Determination of the Structural Role of the Internal Guanine−Cytosine Base Pair in Recognition of a Seven-Base-Pair Sequence Cross-Linked by Bizelesin†

Andrew S. Thompson, Jun-Yao Fan, Daekyu Sun, Mark Hansen, and Laurence H. Hurley*

Drug Dynamics Institute, College of Pharmacy, The University of Texas at Austin, Austin, Texas 78712

Received March 23, 1995; Revised Manuscript Received June 20, 1995®

ABSTRACT: Bizelesin (formerly U77,779, The Upjohn Co.) is a bifunctional DNA cross-linking antitumor antibiotic consisting of two open-ring homologs of the (+)-CC-1065 cyclopropa[*c*]pyrrolo[3,2-*e*]indol-4(*5H*)-one (CPI) subunits connected by a rigid linking moiety. Previous studies have shown that Bizelesin most often forms an interstrand cross-link through the N3 of two adenines 6 base pairs (bp) apart (inclusive of the modified adenines). However, gel electrophoresis studies have also indicated that Bizelesin forms 7-bp cross-links in specific sequences. In most of these sequences the cross-linked adenines represent the only possible cross-link site (i.e., no 6-bp site is available); however, in several sequences, a 7-bp sequence is selected in overwhelming preference to a possible 6-bp sequence. In this study, we demonstrate the unique requirement for a G•C base pair within this sequence and the critical presence of the exocyclic 2-amino group of guanine. In a subsequent two-dimensional ^{1}H-NMR study that concentrates on the 7-bp cross-link formed with the sequence 5′-TTAGTTA-3′, the role of the central G•C base pairs in the formation of a 7-bp cross-link is probed. ^{1}H-NMR analysis coupled with restrained molecular dynamics (rMD) provides evidence for distortion around the covalently modified adenines. Because of this distortion, the modified bases are twisted toward the center of the duplex adduct, effectively reducing the cross-linked distance. The rMD study also indicates that a hydrogen bond is formed between the exocyclic amine of the central guanine and the carbonyl of the ureylene linker. On the basis of the observation of the distortion in the duplex and the hydrogen bonding between the drug and DNA, it is possible to speculate on the role of the central G•C bases in this sequence preference and propose a mechanism by which Bizelesin forms a 7-bp rather than a 6-bp cross-link with this sequence.

The antitumor antibiotic (+)-CC-1065 is produced by *Streptomyces zelensis* and was first isolated by scientists at The Upjohn Co. (Hanka et al., 1978; Martin et al., 1981). Structurally, (+)-CC-1065 consists of three repeating pyrroloindole subunits, the first of which contains a DNA-reactive cyclopropapyrroloindole (CPI)[1] function (Figure 1). (+)-CC-1065 reacts with double-stranded DNA through the N3 of a reactive adenine, forming a covalent adduct that overlaps a 5-bp region in the minor groove (Swenson et al., 1982; Hurley et al., 1984; Scahill et al., 1990). Despite promising activity in a variety of tumor cell lines *in vitro* (Li et al., 1982; Bhuyan et al., 1982) and *in vivo* (Martin et al., 1981), the development of (+)-CC-1065 as an antitumor antibiotic was abruptly terminated when it was discovered that the ligand caused an unusual hepatoxicity, which led to delayed death in mice at therapeutic doses (McGovren et al., 1984). Extensive SAR studies revealed the structural features of (+)-CC-1065 that lead to the observed toxicity (Warpehoski & Bradford, 1988; Hurley et al., 1988, Warpehoski & Hurley, 1988), and more efficacious analogs were synthesized. Adozelesin (Figure 1) represented the first of these compounds to enter clinical trials (Fleming et al., 1992; Burris et al., 1992), and the interest in this compound led to the development of second-generation bifunctional analogs. The first bifunctional CPI analogs were linked by flexible linkers, and the initial success *in vitro* with these agents rapidly led to the development and synthesis of Bizelesin (Mitchell et al., 1991), which consists of two DNA-reactive

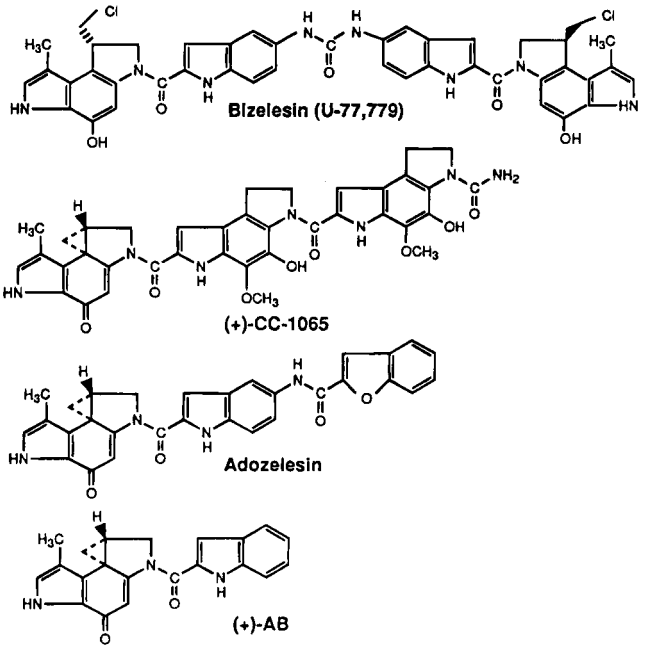

FIGURE 1: Structures of (+)-CC-1065 and the cross-linking analog Bizelesin and the monoalkylation compounds Adozelesin and (+)-AB.

† This research was supported by grants from the U.S. Public Health Service (CA-49751), the Welch Foundation, and The Upjohn Co.

* Address correspondence to this author.

® Abstract published in *Advance ACS Abstracts,* August 1, 1995.

[1] Abbreviations: CPI, cyclopropapyrroloindole; bp, base pair; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, two-dimensional NOE correlated spectroscopy; rMD, restrained molecular dynamics; FID, free induction decay; DMF, dimethylformamide; EDTA, ethylenediaminetetraacetic acid; HPLC, high-performance liquid chromatography; DDW, double-distilled water.

0006-2960/95/0434-11005$09.00/0

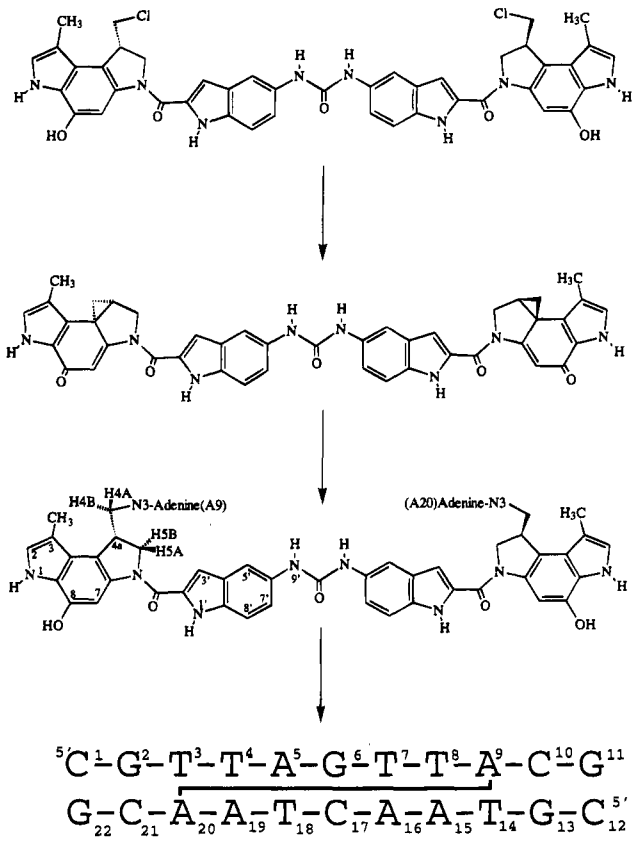

FIGURE 2: Scheme depicting the cyclization of Bizelesin to give the cyclopropyl derivative followed by the reaction of adenines A9 and A20 on the opposite DNA strands to ultimately form the interstrand cross-linked duplex adduct (including numbering systems for the duplex and Bizelesin used in the high-field NMR study).

CPI subunits linked via a rigid bis(indolecarboxylic acid) linker.

Bizelesin generally forms interstrand cross-links with adenines primarily spaced 6 bp apart (including the covalently modified adenines) while occupying the intervening minor groove (Figure 2) (Ding & Hurley, 1991). Previous gel electrophoresis studies have shown several preferred sequences for the formation of DNA−DNA interstrand cross-links (Sun & Hurley, 1993; Lee & Gibson, 1993a,b). The most-preferred sequence spans six nucleotides and mimics the sequence preference for the monoalkylating (+)-CC-1065 analogs. This preferred target sequence contains two 5′-TTA*-3′ sequences (* indicates covalently modified base) within the 6-bp cross-link site 5′-TAATTA*-3′. Unexpectedly, NMR studies of a Bizelesin cross-linked adduct formed with this sequence have indicated the presence of two adduct conformations, both containing unusual base pairing within the central A·T base pairs (Seaman & Hurley, 1993, 1995). In the first case (40% of the observed species), the bases are arranged in an open conformation, in contrast to the second case (60% of the observed species), in which the bases are Hoogsteen base paired. The unusual conformations observed within this sequence are proposed to be partially the result of the unusually low base-pair stability at the A·T step in the unmodified duplex, and it has been proposed that Bizelesin may trap out and stabilize the open and Hoogsteen conformations prior to cross-link formation (Seaman & Hurley, 1993, 1995). The second most preferred 6-bp cross-linking sequence is the intrinsically bent A-tract-containing sequence 5′-TTTTTA*-3′. Nondenaturing gel electrophoresis (PAGE) experiments have indicated that upon cross-link formation the intrinsic A-tract bending was eliminated, while a high-field NMR study of the adduct formed with the A-tract 5′-TTTTTA*-3′ confirmed the loss of A-tract character, including propeller twisting and junctions at the 3′ and 5′ ends of the A-tract (Thompson et al., 1995; Thompson & Hurley, 1995).

Chart 1. Sequences of Oligomers Used in This Study

I	5′-TGGCCGCAATTAATTACGGGT-3′ 3′GGCGTTAATTAATGCCCAACC-5′
II	5′-TGGCCGCAGTTAGTTACGGGT-3′ 3′-GGCGTCAATCAATGCCCAACC-5′
III	5′-TGGCCGCAGTTAITTACGGGT-3′ 3′-GGCGTCAATCAATGCCCAACC-5′
IV	5′-CGTTAGTTACG-3′ 3′-GCAATCAATGC-5′

Molecular modeling studies and gel electrophoresis studies indicate that Bizelesin is most suited to cross-linking two adenines 6 bp apart. There are, however, examples of 7-bp cross-link sites, and in several of these, a 7-bp site is selected over a possible 6-bp cross-link site. In this study, we identify the importance of the central G·C base pairs in the center of the sequence 5′-TTAGTTA-3′ in mediating a 7-bp cross-link. This observation leads to two important questions. First, since Bizelesin is too short to alkylate two adenines 7 bp apart in normal B-form DNA, what is the location and nature of Bizelesin-induced distortion of this duplex? Second, what is the role of the central guanine in facilitating the formation of a 7-bp cross-link? We have addressed these questions using a combined gel electrophoresis and high-field NMR study. In the gel electrophoresis study, we have demonstrated both the importance of the central G·C base pair and, in particular, the 2-NH_2 group of guanine in mediating a 7- rather than a 6-bp cross-linkage. In addition, using the monoalkylating CPIs, Adozelesin, and (+)-AB (Figure 1), we have demonstrated the unique reactivity of the 5′-TAA* sequence to the cross-linking CPIs. High-field NMR was used to provide the solution conformation of the 7-bp cross-linked sequence and pinpoint the presence of distortion in the duplex adduct around the covalently modified adenines. Finally, on the basis of these observations, we make some conclusions as to the molecular mechanism for cross-linking this 7-bp cross-link site.

EXPERIMENTAL PROCEDURES

Chemicals and Enzymes. Bizelesin, (+)-CC-1065, Adozelesin, and (+)-AB compounds were a gift from The Upjohn Co. and were used without further purification. Reagents used to prepare the NMR buffer, sodium dihydrogen phosphate (99.99%), sodium chloride (99.99%), and EDTA (99.99%), were purchased from Aldrich. HPLC water, methanol, ethyl acetate, and acetonitrile were purchased from Baxter. Ammonium acetate (99.99%), sodium hydroxide (99.99%), glacial acetic acid (99.99%), and hydrochloric acid were purchased from Aldrich. Electrophoretic reagents [acrylamide, *N,N,N′,N′*-tetramethylethylenediamine, ammonium persulfate, and bis(acrylamide)] were purchased from Bio-Rad. T_4 polynucleotide kinase was from United States Biochemical Co.; [γ-^{32}P]ATP was from ICN; and X-ray film, intensifying screens, and developing chemicals were from Kodak.

Preparation of Oligonucleotides for Gel Electrophoretic Studies. A series of oligonucleotides (Chart 1) were

synthesized on an automated DNA synthesizer (Applied Biosystems 381A) by the phosphoramidite method. The oligomers were deprotected with saturated ammonium hydroxide at 55 °C overnight. The ammonium hydroxide solution was evaporated under vacuum, and the dried pellets were redissolved in DDW.

Preparation of the 5′-^{32}P End-Labeled Oligomer Duplex. Individual strands of the 21-mer were labeled with [γ-^{32}P]-ATP using T_4 polynucleotide kinase and hybridized to an excess of unlabeled complementary strand, and the resulting duplex DNA was gel-purified by 12% nondenaturing polyacrylamide gel electrophoresis. Duplex DNA was located on the gel using autoradiography, cut from the gel, minced with a blade, and extracted with annealing buffer [10 mM Tris-HCl (pH 7.6) and 100 mM NaCl].

Drug Treatment of Oligomer Duplexes and Determination of Drug-Bonding Sites on Electrophoretic Gels. The 5′ end-labeled oligomer duplexes were reacted with drug in 100 μL of annealing buffer for 24 h. Drug concentrations were 100 μM for Bizelesin, 10 μM for (+)-AB, and 10 μM for Adozelesin. The 10-fold higher concentration by Bizelesin was used to compensate for the overall lower reactivity of Bizelesin. Unbound drug molecules were removed by phenol/chloroform extraction, and DNA molecules were precipitated by ethanol precipitation. DNA pellets were dried, redissolved in DDW, and subjected to the thermal strand breakage assay (Reynolds et al., 1985) to induce DNA breakage at the drug-bonding sites.

Separation of the Cross-Linked Duplex from Monoalkylated and Unreacted Species Using Denaturing Polyacrylamide Gel Electrophoresis. Duplex oligomer (about 100 ng) was reacted with Bizelesin (100 μM) in 100 μL of annealing buffer for the indicated period of time. Unbound drug molecules were removed by phenol/chloroform extraction, and DNA was precipitated by ethanol precipitation. DNA pellets were dried, redissolved in 20 μL of alkaline dye (80% formamide and 10 mM NaOH), and incubated at 80 °C for 1 min just before the samples were loaded on a prewarmed 12% denaturing polyacrylamide gel. Electrophoresis was carried out at 1200 V for 2 h to separate the Bizelesin cross-linked duplex from the monoalkylated and unreacted species (Sun et al., 1994).

Oligonucleotide Preparation and Purification for NMR Studies. The two strands for the 11-mer [d(CGTTAGTT-ACG)·d(GCAATCAATGC)] were synthesized on a 10-μmol scale by using automated solid-phase phosphotriester and phosphoramidite chemistry (Gait, 1984) on an Applied Biosystems automated DNA synthesizer (Model 381A), leaving the trityl protecting group on each of the single strands. The two strands were purified by reverse-phase HPLC, as outlined previously (Thompson & Hurley, 1995). Equal quantities of the two strands (measured by UV) were mixed and annealed in NMR buffer (10 mM NaH_2PO_4/100 mM NaCl/0.1 mM EDTA, pH 6.85) at 80 °C for 1 h and cooled slowly. The annealed sample was purified on a hydroxylapatite column using a mobile solvent (NaH_2PO_4 buffer, pH 6.85) gradient increasing from 10 to 200 mM. The pure duplex DNA was desalted on a G-25 column. After being lyophilized to dryness, the sample was redissolved in 600 μL of NMR buffer, and a one-dimensional ^{1}H-NMR of the sample was examined. Only duplex signals were observed.

Adduct Preparation and Purification for NMR Studies. The Bizelesin adduct was prepared by stirring 10 mg of Bizelesin in 0.2 mL of DMF solution with 0.03 mL of 1 M sodium bicarbonate for 20 min before adding 30 mg of purified 11-mer in 0.75 mL of twice-concentrated NMR buffer containing 20 mM sodium dihydrogen phosphate and 200 mM sodium chloride (pH 6.85). The reaction mixture was stirred for 48 h in the dark, lyophilized to dryness overnight, and desalted, and the excess drug was removed on C^{18} Sep-Pak cartridges (Waters). The sample was examined by ^{1}H-NMR in buffer (10 mmol of NaH_2PO_4/100 mmol of NaCl/0.1 mM EDTA, pH 6.85) and was found to contain a mixture of duplex and duplex adduct, and hence, a further 2.5 mg of Bizelesin was added (as above) and the reaction allowed to proceed for a further 48 h. Once completed, the sample was again lyophilized to dryness and desalted, and the excess drug was removed on C^{18} Sep-Pak cartridges. One-dimensional ^{1}H-NMR then indicated that the reaction had gone to completion and no further purification was required.

Proton NMR Experiments. One- and two-dimensional 500 MHz ^{1}H-NMR data sets in H_2O and D_2O buffered solution were recorded on a Bruker AMX 500 FT NMR spectrometer. Proton chemical shifts were recorded in parts per million (ppm) and referenced relative to the water signal.

Phase-sensitive two-dimensional NOESY spectra (Bruker) were obtained for two mixing times, 150 and 250 ms. All spectra were acquired with 16 scans at each of 1024 t_1 values, at a spectral width of 10.002 ppm and a relaxation delay of 10 s between scans. During data processing, a 90°-shifted squared sine-bell function was used in both ω_1 and ω_2 dimensions. The FID in ω_1 was zero-filled to 2 K prior to Fourier transformation to give a 2 K × 2 K spectrum. Two-dimensional NOE spectra in 90% H_2O at 150 ms mixing time were recorded using the 1-1 echo read pulse sequence (Sklenar & Bax, 1987; Blake & Summers, 1990) with a 2.5 s pulse repetition time and a sweep width of 24.396 ppm. In addition, select ^{13}C resonances were identified using a heteronuclear multiple-quantum coherence experiment in D_2O (Bax et al., 1983).

Restrained Molecular Dynamics. Interproton distances were derived using the program MARDIGRAS (Borgias & James, 1990) from NOESY experiments using mixing times of 150 and 250 ms. A complete two-dimensional NOE relaxation matrix was set up using the geometry of a starting structure to provide interproton NOEs not available from the experimental data sets. Alternative starting structures were (1) Bizelesin 11-mer adduct nonrestrained molecular dynamics products and (2) minimized B-DNA adducts. The complete NOE matrix was calculated for the starting structure using CORMA (Borgias et al., 1989; Borgias & James, 1988). MARDIGRAS calculates upper and lower boundaries for the interproton distances (Borgias & James, 1990), and subsequent procedures followed those described previously (Kerwood et al., 1991). Interproton distances were incorporated into the rMD and restrained dynamics calculations of the solvated system. The rMD calculations were performed using the SANDER module of AMBER (Pearlman et al., 1991), version 4.0, on an SGI 4D-35G Personal Iris workstation. The AMBER force field pseudoenergy terms for the interproton distances were given the form of flat wells with parabolic sides within a defined distance. The molecular dynamics calculations were started at 0 K and ramped to 20 kcal/(mol·Å^2) over 15 ps to 300 K. During the first 10 ps of this period, the 189 NMR-derived distance restraints were applied to the system along with additional base-pair

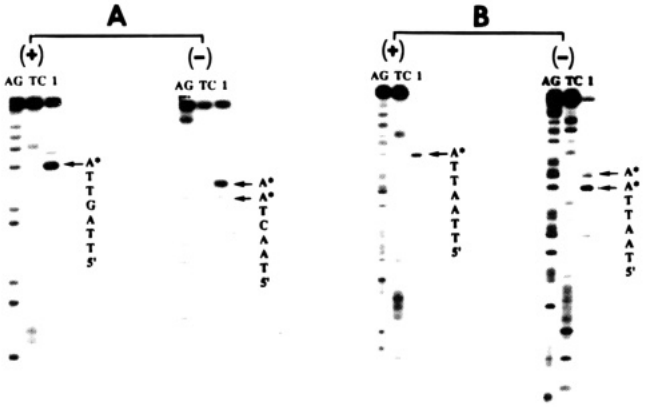

FIGURE 3: Comparison of the Bizelesin alkylation sites on the (+) (upper) and (−) (lower) strands of the 21-mer oligomer duplexes (I and II in Chart 1) containing the sequences 5′-TTAGTTA-3′ (A) and 5′-TTAATTA-3′ (B). AG and TC represent the purine- and pyrimidine-specific cleavage reaction, respectively. In each case, lane 1 corresponds to the thermal strand breakage products (Reynolds et al., 1985) after modification of oligomer DNA with Bizelesin. Arrows correspond to the modified adenine (*) on the sequences.

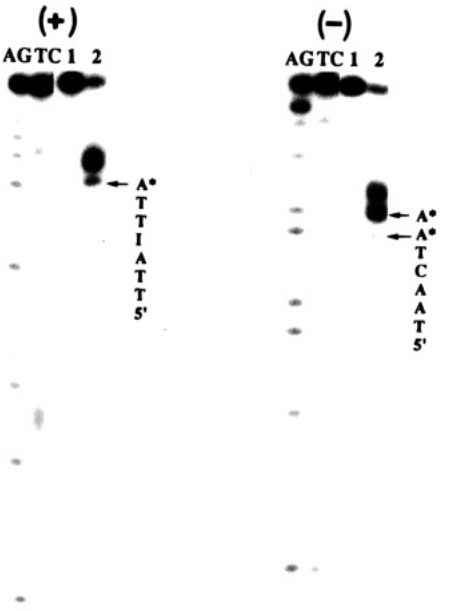

FIGURE 4: Bizelesin alkylation sites on the (+) (upper) and (−) (lower) strand of the 21-mer oligomer sequence (III in Chart 1) containing 5′-T•T•AITTA*-3′ as the cross-linking sequence. Lanes 1 and 2 correspond to unmodified and Bizelesin-modified DNA, respectively. In contrast to the results in Figures 3 and 5, the thermal cleavage reaction, which normally gives rise to products that migrate alongside Maxam and Gilbert sequence reactions, is incomplete, and therefore the major products run at higher molecular weight.

distance and angle restraints, which were applied to the terminal bases to prevent duplex fraying at elevated temperatures. The system was then ramped to 500 K over 20 ps, held for 10 ps, and returned over 20 ps to 300 K. The solvated system was then held at 300 K for 50 ps, and the data from the last 50 ps were used to generate averaged structures (see later) and hydrogen-bonding information.

RESULTS

(A) Gel Electrophoresis Studies

Bizelesin Shows a Unique Cross-Linking Reactivity with a 7-bp Sequence That Has a Requirement for a Central G·C Base Pair. Although Bizelesin is generally known to form interstrand cross-links with adenines primarily spaced 6 bp apart, several 7-bp cross-link sites have been identified from the previous studies (Sun & Hurley, 1993). In particular, in the sequence 5′-T•TAGTTA*-3′ (* indicates one of the cross-linking alkylation sites and • indicates the thymine opposite the other cross-linking site), a 7-bp cross-link site is highly favored over the alternative 6-bp cross-link site (5′-T•TAGTTA*-3′). In the present study, two 21-bp oligomer duplexes (I and II in Chart 1) were designed to compare both the Bizelesin reactivity and the cross-linking sequence specificity in the purely AT sequence 5′-TTAATTA-3′ with the mixed AT/GC-containing sequence 5′-TTAGTTA-3′.[2] As shown in Figure 3, Bizelesin forms almost exclusively a 7-bp cross-link with the sequence 5′-T•TAGTTA*-3′ (A), while the 6-bp cross-link is predominant in the sequence 5′-T•TAATTA*-3′ (B). These results show the apparent important role of the central guanine in facilitating the formation of the 7-bp cross-link. Overall, the alkylation reactivity of duplex I is less than duplex II.

The Exocyclic 2-NH_2 Group of Guanine Directs Cross-Linking Specifically at the 7-bp Cross-Linking Site. In order to test the importance of the exocyclic 2-amino group of guanine in facilitating the formation of a 7-bp cross-link, the 21-bp oligomer duplex containing the sequence 5′-TTAITTA-3′ (III in Chart 1) was constructed and reacted with Bizelesin to measure the amount of the 7- vs 6-bp cross-link. As shown in Figure 4, the replacement of the central guanine with inosine, which lacks the exocyclic 2-amino group substituent, leads to formation of approximately *equal* amounts of the 6- and 7-bp cross-link on the sequence 5′-TTAITTA-3′. This result suggests that the exocyclic 2-NH_2 group of guanine has an important role in directing the cross-linking at adenines spaced 7 bp apart.

Alkylation at 5′-TAA Is Unique to the Cross-Linking CPIs, and Monoalkylation Occurs at the 5′-TTA* Sequence Prior to Cross-Linking at 5′-TAA*.* The underlying molecular mechanism of the 7-bp cross-link was further explored by a comparison of the Bizelesin drug-modification sites on the sequence 5′-TTAGTTA-3′ (oligomer II in Chart 1) with the monoalkylating analogs Adozelesin and (+)-AB (Figure 1). As expected, Adozelesin and (+)-AB showed much higher monoalkylation reactivity with the sequence 5′-TTA* in comparison to the sequence 5′-CTA* (compare lanes 2 and 3 in left and right panels of Figure 5). Furthermore, monoalkylation at the 5′-TAA* sequence was not observed. In contrast, Bizelesin cross-links at the 5′-TTA* and 5′-TAA* sequences (lane 1 in right and left panels of Figure 5). On the basis of a comparison of the results of the cross-linking by Bizelesin and the monoalkylating by Adozelesin and (+)-AB, it can be concluded that for Bizelesin the monoalkylation occurs at the 5′-TTA* sequence and cross-linking at the 5′-TAA* sequence.

Using denaturing gel analysis, the time−course of monoalkylation and cross-linking alkylation produced during the reaction of Bizelesin with oligomer II (Chart 1) was monitored. The results shown in Figure 6 reveal that the formation of cross-linked product apparently occurs without

[2] The base sequence to the 5′-side of the target sequence (TTANTTA) in oligomers II and III was also changed to eliminate the possible competitor 6-bp (A,G) *cross-linked* and *monoalkylation* sites.

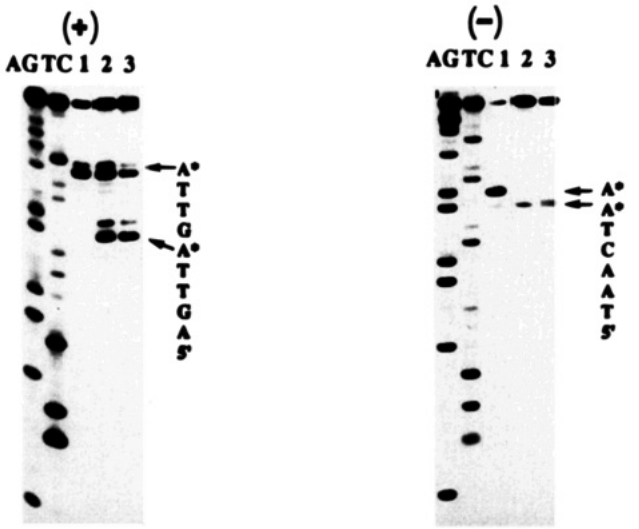

FIGURE 5: Comparison of the Bizelesin and monoalkylating analogs' alkylation sites on (+) (upper) and (−) (lower) strands of the 21-mer oligomer duplex (II in Chart 1) containing the sequence 5′-TTAGTTA-3′. Lanes 1−3 correspond to the thermal strand breakage products of the oligomer DNA modified with Bizelesin, Adozelesin, and (+)-AB, respectively.

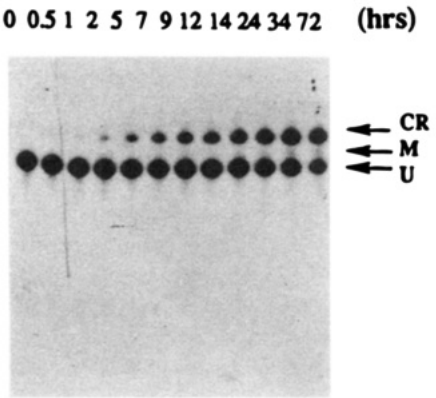

FIGURE 6: Kinetics of the monoalkylation and cross-linking adduct formation by Bizelesin on the 21-mer oligomer duplex containing the sequence 5′-TTAGTTA-3′. The oligomer duplex was reacted with Bizelesin for the indicated period of time at the lane heading before loading onto the denaturing polyacrylamide gel to separate cross-linked DNA from the monoalkylated and unmodified DNAs. CR, M, and U represent cross-linked, monoalkylated, and unmodified species, respectively.

accumulation of detectable amounts of monoalkylation product. This suggests that the cross-linking reaction by Bizelesin on this sequence is very fast compared to the monoalkylation reaction.

(B) High-Field NMR Studies

The gel electrophoresis studies show that the 7-bp cross-linking of the sequence 5′-TTAGTTA by Bizelesin is dependent upon the exocyclic 2-amino group of guanine. In order to investigate the structural role of this minor groove amino group of guanine in mediating the 7- vs the expected 6-bp cross-link, a two-dimensional high-field ^{1}H-NMR study was performed.

Assignment of the Proton Resonances for the Bizelesin 11-mer Duplex Adduct. (i) Duplex Base and Deoxyribose Nonexchangeable Proton Assignments. The general strategy for ^{1}H-NMR assignments of the non-self-complementary duplex and cross-linked Bizelesin adduct was followed as outlined in previous (+)-CC-1065 and Bizelesin studies (Lin et al., 1991; Seaman & Hurley, 1993; Thompson & Hurley, 1995). A single set of methyls and single aromatic H6/H8 to H1′ walks is observed for each of the noncomplementary strands, indicating that only a single adduct is formed (Figure 7). Previous studies on the cross-linked Bizelesin duplex adduct formed with 5′-TAATTA-3′ showed the presence of Hoogsteen and open base-pair conformation within the central A•T step (underlined). To eliminate the possibility of mistaking adenine H2 protons with the aromatic H8 protons and misinterpreting any unusual base-pair conformations between the cross-linked adenines, the ^{13}C shifts for the carbons connected to the aromatic protons were measured and assigned by the use of ^{1}H−^{13}C correlation experiments (unpublished data). Due to the differences within the electronic environment, a carbon connected to an adenine H2 will typically have a chemical shift between 148 and 155 ppm; in contrast, a carbon connected to a PyH6 or PuH8 aromatic proton will have a chemical shift between 130 and 145 ppm. The values obtained for the ^{13}C shifts confirm the assignments of the adenine H2 protons, and on the basis of the connectivities from the adenine H2 protons into duplex and Bizelesin protons located within the minor groove, the possibility of Hoogsteen and non-base-paired bases (previously observed within the Bizelesin duplex adduct formed with 5′-TAATTA*) can be eliminated. A summary of the DNA nonexchangeable proton NMR assignments for the duplex adduct and a comparison to the duplex are given in Table 1A.

(ii) Bizelesin Nonexchangeable Proton Assignments. Nonexchangeable ^{1}H-NMR signals for the Bizelesin molecule are divisible into (+)-CPI subunit and indole subunit clusters (Figure 2). Many of the drug protons are located on the inside edge of the bound ligand and generate numerous drug-to-DNA NOESY cross-peaks. COSY cross-peaks for ortho-related aromatic protons and NOEs (H_2O) between indole H1′ proton resonances and their neighboring H8′ confirm the H3′, H5′, H7′, and H8′ assignments (Table 1B). The CPI subunit H7 protons are assigned on the basis of weak connectivities into the H4′ of the covalently modified bases, connectivities into the CPI phenol moiety, and the ^{13}C shifts from the ^{13}C−^{1}H correlation experiment.

(iii) DNA and Bizelesin Exchangeable Proton Assignments. Two-dimensional NOE (H_2O) experiments were conducted, and the exchangeable imino signals were assigned on the basis of their connectivity networks into the aromatic region and connectivities into the imino protons of the neighboring bases. All exchangeable imino signals for the duplex and duplex adduct were detected except for those associated with the terminal bases, G11 and G22. Consistent with previously studied CC-1065 (Lin & Hurley, 1990) and U-77,779 (Seaman & Hurley, 1993) adducts, the imino signals for T3 and T14 (11.30 and 10.93 ppm, respectively) were significantly upfield-shifted in the duplex adduct, partly confirming the location of the covalently modified adenines (Figure 8). The CPI subunit NH1 and indole subunit NH1′ protons for both ends of the ligand were assigned on the basis of their respective connectivities to the CPI subunit H2 and indole subunit H8′. ^{1}H-NMR signals for the CPI-phenol protons (H8) were assigned on the basis of connectivities into the Bizelesin CPI subunit H7 proton and duplex H4′ protons (Table 2). The resonance signals for the two phenol protons were observed at 11.06 and 11.10 ppm (Figure 8), contrasting

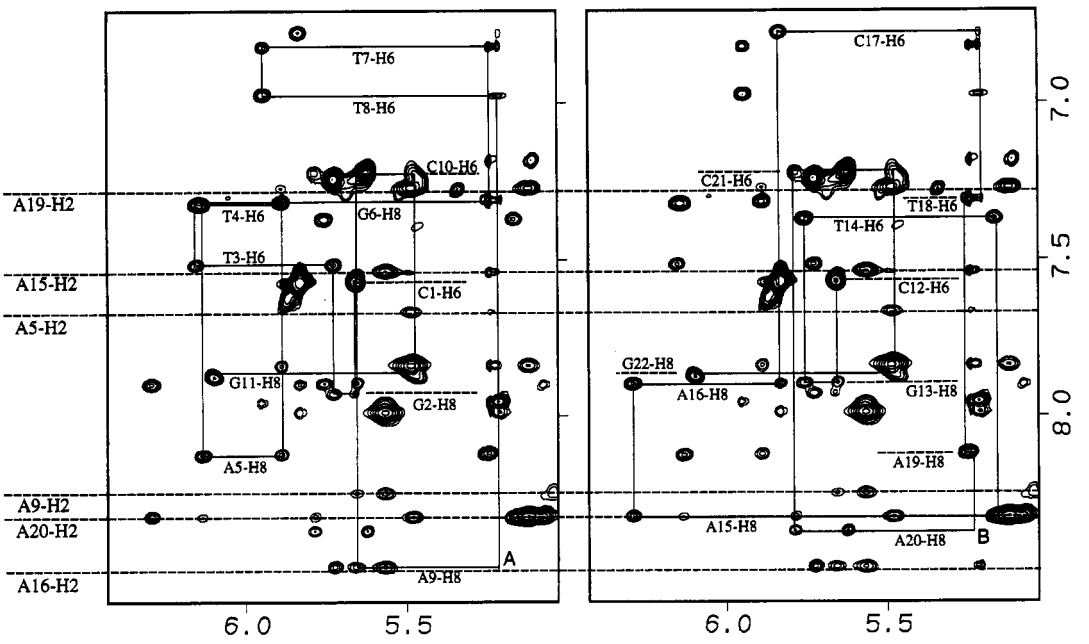

FIGURE 7: Two-dimensional phase-sensitive NOESY spectrum (250 ms mixing time) expanded contour plot of the Bizelesin cross-linked duplex adduct in buffered D_2O solution, pH 6.85, at 300 K, displaying connectivities for aromatic PuH8/PyH6 to deoxyribose H1′ and to deoxyribose H1′ protons of the 5′-neighbor (strands are shown separately in left and right panels). Missing connectivities between the H8 protons of adenine A9 and adenine A20 are indicated as A and B, respectively.

Table 1

(A) Chemical Shifts (ppm) of Nonexchangeable Protons in the [d(CGTTAGTTACG)·d(GCAATCAATGC)] Bizelesin Adduct at Room Temperature[a]

	H8		H6		H2		H5		H1′		H2′		H2″		H3′		H4′		H5′/H5″
C1			7.58	*(−0.09)*			5.84	*(−0.08)*	5.67	*(−0.13)*	2.30	*(+0.24)*	1.89	*(−0.57)*	4.65	*(−0.07)*	4.04	*(−0.33)*	2.65/3.65
G2	7.93	*(−0.07)*							5.73	*(−0.32)*	2.45	*(−0.27)*	2.67	*(−0.18)*	4.94	*(−0.06)*	4.30	*(−0.10)*	3.90/4.10
T3			7.52	*(+0.21)*			1.62	*(+0.18)*	6.16	*(+0.09)*	2.47	*(+0.34)*	2.26	*(−0.32)*	4.90	*(+0.03)*	4.34	*(−0.06)*	4.12/4.17
T4			7.32	*(−0.05)*			1.78	*(+0.10)*	5.90	*(+0.15)*	2.91	*(+0.78)*	2.34	*(−0.16)*	4.62	*(−0.28)*	1.84	*(−2.60)*	3.22/3.28
A5	8.14	*(−0.09)*			7.67	*(+0.50)*			6.14	*(+0.08)*	2.67	*(−0.06)*	2.13	*(−0.79)*	4.92	*(−0.13)*	2.68		
G6	7.31	*(−0.24)*							5.24	*(−0.59)*	2.24	*(−0.20)*	1.92	*(−0.79)*	4.62	*(−0.25)*	3.70	*(−0.70)*	4.01/4.13
T7			6.83	*(−0.36)*			0.90	*(−0.31)*	5.95	*(−0.01)*	2.38	*(+0.32)*	1.75	*(−0.79)*	4.58	*(−0.25)*	4.09	*(−0.13)*	2.50/3.59
T8			6.97	*(−0.40)*			1.44	*(−0.20)*	5.22	*(−0.56)*	1.12	*(−0.95)*	1.31	*(−1.18)*	4.38	*(−0.52)*	1.75		4.20/3.61
A9	8.48	*(+0.16)*			8.25	*(+0.70)*			5.67	*(−0.52)*	2.57	*(−0.14)*	3.07	*(+0.22)*	4.98	*(−0.06)*	4.57	*(+0.13)*	4.17/3.93
C10			7.23	*(−0.05)*			5.73	*(+0.38)*	5.47	*(−0.19)*	2.16	*(+0.31)*	1.66	*(−0.63)*	4.57	*(−0.21)*	2.10		3.33/3.75
G11	7.86	*(−0.04)*							6.10	*(−0.04)*	2.36	*(−0.00)*	2.64	*(+0.06)*	4.72[b]	*(+0.07)*	4.27	*(+0.11)*	3.86/3.86
C12			7.58	*(−0.07)*			5.84	*(−0.08)*	5.66	*(−0.12)*	2.30	*(+0.28)*	1.89	*(−0.55)*	4.67	*(−0.03)*	4.04	*(−0.33)*	2.65/3.65
G13	7.90	*(−0.08)*							5.76	*(−0.20)*	2.49	*(−0.17)*	2.63	*(−0.15)*	4.92	*(−0.06)*	4.29	*(−0.07)*	3.91/4.05
T14			7.37	*(+0.13)*			1.64	*(+0.13)*	5.17	*(−0.41)*	2.41	*(+0.39)*	2.23	*(−0.12)*	4.87	*(0.00)*	4.22	*(+0.03)*	4.10/4.10
A15	8.33	*(+0.09)*			7.51	*(+0.47)*			6.29	*(−0.32)*	2.95	*(+0.20)*	2.78	*(−0.14)*	5.07	*(+0.02)*	4.47		4.14/4.18
A16	7.91	*(−0.20)*			8.48	*(+0.93)*			5.83	*(−0.23)*	2.70	*(+0.07)*	2.18	*(−0.62)*	4.55	*(−0.46)*	1.73	*(−2.74)*	3.27/3.40
C17			6.79	*(−0.40)*			4.80	*(−0.31)*	5.23	*(−0.42)*	1.85	*(−0.01)*	1.23	*(−1.17)*	4.38	*(−0.21)*	2.20	*(−1.97)*	3.73/3.85
T18			7.31	*(−0.01)*			1.42	*(−0.13)*	5.25	*(−0.38)*	2.24	*(+0.20)*	1.85	*(−0.55)*	4.61	*(−0.26)*	4.15		
A19	8.12	*(−0.12)*			7.26	*(+0.32)*			5.24	*(−0.64)*	2.12	*(−0.87)*	2.40	*(−0.77)*	4.49	*(−0.57)*	3.38	*(−1.01)*	4.02/4.09
A20	8.37	*(+0.26)*			8.32	*(+0.66)*			5.79	*(−0.28)*	2.58	*(+0.02)*	2.82	*(+0.02)*	4.97	*(−0.04)*	4.54	*(+0.11)*	3.60/4.22
C21			7.22	*(+0.02)*			5.63	*(+0.39)*	5.48	*(−0.15)*	2.16	*(+0.34)*	1.66	*(−0.60)*	4.57	*(−0.30)*	1.85	*(−2.55)*	3.04/3.64
G22	7.85	*(−0.02)*							6.10	*(−0.04)*	2.36	*(+0.01)*	2.64	*(+0.07)*	4.72[b]	*(+0.08)*	4.27	*(+0.12)*	3.86/3.86

(B) Proton Chemical Shifts (ppm) for the Bound Bizelesin in the Duplex Adduct at Room Temperature[a]

H2	H3Me	NH1	H4a	H4A	H4B	H5A	H5B	H7	H3′	H5′	H7′	H8′	NH1′	NH9′
Bizelesin CPI and Indole Subunit (Linked to Adenine 9)														
7.38	2.75	10.45	4.50	4.64	4.97	5.58	4.94	7.68	7.99	7.95	7.87	7.57	10.99	
Bizelesin CPI and Indole Subunit (Linked to Adenine 20)														
7.40	2.70	10.52	5.13	4.64	4.50	5.49	4.92	7.85	7.84	7.54	7.59	7.38	11.22	

[a] Chemical shift differences, $\delta_{DNA-drug} - \delta_{DNA}$, are shown in italics; shift differences greater than 0.25 ppm are underlined). Sample was dissolved in 0.5 mL of D_2O buffer containing 0.1 M sodium chloride and 10 mM sodium dihydrogen phosphate, pH 6.85, at 300 K. [b] Signal not visible due to H_2O suppression.

to previous studies on a (+)-CC-1065 adduct (Lin et al., 1991) and a Bizelesin adduct (Thompson & Hurley, 1995), in which the signals for the phenol protons were observed significantly downfield to expected values [13.62 ppm within the (+)-CC-1065 adduct and 13.05 and 13.80 ppm within the cross-linked Bizelesin adduct]. On the basis of the unusually downfield shifts observed within the (+)-CC-1065 adduct, it was predicted and confirmed by the use of ^{1}H-NMR coupled with ^{17}O-labeled water and phosphates that the phenolic proton in the (+)-CC-1065 duplex adduct was hydrogen bound (through an ordered water molecule) to the phosphate backbone (Lin et al., 1991). On the basis of the

Table 2: Intensities[a] of the NOE Connectivities between Bizelesin and the Covalently Modified Duplex

	(A) Bizelesin and Indole Subunit (Linked to Adenine 9)										
	H3Me	H4a	H4A	H4B	H5A	H5B	H3′	H5′	H7	H7′	H8′
A9-H2	S	S	S	S	M	M					
A15-H2	M	W	M	M	S	S	M				
A16-H2							S	S			
T8-H1′							M	S			
A9-H1′									W		
A9-H4′									W		
C10-H1′	S										
C10-H4′									W		
A16-H1′							W				

	(B) Bizelesin and Indole Subunit (Linked to Adenine 20)											
	H3Me	H4a	H4A	H4B	H5A	H5B	H2	H3′	H5′	H7	H7′	H8′
A20-H2	S	S	S	S	M	M		W				
A19-H2	M	M			S	S		M	W			
A5-H2					M	M		S	S			
T3-H1′	S											
T3-H4′										W		
T4-H1′	M							W				
T4-H4′	S						W					
T4-H5′/5″	S											
A18-H4′												M
A20-H1′	W		S	S								
A20-H4′			M	M								
C21-H1′	S		W									
C21-H4′							W					

[a] W = weak, M = medium, and S = strong (relative to cytosine H5 to H6 cross-peak intensities).

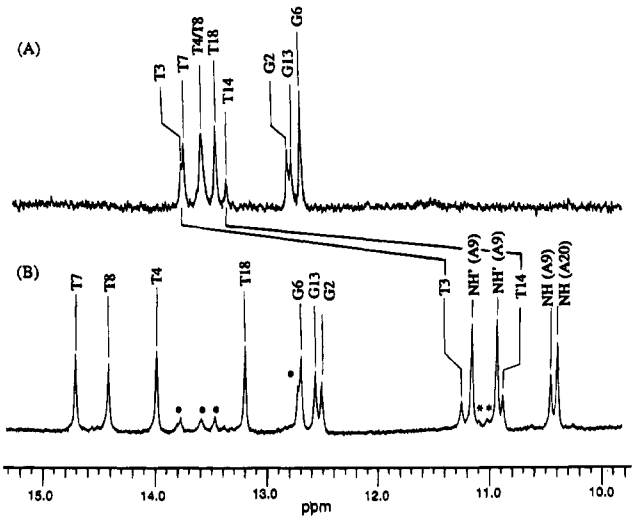

FIGURE 8: One-dimensional proton NMR spectra (500 MHz; 8.5–15 ppm, downfield region) of the 11-mer duplex (A) and cross-linked Bizelesin adduct (B). Signals from the unreacted duplex are indicated by a •, and broad signals for two CPI phenol protons are indicated by an *. The upfield-shifted amino signals for thymine T3 and T14 are highlighted.

chemical shifts for the two phenolic protons in this Bizelesin duplex adduct, it would appear that there is no significant hydrogen bonding into the phosphate backbone. Signals for the amide NH protons within the ureylene linking moiety could not be found in either the one- or two-dimensional ^{1}H-NMR data.

Bizelesin Occupies the Minor Groove and Is Covalently Attached Through A9 and A20. There are sufficient DNA and Bizelesin protons within the minor groove to unambiguously define the location of the Bizelesin molecule and the sites of covalent attachment. Connectivities between the adenine H2 protons along the length of the duplex into the CPI linkers' H4A and H4B protons, the CPI subunits' H5A and H5B protons, and the indoles' H3′ and H5′ protons confirm not only the location and orientation of the ligand but also the assignment of the six adenine H2 protons (connectivities from A20, A19, and A5 are shown in Figure 9; all drug to duplex connectivities are shown in Table 2). The intense cross-peaks from the H2 protons of A9 and A20 into the CPI subunits' H4a, H4A, H4B, H5A, and H5B protons (shown for A20 in Figure 9), in addition to strong connectivites into the H2 and methyl resonances of the CPI subunits (A and B in Figure 10), confirm that A9 and A20 are the covalently modified adenines. Cross-peaks are not observed from the duplex protons within the base of the minor groove to the ligand aromatic protons H7, H7′, and H8′. However, weak cross-peaks exist from the CPI subunit H7 and the indole subunits H7′ and H8′ into duplex H4′ protons (Table 2) located high on the walls of the minor groove, confirming the assignment of the ligand proton and indicating that the Bizelesin is oriented edge on into the minor groove.

Proton NMR Studies Indicate the Presence of Distortion in the Duplex Adduct around the Covalently Modified Adenines. Initial molecular modeling studies indicated that in order for the Bizelesin to cross-link 7 bp there would have to be a considerable amount of distortion introduced in the DNA. The ^{1}H-NMR data indicate that the modified adenines (A9 and A20) and the 5′-neighboring bases are distorted and propeller-twisted, which results in N3 being pulled toward the center of the duplex. Strong connectivities between the H8 of the modified adenines and their own H2′ and H2″ can be observed for adenines A9, A19, and A20. However, for the same bases, unusually weak connectivities are seen between the H8 protons and their respective 5′ H2′ and H2″ neighbor protons (A20, A19, and A9 aromatic H8 to H2′ and H2″ connectivities are shown in Figure 10). Furthermore, the cross-peaks between the H8 of the modified adenines (A9 and A20) and their respective neighbors' H1′

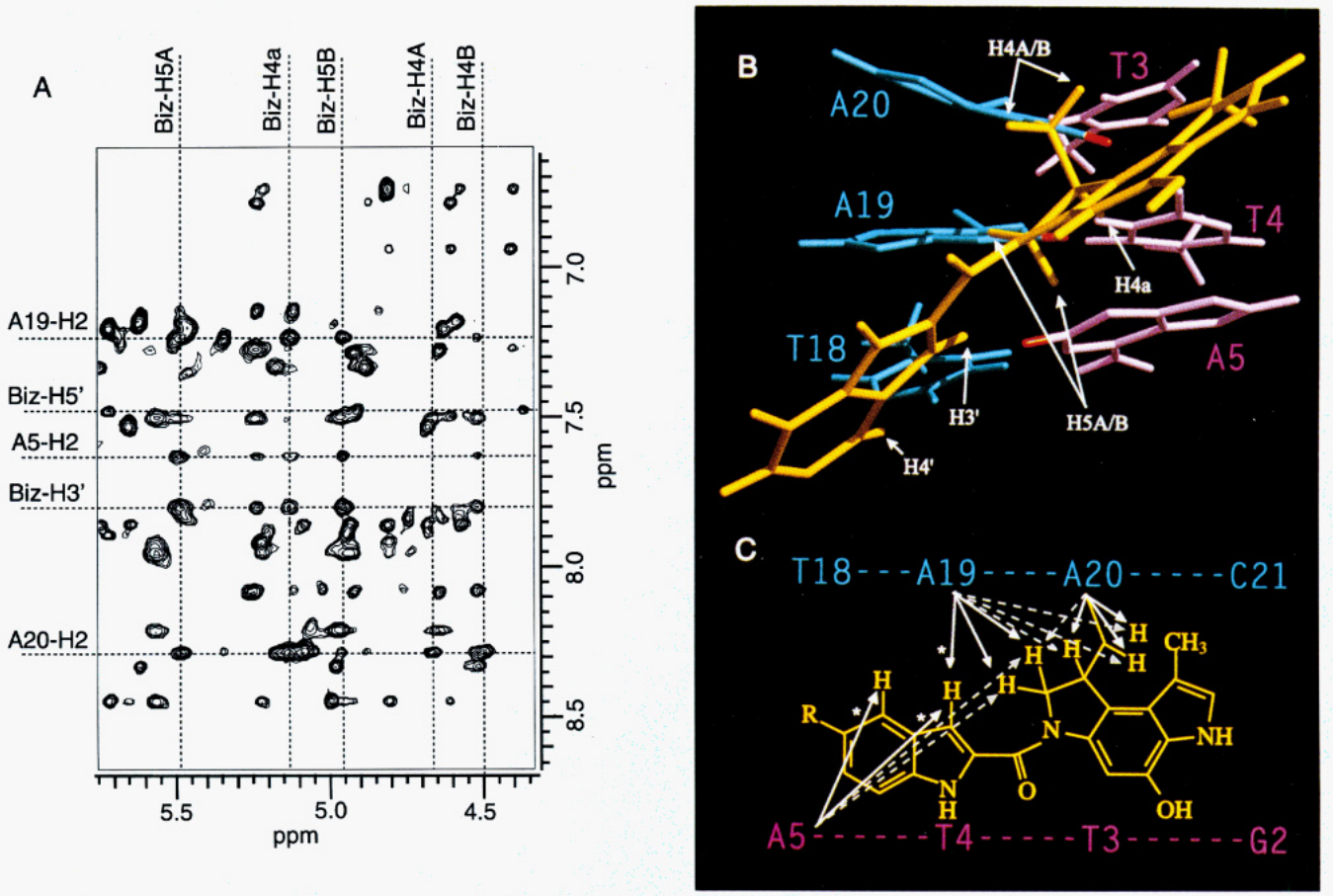

FIGURE 9: NOE connectivities between select adenine H2 protons and the Bizelesin geminal H4A, H4B, H5A, H5B, and H4a protons. (A) Contour plot of the NOESY data showing the various adenine H2 to ligand H4A, H4B, H5A, H5B, and H4a cross-peaks for the alkylated adenine at A20. (B) Relative positions of the adenine H2 protons (red) and the key protons on the ligand (averaged rMD structure). (C) Relative intensities of the adenine H2 connectivities (intense cross-peaks are shown in solid white; less intense cross-peaks, in dashed white).

are exceptionally weak and missing in the aromatic H6/H8 to H1′ walks for each of the noncomplementary strands (A and B in Figure 7). The relative weakness in the connectivities between the H8 and the sugar protons of the 5′-neighboring base reflects an increase in distance resulting from the twist and distortion in the covalently modified bases and their immediate 5′-neighbors (see panels B and C in Figure 10).

Restrained Molecular Dynamics Studies on the Bizelesin Cross-Linked Duplex Provide a Model for Locating the Distortion of the Duplex and Insight into the Role of the Linker Moiety in Determining the Sequence Specificity of Bizelesin toward This Sequence. The initial rMD studies using a minimized B-form adduct as a starting structure repeatedly failed because the duplex adduct has a tendency to lose base pairing at one or both of the covalently modified adenines. On the basis of the exchangeable ^{1}H-NMR data, in which imino proton signals are observed for the paired thymines of both covalently modified adenines, additional distance and angle restraints were applied to maintain base pairing and stabilize the initial steps of the rMD calculations. Once the initial calculations had been performed and the duplex adducts' structure had stabilized at 300 K, these restraints were ramped down over a period of 15 ps, and the duplex adduct was found to behave normally, maintaining Watson−Crick base pairing throughout the rest of the molecular dynamics study.

The molecular dynamics studies provide an average structure that can be used as a model to explain the unusual ^{1}H-NMR data, indicating distortion in the duplex adduct. The distortion of the helix is focused around the two covalently modified adenines (A9 and A20). In order to accommodate a ligand that is too short to span the five base pairs between the two covalently modified adenines, the immediate 5′-neighbors are twisted toward the center of the duplex, thus effectively reducing the distance between the cross-linked bases.

Within the first few picoseconds of the rMD study, a strong hydrogen bond is formed between the carbonyl of the ureylene linker and the exocyclic amine of the central guanine (G6). This hydrogen bond, once established, showed 100% occupancy within the duplex adduct, even at elevated temperatures (500 K), and restricted the movement of the linker moiety. Within the averaged structure, the ureylene linker was found to be twisted to almost 90° out of plane from the two indole rings flanking it (84° at A9 and 85° at A20 ends). This provides the molecule with a curved conformation closely fitting the shape of the minor groove of the DNA. The carbonyl of the ureylene linker was in exactly the same plane as the central guanine (G6) in the adduct, giving hydrogen-bonding distances to the exocyclic $-NH_2$ of 1.56 Å to the non-hydrogen-bounded proton and 2.48Å to the proton involved in base pairing. Since both of these distances are within the normal hydrogen-bonding

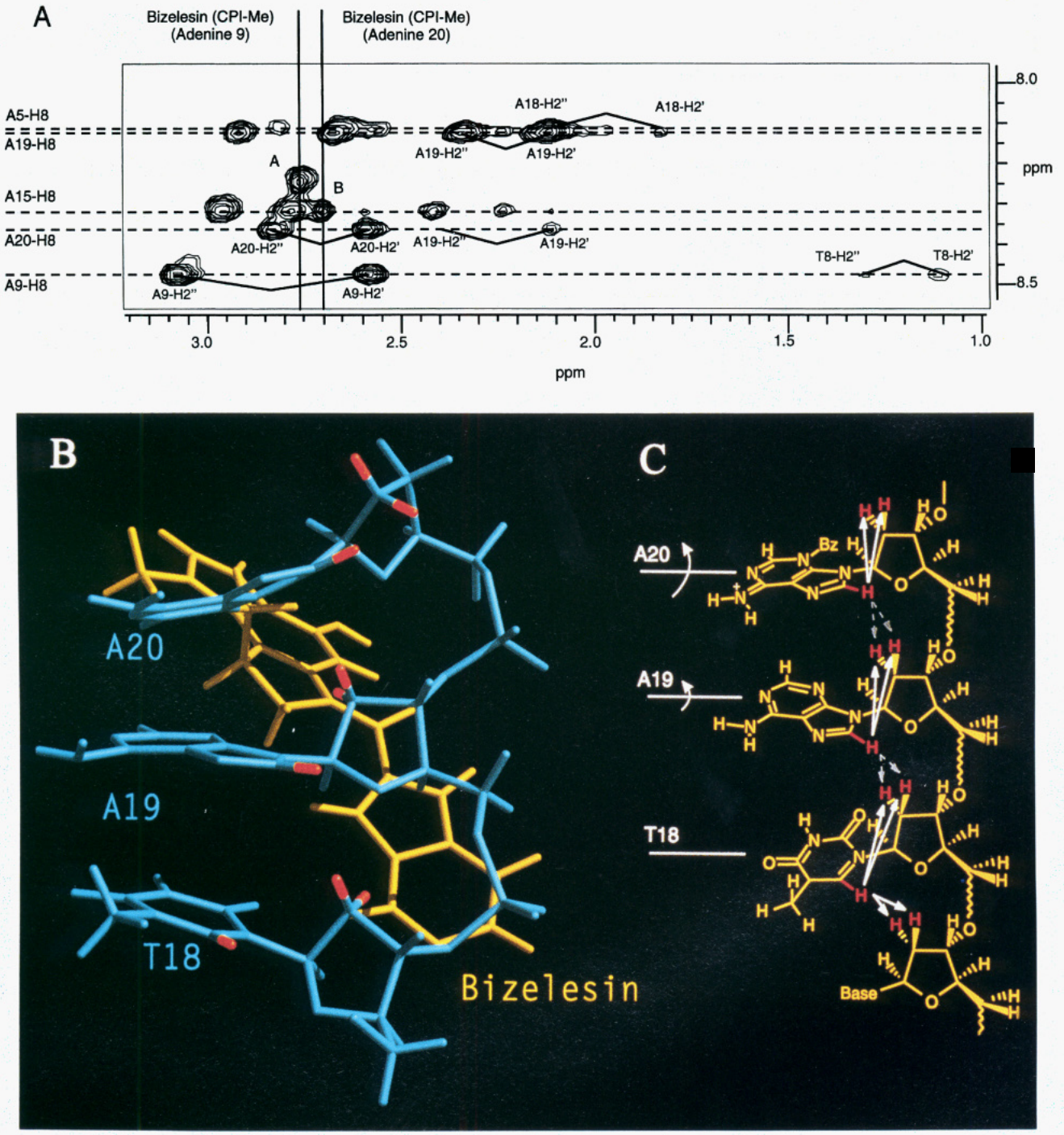

FIGURE 10: Connectivities between the base H8 protons (A20 and A19) and the H2′ and H2″ of the associated sugar and the sugar of the immediate 5′-neighbor. (A) Contour plot of the NOESY data showing the adenine H8 (A20, A19, and A9) to sugar H2′ and H2″ protons. (B) Relative positions of the adenine H6/H8 protons of the T18 to A20 bases (red) and the H2′ and H2″ sugar protons (red) (averaged rMD structure). (C) Relative intensities of the adenine H8 connectivities (intense cross-peaks are shown in solid white; less intense cross-peaks, in dashed gray).

range (1.0−2.5 Å), an unbalanced three-center donor−acceptor−donor arrangement could be envisaged, although the significance of the longer bond would be minimal. There is also a second transient hydrogen bond between the amido protons of the ureylene linker and the deoxyribose oxygen of thymine (T18). However, the significance of this hydrogen bond is unclear, because the hydrogen bond distance is relatively large and the deoxyribose oxygen is a relatively poor hydrogen bond acceptor.

The last 50 ps (300 K) of the rMD calculation was averaged to give the structure shown in Figure 11; the insert shows the central three bases and the hydrogen bond formed between the carbonyl in the ureylene linker and the exocyclic amino group of the central guanine (G6).

DISCUSSION

One- and two-dimensional ^{1}H- and ^{13}C-NMR analyses of the cross-linked Bizelesin adduct reveal that a 7-bp region of the central duplex region is cross-linked to yield one adduct species in which normal Watson−Crick base pairing is maintained. NOE connectivities between nonexchangeable protons on the bound ligand and the duplex determine both the location and orientation of the drug within the minor groove and indicate that the ligand is bound edge on into

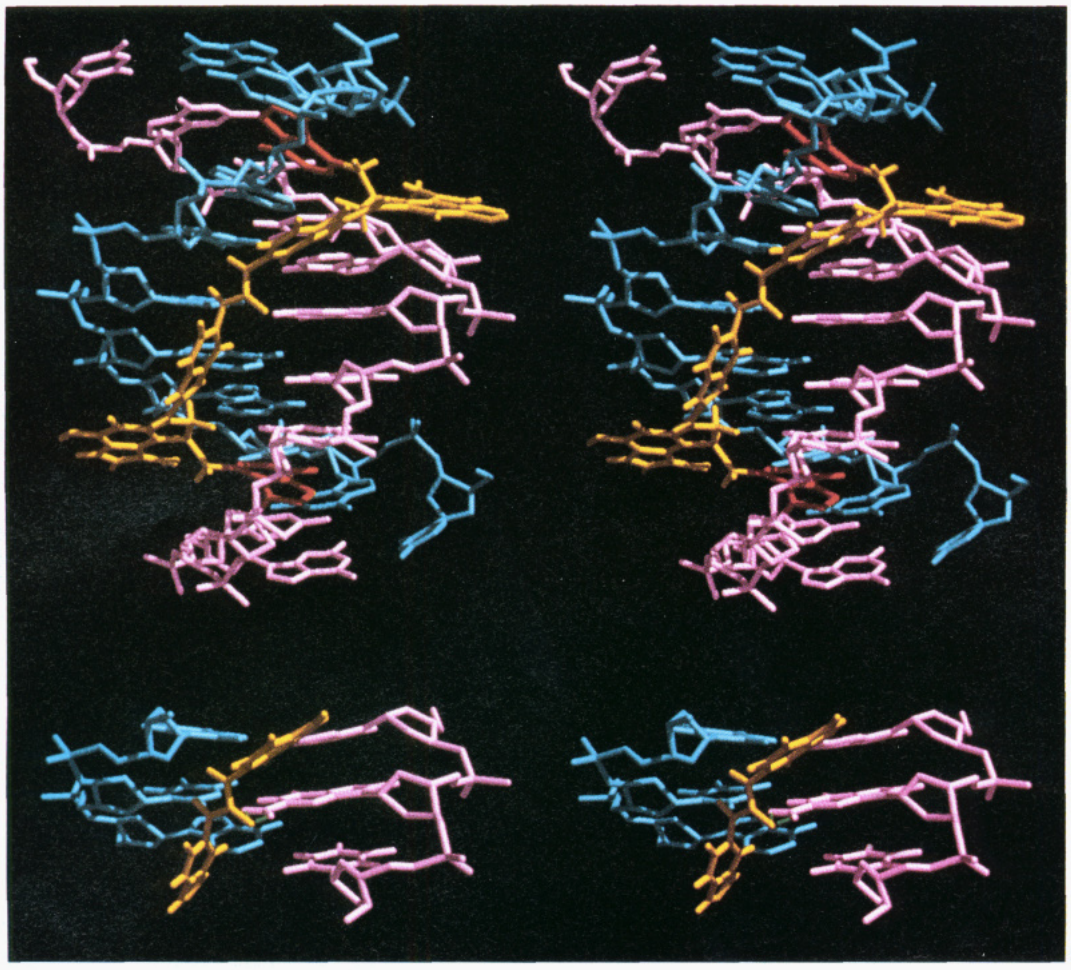

FIGURE 11: Stereo figures showing the average structure from the last 50 ps of the rMD calculation for the Bizelesin 11-mer duplex adduct (300 K). Bizelesin is shown in yellow, the duplex is shown in cyan and magenta, and the modified adenines are shown in red (top). Expansion shows the hydrogen bonds (green) between the carbonyl of the ureylene linker and the protons on the exocyclic amine of guanine (G6) (bottom).

the minor groove (Figure 11). In addition, the adenine H2 connectivities into the protons on the ureylene linking moiety and Bizelesin methyl protons confirm the location of the covalently modified adenines and the site (N3) of covalent attachment. The chemical shifts for the two CPI-phenol protons indicate an absence of hydrogen bonding into the phosphate backbone in this duplex adduct. This is in contrast to previous studies using (+)-CC-1065 (Lin et al., 1991) and Bizelesin (Thompson & Hurley, 1995), in which the phenolic protons on the CPI subunit are shifted downfield as a result of hydrogen bonding into the phosphate backbone.

Bizelesin cross-linking sites can be conveniently divided into three categories. The first is typified by the sequence 5′-ṪAATTA*-3′ and represents a series of cross-link sites in which 6 bp are cross-linked. Studies with the sequence 5′-TAATTA*-3′ indicated that the central A·T base pairs were trapped into Hoogsteen and open base-pair conformations, although there is no evidence that this is a common feature in all 6-bp cross-linked sequences. The second type of sequence cross-linked by Bizelesin contains adenine tracts (e.g., 5′-ṪAAAAA*-3′). These highly reactive sites, when cross-linked by Bizelesin, lose the intrinsic bending observed within the unmodified duplexes (Thompson et al., 1995; Thompson & Hurley, 1995). The third group of sequences contains two adenines 7 bp apart, and a representative of this group is studied here. In most of these 7-bp sequences, the two adenines cross-linked represent the only possible cross-link site (i.e., no 6-bp site available). Others, like the sequence studied here, contain possible 6- *or* 7-bp cross-link sites, and due to structural features within the sequence, the 7-bp sequence is targeted in overwhelming preference to the potential 6-bp site.

This study demonstrates the importance of the central G·C base pair in the formation of the 7-bp cross-link in the sequence 5′-TTAGTTA-3′. We also address the nature of the distortion of the duplex that is required for the 7-bp cross-link to form and propose a possible mechanism for the formation of the 7-bp cross-link in the sequence.

Distortion around the Covalently Modified Adenines Leads to a Reduction in the Distance between the Cross-Linked Bases, Allowing Bizelesin To Form the Extended Cross-Link. Previous NMR and molecular modeling studies have shown that the Bizelesin molecule is almost the exact length required to cross-link two adenines 6 bp apart (Seaman & Hurley, 1993; Thompson & Hurley, 1995). On the basis of this observation, it was clear that, in order for the ligand to cross-link 7 bp, distortion would have to be induced into the duplex. Analysis of the NMR data provides evidence for distortion around both of the covalently modified adenines (A9 and A20). These adenines are pulled toward the center

of the duplex, resulting in a high degree of propeller twist within the drug-modified base pairs. This base-pair distortion is also clearly transmitted to the immediate internal base pairs, particularly adenine A19, which appears to be stacked below the twisted adenine A20. From the molecular dynamics study it is clear that the covalently modified adenines are pulled into the minor groove and toward the center of the duplex, decreasing the distance between the cross-linking sites by around half a base pair. This effectively reduces the distance between the respective N3s of each modified adenine to about that found in the 6-bp cross-linked adducts. The induced distortion may be responsible for the subsequent disruption of hydrogen bonding normally seen between the CPI-phenol proton and the phosphate backbone in the duplex adduct. If the ligand is displaced by around half a base pair at either end of the duplex adduct, hydrogen bonding would no longer be possible because the phenolic proton would be positioned over the sugar rather than the phosphate of the backbone.

Sequence Selectivity Studies Reveal That the Central Guanine Is Essential in the Formation of a 7-bp Cross-Link with This Sequence. Although the distortion in the modified duplex explains how Bizelesin is able to cross-link the additional distance in a 7-bp sequence, it does not immediately reveal a mechanism that explains how the 7-bp cross-link is formed *in preference* to the more usual 6-bp cross-linked adduct (i.e., 5′-$\dot{T}$TAGTTA* rather than 5′-T$\dot{T}$AGTTA*; see Figure 3). The critical importance of the guanine (G6) in the center of the 7-bp cross-linked sequence is demonstrated in the gel electrophoresis studies. Replacing the central guanine (G6) with an adenine results in a 6-bp cross-link formed between the equivalents of adenine A9 and adenine A19. Furthermore, replacement of the guanine with inosine results in a mixture of about equal amounts of the 6- and 7-bp cross-linked sequences (Figure 4). These results demonstrate that it is the exocyclic 2-amino group of guanine located in the minor groove of DNA that apparently facilitates the ability of the Bizelesin to distort this sequence so that a 7-bp cross-link can be almost exclusively achieved.

The molecular dynamics studies on the Bizelesin cross-linked 7-bp duplex adduct indicate that a stable hydrogen bond is quickly formed between the carbonyl of the ureylene linker and the exocyclic amine of the central guanine within the duplex adduct (Figure 11, bottom). We propose that the formation and stability of this hydrogen bond provides critical insight into the origin of the sequence preference for this 7-bp guanine-containing sequence. Prior to cross-link formation, a monoadduct must be formed with alkylation at either A9 or A20. The gel electrophoretic studies show that the monoalkylating (+)-CC-1065 analogs [Adozelesin and (+)-AB] have a sequence preference for 5′-TTA* and do not alkylate 5′-TAA* (see Figure 5). Therefore, the 5′-TTA* sequence is the presumed site of the monoalkylation prior to cross-linking, which subsequently occurs very rapidly (see Figure 6). Once the monoalkylated Bizelesin adduct has formed, the carbonyl of the ureylene linker will be positioned in the minor groove close to the exocyclic 2-amino group of the central guanine (G6). However, on the basis of the model of the Bizelesin monoalkylated duplex adduct, without distortion around the covalently modified adenine (A9), the linker would not be positioned directly over the exocyclic 2-amino group. As the hydrogen bond between the ureylene carbonyl and the exocyclic amino group of guanine is formed, Bizelesin is pulled down the minor groove, producing distortion in the modified adenine.[3] The distortion produced at adenine A9 would then mimic that observed in the cross-linked adduct and would consequently position the second alkylating CPI moiety closer to adenine A20 than to A19, thus leading to the preferential formation of a 7-bp rather than a 6-bp cross-link.

The results of this study coupled with those of the previous study (Seaman & Hurley, 1993) of the Bizelesin duplex adduct formed with 5′-TAATTA-3′ clearly indicate that the ureylene linker can play a critical role in the sequence selectivity of Bizelesin. Within the 5′-TAATTA-3′ sequence, inherent instability within the central A·T bases leads to base-pair opening. Once these base pairs are opened, hydrogen bonding between the ureylene linker amido protons and the bases stabilizes the normally transient Hoogsteen and open base-pair conformations. In parallel, the present study of the duplex adduct formed with 5′-TTAGTTA-3′ also demonstrates the importance of ureylene linking in interaction with the central guanine in the formation of the 7-bp cross-link. In contrast to these sequences, in which the ureylene linkage plays an important stabilizing H-bonding role, the ureylene linkage appears to play merely a spacer role in A-tract sequences. On the basis of these observations, it may be possible, by the careful choice of linkers, to subtly change the sequence specificity and reactivity of these dimeric compounds, with resulting modulation of biological consequences. The sequence specificity of other cross-linking CPIs with different linkers will be reported in due course.

[3] It is at this step that we propose the H-bonding between the 9-phenolic proton of the CPI moiety and the phosphate involved in the initial catalysis (Lin et al., 1991) is lost.

ACKNOWLEDGMENT

We thank David Bishop for preparing, proofreading, and editing the manuscript and an anonymous reviewer for helpful comments.

REFERENCES

Aristoff, P. A., & McGovren J. P. (1993) *Drug News and Perspectives 6*, 229−234.

Bax, A., Griffey, H., & Hawkins, B. (1983) *J. Magn. Reson. 55*, 301−315.

Bhuyan, B. K., Newell, K. A., Crampton, S. L., & Von Hoff, D. D. (1982) *Cancer Res. 42*, 3532−3537.

Blake, P. R., & Summers, M. F. (1990) *J. Magn. Reson. 86*, 622−624.

Borgias, B. A., & James, T. L. (1988) *J. Magn. Reson. 79*, 493−512.

Borgias, B. A., & James, T. L. (1990) *J. Magn. Reson. 87*, 475−487.

Borgias, B. A., Thomas, P. D., & James, T. L. (1989) *Complete relaxation analysis (CORMA)*, University of California, San Francisco.

Borgias, B. A., Gochin, M., Kerwood, D. J., & James, T. L. (1990) *Spectroscopy 22*, 83−100.

Burris, H., Earhart, R., Kuhn, J., Shaffer, D., Smith, L., Weiss, G., Kasunic, D., Radbury, G., Campbell, L., & Von Hoff, D. D. (1992) *Proc. Am. Assoc. Cancer Res. 33*, 520.

Ding, Z.-M., & Hurley, L. H. (1991) *Anti-Cancer Drug Des. 6*, 427−452.

Fleming, G. F., Ratain, M. J., O'Brien, S. M., Vogelzang, N. J., & Earhart, R. H. (1992) *Proc. Am. Assoc. Cancer Res. 33*, 265.

Gait, M. J., Ed. (1984) *Oligonucleotide Synthesis—A Practical Approach*, IRL Press, Oxford, England.
Hanka, L. J., Dietz, A., Gerpheide, S. A., Kuentzel, S. L., & Martin, D. G. (1978) *J. Antibiot. 31*, 1211−1217.
Hurley, L. H., Reynolds, V. L., Swenson, D. H., Petzold, G. L., & Scahill, T. A. (1984) *Science 226*, 843−844.
Hurley, L. H., Lee, C.-S., McGovren, J. P., Warpehoski, M. A., Mitchell, M. A., Kelly, R. C., & Aristoff, P. A. (1988) *Biochemistry 27*, 3886−3892.
Kerwood, D. J., Zon, G., & James, T. L. (1991) *Eur. J. Biochem. 197*, 583−595.
Lee, C.-S., & Gibson, N. W. (1993a) *Biochemistry 32*, 2592−2600.
Lee, C.-S., & Gibson, N. W. (1993b) *Biochemistry 32*, 9108−9114.
Lee, C.-S., Sun, D., Kizu, R., & Hurley, L. H. (1991) *Chem. Res. Toxicol. 4*, 203−213.
Li, L. H., Swenson, D. H., Schpok, S. L., Kuentzel, S. L., Dayton, B. D., & Krueger, W. C. (1982) *Cancer Res. 42*, 999−1004.
Li, L. H., Kelly, R. C., Warpehoski, M. A., McGovren, J. P., Gebhard, I., & DeKoning, T. F. (1990) *Invest. New Drugs 9*, 137−148.
Lin, C. H., & Hurley, L. H. (1990) *Biochemistry 29*, 9503−9507.
Lin, C. H., Beale, J. M., & Hurley, L. H. (1991) *Biochemistry 30*, 3597−3602.
Martin, D. G., Biles, C., Gerpheide, S. A., Hanka, L. J., Krueger, W. C., McGovren, J. P., Miszak, S. A., Neil, G. L., Stewart, J. C., & Visser, J. (1981) *J. Antibiot. 34*, 1119−1125.
McGovren, J. P., Clarke, G. L., Pratt, E. A., & DeKoning, T. F. (1984) *J. Antibiot. 37*, 63−70.
Mitchell, M. A., Kelly, R. C., Wicnienski, N. A., Hatzenbuhler, N. T., Williams, M. G., Petzold, G. L., Slightom, J. L., & Siemieniak, D. R. (1991) *J. Am. Chem. Soc. 113*, 8994−8995.
Pearlman, D. A., Case, D. A., Caldwell, J., Seibel, G. L., Singh, U. C., Weiner, P. K., & Kollman, P. A. (1991) *AMBER 4.0 (UCSF)*, University of California, San Francisco.
Reynolds, V. L., Molineaux, I. J., Kaplan, D. J., Swenson, D. H., & Hurley, L. H. (1985) *Biochemistry 24*, 6228−6237.
Scahill, T. A., Jenson, R. M., Swenson, D. H., Hatzenbuhler, N. T., Petzold, G., Wierenga, W., & Brahme, N. D. (1990) *Biochemistry 29*, 2852−2860.
Seaman, F. C., & Hurley, L. H. (1993) *Biochemistry 32*, 12577−12585.
Seaman, F. C., & Hurley, L. H. (1995) *Nature Struct. Biol.* (submitted for publication).
Sklenar, V., & Bax, A. (1987) *J. Magn. Reson. 74*, 469−479.
Sun, D., & Hurley, L. H. (1993) *J. Am. Chem. Soc. 115*, 5925−5933.
Sun, D., Park, H.-J., & Hurley, L. H. (1994) *Chem. Res. Toxicol. 6*, 889−894.
Swenson, D. H., Li, L. H., Hurley, L. H., Rokem, J. S., Petzold, G. L., Dayton, B. D., Wallace, T. L., Lin, A. H., & Krueger, G. L. (1982) *Cancer Res. 42*, 2821−2828.
Thompson, A. S., & Hurley, L. H. (1995) *J. Mol. Biol.* (in press).
Thompson, A. S., Sun, D., & Hurley, L. H. (1995) *J. Am. Chem. Soc. 117*, 2371−2372.
Warpehoski, M. A., & Bradford, V. S. (1988) *Tetrahedron Lett. 29*, 131−134.
Warpehoski, M. A., & Hurley, L. H. (1988) *Chem. Res. Toxicol. 1*, 315−333.

BI950662E