Manipulative Interplay of the Interstrand Cross-Linker Bizelesin with d(TAATTA)$_2$ To Achieve Sequence Recognition of DNA

Frederick C. Seaman* and Laurence Hurley*

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

Received June 7, 1996[⊗]

Abstract: As expected, Bizelesin, which is a biscyclopropa[*c*]pyrrolo[3,2-*e*]indol-4(5*H*)-one [(+)-CPI]-derived DNA−DNA cross-linker, has a high interstrand cross-linking reactivity with the palindromic sequence 5′-d(CGTAATTACG)$_2$ (˙ and _ indicate (+)-CPI alkylation sites at adenines on the same and opposite strands, respectively). Contrary to expectations, the target duplex is rearranged to yield two products: one (major product) contains an AT step wherein both adenines are *syn*-oriented and hydrogen bonded to thymines forming a stable Hoogsteen base-paired region flanked by Watson−Crick base-paired regions (5HG); the other (minor product) contains *anti*-oriented AT-step adenines that show no evidence of hydrogen bonding with pairing thymines (5OP) in an otherwise normally base-paired duplex. In another unexpected outcome, the reaction of two "uncoupled" monoalkylating (+)-CPI "halves" of Bizelesin with the same duplex alkylates *same*-strand adenines three base pairs apart [5′-d(CGTAATTACG)$_2$] rather than the anticipated *opposite*-strand adenines six base pairs apart (which would mimic Bizelesin). To probe the molecular mechanism that leads to Bizelesin's unusual DNA rearrangement, which appears to be a requirement for DNA−DNA interstrand cross-linking, we have carried out conformational exchange analyses (NOESY and ROESY) and restrained molecular dynamics simulations of these adducts. These studies suggest that Bizelesin controls the rearrangement of the six-base-pair target prior to cross-linkage and restricts the cross-linked DNA adduct's range of motion, freezing-out adduct conformers defined by alternative drug−DNA hydrogen-bond regimes. The two competing cross-linkage pathways share a common first step, the opening of the central AT-step base pairs, an event that is facilitated by the energetics of monoadduct-induced DNA bending distortion. One pathway (to 5HG) stabilizes these open bases by reorganizing the AT-step region into two Hoogsteen base pairs, the thymine bases of which also hydrogen bond with Bizelesin's ureadiyl subunit. A second pathway (to 5OP) directly stabilizes the open bases by forming a hydrogen-bonding complex between the AT-step thymines and Bizelesin's ureadiyl subunit. Cross-linked DNA motion drives both of the 5HG and 5OP adducts from one ephemeral hydrogen-bonding regime to another, a process documented in the NOESY conformational exchange data and simulated in restrained molecular dynamics trajectories. These results, together with the analysis of other six-base-pair Bizelesin cross-linked species, suggest a novel mechanism for sequence recognition by this cross-linker where monoalkylation distortive stress associated with a bent DNA conformation must be dissipated by a cooperative interaction between drug and duplex to produce a straight B-form-like structure before cross-linking can proceed. This example provides a new mechanism for DNA sequence recognition involving a "drug-induced rearrangement" of DNA that critically depends upon the interplay of drug and sequence recognition elements.

Introduction

The DNA-alkylating prodrug dimer Bizelesin consists of two open-ring homologs of the (+)-CC-1065 (+)-CPI[1] subunit connected by a rigid diindole−urea linker moiety (Figure 1A).[2] Presently in phase 1 clinical trials, Bizelesin displays impressive *in vitro* cytotoxic potency and excellent *in vivo* efficiency in comparison to its (+)-CPI monoalkylating parent compound (+)-CC-1065 and related compounds, Adozelesin and Carzelesin, which are both currently in phase 1 and 2 clinical trials in the United States and Europe, respectively.[2,3] Bizelesin most frequently cross-links adenines (Figure 1B, right panel) on opposite DNA strands six base pairs apart (inclusive of the adenines) while being cradled in the intervening minor groove. Although previous studies have determined that 5′-TAATTA is the cross-linking sequence most reactive toward Bizelesin,[4] these studies left the underlying basis for the high reactivity of this favored sequence unresolved. This sequence preference has been attributed to the juxtaposition of two opposite-strand 5′-TTA triplets (Figure 2A).[4d] This triplet is the preferred target for the parent drug, (+)-CC-1065, and other (+)-CPI-based monoalkylators.[5] It was argued that Bizelesin's six-base-pair sequence preference was an additive process combining two favored alkylation events each of which mimics (+)-CC-1065's 5′-TTA triplet preference. In contrast to monoalkylation by (+)-CPI drugs, which results in a bent DNA structure,[4c,6] Bizelesin

⊗ Abstract published in *Advance ACS Abstracts,* October 15, 1996.

(1) Abbreviations: (+)-CPI, cyclopropa[*c*]pyrrolo[3,2-*e*]indol-4(5*H*)-one; (+)-CPI-I, 2-(indole-2-ylcarbonyl)-1,2,8,8a-tetrahydro-7-methylcyclopropa-[*c*]pyrrolo[3,2-*e*]indol-4(5*H*)-one (U72779); NMR, nuclear magnetic resonance; rMD, restrained molecular dynamics; RMSD, root mean square difference; NOE, nuclear Overhauser effect; PE COSY, primitive exclusive correlated spectroscopy; 2QF-COSY, double-quantum-filtered correlated spectroscopy; ROESY, rotating frame Overhauser spectroscopy; ppm, parts per million; FID, free-induction decay; ^{1}H−^{31}P COSY, hydrogen−phosphorus correlated spectroscopy; 5HG, self-complementary 10-mer cross-linked product with the AT-step adenine of each strand Hoogsteen base paired to its thymine partner; 5OP, self-complementary 10-mer cross-linked product with the AT-step adenine of each strand not hydrogen bonded (open) to its thymine partner.

(2) Mitchell, M. A.; Kelly, R. C.; Wicnienski, N. A.; Hatzenbuhler, N. T.; Williams, M. G.; Petzold, G. L.; Slightom, J. L.; Siemieniak, D. R. *J. Am. Chem. Soc.* **1991**, *113*, 8994−8995.

(3) (a) Aristoff, P. A.; McGovren, J. P. *Drug News Perspect.* **1993**, *6*, 229−234. (b) Lee, C.-S.; Gibson, N. W. *Cancer Res.* **1991**, *51*, 6586−6591.

(4) (a) Ding, Z.-M.; Hurley, L. H. *Anti-Cancer Drug Des.* **1991**, *6*, 427−452 (b) Sun, D.; Hurley, L. H. *J. Am. Chem. Soc.* **1993**, *115*, 5925−5933. (c) Sun, D.; Lin, C. H.; Hurley, L. H. *Biochemistry* **1993**, *32*, 4487−4495. (d) Lee, C.-S.; Gibson, N. W. *Biochemistry* **1993**, *32*, 2592−2600. (e) Lee, C.-S.; Gibson, N. W. *Biochemistry* **1993**, *32*, 9108−9114.

S0002-7863(96)01924-5 CCC: $12.00

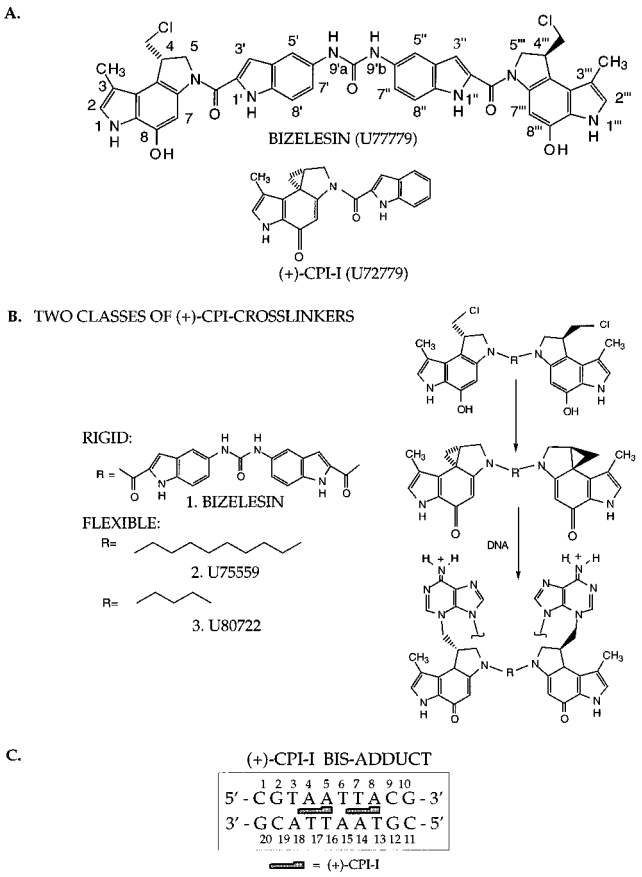

Figure 1. Structures and reaction of (+)-CPI cross-linkers and (+)-CPI-I monoalkylators with duplex DNA. (A) Structures of Bizelesin and (+)-CPI-I, also showing numbering system used for Bizelesin. (B) Structures of the rigid and flexible cross-linkers and their conversion to the cyclopropyl derivatives and reaction with duplex DNA. (C) Pattern of same-strand alkylation by (+)-CPI-I of the 10-mer sequence.[10]

cross-linking of 5′-TAATTA̲ does not result in a bent duplex structure.[4a,4d] The absence of bending in the drug-modified DNA containing this six-base-pair sequence was interpreted as suggesting that bending was not required for the formation of an interstrand cross-link.[4d] This rationale for cross-linkage was consistent with the generalization that the sequence preference of DNA interstrand cross-linkers was usually correlated with *minimal reorganization* of DNA during the cross-linking of the monoalkylation adduct.[7] This minimal reorganization/sequence specificity relationship typifies the reactions of many cross-linkers.[7,8] For cross-linking agents that are nondistortive and have a limited capacity to influence DNA structure, minimal DNA reorganization translates into a minimal energy transition state for conversion of monoadduct to cross-link.[7]

The first (+)-CPI-based interstrand cross-linkers joined two (+)-CPI alkylating moieties with flexible methylene chains of varying lengths (e.g., Figure 1B, structures **2** and **3**). These structures successfully cross-linked across three, four, and five base pairs, but showed reduced cross-linking capacity and cytotoxicity as the methylene chains were extended to allow cross-linkage across greater distances.[9] Thus, Bizelesin, with its rigid diindole−urea linker unit, was designed to target these longer DNA sequences, and 5′-TAATTA̲ was expected to be the favored target.[2] It became clear that the introduction of the rigid linker moiety belied the interpretation that based Bizelesin preference for this sequence on a simple additive preference for adjoining antiparallel 5′-TTA̲ triplets. Two independent observations suggested that this simple explanation provided by the earlier study[4d] could not account for the high reactivity of this 5′-TAATTA cross-linking sequence. First, bisalkylation of 5′-TAATTA with two (+)-CPI-I monoalkylators (U72779, The Pharmacia Upjohn Co., Figure 1A) does not lead to the expected symmetrical bisadduct with an indole-to-indole two-drug configuration resembling the interstrand Bizelesin cross-linked adduct.[10] If bisalkylation is driven by twin 5′-TTA̲ triplet monoalkylation events, then (+)-CPI-I should mimic its parent compound (+)-CC-1065 and produce this symmetrical bisadduct. Instead, (+)-CPI-I produced the same-strand bisalkylation product (Figure 1C). Second, NMR analysis of the Bizelesin cross-linking of the 5′-TAATTA̲ sequence[11] shows that the major product (60%) AT-step adenine bases are *syn*-oriented and Hoogsteen base-paired (5HG; Figure 2A,B), while the minor product (40%) AT step adenine and thymine bases are *anti*-oriented, non-base-paired, and displaced toward the major groove (5OP; Figure 2A,C). If as predicted by early drug design efforts,[2] Bizelesin fits this six-base-pair B-form cross-linkage span, and its cross-linkage mimics twin (+)-CC-1065 alkylation events, why must Bizelesin always reorganize 5′-TAATTA prior to cross-linkage?

Another complicating factor not addressed in attempts to equate interstrand cross-linkage with twin monoalkylating events is the propagation of duplex bending distortion to the 5′-side of the covalently modified adenine by (+)-CC-1065 and other related monoalkylators.[6] While the unreacted duplex displays no significant bending,[4a,d] what are the dynamic bending properties of the Bizelesin monoalkylated 5′-TAATTA̲ intermediate? What is the impact of this monoalkylated duplex bending distortion and associated dynamics on the cross-linking reaction?

These inconsistencies between our results[10,11] and the previous predictions of Bizelesin's interstrand cross-linkage of its preferred sequence[2] prompted us to examine this reaction through previously unexplored avenues. Our initial goal was to resolve the incongruities between our data[10,11] and the previous prediction[2] and interpretation.[4d] The second goal was to propose a mechanism of sequence recognition and cross-linking that encompasses all the available data. Relevant corollary information follows from the modeled structures of two other Bizelesin cross-linked sequences that have recently been published. One is a six-base-pair 5′-TAAAAA̲ cross-linked A tract, which is reported to be a straight DNA structure,[12] and the other is a seven-base-pair 5′-TTAGTTA̲ cross-linked sequence.[13]

Results

NMR Analysis of the Unmodified 10-Mer Duplex: A. Evidence for Region-Specific Conformational Features of

(5) (a) Reynolds, V. L.; Molineaux, I. J.; Kaplan, D. J.; Hurley, L. H. *Biochemistry* **1985**, *24*, 6228−6237. (b) Hurley, L. H.; Lee, C.-S.; McGovren, J. P.; Warpehoski, M. A.; Mitchell, M. A.; Kelly, R. C.; Aristoff, P. A. *Biochemistry* **1988**, *27*, 3886−3892. (c) Warpehoski, M. A.; Hurley, L. H. *Chem. Res. Toxicol.* **1988**, *1*, 315−333.

(6) (a) Lee, C.-S.; Sun, D.; Hurley, L. H. *Chem. Res. Toxicol.* **1991**, *4*, 203−213. (b) Lin, C. H.; Sun, D.; Hurley, L. H. *Chem. Res. Toxicol.* **1991**, *4*, 21−26.

(7) Hopkins, P. B.; Millard, J. T.; Woo, J.; Weidner, M. F.; Kirchner, J. J.; Sigurdsson, S. T. *Tetrahedron* **1991**, *47*, 2475−2489.

(8) (a) Wang, J.-J.; Hill, G. C.; Hurley, L. H. *J. Med. Chem.* **1992**, *35*, 2995−3002. (b) Bose, D. S.; Thompson, A. S.; Ching, J.; Hartley, J. A.; Bernardini, M. D.; Jenkins, T. C.; Neidle, S.; Hurley, L. H.; Thurston, D. E. *J. Am. Chem. Soc.* **1992**, *114*, 4939−4941. (c) Mountzouris, J. A.; Wang, J.-J.; Thurston, D.; Hurley, L. H. *J. Med. Chem.* **1994**, *37*, 3132−3140.

(9) Mitchell, M. A.; Johnson, P. D.; Williams, M. G.; Aristoff, P. A. *J. Am. Chem. Soc.* **1989**, *111*, 6428−6429.

(10) Seaman, F. C.; Chu, J.; Hurley, L. H. *J. Am. Chem. Soc.* **1996**, *118*, 5383−5395.

(11) Seaman, F. C.; Hurley, L. H. *Biochemistry* **1993**, *32*, 12577−12585.

(12) (a) Thompson, A. S.; Hurley, L. H. *J. Mol. Biol.* **1995**, *252*, 86−101. (b) Thompson, A. S.; Sun, D.; Hurley, L. H. *J. Am. Chem. Soc.* **1995**, *117*, 2371−2372.

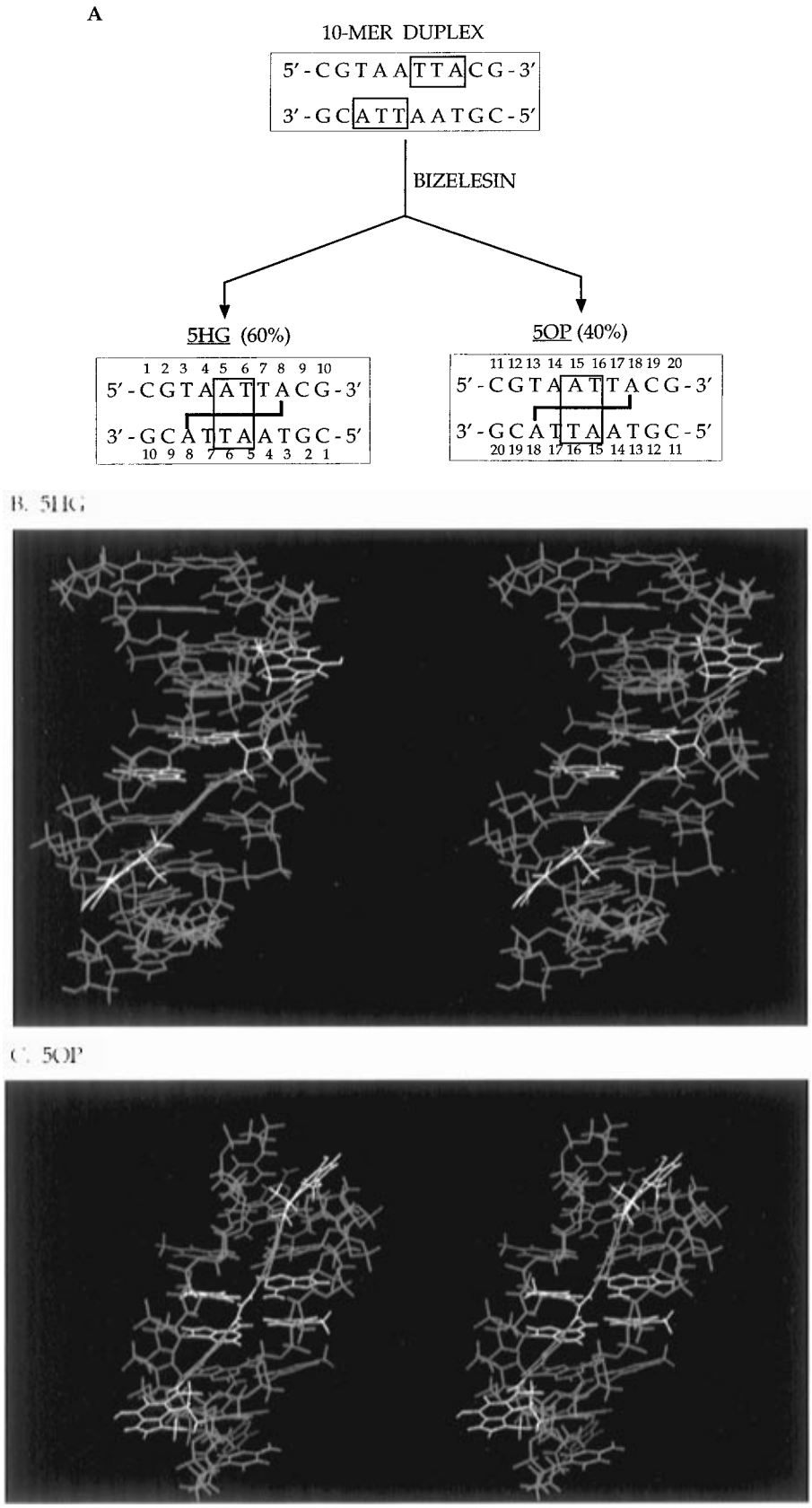

Figure 2. Reaction of Bizelesin with the 10-mer duplex to yield the two products 5HG and 5OP.[11] (A) (+)-CPI alkylating 5′-TTÀ target sequences are indicated by rectangular boxes. The 5HG AT-step box refers to the region containing two Hoogsteen base pairs, and the 5OP AT-step box refers to the region of two open base pairs. The 5HG and 5OP DNA duplex numbering systems are included. (B) Stereoview of 5HG. Colors are as follows: white = Bizelesin (+)-CPI, cyan = indole, and yellow = ureadiyl subunits; red = 10-mer thymine, green = *anti*-oriented adenine, yellow = AT-step adenine, blue = guanine, orange = cytosine, and magenta = backbone. (C) Stereoview of 5OP, excluding the terminal base pairs. Colors are as in B.

the 10-Mer Duplex Based on Two-Dimensional NMR Analysis. The single set of proton chemical shift assignments (Table 1) for the self-complementary sequence indicates a symmetrical duplex structure. All thymine H3 and guanine H1 imino proton signals (12.0−15.0 ppm) are observed at temperatures below 45°, except the terminal guanine H1 signal, which becomes too

Table 1. ^{1}H NMR Chemical Shifts, Deoxyribose Coupling Constants, Conformations of the S-Type Conformers Derived from *J*-Couplings, and H1′ × H4′ NOESY Cross-Peak Intensity Data (200 ms) of the Unmodified 10-Mer

	chemical shifts, ppm (coupling constants, Hz)										deoxyribose conformation			
base	H1′	H2′	H2″	H3′	H4′	H5′	H5″	PuH8/ PyH6	TMe/ CH5/ AH2	PuH1/ PyH3	phase angle, deg	amplitude, deg	*S*, %	NOESY H1′ × H4′ intensity
C	5.73 (5.8/7.9)[a]	1.97 (14.4/6.4)[b]	2.39 (3.2)[c]	4.66 (5.2)[d]	4.02	3.69	3.68	7.60	5.85		129[e]	45	71	130[f]
G	5.91 (4.7/9.0)	2.62 (14.9/5.7)	2.71 (2.2)	4.93 (3.7)	4.32	4.06	3.96	7.95		12.61	137	47	83	138
T	5.56 (5.1/9.5)	2.04 (13.1/7.5)	2.37 (2.6)	4.84 (5.0)	4.13	4.14	4.08	7.22	1.45	13.48	119	46	86	286
A	5.96 (5.8/9.5)	2.73 (15.0/5.5)	2.91 (1.5)	5.02 (3.6)	4.04	4.13	4.04	8.23	6.80		143	39	89	150
A	6.10 (6.1/7.8)	2.55 (15.3/6.7)	2.89 (2.8)	4.96 (4.6)	4.44	4.25	4.22	8.12	7.49		135	37	75	155
T	5.82 (5.4/8.6)	1.91 (15.5/6.8)	2.48 (2.7)	4.77 (5.9)	4.14	4.29	4.12	7.10	1.23	13.29	118	51	77	449
T	5.77 (5.2/9.3)	2.18 (15.6/5.9)	2.46 (2.4)	4.86 (5.0)	4.12	4.12	4.05	7.32	1.56	13.33	126	50	82	347
A	6.13 (5.0/9.0)	2.63 (15.2/5.6)	2.81 (2.1)	4.99 (4.2)	4.39	4.14	4.10	8.27	7.45		134	47	82	151
C	5.58 (5.5/8.2)	1.81 (13.3/6.9)	2.23 (4.0)	4.73 (5.9)	4.11	4.22	4.11	7.25	5.27		120	50	68	220
G	6.08 (6.6/8.1)	2.54 (14.7/5.8)	2.32 (3.1)	4.61 (5.1)	4.13	4.04	4.02	7.86		12.82	136	40	71	132

[a] Coupling constants preceding and following the slash are $J_{H1',H2''}$ and $J_{H1',H2'}$, respectively. [b] Coupling constants preceding and following the slash are $J_{H2',H2''}$ and $J_{H2',H3'}$, respectively. [c] Coupling constant is $J_{H2'',H3'}$. [d] Coupling constant is $J_{H3',H4'}$. [e] In the PSEUROT calculation, the N-type conformers were restricted to a phase angle of 9° and a pucker amplitude of 36°. [f] Relative cubic units of volume calculated by FELIX 2.10.

broad to detect at temperatures above 27°. The NOESY spectrum contains several unique AT step features: (1) Intense cross-peaks link the AT-step adenine H2 proton with the 3′-side thymine H1′ (intrastrand) and the thymine H1′ one base pair to the 3′-side on the opposite strand (interstrand). (2) The AT-step thymidine H5′ is downfield-shifted beyond any of the other deoxyribose H5′ and H5″ proton signals (Table 1). These NOESY and additional COSY cross-peak data for the 10-mer internal 5′-AATT region resemble those of the corresponding region of the Dickerson dodecamer, [d(CGCGAATTCGCG)$_2$].[14] Coupling constants (Table 1) for deoxyribose H1′−H2′, H1′−H2″, H2′−H3′, H2″−H3′ (PE COSY), and H3′−H4′ (partially decoupled 2QF-COSY) were used to calculate (PSEUROT, see the Experimental Section) the major S-type phase angle of pseudorotation while the minor N-type conformer was constrained to $P = 9°$ and $F_m = 36°$. Deoxyribose conformational equilibrium mixtures (N- and S-types) were calculated for each 10-mer sugar (Table 1). On the basis of these data and corroborating NOESY distance calculations (H1′ × H4′ NOESY cross-peak intensities, Table 1), the terminal nucleotides 1C and 10G and the AT-step nucleotides 5A and 6T display properties of an equilibrium mixture with a significant proportion of C3′-endo conformations (Table 1). For the internal region, the AT-step sugar geometries display the greatest proportion of N-type members and the lowest *P* values for the S-type members. For nonterminal base pairs, the most aberrant pseudorotation phase angle and percent N-type properties are associated with the AT-step thymidine just as in the case of the Dickerson dodecamer.[15] Although the significance of these calculations has been variously interpreted,[16] abrupt discontinuities in sugar conformations based on coupling constant values suggest conformation heterogeneity in the backbone sugars of the duplex.

B. Base-Pair Opening Rates. Ammonia catalyst proton exchange experiments (see the Experimental Section) indicate that the 5′-AATT base pairs of the unmodified 10-mer display relatively high opening rates at 23°, i.e., 5A/6T, 35 ms; 4A/7T, 6 ms. These results are consistent with reported high base-pair-opening rates for ammonia- and Tris base-catalyzed exchange studies of other sequences incorporating 5′-AATT.[17]

Two-Dimensional NMR and Molecular Dynamics Analysis of the Bizelesin Cross-Linked 10-Mer Duplex Adduct. Reaction of Bizelesin with the 10-mer duplex resulted in a mixture of two principal cross-linked conformers and the slow buildup of DNA strand breakage products. C18 reverse phase chromatography of the reaction products yielded a single major peak containing the conformer mixture and minor peaks identified by NMR as strand breakage products. Modification of the solvent program failed to achieve separation of the conformers. Continued strand breakage of the cross-linked products diminished the effective life span of the sample to less than 1 week and reduced the reliability of NOE buildup rates for consecutive NOESY spectra collected at different mixing times. Consequently, the NOE-derived distance restraints were not suitable for determination of highly resolved structures.

A. NOESY Cross-Connectivity Networks Show That Bizelesin Forms Two Symmetrical Cross-linked Adducts (5HG and 5OP) with 5′-TAATTA in the 10-Mer Sequence. NOESY, ROESY, COSY, and ^{1}H−^{31}P COSY experiments were conducted on the Bizelesin reaction products of the self-complementary oligomer (10-mer) d(CGTAATTACG)$_2$ (Figures 1−8 in the Supporting Information), and a partial description of the 5HG and 5OP cross-peak connectivity networks was presented in an earlier publication.[11] Cross-linkage-induced DNA properties appear to be independent of the duplex length in that the same results were observed in an NMR analysis of the cross-linked product of the 5′-TAATTA sequence embedded in a 16-mer duplex, 5′-GCGATCTGTAATTACG-

(13) Thompson, A. S.; Fan, J.-Y.; Sun, D.; Hansen, M.; Hurley, L. H. *Biochemistry* **1995**, *34*, 11005−11016.
(14) (a) Nerdal, W.; Hare, D. R.; Reid, B. R. *Biochemistry* **1989**, *28*, 10008−10021. (b) Lane, A. N.; Jenkins, T. C.; Brown, T.; Neidle, S. *Biochemistry* **1991**, *30*, 1372−1385.
(15) Bax, A.; Lerner, L. *J. Magn. Reson.* **1988**, *79*, 429−438.
(16) (a) Kim, S.-G.; Lin, L.-J.; Reid, B. R. *Biochemistry* **1992**, *31*, 3564−3574. (b) Kim, S.-J.; Reid, B. R. *Biochemistry* **1992**, *31*, 12103−12116. (c) Weisz, K.; Shafer, R. H.; Egan, W.; James, T. L. *Biochemistry* **1992**, *31*, 7477−7487.
(17) Leroy, J.-L.; Charretier, E.; Kochoyan, M.; Guéron, M. *Biochemistry* **1988**, *27*, 8894−8898.

Table 2. ^{1}H NMR Chemical Shifts of the Major and Minor Interconverting Conformers of the 5HG and 5OP Classes

		chemical shifts, ppm	
class	^{1}H NMR assignment[a]	major isomer	minor isomers
5HG	4A H8	8.27	7.00
	4A H2	7.37	8.12 or 6.99[b]
	5A H8	7.49	7.71
	5A H2	7.60	7.85
	6T H6	6.76	6.89
	7T H6	7.32	7.22
	8A H8	8.45	8.38
	8H3′	7.22	7.13[c]
	6T H3	12.42	13.25
	7T H3	14.74	12.87
	8H1′	11.33	11.52
5OP	14A H2	6.96	7.04
	14A H8	8.30	8.23
	15A H2	7.36	8.12 or 6.99[b]
	15A H8	7.92	7.50
	16T H6*	6.65	6.36/6.77
	16T Me*	1.09	0.88/0.70
	17T H6*	7.12	7.12[c]
	17T Me*	1.40	1.49/1.41
	17T H3	14.25	14.59
	18H5′	8.28	8.55
	N2H/N3H	7.71	not observed

[a] An asterisk (*) indicates that a minor conformer exists in sufficiently high concentrations such that these signals are associated with an independent partial set of cross-peaks. [b] 4A H2 (5HG) and 15A H2 (5OP) overlap. [c] H3′ (5HG) and 17T H6 (5OP) overlap.

3′/3′-CGCTAGACȦTTAATGC-5′ (F. C. Seaman and L. H. Hurley, unpublished results). Despite discernible AT-step structural differences, many parallel 5HG and 5OP drug−DNA NOESY cross-peaks (summarized in Figure 9 of the Supporting Information) indicate that Bizelesin is similarly positioned in the minor grooves of both products.

B. Both of the 5HG and 5OP 10-Mer Duplex Cross-Linked Adducts Encompass Multiple Positional Isomers. Both the 5OP and 5HG NOESY and ROESY spectra contain numerous conformational exchange cross-peaks that, in the instances of nonexchangeable protons, correspond to ROESY "negative" cross-peaks (Table 2). These intense NOESY/ROESY exchange cross-peaks occur between a relatively intense signal of a major positional isomer and the weaker signals of a minor isomer. Thus, conformational exchange cross-peaks allow assignment of additional chemical shifts (minor isomers) to protons whose motion during the conformational exchange process positions them in more than one electronic chemical environment.[18] Because of the weakness of minor isomer signals, their cross-peak connectivity networks are difficult to discern.

Conformational exchange was observed between interconverting positional isomers within both 5HG and 5OP networks. These conformational exchange properties were compared to the findings of restrained molecular dynamics (rMD) trajectory analyses (restraints derived from nonexchangeable proton NOESY cross-peaks). From the results of rMD trajectory hydrogen-bonding analysis, hydrogen-bonding patterns that stabilize interconverting positional adduct conformers were identified. If rMD trajectory analysis yields interconverting hydrogen bonding patterns in the regions of conformational exchange, then the level of agreement between modeled positional isomer exchange and experimentally observed conformational exchange can be evaluated. The correspondence of these two sets of data argues persuasively for the proposed conformer models. The characterization of these positional isomers and their relationship to each other will be the focus of the following discussion.

C. Isomerization of the 5HG Species Results from Concerted Movement of the Ureadiyl Subunit and the Neighboring Minor Groove 5′-AATT Region. (i) Structural Models Derived from NOESY/ROESY Conformational Exchange Data. The duplex portion of 5HG is stabilized throughout by either Watson−Crick or Hoogsteen base pairing. Nevertheless, ureadiyl amido and DNA base proton NOESY and ROESY conformational exchange cross-peaks show that the system moves from one hydrogen-bonding-restrained positional isomer to another. Presumably, the alternative chemical shifts and conformational exchange properties of the central AT base-pair protons result from the shifting of the base protons into different ring current electronic environments within the stacked bases of different positional isomers.

Virtually all 5HG 5′-AATT base H6/H8 (nonexchangeable) and H3 (exchangeable) proton signals display prominent conformational exchange cross-peaks (Table 2). Two examples of NMR time scale conformational exchange help to characterize the motion in this region: (1) Major and minor conformer signals exist for both 7T H3 (major, 7T H3f; minor, 7T H3e; "f" and "e" indicate face-on and edge-on orientation, respectively, of the drug ureadiyl subunit relative to the minor groove) and 6T H3 (major, 6T H3f; minor, 6T H3e), and these major and minor signal pairs share intense conformational exchange cross-peaks (Figure 3A,B). Both major and minor conformer 6T and 7T H3 signals exist in the region (12.0−15.0 ppm) typical of hydrogen-bonded thymine bases. This indicates that, while 6T H3 and 7T H3 major and minor conformer protons have different chemical shifts, these 5′-AATT base pairs remain hydrogen bonded (Figure 3B). (2) The Bizelesin indole subunit's NMR data indicate that this portion of the drug also experiences isomerization. This interpretation is supported by the exchangeable NH1′ assignments, which likewise consist of a major and minor signal (H1′f and H1′e, 11.33 and 11.52 ppm, respectively; Figure 3A,B) sharing an intense conformational exchange cross-peak.

Given the symmetrical behavior of equivalent protons in this cross-linked adduct, it is assumed that an isomer's two ureadiyl H9′ protons, H9′a and H9′b (Figure 1A), produce overlapping (equivalent) signals. Two different H_2O-exchangeable ureadiyl H9′ signals (8.74 and 8.81 ppm; each presumably representing a ureadiyl subunit H9′a and b pair) and their cross-peaks were identified (Figure 4A). Chemical shift similarity hinders analysis of possible conformational exchange cross-peaks between these two signals. The most intense ureadiyl amido signal, H9′f ("f" for face-on ureadiyl orientation relative to DNA, 8.74 ppm), generates moderate to intense cross-peaks with neighboring indole protons, 8H7′ and 8H5′, and duplex backbone protons, 6T H1′, 7T H1′, 7T H5″, 6T H2″, and 6T H2′ (Figure 4A,B; H9′f cross-peaks A−G). The other prominent amido signal, 8H9′e ("e" for edge-on orientation, 8.81 ppm), shows a different connectivity pattern consisting of cross-peaks to 5A H8 and 6T H1′ (Figure 4A,B; cross-peaks A−B). Thus, within the 5HG symmetrical adduct, the 8H9′f cross-peaks indicate that ureadiyl H9′a and H9′b are oriented toward the duplex backbone ("face-on" orientation relative to the minor groove floor; Figure 4B) in one isomer, and the 8H9′e cross-peaks indicate that this H9′a and b pair are oriented toward the floor of the minor groove in another isomer ("edge-on" ureadiyl subunit orientation; Figure 4B).

As a first step toward formulating a 5HG conformational model, it is necessary to incorporate these cross-peak data for contrasting positional isomers into a Bizelesin cross-linked adduct structure in which the internal AT-step Hoogsteen base

(18) Choe, B.; Cook, G. W.; Krishna, N. R. *J. Magn. Reson.* **1991**, *94*, 387−393.

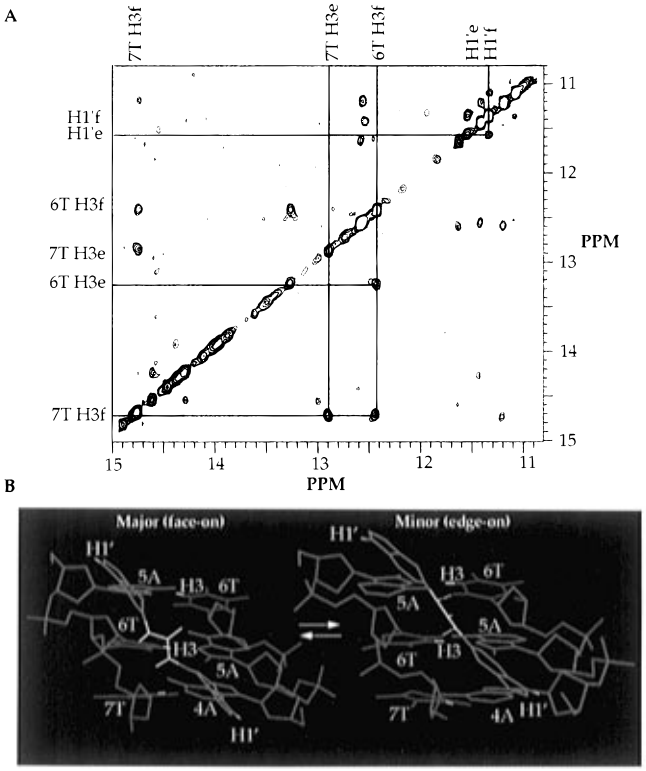

Figure 3. Conformational exchange of the interconverting 5HG positional isomers. (A) Two-dimensional NOESY (150 ms mixing time) expanded contour plot of the Bizelesin−10-mer duplex adduct (90% H_2O/10% D_2O) showing major conformer thymine imino signals, 7T H3f and 6T H3f, dipolar cross-peaks with each other and conformational exchange cross-peaks with the corresponding minor conformer's 7T H3e and 6T H3e signals and conformational exchange cross-peak of Bizelesin indole exchangeable H1′f signal (major conformer) with the minor conformer's H1′e signal. (B) View from the major groove of the 5HG interconverting isomers: (left) the major face-on conformer and (right) the edge-on conformer. The ureadiyl subunit H9′ and indole subunit H1′ substituents and the neighboring AT-step thymine H3 substituents are colored white in order to depict their relative positions in the two isomers. Other colors are as in Figure 2, except that all adenines are colored green.

pairs remain hydrogen bonded. As depicted in the proposed model (face-on "A" and "B" in Figure 4C), the stronger signal, 8H9′f, is a product of two different conformational forms, face-on "A" (amido protons directed toward strand two) and "B" (amido protons directed toward strand one), wherein the two amido H9′f protons, H9′a and H9′b, alternate between two equivalent chemical environments. Rapid symmetrical flipping motion results in an averaged signal (8.74 ppm) for these two protons. Interruption of this motion in the edge-on form restrains the equivalent amido protons, H9′a′ and H9′b′ (H9′e signal at 8.81 ppm), in symmetrical positions relative to the duplex structure (edge-on in Figure 4C). The absence of conformational exchange data for these H9′e and H9′f protons precludes determining if they are exchanging at rates equivalent to the neighboring base protons.

(ii) Models Derived from Restrained Molecular Dynamics for the 5HG Conformers. Restrained molecular dynamics (Amber 4.0)[19] calculations using 156 NMR distance restraints for 5HG (117 DNA−DNA and 39 drug−DNA distance restraints) permit modeling of the AT-step hydrogen-bonding regimes of the 5HG adduct. Because of the Bizelesin cross-linked adduct properties (e.g., conformational exchange, DNA strand breakage, and sample decomposition), a highly refined solution structure is judged to be beyond the scope of the present study. Realistically, the overall goal of this rMD analysis is not to generate a highly refined conformation but to establish that the experimental data are consistent with the major conformational properties proposed for the 5HG and 5OP structures. Once a stable low-energy structure is generated, it is then used as a model for evaluating hydrogen-bonding properties.

The first stage of rMD analysis involves *in vacuo* calculations using A-form and B-form starting structures with Hoogsteen base-paired AT steps and the analysis of convergence of these structures (see the Experimental Section and Table 3A). The RMSD calculations for rMD products derived from A- and B-form starting structures (Table 3A) show that the 5HG final rMD structures converged satisfactorily (average RMSD: 5HG, 1.14 Å).[20] Then, B-form starting structures (with Hoogsteen base-paired AT-step nucleotides) derived from averaged *in vacuo* rMD products (RMSD study) were solvated and used in the next stage of hydrogen-bonding analysis. During this second stage, the temperature for each rMD trajectory is ramped to 500 K, lowered to a plateau temperature of 300 K, and maintained at this temperature for a period of 175 ps. For this 175-ps period, adduct structure coordinate sets are generated each 0.5 ps and are screened for drug−DNA hydrogen-bond donor/acceptor pairs. From one 0.5 ps sampling point to the next, motion of the adduct structure can cause the ureadiyl amido hydrogen donors and the AT-step thymine O2 acceptors to move in and out of the hydrogen bond distance and angle limit ranges.

Repeated 5HG rMD hydrogen-bonding analyses yield the same results in terms of ureadiyl amido hydrogen bonding: Each amido hydrogen atom donor pairs with the nearest of two AT-step thymine O2 acceptors, creating symmetrical edge-on hydrogen bonding that stabilizes the Hoogsteen base pairing of the AT-step region (Figure 5). In contrast to the rMD results, NOESY data indicate that over the much longer time frame of the NMR experiment the stable edge-on form interconverts with a long-lived alternative face-on form (Figure 4B,C). Why this alternative form is not observed during the much briefer (200−300 ps) rMD experiments is difficult to determine. One possibility is that the hydrogen-bonding or other electrostatic properties of the residues contained in the central region of the 5HG rMD model are not accurately estimated, resulting in the "edge-on" hydrogen-bonding regime being favored over alternative "face-on" forms.

The 5HG edge-on isomer is a highly unconventional conformation. Its drug−DNA hydrogen-bonding association is feasible only because of the unique minor groove properties of the symmetrical Hoogsteen base-paired AT-step region; i.e., (1) the AT-step interstrand C1′−C1′ distance is shortened by 1.5−2.0 Å relative to Watson−Crick B-form geometry; (2) AT-step thymine O2 moieties are positioned in the middle (floor) of the minor groove, unlike Watson−Crick B-form structures wherein these O2 substituents are displaced farther toward opposite walls of the minor groove; and (3) as a consequence of the Hoogsteen base-pairing mode, the minor groove is narrowed in the AT-step region. These changes in the Hoogsteen AT-step region move the two thymine O2 moieties to sites suitable for simultaneous hydrogen bonding with the two ureadiyl NH donor

(19) Pearlman, D. A.; Case, D. A.; Caldwell, J. C.; Seibel, G. L.; Singh, C.; Weiner, P. K.; Kollman, P. A. *AMBER 4.0*; University of California: San Francisco, 1991.

(20) RMSD's between starting structures degraded rapidly when the first true A- and B-form starting structures (SS 1, RMSD 4.21 Å) were submitted to brief MD periods (1−4 ps) to generate the other three starting structures. A-form starting structures 2−4 generated during such periods show substantially reduced RMSD's with the B-form starting structures (Table 3), indicating that for 5HG, the A-form starting structures 2−4 display rapid conformational change in the direction of the B-form conformation. In contrast, all 5HG B-form starting structures showed consistently small RMSD's with their final products (average 1.28 Å).

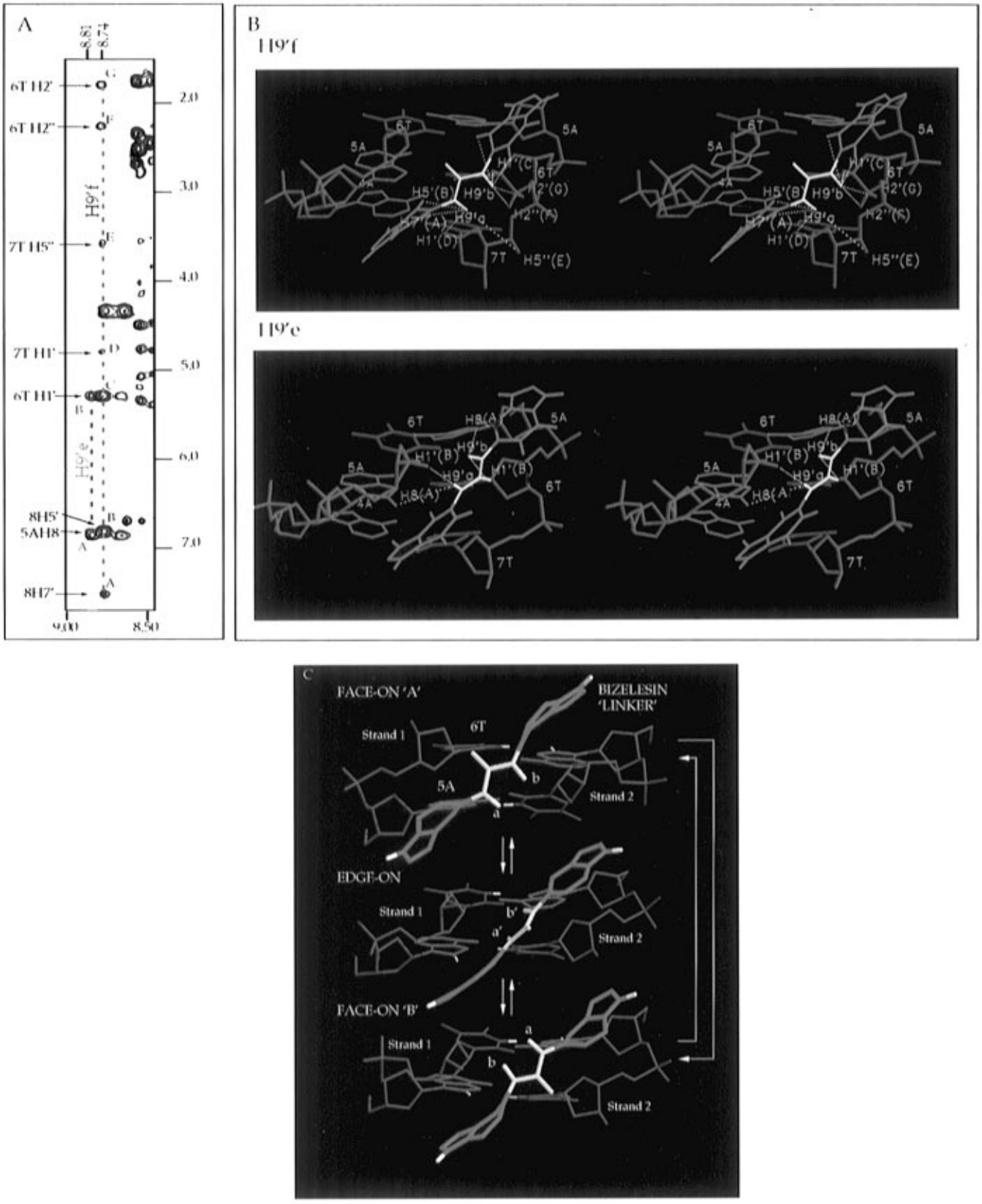

Figure 4. NOESY connectivities, stereodiagrams of the rMD generated structures, and proposed interconversion of the various cross-linked 5HG species. (A) Shown is the two-dimensional NOESY (150 ms mixing time) expanded contour plot of the Bizelesin−10-mer duplex adduct (90% H_2O/10% D_2O) showing the 5HG's Bizelesin ureadiyl subunit H9′ signal cross-peaks with duplex minor groove constituents. Cross-peaks for the edge-on isomer's Bizelesin ureadiyl H9′e amido signal are (A) 5A H8 and (B) 6T H1′ (indicated by red dashed line). Cross-peaks of the face-on isomer's ureadiyl H9′f amido signal are (A) 8H7′, (B) 8H5′, (C) 6T H1′, (D) 7T H1′, (E) 7T H5″, (F) 6T H2″, and (G) 6T H2′ (indicated by blue dashed line). (B) Stereoviews of the central AT-step region depicting the different chemical environments (top and bottom) experienced by the H9′f and H9′e ureadiyl amido protons, respectively. H9′f and H9′e cross-peaks to neighboring protons (shown in Figure 5A) are indicated by dotted lines. Colors are as indicated in Figure 2, except that all adenines are colored green. (C) Proposed interconversion scheme for 5HG in which equivalent face-on oriented ureadiyl subunits (top and bottom) flip back and forth, being occasionally trapped in the edge-on oriented form (center). Ureadiyl subunit H9′f (equivalent H9′a and H9′b protons) and H9′e (equivalent H9′a′ and H9′b′ protons), indole subunit H1′, adenine H8, and thymine H3 are white. Other colors are as Figure 2, except that all adenines are green.

Table 3. RMSD Calculations Derived from Comparison of the Starting Structures and the Average of Last 15 ps (300 K) of 100 ps rMD Trajectories from Four A-form and B-form Starting Structures (AMBER 4.0)[19]

A. Bizelesin Cross-Linked 10-Mer with Hoogsteen AT Step (5HG)

	SS origin,[b] ps		RMSD, Å			
5HG SS[a]	A	B	A & B	A & SS	B & SS	A & B SS
1	0	0	1.02	3.24	1.69	4.21
2	1	4	1.15	2.81	1.33	2.80
3	2	1	1.36	2.57	1.12	2.60
4	0.6	2	1.05	2.92	1.00	2.79
average, Å			1.14	2.88	1.28	3.10

B. Bizelesin Cross-Linked 10-Mer with Open AT Step (5OP)

	SS origin, ps		RMSD, Å			
5OP SS	A	B	A & B	A & SS	B & SS	A & B SS
1	0	0	1.18	2.82	2.73	4.20
2	2	1	1.58	1.42	1.94	3.51
3	1	2	1.19	1.74	1.67	2.94
4	3	4	1.18	1.78	1.12	2.48
average, Å			1.28	1.94	1.86	3.28

[a] SS, starting structure. [b] Starting structures were derived by extracting coordinate sets from the early stage of an initial *in vacuo* molecular dynamics (unconstrained) trajectory of a NUCGEN (AMBER 4.0) generated A- or B-form duplex structure docked and cross-linked to Bizelesin.

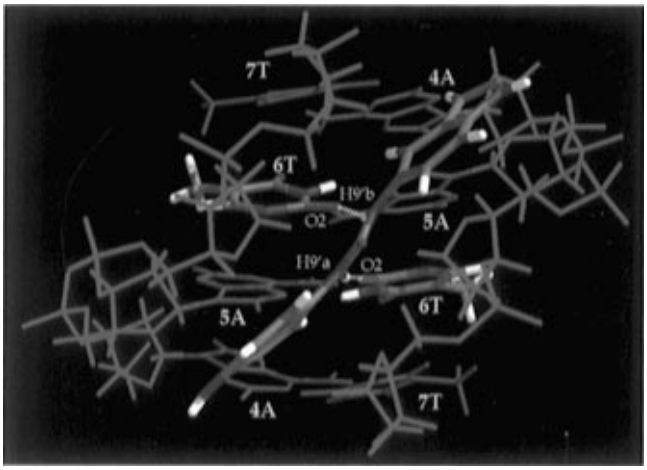

Figure 5. AT-step region of 5HG rMD-product displaying a symmetric hydrogen-bonding pattern involving the two ureadiyl amido donor protons and the AT-step's two thymine O2 acceptors (yellow dotted lines). Colors of the diindole−urea portion of Bizelesin and AT-step thymine bases according to atom type: blue for nitrogen, red for oxygen, white for hydrogen, and gray for carbon. Otherwise, colors are as follows: magenta for adenine, red for non-AT-step thymine, and green for backbone.

moieties (Figure 5). This arrangement contrasts with a Watson−Crick base-paired AT-step region wherein only a single thymine O2 acceptor is suitably oriented at a given time for hydrogen bonding with ureadiyl NH donors. In order to achieve simultaneous hydrogen bonding of the two thymine O2 acceptors, only AT-step Hoogsteen base pairing positions drug hydrogen-bond donors and thymine acceptors at appropriate sites within the minor groove's hydrophobic environment.

(D) Isomerization of the 5OP Species Is a Product of Concerted Bizelesin Ureadiyl Subunit and AT-Step Thymine Base Motion in the Open Base Pair AT-Step Region. (i) Structural Models Derived from NOESY/ROESY Conformational Exchange Data. Displacement of the 5OP AT-step bases away from a Watson−Crick base-paired B-form condition is supported by two additional lines of evidence to that previously described:[11] (1) An intense cross-peak expected between the 5OP 17T H3 proton (14.23 ppm) and the 16T H3 proton (10.41 ppm) is absent. In fact, the far-upfield 16T H3 broad signal, which yields only a single cross-peak with 15A H2, occurs at a chemical shift similar to those of unpaired thymines contained within loops.[21] (2) A break occurs at the 16T H6 to 15A H3′ step in the 5OP 11C → 20G base H8/H6 to H3′ cross-connectivity walk (Figure 4, Supporting Information).

Despite the AT-step "open" state, conformational exchange data indicate that these central bases are not free to move randomly throughout their potential range of motion. Conformational exchange cross-peaks are found for the 14A, 15A, 16T, and 17T base protons of major and minor isomers (Table 2). These conformational exchange cross-peaks for the 16T methyl and H6 protons reveal that AT-step base motion is constrained in either a major symmetrical hydrogen-bonding pattern (5OP$_{SYM}$: Figure 6, 16Ta Me × 17Ta Me cross-peak A, and Figure 7A) or a minor hydrogen-bonding pattern wherein the two AT-step thymines occupy different chemical environments (e.g., 5OP$_{ASYM}$: Figure 6, 16Tb Me × 17Tb Me cross-peak B and 16Tc Me × 17Tc Me cross-peak C, and Figure 7B). Because of the symmetric association of the 5OP$_{SYM}$ AT-step thymine O2 acceptors and the ureadiyl amido hydrogen donors, one major 16T Me × 17T Me cross-peak (Figure 6, cross-peak A) is produced for both AT-step thymine methyl functionalities and their neighboring 17T methyl groups. Each of the 5OP$_{SYM}$ 16T and 17T methyl signals produces a conformational exchange cross-peak with the alternative asymmetric conformer's corresponding methyl signals. Hence, the 16Ta methyl signal yields two exchange cross-peaks, one with 16Tb Me and the other with 16Tc Me (Figure 6A, cross-peaks E1 and E2, respectively). The adjacent 17Ta Me produces an exchange cross-peak with 17Tb Me (Figure 6, cross-peak E3) but not with the 17Tc Me cross-peak, due to their signal overlap (17Ta Me, 1.40 ppm; 17Tc Me, 1.41 ppm). A parallel series of conformational exchange cross-peaks are produced between 16T H6a and 16T H6b and H6c.

While only a single structure (5OP$_{SYM}$, Figure 7A) fits the symmetric "a" isomer NMR data, more than one asymmetric hydrogen bonding regime satisfies the ^{1}H-NMR-derived constraints. In one hypothetical example (5OP$_{ASYM}$), disruption results in a hydrogen-bonding association between the two ureadiyl subunit donors and only one AT-step thymine O2 acceptor (Figure 7B). In 5OP$_{ASYM}$, the 16Tb and 16Tc methyl signals (Figure 6) derive from thymine bases on opposite strands of an adduct lacking the 5OP$_{SYM}$ symmetry properties. The assignment of one pair of the minor isomer's 16T methyl and H6 signals to strand one and the other pair to strand two is supported by two lines of evidence. First, the intensities of the minor 16Tb and 16Tc methyl signals (and corresponding H6 signals) are equivalent, and second, conformational exchange cross-peaks between 16Tb and 16Tc methyl signals (and between 16Tb and 16Tc H6 signals) are absent. These differences in 5OP 16T methyl and H6 chemical shifts and conformational exchange cross-peaks can be explained by assuming that alternating 5OP hydrogen bond regimes persist long enough to permit the detection of different NMR signals for corresponding protons of interconverting symmetric and asymmetric isomers. The one detectable 5OP ureadiyl amido signal, 18H9′ (7.66 ppm, not shown), displays cross-peaks with 15A H2 (strong), 16T H1′ (moderate), and 18H5′ (weak), suggesting that the drug amido hydrogens are oriented edge-on toward the floor of the minor groove.

(ii) Models Derived from 5OP Restrained Molecular Dynamics for the 5OP Conformers. Paralleling the 5HG

(21) Gao, X.; Patel, D. *Biochemistry* **1988**, *27*, 1744−1751.

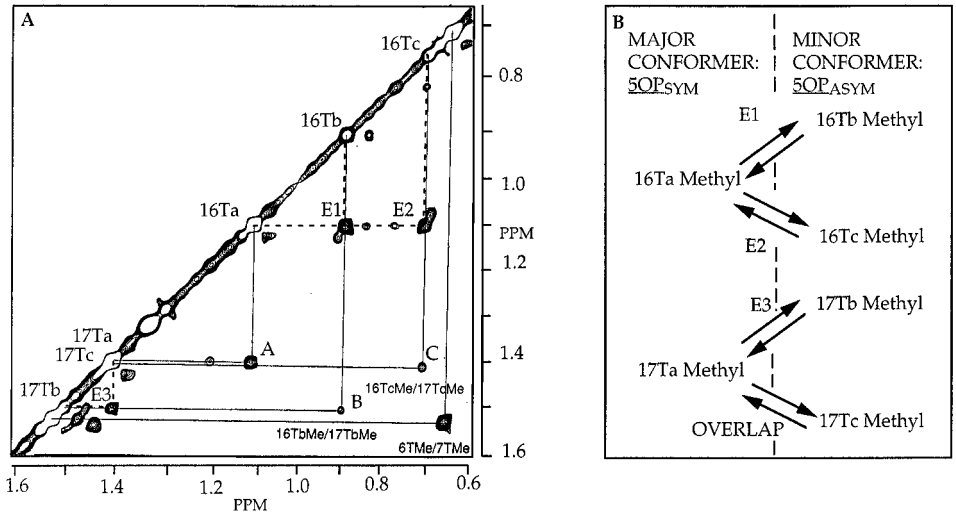

Figure 6. Conformational exchange of the 16T and 17T methyl signals in the 5OP structure. (A) Two-dimensional NOESY (200 ms mixing time) expanded contour plot of the Bizelesin−10-mer duplex adduct showing 5HG 6T and 7T and 5OP 16T and 17T methyl-to-methyl internucleotide cross-peaks and direct conformational exchange cross-peaks. The most intense methyl−methyl cross-peak (16Ta−17Ta) (A) corresponds to the symmetrical isomer, 5OP$_{SYM}$, in which the strand one and two 16T bases are held in equivalent environments, producing the same chemical shifts for the two methyl groups. Focusing of hydrogen bonding on a single thymine acceptor O2 (5OP$_{ASYM}$) results in this thymine (16Tb or c) methyl occupying a different chemical environment from the non-hydrogen-bonded thymine (16Tb or c). 16Tb and 16Tc methyl signals produce cross-peaks with 17Tb (B) and 17Tc (C) methyl signals, respectively, and conformational exchange cross-peaks with 16Ta (E1 and E2, respectively). (B) Diagram depicting the pattern of methyl exchange between the major (16Ta) and minor (16Tb and 16Tc) 5OP isomers. E1−E3 refer to exchange cross-peaks shown in Figure 8A (Supporting Information).

study, the first stage of 5OP rMD analysis involves the examination of convergence properties of *in vacuo* calculations using A- and B-form starting structures (Table 3B). The four pairs of starting structures converge satisfactorily (1.28 Å). Unlike 5HG, wherein the A-form starting structure rapidly converges on a B-form structure that does not differ appreciably from the B-form starting structures, both the A- and B-form 5OP converged structures display parallel levels of divergence from their starting structures (A, 1.94 Å; B, 1.86 Å). The final 5OP rMD product invariably diverges from the B- and A-form starting structures to give a structure with an "open" central AT-step region showing thymine O2 hydrogen bonding into the drug ureadiyl subunit.

The results of solvated 5OP rMD studies (162 NOESY-derived distance constraints: 124 DNA−DNA, 20 drug−DNA, 7 intradrug distances) and subsequent drug−DNA hydrogen-bonding analyses for the rMD trajectories (Figure 11, Supporting Information) yield hydrogen-bonding patterns that are congruent with the NMR-based models summarized above. While 5OP rMD trajectories vary more than the 5HG trajectories, there is a predictable trend: In each trajectory, the starting structure's ureadiyl subunit initiates hydrogen bonding with either one of the two AT-step thymine O2 acceptors (e.g., the hydrogen bonding regime depicted in Figure 7B). During this early part of the trajectory's constant temperature period, the donor pair of hydrogens commonly reorients from one AT-step thymine O2 acceptor to the other. As the rMD trajectories progress, the initially B-form-oriented bases rearrange into conformations resembling a "staggered" pattern of AT-step base positioning. When the AT-step thymines adopt a roughly symmetrical orientation vis-a-vis the ureadiyl subunit, the two thymine O2 acceptors are positioned equidistantly from the two amido hydrogen donors (5OP$_{SYM}$, Figure 7A). This major hydrogen-bonding system (5OP$_{SYM}$) common to these later periods of the rMD trajectories shows a maximum interaction arrangement of hydrogen-bond donors and acceptors. However, the ephemeral nature of this symmetrical hydrogen-bonding regime is evident in its frequent interconversion with the single O2 target acceptor regime (5OP$_{ASYM}$, Figure 7B). This 5OP variability contrasts with the 5HG results, portraying a relatively stable symmetrical hydrogen-bonding regime (Figure 5). These interconverting 5OP rMD hydrogen-bonding patterns are consistent with the NMR data for a symmetrical complex of dual thymine O2 and ureadiyl bifurcated hydrogen bonds (Figure 7A), interconverting with an asymmetrical system, wherein a three-membered H-bond draws only one thymine close into a hydrogen-bonding complex, while the other thymine, unconstrained by hydrogen bonding, resides nearer the major groove (Figure 7B).

Discussion

The idea that Bizelesin's cross-linking preference for the 5′-TAATTȦ sequence stems simply from the juxtaposition of two (+)-CC-1065 consensus sequences (i.e., 2 × 5′-TTȦ) is clearly untenable, in view of the results reported here and in previous publications.[10,11] Our objectives in this contribution are to delineate those factors that lead to this optimum sequence recognition and arrive at some basic rules for the design of sequence recognition molecules that cross-link at spans of six or more base pairs. This discussion is organized to examine why the simple additive model is misleading, to expose the cooperative role of the drug and sequence in the associated DNA reorganization and sequence recognition, and to delineate any lessons we can learn from these studies for the design of future therapeutically important DNA−DNA interstrand cross-linkers.

Consecutive alkylation by (+)-AB leads to same-strand alkylation, while bisalkylation by Bizelesin leads to the interstrand cross-linked product. A simple cross-linkage model equates cross-linkage by Bizelesin with two steps, each of which mimics (+)-CC-1065 monoalkylation. If, as suggested by this model, interstrand cross-linkage by Bizelesin entails consecutive 5′-TTȦ monoalkylation events, then bisalkylation by (+)-CPI-I (Figure 1A), the monoalkylating "half" of the Bizelesin dimer with (+)-CC-1065-like sequence preference properties,[4b,c,5] should produce a symmetrical bisadduct with a two-drug configuration resembling the interstrand cross-linked adduct. The unexpected (+)-CPI-I same-strand bisalkylation results (Figure 1C) show that while the first alkylation rapidly targets a 5′-TTȦ site (A^1 of 5′-TAA^{2}TTA^1), the much slower second alkylation occurs almost exclusively at A^2.[10] The

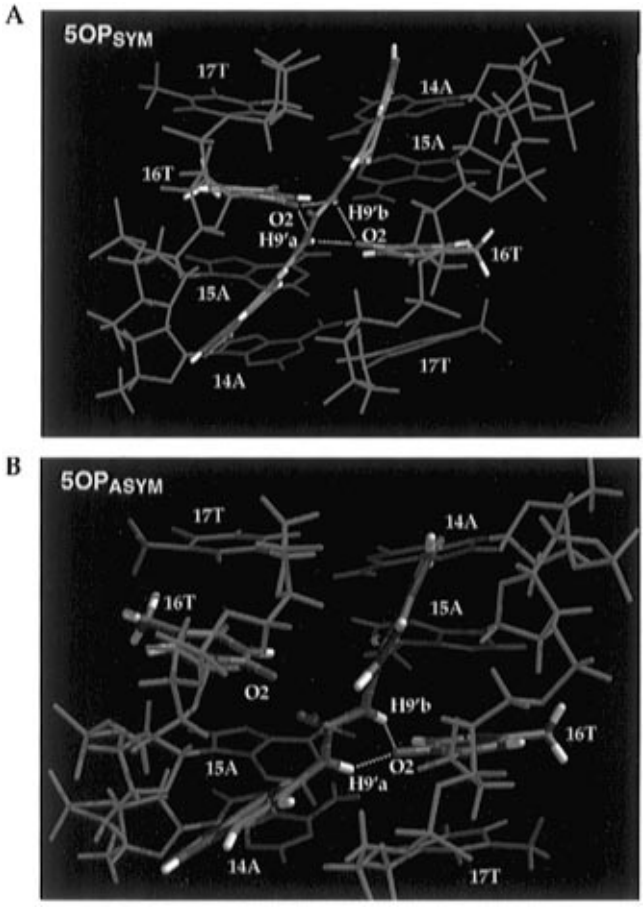

Figure 7. Symmetrical and asymmetrical hydrogen-bonding patterns for the AT-step region of rMD-product 5OP structure. (A) A symmetrical hydrogen-bonding pattern involving the two ureadiyl amido donor protons and the AT-step's two thymine O2 acceptors. (B) An asymmetric hydrogen-bonding pattern, 5OP$_{ASYM,}$ involving the two ureadiyl amido donor protons and one of the two AT-step's thymine O2 acceptors. Colors of Bizelesin and 16T base are according to atom type: blue for nitrogen, red for oxygen, white for hydrogen, and gray for carbon. Colors of other bases are as follows: purple for adenines and red for thymines.

conclusions drawn from this study are that, generally, the second (+)-CPI-I encounters a hierarchy of preferred alkylation targets qualitatively different from those encountered by the first. Specifically, distortion induced by the strand one 5′-TTȦ monoalkylation reaction[6] is sufficient to prevent the strand two 5′-TTȦ cross-linkage reaction.

These results suggest that monoadduct-generated bending distortion affects subsequent neighboring alkylations. Although no experimental data exist for the Bizelesin monoadduct intermediate, (+)-CPI alkylation-induced bending distortion that is propagated to the 5′-side of the covalently modified adenine[6] can be expected to disrupt the reaction of the second adenine N3 alkylation target relative to the remaining Bizelesin (+)-CPI-subunit's cyclopropyl moiety. The covalent reaction of the second (+)-CPI subunit just like the first presumably requires both DNA conformational flexibility[4c] and catalytic activation.[22] Comparison of Bizelesin cross-linkage and (+)-CPI-I bis-alkylation results[10] indicate that, following strand one monoalkylation, the conformation of the strand two 5′-TTȦ second alkylation site is too distorted to allow suitable reaction conditions for drug and DNA *as long as the DNA retains its Watson−Crick fully base-paired state.* That the second (+)-CPI-I alkylation occurs within a minor groove environment

(22) Lin, C. H.; Beale, J. M.; Hurley, L. H. *Biochemistry* **1991**, *30*, 3597−3602.

Table 4. Six-Base-Pair Bizelesin Cross-Link AT Sequences Showing the Percent of Monoalkylated and Cross-Linked Sequences Produced after 20 h (from Lee and Gibson)[4d]

	sequence[a]	monoalkylated, %	crosslinked, %
I	TAATTȦ ȧTTAAT	1.4	47.3
II	TTTTTȦ ȧAAAAT	27.0	9.6
III	TTTAAȦ ȧAATTT	15.1	2.1
IV	TTAAAȦ ȧATTTT	14.9	1.8
V	TATTTȦ ȧTAAAT	1.3	10.7

[a] The symbols ˙ and ° represent monoalkylation and cross-linked sites, respectively.

distinct from that targeted by the first (+)-CPI-I ligand raises the central question relating to 5′-TAATTȦ cross-linkage: What does Bizelesin do that two uncoupled (+)-CPI-I alkylators cannot do to dissipate the monoalkylation engendered duplex bending distortion and symmetrically cross-link the two 5′-TTȦ sites?

The rate of conversion of monoalkylated to cross-linked species is dependent upon the differential abilities of the selected duplexes to dissipate the bending-induced distortion associated with monoalkylation. The importance of the sequence context in determining the relative amount of monoalkylated vs cross-linked sequence is demonstrated in Table 4. Although all five sequences are entirely AT-derived, they differ in their constituent sequences. Sequences II, III, and IV all show significant bending[4d] and upon monoalkylation would presumably form drug-enforced and stabilized bent DNA structures. These sequences also show significant accumulation of monoalkylated product (14−27%) with minimal conversion to the cross-linked product (2−10%). In contrast, sequences I and V show only minor amounts of monoalkylation product (1.3−1.4%), with the majority of the product being in the form of the cross-linked species (11−47%). Duplex I appears to be a special case in that conversion to the straight form suitable for cross-linking occurs very rapidly and is facilitated not only by the unique characteristics of this sequence (see before) but also by the ureadiyl linkage of Bizelesin, since replacement of this linkage with a guanidino linkage leads to a very slow cross-linking reaction without the associated base pairing rearrangement of the DNA.[23] Duplex V may also be a special case, perhaps mimicking a duplex I rearrangement but with lower effeciency. Finally, in the two cases (I and II) examined by NMR and gel electrophoresis, the cross-linked product is a straight DNA structure in contrast to the bent monoalkylated product. In these two cases, for cross-linking to occur with Bizelesin, the bent monoalkylated product must be converted to straight DNA before cross-linking can occur.

Elucidating How the Bizelesin-Directed DNA Rearrangement Occurs. The relatively efficient cross-linking of the 5′-TAATTȦ sequence by Bizelesin leads to the conclusions that there is inherent in the 5′-TAATTȦ−Bizelesin complex a unique solution to the problem posed by the bending distortion-based obstacle to cross-linkage and that Bizelesin's rearrangement of this preferred sequence removes this obstacle to the second 5′-TTȦ alkylation event. In order to explore the details of this

(23) Seaman, F.; Lee, S.-J.; Hurley, L. H. Unpublished results.

rearrangement, we sought answers to three questions: (1) What initiates the process, and where does the driving force come from for this step? (2) Given the range of possible outcomes, what restricts the course of the reaction to two pathways leading to 5HG and 5OP conformer products? (3) Within the 5HG and 5OP clasess, what drives the interconversion between positional isomers corresponding to the conformational exchange patterns? These three questions are dealt with in the next section, which also provides a viable explanation for how the duplex reorganization occurs.

(a) The rearrangement process is initiated by AT-step base-pair opening, a low-energy process driven by dissipation of the energy contained in the bent monoalkylated product. Because bent Watson−Crick base-paired cross-linked product is not detected, it follows that cross-linkage proceeds only after the effects of this monoadduct-associated bending have been nullified. Bending is dissipated by central AT-step base-pair opening and base displacement toward the major groove, in compliance with previously documented bending/base-pair-opening linkage.[24] This study shows that a relatively low energy barrier is associated with AT-step base-pair opening when coupled with the energy contained in a bent DNA structure. The base-pair opening event is facilitated by two AT-step properties of the unmodified duplex: (1) AT-step flexibility resembling that of the Dickerson dodecamer's central region[14b,25] and (2) relatively fast base-pair-opening rates for the four central base pairs. In addition to this innate duplex behavior, another contributor to the base-pair-opening event is the monoadduct's drug−DNA binding interactions. Once covalently immobilized at one end (monoadduct), Bizelesin uses complex electrostatic and hydrogen-bonding interactions to "fish" for any DNA behavior that will set in motion the distortion-removing reorganizational process. Expressed differently, covalent bonding at one site in the drug−DNA complex provides the opportunity (via the minor groove proximity of two linked reactants) to not only fully explore conformational space in search of conformations capable of cross-linkage, but also to change the conformational topography of this space.

(b) In the monoalkylated intermediates, coordinated hydrogen bonding between the ureadiyl amido hydrogen bond donors and thymine O2 hydrogen bond acceptors directs the cross-linkage reactions to the 5OP and 5HG products. Bizelesin adducts of 5′-T̲AAAAA̲,[12] 5′-T̲AATTA̲,[11] and 5′-T̲TAGTTA̲[13] differ substantially in the course of the reaction and the nature of the products. These reactions are differentiated by the sequence-specific DNA properties that set the ground rules for the interaction with the "recognition" components of the bound drug. The most obvious 5′-TAATTA properties are (1) the unique array of thymine O2 hydrogen bond acceptors that serve as likely targets for the hydrogen bond donor pair in the central ureadiyl subunit and (2) the base-pair-opening properties of the central AT base pairs.

Opening and displacement of the bases of a central AT-step base pair toward the major groove remove a physical obstruction to the association and coordination of the drug hydrogen bond donor and DNA acceptor groups in the minor groove. Stabilization of intermediate and final products of this process can be attributed to the hydrogen bonding between Bizelesin's ureadiyl hydrogen bond donors and the AT-step thymine O2 acceptors. Complex patterns of hydrogen bonding between drug donor substituents and minor groove DNA acceptors have been reported for many minor groove binding drugs.[26] A common feature of such complexes is a hydrogen bond network capable of extraordinarily long donor−acceptor distances resulting from the combination of the largely electrostatic character of the bond's energy and the embedding of the hydrogen bond within the hydrophobic environment of the minor groove floor.[26b,27] Because minor groove drug−DNA hydrogen bonds are more stable than those in regions accessible by solvent, hydrogen bond networks provide the means for stabilizing the intermediate and final products of Bizelesin-induced DNA reorganization.

The initiation of such binding interactions requires that the ureadiyl portion of the drug be capable of adopting an edge-on orientation relative to the minor groove floor, an event that modeling indicates cannot precede the first base-pair-opening step. Following the first base-pair opening, the remaining AT-step base pair opens, presumably as a consequence of the competition between the ureadiyl hydrogen donors and the adenine of this remaining base pair for the pairing thymine. Once both AT-step base pairs are open, the full range of hydrogen-bonding interactions between Bizelesin's edge-on oriented ureadiyl amido hydrogens and AT-step thymine bases is possible. The next step of this sequence-dependent interplay of ureadiyl hydrogen donors and AT-step base acceptors proceeds along two alternative paths, either stabilization of the AT-step open state (pre-5OP) or rotation of the two adenines from *anti*- to *syn*-orientation (pre-5HG). Opening of both pre-5HG AT-step base pairs coupled with the rapid base-pair-opening rates of adjacent AT base pairs in the central 5′-AATT region widens the path through which the AT-step adenines can undergo rotation about their glycosidic bonds. Once rotated, the adenines are stabilized in the *syn*-orientation by both the Hoogsteen base-pair hydrogen bonds and the drug−DNA hydrogen bonding described earlier.

Both the pre-5HG monoadduct Hoogsteen base pairing-facilitated drug−DNA hydrogen bonding and the pre-5OP monoadduct bifurcated hydrogen bonding systems are binding mechanisms that stabilize monoadducts in a rearranged, nonbent conformation suitable for the final cross-linking alkylation reaction. Once cross-linked, these alternative monoadduct hydrogen-bonding regimes yield conformers typified by four major patterns: the edge-on 5HG pattern involving simultaneous drug hydrogen bonding by both AT-step thymines (Figures 4C and 5), the face-on 5HG pattern of unknown hydrogen bonding properties (Figure 4C), the edge-on 5OP symmetrical dual bifurcated pattern (Figure 7A), and the 5OP asymmetrical pattern exemplified by $5OP_{ASYM}$'s drug hydrogen bonding to a single AT-step thymine (Figure 7B).

(c) Conformers interconvert between positional isomers characterized in part by different drug−DNA hydrogen-bonding patterns. Experimentally, these positional isomers are defined by conformational exchange patterns (see, for example, Figures 3, 4, and 6 and Table 2) linking member isomers within the 5HG and 5OP classes. Conformational exchange on the NMR time scale suggests a scenario wherein hydrogen-bonding regimes that are too weak to resist disruption by molecular motion experience transitions from one hydrogen-bonding pattern to the other (see, for example, Figures 4C and 7A,B).

Lessons from Bizelesin-Induced Rearrangement of DNA for the Design of Second-Generation Sequence-Specific

(24) Ramstein, J.; Lavery, R. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 7231−7235.

(25) Sarai, A.; Mazur, J.; Nussinov, R.; Jernigan, R. L. *Biochemistry* **1989**, *28*, 7842−7849.

(26) (a) Pelton, J. G.; Wemmer, D. E. *Biochemistry* **1988**, *27*, 8088−8096. (b) Neidle, S.; Pearl, L. H.; Skelly, J. V. *Biochem. J.* **1987**, *243*, 1−13. (c) Kumar, S.; Bathini, Y.; Joseph, T.; Pon, R. T.; Lown, W. *J. Biomol. Struct. Dyn.* **1991**, *9*, 1−21. (d) Quintana, J. R.; Lipanov, A. A.; Dickerson, R. E. *Biochemistry* **1991**, *30*, 10294−10306. (e) Brown, D. G.; Sanderson, M. R.; Skelly, J. V.; Jenkins, T. C.; Brown, T.; Garman, E.; Stuart, D. I.; Neidle, S. *EMBO J.* **1990**, 9, 1329−1334.

(27) Kopka, M. L.; Yoon, C.; Goodsell, D.; Pjura, P.; Dickerson, R. E. In *Structures and Motion: Membranes, Nucleic Acids and Proteins*; Clementi, E., Corongiu, G., Sarma, M. H., Sarma, R. H., Eds.; Adenine: Guilderland, NY, 1985; p 461.

DNA−DNA Cross-Linkers. At first glance, optimum cross-linking of the sequence 5′-TAATTA̲ appears to depend on the repetition of the monoalkylation sequence preference for 5′-TTA̲ during both alkylation events. This interpretation would be consistent with the principle of minimal reorganization of DNA during the cross-linking of the monoalkylation adduct[7,8] if the Bizelesin cross-linkage did not reorganize the target DNA. Bizelesin clearly does not fit into this common pattern of sequence recognition for cross-linkers. What appears to set up a distinctly different pathway to the products of Bizelesin cross-linking is the bent and distorted nature of the monoalkylated intermediate. In order for the second alkylation reaction (cross-linking) to proceed, the DNA must be reorganized from its distorted form to a nonbent DNA duplex that brings into proximity all functional groups involved in the cross-linking alkylation step.

Bizelesin meets the requirements for a second-generation cross-linker where monoalkylation duplex distortion (leading to a conformation unsuitable for cross-linking) works in tandem with the manipulative properties of the ureadiyl drug subunit in trapping out a conformational form suitable for cross-linking. Depending on the duplex sequence, the precise route to the cross-linked product is quite variable. For 5′-TAAAAA̲, Bizelesin takes advantage of the conformational equilibrium between bent and straight forms of DNA.[12] For 5′-TAATTA̲, the pathway is more complex, involving participation of the ureadiyl hydrogen bond donors and a unique rearrangement of the central AT base pairs.

In the accompanying companion paper,[28] we have demonstrated that more complex linker subunits providing multiple donor−acceptor pairs when coupled with minor groove base substituents dictate mixed AT/GC sequence selectivities. What is intrinsically different about this example is that the overall structure, which is still unknown in detail, is grossly distorted at the central part of the cross-linked duplex. Is the design and synthesis of DNA−DNA cross-linkers that have distortive monoalkylation intermediates a generally applicable approach to increasing sequence selectivity and therapeutic selectivity? Clearly, designing sequence selectivity is achievable, but it is still beyond the capabilities of students of molecular recognition to predict both the conformational variability in stretches of six or more base-pair sequences and how the interplay of hydrogen bond donors and acceptors on the drug and DNA can facilitate unusual rearrangements, such as those documented here. In preclinical evaluation, Bizelesin is not only more potent as a cytotoxic agent than the monoalkylating clinical candidates Adozelesin and Carzelesin, but also appears to be more efficacious in animal tumor models.[2,3] Whether this translates into clinical improvement is still an open question, but this important corollary should be answered by the results of phase 1 and 2 clinical trials.

Conclusions

1. Bizelesin cross-linkage of its preferred sequence, 5′-TAATTA, can only proceed after the DNA is reorganized by a series of drug-induced base-pair opening, rearrangement, and hydrogen-bonding steps.

2. These results demonstrate how "drug-induced reorganization" of DNA (a specialized type of induced fit[26b]) can have a role in extending the range of sequence specificity of cross-linkers.

Experimental Procedures

(a) Chemicals. Bizelesin was a gift from The Pharmacia Upjohn Co. Reagents used to prepare the NMR buffer, sodium phosphate (99.99%) and sodium chloride (99.99%), were purchased from Aldrich. HPLC water and methanol were purchased from Baxter Scientific and Fisher, respectively.

(b) Oligonucleotide Preparation and Purification. Synthesis and purification of the self-complementary 10-mer [d(CGTAATTACG)$_2$] was previously described.[11]

(c) Adduct Preparation and Purification. Preparation of the Bizelesin adduct was previously described.[11] Additional purification was achieved by reverse-phase HPLC using Rainin C18 and C8 Dynamax-300A preparative columns (21.4 mm × 250 mm). The solvent gradient progressed from a buffer solution of 15 mM $NaHPO_4$ toward a solvent of CH_3CN (100%) with a flow rate of 5 mL/min. The percentage of the latter solvent increased according to a regimen of 0.0% (0 min), 45% (5 min), 65% (50 min), and 100% (60 min). 5HG and 5OP conformers were inseparable by HPLC using these and several other solvent systems.

(d) Proton NMR Experiments. One- and two-dimensional 500 MHz ^{1}H and 202.44 MHz ^{31}P NMR data sets in H_2O and D_2O buffered solution (pH 6.8−7.0) were recorded on General Electric GN-500 and Bruker AMX 500 FT NMR spectrometers. Proton chemical shifts of the ca. 6 mM buffered solution were recorded in parts per million (ppm) and referenced relative to external TSP (1 mg/mL) in D_2O (HOD signal was set to 4.751 ppm).

NOESY. Phase-sensitive two-dimensional NOESY spectra (Bruker) were obtained at 27 °C (TPPI) for two mixing times, 100 and 200 ms, using a presaturation pulse to suppress the HOD signal. All spectra were acquired with 16 scans at each of 1024 t_1 values, a spectral width of 10.002 ppm, and a repetition time of 10 s between scans. During data processing, a shifted squared sine bell function (shift = 90°) was used in both ω_1 and ω_2 dimensions. The FID in ω_1 was zero-filled to 2K prior to Fourier transformation to give a 2K × 2K spectrum. Two-dimensional NOE spectra in 90% H_2O at 150 ms mixing time were recorded at 27 °C using the 1−1 echo read pulse sequence[29] with a 2.5 s pulse repetition time, a sweep width of 24.396 ppm, and a 90° pulse width of 12 μs.

ROESY. Exchange processes were studied by ROESY[30] in D_2O with a spectral width of 5000 Hz in both dimensions. Mixing times of 100 and 200 ms were used to collect 1K FIDs, each of 32 scans, and to generate 1024 × 2048 data matrices.

Ammonia Catalyst-Mediated Hydrogen Exchange Experiments. Ammonia catalyst-mediated base-pairing hydrogen exchange experiments[31] were conducted on the unmodified 10-mer.

Fast PE COSY. Two-dimensional fast PE COSY experiments[15,32] yielded 2 × 512 × 2048 data matrices with acquisition times of 152 and 304 ms in the t_1 and t_2 dimensions, respectively. Zero-filling was used in both dimensions. Sixty-four scans were acquired per t_1 value, with a delay time of 2.2 s between scans, resulting in a total measuring time of 20 h. The one-dimensional reference spectrum was acquired in a 4K data table with 512 scans. Other conditions and parameters were as described previously.[32]

Two-Dimensional Partially Decoupled *J*-Scaled 2QF-COSY. Two-dimensional partially decoupled *J*-scaled 2QF-COSY[15] provides an estimate of the coupling between deoxyribose H3′ and H4′ using a modified COSY experiment wherein the couplings between H3′ and H2′, H2″, and ^{31}P were removed in the F_1 dimension. The two-dimensional spectrum was derived from a 2 × 256 × 1024 data matrix with acquisition times of 512 and 160 ms in the t_1 and t_2 dimensions, respectively. The transmitter was placed at 4.6 ppm and the duration of the soft 180° pulse was 1 ms. The data were processed with a sine-bell window function phase-shifted by 30° and zero-filled to a 2K × 2K matrix.

DNA Sugar Conformation Calculations. PE COSY data and partially decoupled 2QF-COSY data were used to calculate coupling constants for deoxyribose H1′−H2′, H1′−H2″, H2′−H3′, H2″−H3′, and H3′−H4′ (Table 1). These coupling constants were input into

(28) Park, H.-J.; Kelly, R.; Hurley, L. H. *J. Am. Chem. Soc.* **1996**, *118*, 10041−10051.

(29) (a) Sklenar, V.; Bax, A. *J. Magn. Reson.* **1987**, *74*, 469−479. (b) Blake, P. R.; Summers, M. F. *J. Magn. Reson.* **1990**, *86*, 622−624.

(30) Bothner-By, A. A.; Stevens, R. L.; Lee, J. T.; Warren, C. D.; Jeanloz, R. W. *J. Am. Chem. Soc.* **1984**, *106*, 811−813.

(31) (a) Moe, J. G.; Russu, I. M. *Nucleic Acids Res.* **1990**, *18*, 821−827. (b) Guéron, M.; Kochoyan, M.; Leroy, J.-L. *Nature* **1987**, *328*, 89−92.

(32) Mueller, L. *J. Magn. Reson.* **1987**, *72*, 191−196.

PSEUROT[33] in order to calculate the phase angle of pseudorotation, P, and maximum pucker amplitude, F_M, from the vicinal proton spin-spin coupling constants (vicinal^3JHH) involving deoxyribose H1′, H2′, H2″, H3′, and H4′ (10-mer). In order to calculate the major S-type conformation, the minor N-type conformer was constrained to $P = 9°$ and $F_m = 36°$. Deoxyribose conformational equilibrium mixtures (N- and S-types) were calculated for each sugar.

(e) Restrained Molecular Dynamics. Interproton distance restraints were derived via the program MARDIGRAS[34] from NOESY experiments involving mixing times of 100 and 200 ms. The symmetry of the complex dictated that the single set of cross-peak intensities be applied to strand one intrastrand distances and a duplicate set be applied to strand two. Interstrand cross-peaks were evaluated case by case to exclude possible intra/interstrand ambiguity near the axis of symmetry. A complete two-dimensional NOE relaxation matrix was set up using the geometry of a C2-symmetrical 10-mer starting structure to provide interproton NOEs not available from the experimental datasets. Alternative starting structures were (1) Bizelesin 10-mer adduct nonrestrained molecular dynamics product and (2) energy-minimized (to energy convergence criterion of 0.001 kcal/mol for successive steps) B-DNA adduct (Hoogsteen modified in 5HG). The complete NOE matrix was calculated for the starting structure using CORMA.[35] MARDIGRAS calculates upper and lower bounds, r_2 and r_3, for the interproton distances depending upon the agreement between the experimental and the converged MARDIGRAS cross-peak intensities as well as signal-to-noise ratio.[35b] In these calculations, isotropic motion was assumed and a single correlation time of $\tau_C = 2.0$ ns was used. Subsequent procedures followed those described previously.[36] An Amber 4.0 restraint file was prepared using the MARDIGRAS NMRCNST module. Several nonstandard residues had to be produced for the 5HG and 5OP model structures. The 5HG model structure Hoogsteen AT base pair's adenosine was derived simply by rotating the standard Amber 4.0 adenosine residue's adenine approximately 180° about its glycosidic bond. The 5OP open region's pairing behavior was approximated by redefining the AT-step N3 hydrogen atom type to a non-hydrogen bonding form (atom type H to H5) and leaving the base charges unchanged. Parameterization of the Bizelesin (+)-CPI subunit for use with the Amber 4.0 force field follows from those derived earlier for (+)-CC-1065.[37] The indole subunit's parameterization followed from the tryptophan standard residue's parameters.[19] The initial coordinates and charges for Bizelesin and a pair of N9-methylated adenine base ligands were calculated using MOPAC 5.0 ESP.[38] Interproton distances were incorporated into the restrained molecular dynamics calculations: For the 5HG model, 117 DNA−DNA, 39 drug−DNA, and 6 drug−drug distance constraints and for 5OP 113 DNA−DNA, 28 drug−DNA, and 12 drug−drug distance constraints were derived entirely from nonexchangeable proton cross-peaks. rMD calculations were performed using SANDER (AMBER, version 4.0)[19] on an SGI Indigo2 XZ workstation. The AMBER force field pseudoenergy terms for the interproton distances have the form of flatwells with parabolic sides. RMSD analysis was conducted using CARNAL (Amber 4.1)[39] within and between the two families of A- and B-form starting structures and the average rMD products (Table 4). For the RMSD study (*in vacuo*), the 10-mer AT-10 was expanded to a 14-mer, 5′-CGCGTAATTACGCG, with retention of Bizelesin cross-link and distance restraints at corresponding positions. Large sodium ions (hexahydrated, radius = 5 Å)[40] were placed along the PO_2^- bisector 6.25 Å from the phosphorus atom. Four A-form cross-linked 14-mer starting structures were derived by using NUCGEN (Amber 4.0) to generate one structure, which was used in the initial *in vacuo* rMD analysis. The remaining three starting structures were generated at brief intervals at the beginning of an unconstrained MD trajectory of the first starting structure. The four B-form 14-mer starting structures were generated in a similar fashion. The rMD trajectory followed a temperature program beginning at 0 K and ramping to 800 K over a period of 65 ps. After 10 ps at 800 K, the program was ramped down to 300 K. The restraints are applied as a well with a square bottom with parabolic sides out to a defined distance and then linear beyond there.[19] The flat region of the well is defined by the upper and lower bounds, r_2 and r_3 (see above), calculated using MARDIGRAS. The weights of the hydrogen bond and distance restraints were modulated by multiplying the force constants by a scaling factor. At 800 K, the restraint force constants reached their maximum values of 40 kcal/mol·Å^2 (distance restraints), 20 kcal/mol·Å^2 (hydrogen bond distance restraints), and 20 kcal/mol·rad^2 (hydrogen bond angle restraints) and were reduced by half during the ramping down to 300 K. In the final 20 ps isothermal phase (300 K) of the rMD analysis, an average structure was generated using CARNAL[39] for each of the A- and B-form starting structures.

Solvated rMD was performed on the 5OP and 5HG 10-mer adducts by first positioning counterions (counterion charge = 1.0) 4.5 Å away from phosphorus, surrounding the minimized 10-mer adducts with 64 boxes of 216 Monte Carlo waters, and limiting the number of water molecules to those whose oxygen atoms are within a 5.0 Å cutoff distance (5HG, 416 water molecules; 5OP, 459 water molecules). Following Belly energy minimization for water molecules only, energy minimization to a maximum derivative of 0.01 Å was performed with distance restraints and hydrogen bond and angle restraints (maximum force constant = 10 kcal/mol·Å). Belly dynamics for water only (0−300 K; 10ps; no restraints) was followed by rMD (solvated). rMD was conducted without periodic boundary conditions, with coupling to constant temperature heat bath, and with the SHAKE (removal of bond stretching freedom) option applied to all bonds. The rMD trajectory followed a temperature program beginning at 0 K and ramping to 500 K over a period of 60 ps. After 10 ps at 500 K, the program was ramped down to 300 K. The weights of the hydrogen bond and distance restraints were modulated by multiplying the force constants by a scaling factor. At 500 K, the restraint force constants reached their maximum values of 20 kcal/mol·Å^2 (distance restraints), 10 kcal/mol·Å^2 (hydrogen bond distance restraints), and 10 kcal/mol·rad^2 (hydrogen bond angle restraints) and were reduced by half during the ramping down to 300 K. After stabilization at 300 K, trajectories from rMD analyses were examined (175 ps) for evidence of drug−DNA hydrogen bonds using the CARNAL module of AMBER 4.1. Every 500 steps (0.5 ps) a coordinate set was generated for the solvated 10-mer adduct. For a given coordinate set, the CARNAL HBOND option specifies the H-bond distances and angles for each pair of donors and acceptors that meets a set of predefined criteria (distance = 3.4 Å, angle = 60°). For the two duplex strands, CARNAL defines a total of nine possible hydrogen-bonding schemes involving the Bizelesin ureadiyl subunit's donors and the AT-step thymine O2 acceptors (Figures 6 and 10, Supporting Information). When symmetry properties are considered, hydrogen-bonding patterns 2 and 3 become equivalent, as do patterns 5 and 6.

Acknowledgment. This research was supported by grants from the National Institutes of Health (CA-49751), The Welch Foundation, and The Pharmacia Upjohn Co., who also provided the Bizelesin. We thank David Bishop for proofreading and editing the manuscript and the reviewer of an earlier version of the manuscript for constructive critical comments.

Supporting Information Available: NOESY, ROESY, COSY, and $^1H-^{31}P$ COSY plots of 10-mer d(CGTAATTACG)$_2$ and other relevant NOESY plots (17 pages). See any current masthead page for ordering and Internet access instructions.

JA961924I

(33) de Leeuw, F. A. A. M.; Altona, C. *J. Comput. Chem.* **1983**, *4*, 428−437.

(34) (a) Borgias, B. A.; James, T. L. *J. Magn. Reson.* **1988**, *79*, 493−512. (b) Borgias, B. A.; James, T. L. *J. Magn. Reson.* **1990**, *79*, 475−487.

(35) (a) Borgias, B. A.; Thomas, P. D.; James, T. L. *Complete Relaxation Analysis (CORMA)*; University of California: San Francisco, 1987, 1989. (b) Borgias, B. A.; Gochin, M.; Kerwood, D. J.; James T. L. *Prog. Nucl. Magn. Reson. Spectrosc.* **1990**, *22*, 83−100.

(36) Kerwood, D. J.; Zon, G.; James, T. L. *Eur. J. Biochem.* **1991**, *197*, 583−595.

(37) Yuan, Y.-C.; Seaman, F. C.; Hurley, L. H. Unpublished results.

(38) Besler, B. H.; Merz, K. M.; Kollman, P. A. *J. Comput. Chem.* **1990**, *11*, 431−439.

(39) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E.; Ferguson, D. M.; Seibel, G. L.; Singh, C.; Weiner, P. K.; Kollman, P. A. *AMBER 4.1*; University of California: San Francisco, 1995.

(40) Singh, U. C.; Weiner, S. J.; Kollman, P. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 755−759.