meric form of the covalently modified adenine in the (+)-CC-1065-DNA adduct by ^{1}H and ^{15}N NMR studies. *Biochemistry*, 29, 9503-9507 (1990).

Morden, K. M. and Maskos, K. NMR studies of an extrahelical cytosine in an AT rich region of a deoxyribodecanucleotide. *Biopolymers*, 33, 27-36 (1993).

Park, H. -J., Kelly, R. C., and Hurley, L. H. The Chemical Evolution of DNA-DNA Interstrand Cross-Linkers That Recognize Defined Mixed AT and GC Sequences. *J. Am. Chem. Soc.*, 118, 10041-10051 (1996).

Patel, D. J., Kozlowski, S. A., and Nordheim, A., Rich, A. Right-handed and left-handed DNA: studies of B- and Z-DNA by using proton nuclear overhauser effect and P NMR. *Proc. Natl. Acad. Sci. U.S.A.*, *79*, 1413-1417 (1982).

Patel, D. J., Shapiro L., and Hare, D. R. DNA and RNA: NMR studies of conformations and dynamics in solution. *Q. Rev. Biophys.*, 20, 35-112 (1987).

Petros, A. M., Mueller, L., and Kopple, K. D. NMR identification of protein surfaces using paramagnetic probes. *Biochemistry*, 29, 10041-10048 (1990).

Radhakrishnan, I., Patel, D. J., and Gao, X. Three-dimensional homonuclear NOESY-TOCSY of an intramolecular pyrimidine.purine.pyrimidine DNA triplex containing a central G.TA triple: nonexchangeable proton assignments and structural implications. *Biochemistry*, 31, 2514-2523 (1992).

Reid, B. R. Sequence-specific assignments and their use in NMR studies of DNA structure. *Q. Rev. Biophys.*, 20, 1-34 (1987).

Reynolds, V. L., Molineux, I. J., Kaplan, D. J., Swenson, D. H., and Hurley, L. H. Reaction of the antitumor antibiotic CC-1065 with DNA. Location of the site of thermally induced strand breakage and analysis of DNA. *Biochemistry*, 24, 6228-6237 (1985).

Rinkel, L. J., Tinoco, and I. Jr. A proton NMR study of a DNA dumb-bell structure with hairpin loops of only two nucleotides: d(CACGTGTGTGCCGTGCA). *Nucleic. Acids Res.*, 19, 3695-3700(1991).

Saenger, W. Principles of Nucleic Acid Structure, Springer-Verlag, New York, Berlin, Heidelberg, and Tokyo, pp 324-327(1984).

Seaman, F. C. and Hurley, L. H. Interstrand cross-linking by bizelesin produces a Watson-Crick to Hoogsteen base-pairing transition region in $d(CGTAATTACG)_2$. *Biochemistry*, 32, 12577-12285(1993).

Seaman, F. C. and Hurley, L. H. Manipulative Interplay of the Interstrand Cross-Linker Bizelesin with $d(TAATTA)_2$ To Achieve Sequence Recognition of DNA. *J. Am. Chem. Soc.*, 118, 10052-10064(1996).

Singh, M. P., Hill, G. C., Péoc'h, D., Rayner, B., Imbach, J. L., and Lown, J. W. High-field NMR and restrained molecular modeling studies on a DNA heteroduplex containing a modified apurinic abasic site in the form of cova-lently linked 9-aminoellipticine. *Biochemistry*, 33, 10271-10285(1994).

Subirana, J. A., and Faria, T. Influence of sequence on the conformation of the B-DNA helix. *Biophys. J.*, 73, 333-

338(1997).

Sun, D., Lin, C. H., and Hurley, L. H. A-tract and (+)-CC-1065-induced bending of DNA. Comparison of structural features using non-denaturing gel analysis, hydroxyl-radical footprinting, and high-field NMR. *Biochemistry,* 32, 4487-4495 (1993).

Thompson, A. S. and Hurley, L. H. Solution conformation of a bizelesin A-tract duplex adduct: DNA-DNA cross-linking of an A-tract straightens out bent DNA. *J. Mol. Biol.,* 252, 86-101 (1995a).

Thompson, A. S., Fan, J. -Y., Sun, D., Hansen, M., and Hurley, L. H. Determination of the structural role of the internal guanine-cytosine base pair in recognition of a seven-base-pair sequence cross-linked by bizelesin. *Biochemistry,* 34, 11005-11016(1995b).

Weiss, M. A., Patel, D. J., Sauer, R. T., and Karplua, M. Two-dimensional ^{1}H NMR study of the lambda operator site OL1: a sequential assignment strategy and its application. *Proc. Natl. Acad. Sci. U.S.A., 81,* 130-134 (1984a).

Weiss, M. A., Eliason, J. L., and States, D. J. Dynamic filtering by two-dimensional ^{1}H NMR with application to phage lambda repressor. *Proc. Natl. Acad. Sci. U.S.A.,* 81, 6019-6023 (1984b).

Wemmer, D. E., Chou, S. -H., and Reid, B. R. Sequence-specific recognition of DNA. Nuclear magnetic resonance assignments and structural comparison of wild-type and mutant lambda OR3 operator DNA. *J. Mol. Biol.,* 180, 41-60 (1984).

Wüthrich, K. NMR of Proteins and Nucleic Acids, Wiley-Interscience, New York, (1986).