Cross-Linkage by "Intact" Bizelesin and Bisalkylation by the "Separated Halves" of the Bizelesin Dimer: Contrasting Drug Manipulation of DNA Conformation (5′-TAATTA-3′) Directs Alkylation toward Different Adenine Targets

Frederick C. Seaman,* Jianxiong Chu, and Laurence Hurley*

Contribution from the Drug Dynamics Institute, College of Pharmacy, The University of Texas at Austin, Austin, Texas 78712-1074

Received January 2, 1996[⊗]

Abstract: Gel electrophoresis analysis of CPI-I bisalkylation of a 21-mer duplex containing 5′-TAA^{2}TTA^1-3′ (the palindromic preferred cross-linking sequence of the (+)-CC-1065 analog Bizelesin) shows same-strand (strand one) alkylation of first A^1 and then A^2 instead of the anticipated symmetrical A^1 alkylation of strands one and two. Two-dimensional NMR analyses (NOESY and COSY) confirm the head-to-tail minor groove orientation of the same-strand-bound drugs. CPI-I contrasts sharply with Bizelesin (two CPI-I units linked tail-to-tail by a ureadiyl "linker"), which symmetrically cross-links this sequence at A^1 (strands one and two), but only by first rearranging the duplex structure and consequently removing the duplex distortion stemming from monoadduct formation. CPI-I induces no such major DNA rearrangement prior to or during bisalkylation. Why does CPI-I react with the adenines of only a single strand? Two possible causes for the unexpected strand one A^2 alkylation are, first, retardation of strand two A^1 site's reactivity by focusing of monoadduct conformational distortion on this site and, second, elevation of A^2 reactivity above other competing adenine sites due to unusual monoadduct strand one A^2T-step conformational properties. The relative importance of these two nonmutually exclusive factors was investigated using gel electrophoresis experiments: Time-course CPI-I bisalkylation studies were conducted on the AT-step sequence 5′-TAA^{2}TTA^1-3′ and an A-tract sequence, 5′-TAA^{2}AAA^1-3′, to see if the former sequence's AT-step flexibility, high base-pair opening rate, and unwinding capability (traits not shared by the latter sequence) controlled selection of the second target site. The observed parallel AT-step and A-tract sequence A^1 and A^2 bisalkylation patterns suggest that AT-step properties play at best a secondary role (compared to 5′-end TA-step behavior) in directing the second alkylation reaction to the AT-step site. rMD (solvated) simulations of the AT-step and A-tract monoadducts display distortion that is focused on this 5′-end TA-step site. While two-dimensional ^{1}H NMR spectra of the final bisadduct reveal no significant TA-step conformational distortion, they demonstrate that conformational adjustment at the A^2 ligand attachment site diminishes head-to-tail steric clash of the two drugs. These results suggest that the CPI-I monoadduct propagates bending distortion (to the 5′-side) through five base pairs toward the TA-step junction site. In the AT-step and A-tract sequences, neither adenine straddling this TA-step junction site is alkylated by CPI-I, suggesting that the base pairs forming the junction site are distorted away from a suitable orientation or are unable to assume a conformation suitable for alkylation. Consequently, the second alkylation occurs at a site (AT-step) that requires a modest displacement of the second ligand away from the already attached drug. The results and analysis of the data included in this paper provide important lessons for the design of inter- and intrastrand DNA−DNA cross-linkers.

Introduction

(+)-CC-1065 has served as the inspiration for the design and synthesis of a number of potentially clinically important agents. Originating from The Pharmacia Upjohn Co. program, two synthetic monoalkylating agents are in phase 2 clinical trials (Adozelesin and Carzelesin) and a synthetic bisalkylating compound (Bizelesin) is in preclinical development.[1] Bizelesin displays increased *in vitro* cytotoxic potency and excellent *in vivo* efficiency in comparison to its cyclopropa[*c*]pyrrolo[3,2-*e*]indol-4(5*H*)-one (CPI) monoalkylating parent compound, (+)-CC-1065, Adozelesin (in phase 1 and 2 clinical trials), and U80,722, the most potent of the flexible CPI cross-linkers.[2−4] Bizelesin (Figure 1B), a DNA−DNA interstrand cross-linker, was designed and synthesized by Pharmacia Upjohn scientists[1,4] to cross-link DNA six base pairs apart on opposite strands in the sequence 5′-TAATTA-3′. This duplex sequence contains at the 3′-adenines two consensus bonding sites (5′-TTA) for the two CPI subunits. As predicted, the 5′-TAATTA-3′ sequence was found to be one of the most reactive cross-linking sites for Bizelesin,[6] but unexpectedly, the sequence was rearranged in the cross-linked product.[7] To probe this result, the alkylation reactivity and selectivity of the two "unhooked" constituent halves of Bizelesin (i.e., the two CPI-I structures;[5] see Figure 1) in the same sequence (5′-TAATTA-3′) were determined. On the basis of the cross-linked structure of the Bizelesin 5′-TAATTA-3′ sequence, we anticipated that the CPI subunits of both CPI-I drugs would attach to the 3′-adenines

* To whom correspondence should be addressed. Phone: (512) 471-4841. Fax: (512) 471-2746.

[⊗] Abstract published in *Advance ACS Abstracts,* May 1, 1996.

(1) 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.

(2) Hurley, L. H.; Warpehoski, M. A.; Lee, C.-S.; McGovren, J. P.; Scahill, T.; Kelly, R. C.; Mitchell, M. A.; Wicnienski, N. A.; Gebhard, I.; Johnson, P. D.; Bradford, S. *J. Am. Chem. Soc.* **1990**, *112*, 4633−4649.

(3) Lee, C.-S.; Gibson, N. W. *Cancer Res.* **1991**, *51*, 6586−6591.

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

S0002-7863(96)00017-0 CCC: $12.00

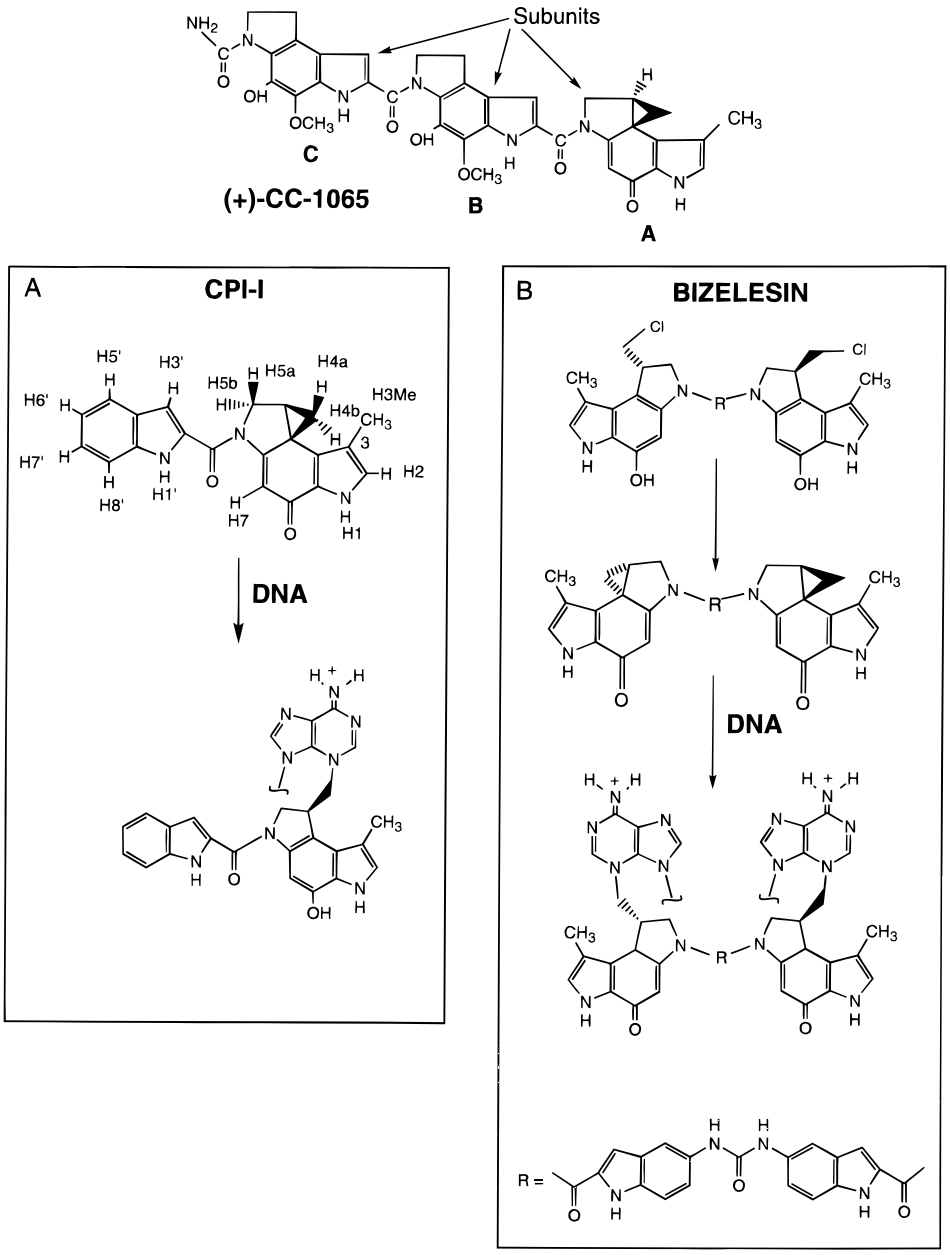

Figure 1. Structures of (+)-CC-1065, CPI-I (U72,779), and Bizelesin (U77,779) and the adenine N3 alkylation by (A) CPI-I and (B) Bizelesin.

and orient in the 5′-direction relative to each covalently modified adenine (Figure 2A and SYMM in Table 1). This would preserve the duplex symmetry just as Bizelesin cross-linking does.[7] The CPI-I (Figure 1) bisadduct was also expected to resemble the Bizelesin cross-linked structure because the predicted CPI-I bisadduct structure differed from the Bizelesin adduct only by the absence of the ureadiyl moiety that joins two CPI-I-like halves to form Bizelesin (Figure 1).

Surprisingly, NMR and gel electrophoresis analyses of CPI-I bisalkylation of DNA sequences containing 5′-TAA^{2}TTA^1-3′ (A^1, first alkylation site on a DNA strand; A^2, second site on that strand) detected same-strand alkylation of first A^1 and then A^2 instead of the anticipated Bizelesin-like symmetrical A^1 alkylation of both strands, but no significant rearrangement of the B-form DNA structure (Figure 2B and ASYM in Table 1). Given that Bizelesin rearranges the six-base-pair sequence prior to cross-linkage, does the CPI-I failure to produce the symmetrical pattern result from an inability to rearrange the duplex in a Bizelesin-like manner?

The possible roles of alkylation-induced bending distortion and other sequence dependent conformational properties in determining the bisalkylation pattern were investigated by comparative time-course gel electrophoresis of CPI-I bisalkylation of 5′-TAA^{2}TTA^1-3′ and the contrasting 5′-TAA^{2}AAA^1-3′ sequence. These and other results suggest that the principal consequence of distortion is introduction of an anomalous 5′-end TA-step junction site. Molecular dynamics (MD) studies of CPI-I mono- and bisadducts of these sequences were used to examine these duplex TA-step distortion phenomena. These unexpected results and the underlying analysis of why this

(5) Abbreviations: CPI, cyclopropa[*c*]pyrrolo[3,2-*e*]indol-4(5*H*)-one; CPI-I, 2-(Indol-2-ylcarbonyl)-1,2,8,8a-tetrahydro-7-methylcyclopropa[*c*]-pyrrolo[3,2-*e*]indol-4(5*H*)-one (U72779); 5- and 8-CPI-I, CPI-I covalently attached at either 5A or 8A, respectively, on a 10-mer sequence; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; DQF-COSY, double-quantum-filtered correlated spectroscopy; PE-COSY, primitive exclusive correlated spectroscopy; ppm, parts per million; FID, free-induction decay; PSEUROT, computer-assisted pseudorotational analysis of five-membered rings by means of vicinal proton spin−spin coupling constants; DEAE, (diethylamino)ethyl; DMF, dimethylformamide; rMD, restrained molecular dynamics; PAGE, polyacrylamide gel electrophoresis; 1D, one-dimensional; 2D, two-dimensional; RMSD, root mean square difference.

(6) Sun, D.; Hurley, L. H. *J. Am. Chem. Soc.* **1993**, *115*, 5925−5933.
(7) Seaman, F. C.; Hurley, L. H. *Biochemistry* **1993**, *32*, 12577−12585.

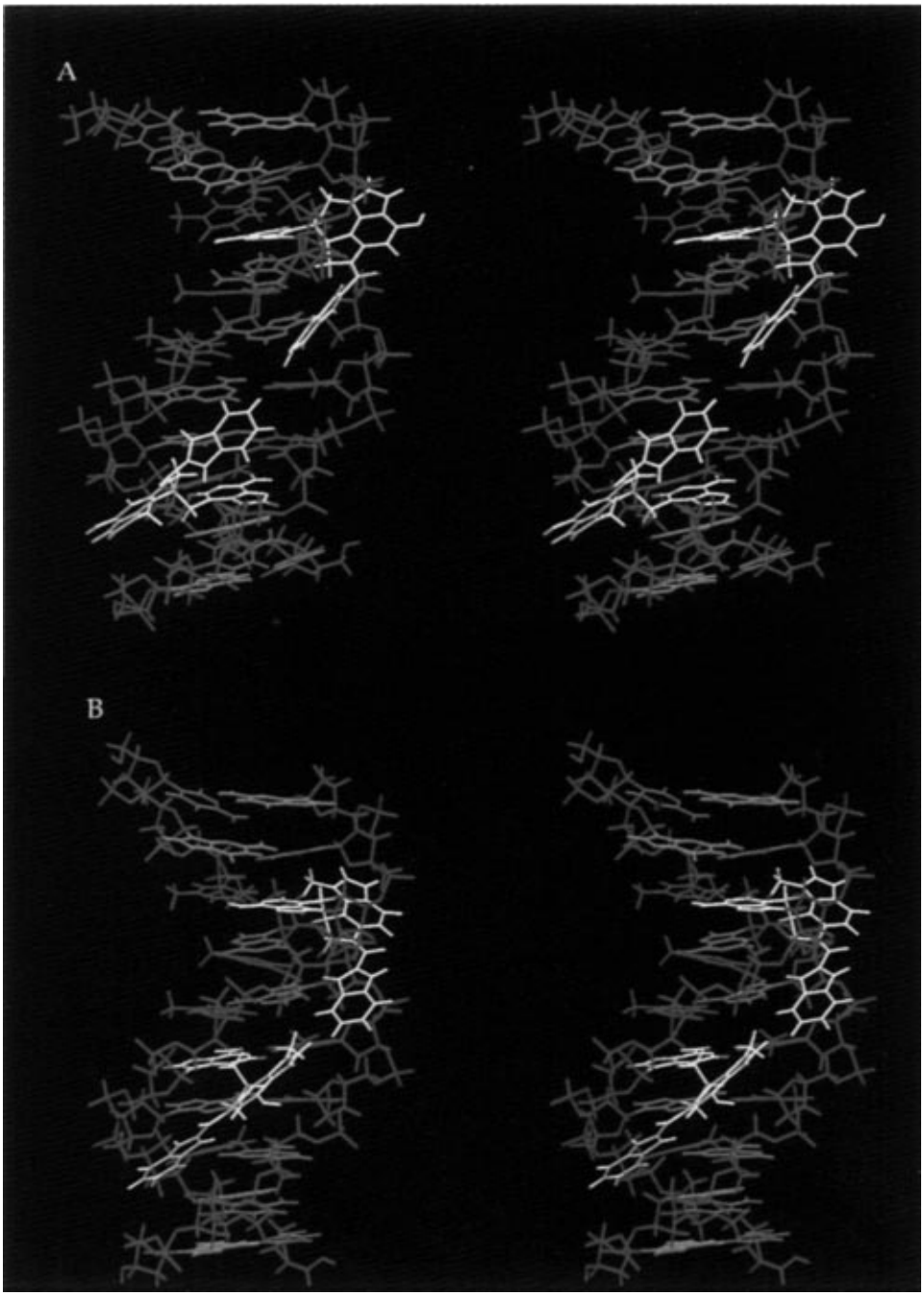

Figure 2. Stereoviews of preliminary modeled bisadduct structures derived from energy minimization of canonical B-form AT-10 with two docked covalent CPI-I moieties: (A) SYMM bisadduct showing the predicted symmetrical 10-mer bisadduct at 8A (strand one) and 18A (strand two); (B) ASYM bisadduct showing the observed asymmetrical 10-mer same-strand bisadduct at 5A and 8A.

pattern of bisalkylation occurs have important implications for the design of both intra- and interstrand DNA−DNA cross-linkers.

Results and Discussion

Time-Course Gel Electrophoresis Analysis of CPI-I Mono- and Bisalkylation. A 48 h PAGE time-course study was conducted of CPI-I reactivity with a 5′-end-^{32}P-labeled 21-mer containing the 5′-TAA^{2}TTA^1-3′ sequence (AT-21, Table 1). Heat-induced strand-breakage monitoring of the reaction at various intervals up to 2 days demonstrates rapid A^1 monoadduct formation followed by slow second alkylation at A^2 three base pairs upstream from the first-targeted adenine (Figure 3A). Densitometric analysis of the gel (Figure 3B) shows that the earliest samples of 5′-end-labeled strand-breakage products contain appreciable A^1 monoadduct (60−65%). After 48 h, the monoadduct band intensity (concentration) diminishes due to the increase in thermal cleavage at the bisadduct's second alkylation site, A^2, and the ratio of monoadduct to bisadduct approaches unity.

Why is only the AT-step adenine targeted for the second drug attack? Two-dimensional NMR results for the Bizelesin cross-link indicate that Bizelesin's symmetrical alkylation only occurs after the duplex B-form structure has been rearranged in the AT-step region.[7] Such a rearrangement presumably dissipates the bending distortion that accompanies Bizelesin monoadduct intermediate formation. These results suggest that monoadduct duplex distortion (or more precisely, the monoadduct's failure to dissipate distortion) within the 5′-TAATTA-3′ sequence plays a role in determining the course of CPI-I bisalkylation. In order to evaluate the role of distortion or structural rearrangement, a

Table 1. DNA Sequences Used in the Gel Electrophoresis and NMR Studies[a]

AT-STEP 21-MER (AT-21)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21			
T	G	G	C	C	G	C	A	C	G	T	A	A	T	T	A	C	G	G	G	T			
			G	G	C	G	T	G	C	A	T	T	A	A	T	G	C	C	C	A	A	A	C
			42	41	40	39	38	37	36	35	34	33	32	31	30	29	28	27	26	25	24	23	22

A-TRACT 21-MER (AA-21)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21			
T	G	G	C	C	G	C	A	C	G	T	A	A	A	A	A	C	G	G	G	T			
			G	G	C	G	T	G	C	A	T	T	T	T	T	G	C	C	C	A	A	A	C
			42	41	40	39	38	37	36	35	34	33	32	31	30	29	28	27	26	25	24	23	22

SYMMETRICAL AT-STEP 10-MER (SYMM; AT-10)

1	2	3	4	5	6	7	8	9	10
C	G	T	A	A	T	T	A	C	G
G	C	A	T	T	A	A	T	G	C
20	19	18	17	16	15	14	13	12	11

ASYMMETRICAL AT-STEP 10-MER (ASYM; AT-10)

1	2	3	4	5	6	7	8	9	10
C	G	T	A	A	T	T	A	C	G
G	C	A	T	T	A	A	T	G	C
20	19	18	17	16	15	14	13	12	11

[a] Double lines indicate the attachment site of CPI-I at a 10-mer adenine. The single line indicates the location of the ureadiyl linker that links two CPI-I units together to form the cross-linking drug Bizelesin.

Table 2. Chemical Shifts of Nonexchangeable and Exchangeable Protons of (A) the Unmodified 10-mer Duplex (AT-10 in Table 1) and (B) the Differences of These Chemical Shifts and Those of the CPI-I Bisadduct[a]

	H1′		H2′		H2″		H3′		H4′		H5′		H5″		PuH8/PyH6		TMe/CH5/AH2		PuH1/PyH3	
base	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B	A	B
C	5.73	−0.10	1.97	−0.19	2.39	−0.08	4.66	0.01	4.02	0.00	3.69	0.00	3.68	0.01	7.60	−0.04	5.85	−0.01	−	−
		−0.05		−0.05		−0.06		0.03		0.04		−0.06		−0.05		−0.01		−0.01		
G	5.91	0.02	2.62	−0.03	2.71	0.00	4.93	0.05	4.32	0.08	4.06	0.04	3.96	0.06	7.95	−0.02	−	−	12.61	−0.20
		−0.09		−0.02		−0.05		0.03		−0.01		0.02		−0.04		−0.05				−0.20
T	5.56	*0.28*[c]	2.04	0.22	2.37	0.21	4.84	0.10	4.13	0.24	4.14	0.09	4.08	0.10	7.22	0.10	1.45	0.13	13.48	0.03
		−0.24		*0.26*		0.14		0.09		0.15		0.04		0.06		0.19		0.18		NO[b]
A	5.96	*−0.53*	2.73	*−0.78*	2.91	0.22	5.02	−0.34	4.04	−0.71	4.13	0.00	4.04	−0.03	8.23	−0.06	6.80	−0.07	−	−
		0.15		−0.07		−0.04		0.02		0.36		0.14		0.18		0.12		0.24		
A	6.10	−0.17	2.55	0.01	2.89	−0.11	4.96	−0.22	4.44	−0.19	4.29	*−0.41*	4.22	*−1.53*	8.12	*0.35*	7.49	*0.84*	−	−
		−0.17		*−0.34*		*−0.32*		*−0.37*		*−2.61*		*−1.12*		*−1.63*		−0.23		*0.63*		
T	5.82	−0.11	1.91	−0.09	2.48	−0.16	4.77	−0.19	4.14	*−2.09*	4.33	*−0.57*	4.12	*−0.85*	7.10	0.12	1.23	0.19	13.29	*0.50*
		−0.67		*−0.49*		*−0.35*		*−0.34*		*−2.16*		*−0.46*		*−0.40*		−0.22		0.01		*2.31*
T	5.77	*−0.58*	2.18	*−0.73*	2.46	*−0.93*	4.86	*−0.39*	4.12	*−1.31*	4.12	−0.19	4.05	−0.16	7.32	−0.02	1.56	−0.04	13.33	*0.79*
		0.05		0.00		0.01		−0.05		0.04		−0.20		*−.72*		−0.04		−0.04		*0.39*
A	6.13	*−0.33*	2.63	−0.05	2.81	0.07	4.99	−0.09	4.39	−0.05	4.14	*−0.33*	4.10	*−1.70*	8.27	*0.26*	7.45	*0.67*	−	−
		−0.14		*−0.34*		−0.23		*−0.39*		*2.26*		*−0.92*		*−1.38*		−0.19		*0.34*		
C	5.58	0.04	1.81	−0.17	2.23	−0.14	4.73	−0.13	4.11	*−1.93*	4.22	*−0.30*	4.11	*−0.52*	7.25	0.02	5.27	*0.35*	−	−
		−0.72		*−0.63*		*−0.71*		*−0.32*		*−1.80*		*−0.29*		*−0.46*		*−0.34*		*−0.26*		
G	6.08	−0.02	2.54	−0.24	2.32	*0.27*	4.61	0.08	4.13	0.10	4.04	0.09	4.02	−0.20	7.86	−0.01	−	−	12.82	−
		0.00		*−0.26*		0.22		0.02		0.04		−0.08		*−0.40*		−0.01				

[a] The upper value in column B is the strand one proton chemical shift difference, and the lower value is the strand two proton chemical shift difference. [b] NO indicates that the signal is not observed. [c] Chemical shift differences greater than 0.25 ppm are italic.

two-dimensional NMR analysis (described below) was conducted on the CPI-I bisadduct of the same 10-mer sequence used in the Bizelesin study.

Proton NMR Assignments for the 10-mer and Its Bisadduct. A. Unmodified 10-mer Duplex. The single set of proton chemical shift assignments (Table 2) for the self-complementary DNA duplex (AT-10 in Table 1) indicates a symmetrical structure. Two unusual features of the NOESY spectrum (unpublished results) indicate a conformationally anomalous AT-step region for the unmodified duplex: First, unusually intense cross-peaks link the AT-step adenine (5A) H2 proton with the 3′-side thymine (6T) H1′. Second, 6TH5′ is downfield shifted beyond any of the other deoxyribose H5′ and H5″ proton signals (Table 2). These NOESY and additional DQF-COSY, PE-COSY, and partially decoupled *J*-scaled DQF-COSY cross-peak data for the 10-mer internal 5′-AATT-3′ region resemble those of the Dickerson dodecamer[8,9] and likewise suggest unusual AT-step flexibility. Further, base-pair opening rates for the 10-mer 5′-TAATTA-3′ region (5A/6T, 35 ms; 4A/7T, 6 ms; unpublished results) suggest that it has a high frequency of opening comparable to other reported opening rates for similar sequences[10,11] and radically different from A-tract sequences of comparable length that have low frequencies of opening.

B. Bis-CPI-I−d(CGTAATTACG)$_2$ Duplex Adduct. Assignment of the Duplex Proton NMR Signals in the Bis-CPI-I−d(CGTAATTACG)$_2$ Duplex Adduct. For each duplex strand,[12] the NOESY (200 ms) spectrum displays weak to moderate cross-peaks linking all nonterminal neighboring base PyH6 and PuH8 protons (supporting information, item 1) except for the undetectable 4A H8 × 5A H8 cross-peak. Base H6/H8 NOESY cross-peaks with H1′ of the attached sugar and H1′ of

(8) Nerdal, W.; Hare, D. R.; Reid, B. R. *Biochemistry* **1989**, *28*, 10008−10021.

(9) Lane, A. N.; Jenkins, T. C.; Brown, T.; Neidle, S. *Biochemistry* **1991**, *30*, 1372−1385.

(10) Moe, J. G.; Russu, I. M. *Nucleic Acids Res.* **1990**, *18*, 821−827.

(11) Leroy, J.-L.; Charretier, E.; Kochoyan, M.; Guéron, M. *Biochemistry* **1988**, *27*, 8894−8898.

(12) The multiplicity of NMR signals is dealt with by assigning different number systems to the two independent 5′-CGTAATTACG-3′ sets of signals, 1C−10G and 11C−20G, corresponding to strand one (doubly covalently modified) and strand two, respectively. Names of drug ^{1}H signals are preceded by 5 or 8 (e.g., 8H3′) depending on the drug's point of attachment, either 5A or 8A.

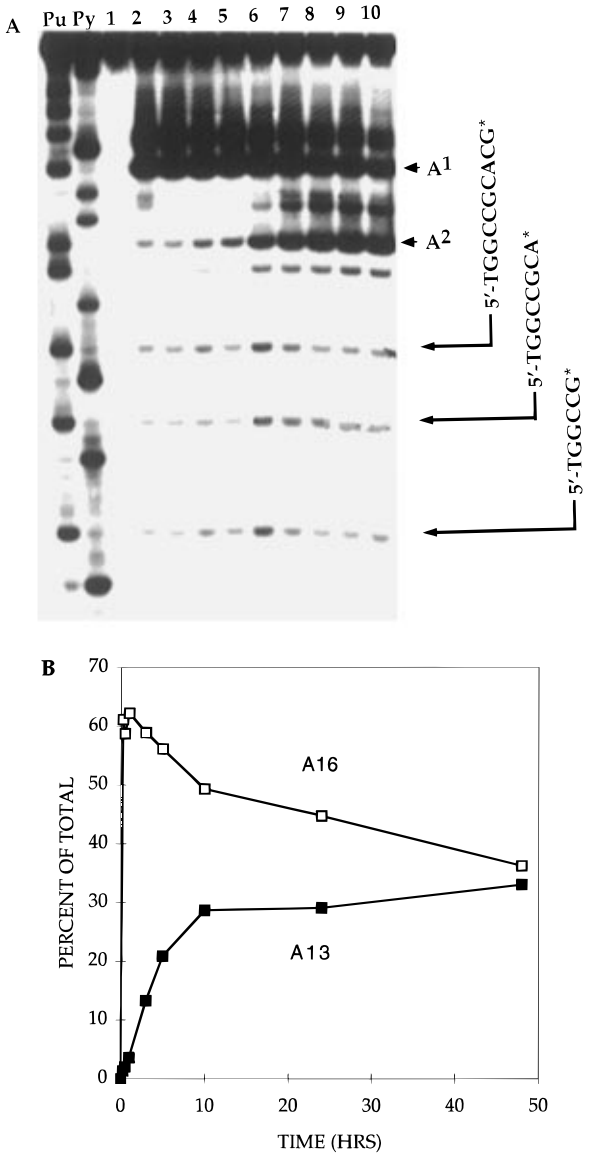

Figure 3. Time-course gel electrophoresis of CPI-I bisalkylation of AT-21. After incubating the 5′-end-labeled DNA (0.40 pmol) with CPI-I (840 pmol) at room temperature for periods of time specified below, 10 μL of calf thymus DNA (1 mg/mL) was added to the reaction mixture (20 μL) and subjected to thermal cleavage (90 °C, 45 min).[32] (A) Autoradiogram of cleavage products: Pu and Py refer to Maxam−Gilbert purine- and pyrimidine-specific strand-break reactions, respectively. Lane 1 is the control, and lanes 2−10 display thermally cleaved strand fragments resulting from different incubation times: 5 min, 15 min, 30 min, 1 h, 3 h, 5 h, 10 h, 1 day, and 2 days, respectively. (B) Densitometric analysis of the autoradiogram shown in (A). The AT-21 bisadduct results are plotted as percentages (vertical axis) of total adduct-related cleavage bands. The A16 curve indicates the band corresponding to cleavage adjacent to the 3′-end adenine (A^1). The A13 curve represents the percentage of total adduct-related cleavage bands consisting of the band corresponding to cleavage adjacent to the AT-step adenine (A^2).

the 5′-side sugar permit the assignment of H1′ signals for strands one and two (supporting information, items 2A and 2B). NOESY cross-peaks link the H6/H8 signals to H2″ and H2′ signals of the attached sugar (intranucleotide) and the sugar on the 5′-side nucleotide (internucleotide) (supporting information, item 3). Other cross-peaks link base H8/H6 signals and the H3′ of the attached sugar and, in most instances, the H3′ of the 5′-side sugar (supporting information, item 4).

NOESY H1′ and H3′ cross-peaks with neighboring deoxyribose H4′, H5′, and H5″ permit the expansion of the cross-connectivity network to include all deoxyribose substituents, and these assignments are confirmed by H4′ × H5′, H4′ × H5″, and H5′ × H5″ NOEs (supporting information, items 5A and 5B; Table 2). Because of difficulty in differentiating H5″ and H5′, the downfield member of the pair is arbitrarily assigned as H5′. Most H1′ × H4′ intraresidue cross-peaks are moderately intense except for undetectable 8A and 15A cross-peaks and very weak 5A, 6T, and 18A cross-peaks. Absent or weak H1′ × H4′ cross-peaks indicate an unusual sugar conformation.[13] The H4′ signals cluster in two chemical shift groups, a tight cluster around 4.5 ppm similar to the unmodified duplex H4′ chemical shifts and a looser grouping around 2.1 ppm. This latter group generates cross-peaks to similarly upfield-shifted neighboring H5′ and H5″ signals, indicating proximity to the CPI-I ring current effects in the minor groove.

Adenine H2 signals are assigned on the basis of their NOESY intra- and interstrand cross-peaks with (1) neighboring adenine H2 substituents, (2) deoxyribose H1′ signals of the attached sugar and the neighboring nucleotide's sugars, and (3) exchangeable thymine H3 signals (Figure 4). (1) Cross-peak intensities indicate that 4A H2 is situated in the minor groove roughly equidistant between the 3′-side 5A H2 (intrastrand cross-peak 5, Figure 4) and 18A H2 (interstrand cross-peak 2, Figure 4), and 14A H2 is positioned analogously with 15A H2 (intrastrand cross-peak 13, Figure 4) and 8A H2 (interstrand cross-peak not shown, Figure 4; supporting information, item 1). (2) Interstrand 8A H2 × 14A H1′, 18A H2 × 4A H1′, and 5A H2 × 17T H1′ and intrastrand 5A H2 × 6T H1′ and 15A H2 × 16T H1′ cross-peaks (cross-peaks 15, 1, 10, 7, and 12, respectively, in Figure 4; supporting information, item 2) reinforce adenine H2 assignments. (3) Irradiation of 7T and 17T H3 imino protons yields a strong NOE to the H2 of the paired adenines 14A and 4A, respectively (cross-peaks 14 and 6, respectively, Figure 4) and a weaker NOE to the 3′-side adenines H2, 8A and 18A, respectively (supporting information, item 6).

Assignment of CPI-I ^{1}H NMR Signals in the Bis-CPI-I Duplex Adduct. Most CPI-I H atoms are localized along the inner edge of the drug (CPI: H2, H3Me, H4a, H4A, H4B, H5A, and H5B; indole: H3′, H5′, and H6′) and generate numerous drug-to-DNA NOESY cross-peaks (Figure 5). Indole proton assignments are confirmed by DQF-COSY cross-peaks for ortho-related aromatic protons H5′, H6′, H7′, and H8′ and 1D-NOEs (H_2O) between the exchangeable indole subunit H1′ and its neighboring H8′. 5- and 8-CPI-I N*H*1 and N*H*1′ exchangeable signals are assigned on the basis of NOEs with the CPI subunit's H2 and indole subunit's H8′, respectively (Table 3). Neither the outer-edge C8 phenolic proton signal nor any of its cross-peaks were observed, presumably due to the rapid exchange of this proton.

DNA and CPI-I ^{1}H NOESY Cross-Peaks Identify Minor Groove Positions of the Two Drug Ligands. The 5-CPI-I methyl signal (2.88 ppm, Figure 6) yields the following duplex cross-peaks (downfield to upfield): 5A H2 , 5A H1′, 17T H1′, and 6T H1′ (Figures 5A and 6), and the 8-CPI-I methyl signal (2.59 ppm, Figure 6A,B) produces corresponding cross-peaks: 8A H2, 14A H2, 14A H1′, and 9C H1′ (Figures 5B and 6A,B). The unusually intense 5-CPI-I Me cross-peak to 17T H1′ and 17T H4′ and correspondingly very weak cross-peak to 6T H1′ (6T H4′ cross-peak unobserved) indicate that this portion of the CPI subunit is displaced toward the noncovalently modified strand two. This displacement is further supported by the observation that the only three 5-CPI-I H2-to-duplex cross-peaks

(13) Kim, S.-G.; Lin, L.-J.; Reid, B. R. *Biochemistry* **1992**, *31*, 3564−3574.

Table 3. CPI-I Drug Chemical Shifts for the 5A- and 8A-Attached Drugs of the ASYM BisAdduct

	chemical shift (ppm)														
	H2	H3Me	NH1	H4a	H4A	H4B	H5A	H5B	H7	H3′	H5′	H6′	H7′	H8′	NH1′
5A	7.29	2.88	10.28	4.43	4.62	5.01	4.44	5.30	7.84	7.22	7.45	8.07	7.60	11.33	8.64
8A	7.29	2.59	10.28	4.60	4.59	5.10	4.59	5.23	7.84	7.75	8.28	8.02	7.56	11.14	7.66

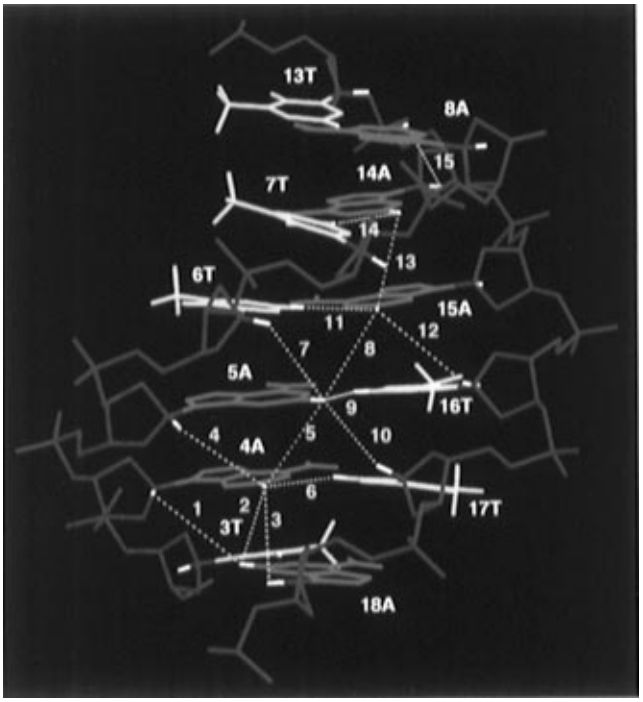

Figure 4. Intra- and interstrand adenine H2, thymine H3, and deoxyribose H1′ NOESY connectivity network for the 5′-TAATTA-3′ portion of the 10-mer CPI-I bisadduct. Strands one and two are numbered 1−10 and 11−20, respectively. Colors are green (adenine), yellow (thymine), and magenta (backbone). Cross-peaks are (1) 18A H2 × 4A H1′, (2) 4A H2 × 18A H2, (3) 4A H2 × 18A H1′, (4) 4A H2 × 5A H1′, (5) 4A H2 × 5A H2, (6) 4A H2 × 17T H3, (7) 5A H2 × 6T H1′, (8) 5A H2 × 15A H2, (9) 5A H2 × 16T H3, (10) 5A H2 × 17T H1′, (11) 15A H2 × 6T H3, (12) 15A H2 × 16T H1′, (13) 14A H2 × 15A H2, (14) 14A H2 × 7T H3, and (15) 8A H2 × 14A H1′.

are with 17T H4′, 17T H5′, and 17T H5″. As predicted by analogy with previously reported Bizelesin adducts,[7] the CPI subunit "linker" methylene 5H4A shows a more intense cross-peak into 5A H2 than 5H4B, while 5H4B yields the more intense cross-peak of the two with 5A H1′ (Figure 5A). Although overlap obscures the relative intensities of cross-peaks of 8A H2 with 8H4A and 8H4B, there is an intense cross-peak linking 8A H1′ and 8H4B (Figure 5B). On the basis of NOESY data, the nearest duplex substituents to the other CPI subunit methylene substituents, 8H5A and H5B and 5H5A and H5B, are 14A H2 (Figure 5B) and 4A H2 (Figure 5A), respectively.

Intense cross-peaks of the 5-CPI-I indole subunit H3′ with 18A H2 and 4A H2 and 8-CPI-I H3′ with 15A H2 and 14A H2 (Figure 5) indicate the indole H3′ substituent positions that are roughly equidistant between these neighboring pairs of adenine H2's. NOESY cross-peaks also confirm two close drug-to-DNA contacts for indole H5′: 18A H2 × 5H5′ and 15A H2 × 8H5′ (parts A and B, respectively, of Figure 5).

CPI-I Alkylates 8A and Then 5A on the Same Duplex Strand. The sequence of the bisalkylation described earlier for the gel electrophoresis analysis of the time course of the reaction was verified by time-course NMR analysis of reaction mixtures. Comparison of thymine methyl signal chemical shift patterns for the unreacted duplex, unreacted duplex and monoadduct mixture, and final bisadduct product (Figure 7) demonstrates that a single 8A monoadduct is produced rapidly and is slowly replaced by the bisadduct. Monoalkylation at 8A rather than 5A is demonstrated by the relative influence of monoalkylation on the chemical shifts of 6T and 16T H6 and methyl chemical shifts. As predicted from model structures, 6T H6 and Me signals are upfield-shifted 0.40 and 0.15 ppm, respectively, in the monoadduct due to steric influence on the strand one 6T base of the 8-CPI-I indole subunit, which drives the 6T methyl and H6 nearer the ring current effects of the 5A base ring system. In contrast, the monoadduct strand two 16T H6 and Me signals experience only 0.05 and 0.04 ppm shifts, respectively. In the bisadduct, alkylation of strand one 5A shifts the neighboring 6T methyl signal 0.34 ppm downfield from its monoadduct position due to the distorted bisadduct 5A base orientation and resulting removal of this base's inductive influence from the 6T substituents. In the bisadduct, the strand two 16T methyl chemical shift remains virtually unchanged.

Drug-to-Drug NOE Cross-Peaks Indicate a Head-to-Tail Orientation of the Two CPI-I Molecules. As predicted from the ASYM model's close proximity of the 5-CPI-I methyl moiety and the 8-CPI-I indole subunit, the relative intensities of the observed 5-CPI-I methyl NOESY cross-peaks with indole substituents follow the sequence 8H6′ > 8H5′ > 8H7′ (Figure 6B,C).

Inductive Effect of CPI-I Indole Substituents on DNA Minor Groove Proton Chemical Shifts. Parallel H5″ and H5′ chemical shift variations (Figure 8) identify regions of 5- and 8-CPI-I-correlated drug shielding influence within the minor groove. The shielding effect is limited to 5A and 8A (strand one) and 15A and 18A H5′ and H5″ (strand two). At the 5A and 8A covalent attachment sites, H5′ and H5″ experience a ring inductive effect on the strand one side of the indole's six-membered ring, and 15A and 18A H5′ and H5″ experience the inductive effect on the strand two side of the CPI six-membered ring. Different strand one (5A and 8A) and strand two (15A and 18A) H5′ chemical shifts suggest variant backbone C5′ orientations for the two strands relative to the CPI and indole subunits. The orientations of the H5′ and H5″ substituents in the stereomodels (Figure 5) correspond to the chemical shift differences: Strand two 15A and 18A H5′ and H5″ are both oriented toward the CPI aromatic system in the minor groove. By contrast, only strand one 5A H5″ and 8A H5″ are directed at the indole subunits in the minor groove, while 5A H5′ and 8A H5′ are oriented away from the minor groove. Coincidentally, the arbitrarily assigned H5′ and H5″ (H5′ defined as the downfield member of the pair) agree with these orientations: The H5″ position nearer the minor groove suggests that it should be more affected by the drug's aromatic rings' inductive effects and, as a consequence, shifted relatively farther upfield. Other deoxyribose protons experience similar inductive effects: 4A, 7T, 16T, and 19C H1′ are upfield shifted (relative to the unmodified duplex) by 0.53, 0.58, 0.67, and 0.72 ppm, respectively, due to their minor groove positions opposite one face (4AH1′ and 7TH1′) or the other (16TH1′ and 19CH1′) of the two indole subunits.

Restrained Molecular Dynamics Analysis of the 5A,8A Bisadduct. NOESY-derived distance constraints were incorporated into two rMD analyses of the 5A,8A bisadduct. An *in vacuo* study compared the RMSD values for averaged products

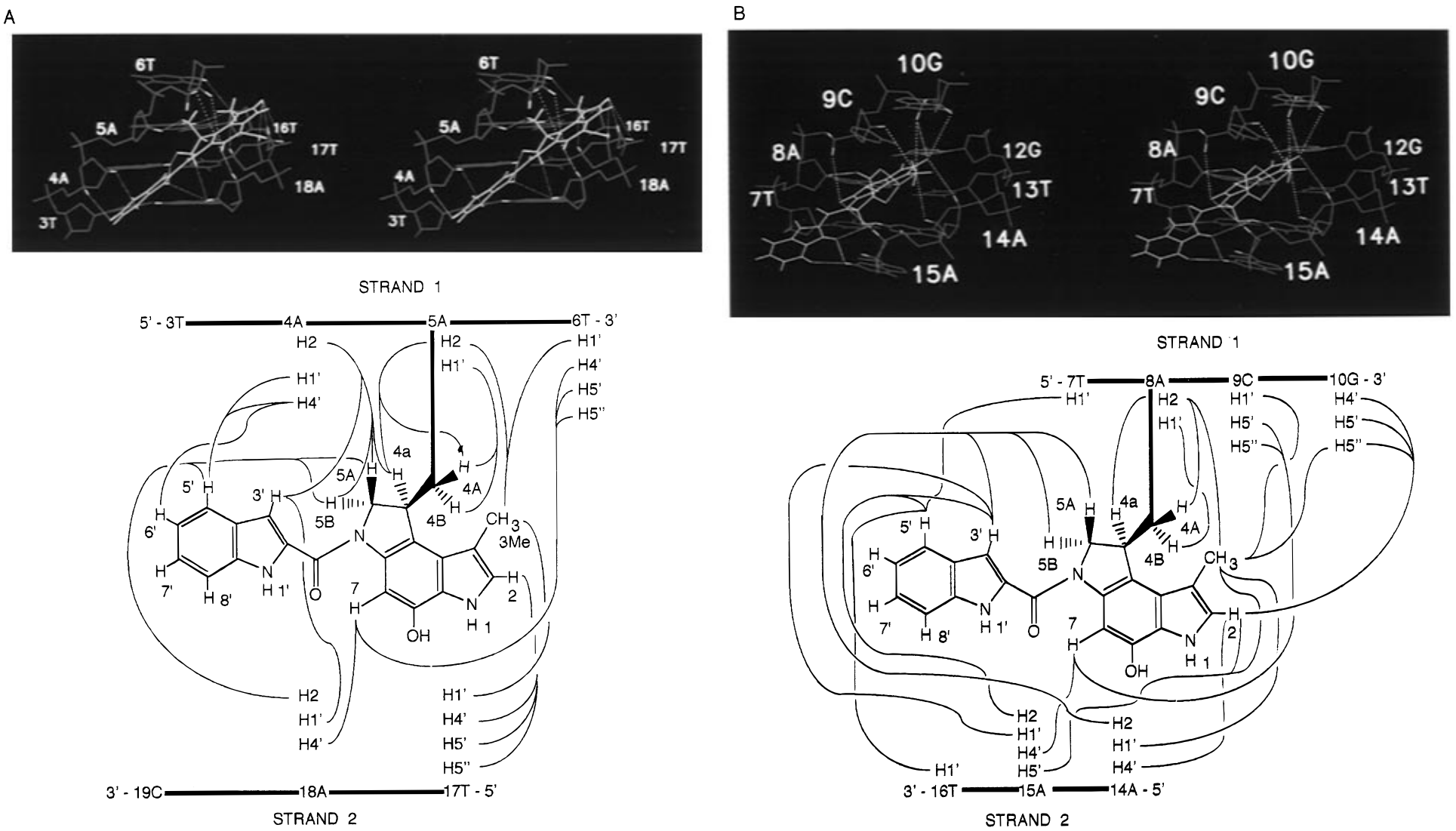

Figure 5. Stereoview and cross-connectivity diagram of the 10-mer bisadduct showing CPI-I cross-connectivity with duplex adenine H2 and deoxyribose H1′, H4′, H5′, and H5″ substituents for (A) the duplex region surrounding 5A-attached CPI-I and (B) the duplex region surrounding 8A-attached CPI-I. Colors are yellow (5-CPI-I and 8-CPI-I), green (adenine), red (thymine), magenta (deoxyribose and phosphate backbone), and white (selected hydrogens).

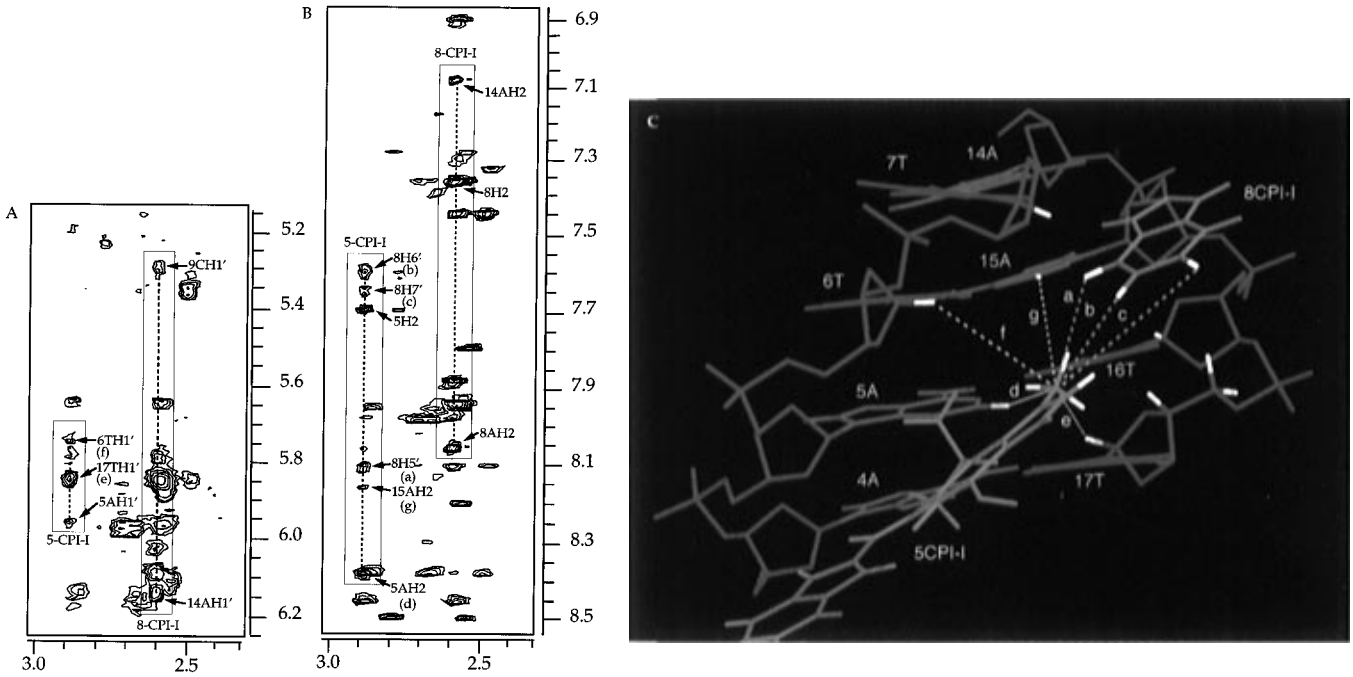

Figure 6. NOESY spectral expansions showing the 10-mer bisadduct 5-CPI-I and 8-CPI-I methyl function cross-peaks in (A) the "H1′ region" and (B) the "aromatic region". Cross-peaks labeled a−g refer to 5-CPI-I methyl cross-peaks depicted in (C). (C) 5-CPI-I methyl to DNA cross-peaks are shown as yellow, while Me to 8-CPI-I cross-peaks are shown as pink. Other colors are cyan (5- and 8-CPI-I), green (adenine), red (thymine), and magenta (backbone).

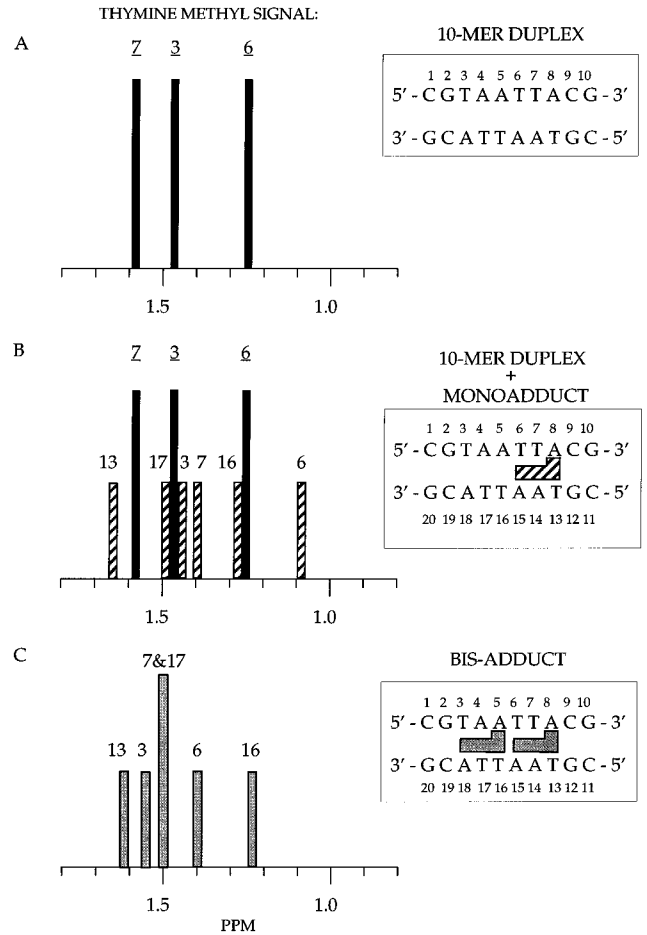

Figure 7. Diagrams showing the 1D NMR thymine methyl region of (A) the unmodified 10-mer duplex, (B) the unmodified 10-mer duplex and its 8-CPI-I monoadduct, and (C) the 10-mer 5- and 8-CPI-I bisadduct.

of rMD trajectories using both A- and B-form starting structures (Table 4). An rMD (solvated) study using a B-form starting structure was conducted in order to evaluate the conformational properties of the bisadduct.

Table 4. RMSD Calculations Derived from Comparison of (Top) the CPI-I Bisadduct of 14-mer (5′-CGCGTAATTACGCG-3′) Starting Structures and the Average of the Last 15 ps (300 K) of 100 ps rMD Trajectories from Four A-Form and B-Form Starting Structures and (Bottom) Pairwise RMSDs for the Four Average Structures Resulting from 100 ps rMD Trajectories (AMBER 4.0)[31]

	SS origin[b] (ps)		RMSD (Å)			
SS[a]	A	B	A and B SS	A and SS	B and SS	**A and B**
1	0.0	0.0	5.41	5.21	1.38	**1.09**
2	0.8	0.8	5.11	4.94	1.24	**0.87**
3	2.0	1.2	4.86	4.54	1.25	**1.30**
4	1.6	1.6	4.85	4.70	1.54	**1.07**
			av: 5.09	4.78	1.35	**1.08**

A-form starting structures					B-form starting structures				
average product	average product (Å)				average product	average product (Å)			
	A1	A2	A3	A4		B1	B2	B3	B4
A1	–	–	–	–	B1	–	–	–	–
A2	0.95	–	–	–	B2	0.81	–	–	–
A3	1.18	1.26	–	–	B3	0.96	1.14	–	–
A4	1.16	0.59	1.43	–	B4	0.66	0.82	1.06	–
av: 1.10					av: 0.91				

[a] SS = starting structure, A = A-form starting structure, B = B-form 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-generated A- or B-form structure.

The RMSD analyses demonstrate a high degree of convergence on a B-form averaged conformation for all four pairs of A- and B-form starting structures (in boldface print, Table 4, top). The average RMSD value for four separate pairs of A- and B-form starting structures was 1.08 Å. Comparison of average RMSDs for the four A-form starting structures and their rMD products (column "A and SS", Table 4, top) and the four B-form starting structures and their rMD products (column "B and SS", Table 4, top) (A-form, 4.78 Å; B-form, 1.35 Å, Table

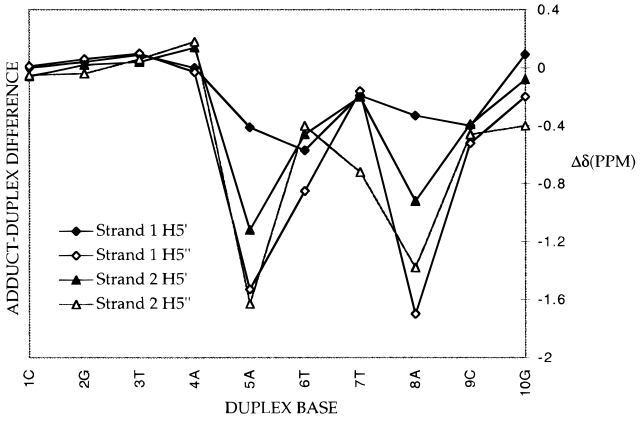

Figure 8. Comparison of the AT-10 unmodified duplex and bisadduct deoxyribose ^{1}H chemical shifts for strand one and strand two H5″ and H5′. The chemical shift difference (vertical axis) is plotted against the duplex sequence position (horizontal axis). CPI adduct H5″ assignments were made arbitrarily to the most downfield member of the H5′ and H5″ pair. In the region of drug proximity, these arbitrary assignments corresponded to the NOE-based discrimination of H5′ and H5″.

4) demonstrate the consistently B-form nature of the rMD products. Pairwise RMSD comparisons of all averaged products of A-form starting structures (average 1.10 Å, Table 4, bottom) and an identical comparison for all B-form starting structures (average 0.91 Å, Table 4, bottom) also display a high degree of convergence. Two structures corresponding to the final 15 ps averaged rMD products (100 ps, *in vacuo*) derived from A- and B-form starting structures (Figure 9) show displacement of 5-CPI-I toward the noncovalently modified strand, suggesting that interdrug steric clash is avoided by positioning the CPI subunit of 5-CPI-I to the side of the 8-CPI-I indole subunit. rMD (solvated) results reflected this same conformational feature and displayed no other significant departure from a B-form conformation.

Why Does CPI-I Bisalkylate the Same Strand Rather Than Form the Symmetrical Product? Bisadduct formation does not rearrange the B-form DNA duplex. Unlike the Bizelesin cross-link, the CPI-I bisadduct of 5′-TAATTA-3′ displays no central base-pair opening, Hoogsteen base pairing, or any other gross conformational response to drug attachment. The only observable change is a modest degree of conformational adjustment by the covalently modified adenines (and their attached ligands), which appears to diminish head-to-tail steric clash of the two drugs. If a Bizelesin-like rearrangement is necessary to symmetrically alkylate the equivalent strand one and two 3′-end adenines, the absence of such a rearrangement suggests that the CPI-I−DNA monoadduct retains the bending distortion that accompanies monoadduct formation. A consequence of the retention of this distortion is that it can affect the reactivity of the remaining adenine alkylation targets. Such distortion in the form of 5′- and 3′-end junction sites occurs in the 5′-AAAAA^1-3′ A-tract sequence[14] and its (+)-CC-1065 A^1 monoadduct wherein a bending-induced junction site occupies the 5′-end of the A-tract four base pairs upstream from the alkylation site.[15] The CPI-I bisalkylation pattern could result from two possible (but not mutually exclusive) distortion-induced mechanisms: (1) an *enhancement effect* whereby an otherwise unattractive target adenine, 5′-TA*A*^{2}TTA-3′, is made more reactive by cooperative effects of bending-induced conformational change and the innate flexibility of the AT-step region and (2) a *retardation effect* whereby ostensibly reactive targets (e.g., strand two 5′-TAATT*A*^{1}-3′) are made less reactive by drug-induced distortion of the minor groove reaction site. The first option can be examined by comparing rates of reaction for two behaviorally different AT-rich regions, the Bizelesin preferred sequence, 5′-TAA^{2}TTA-3′, and a contrasting, but also Bizelesin-targeted sequence, 5′-TAA^{2}AAA-3′. The A-tract A^2 properties contrast sharply with those of the first sequence's A^2. In addition to universal A-tract properties (e.g., very slow base-pair opening rates), it has been shown that Bizelesin A-tract cross-linkage has little effect on the center of the cross-linked region,[15] unlike the profound effect of Bizelesin on the first sequence's AT-step conformation. We reasoned that if the monoadduct's AT-step conformation is modified in a way that elevates reactivity of the AT-step above competing sites, the corresponding monoadduct A-tract position three base pairs upstream from the first covalently modified adenine would not likely experience the same enhancing effect due to the radically different conformational properties. This comparison was conducted as described in the following section by using time-course gel electrophoresis studies of the CPI-I reaction with these two sequences.

Comparative Gel Electrophoresis Analysis of Bisalkylation Shows Similar Patterns in the AT-Step and A-Tract Sequences. The 5′-TAA^{2}TTA^1-3′ AT-step behavior, which resembles that of the *Eco*RI endonuclease recognition sequence,[9] potentially enhances the monoadduct central adenine's reactivity with the second CPI-I. In order to investigate the role of the AT-step properties in producing the ASYM structure, the previously described time-course gel electrophoresis results (Figure 3B) for AT-21 (incorporating the Bizelesin-preferred sequence 5′-TAATTA-3′) were compared to those for AA-21 (Table 1, Figure 10), a sequence that contains the conformationally distinct but Bizelesin-reactive A-tract region 5′-TAA^2-AAA^1-3′. Although this latter sequence was readily cross-linked by Bizelesin,[16] its reactivity with CPI-I was unknown.

Gel electrophoresis time-course analysis of the A-tract sequence AA-21 (Table 1) indicates two competing initial alkylation sites resulting in two different alkylation patterns. The major pattern resembles the AT-21 pattern in that 5′-TAA^2-AAA^1-3′ alkylation occurs first at A^1 (A16) and then at A^2 (A13). Combined 5′- and 3′-end labeling experiments (Figure 10, supporting information, items 7A and 7B) show, in addition to this first bisalkylation pattern, a second pattern arising from a competing initial alkylation at the penultimate adenine, 5′-TAAA*A*^{1}A-3′ (A15). After 2 days, the appearance in the 5′-end-labeled DNA of a very faint band at 5′-T*A*^{2}AAA^{1}A-3′ (A12) suggests that in this second pattern only a trace of this A^1 monoadduct reacts to produce a bisadduct. After 24 h, AA-21 labeled at the 3′-end displays a higher concentration of strand-break products resulting from cleavage of 5′-TAA^{2}AAA^1-3′ at the A^1 site than from cleavage of 5′-TAAA*A*^{1}A-3′ at the A^1 site, indicating that 3′-end adenine alkylation is slightly favored over alkylation at the penultimate adenine. The NMR analysis of the bisadduct resulting from 3-week 10-mer incubation with CPI-I displayed no detectable signals for this product containing a covalently modified penultimate adenine. These NMR results suggest that the 3′-end penultimate adenine N3 alkylation is kinetically favored to a slightly lesser extent than the terminal adenine, but that the latter product is the more thermodynamically favored structure. Consequently, following the 3-week incubation, the only detectable product results from the terminal adenine alkylation.

Retardation of TA-Step Alkylation Rather Than Enhance-

(14) Koo, H. S.; Wu, H. M.; Crothers, D. M. *Nature* **1986**, *320*, 501−506.

(15) Thompson, A. S.; Sun, D.; Hurley, L. H. *J. Am. Chem. Soc.* **1995**, *117*, 2371.

(16) Thompson, A. S.; Hurley, L. H. *J. Mol. Biol.* **1995**, *252*, 86−101.

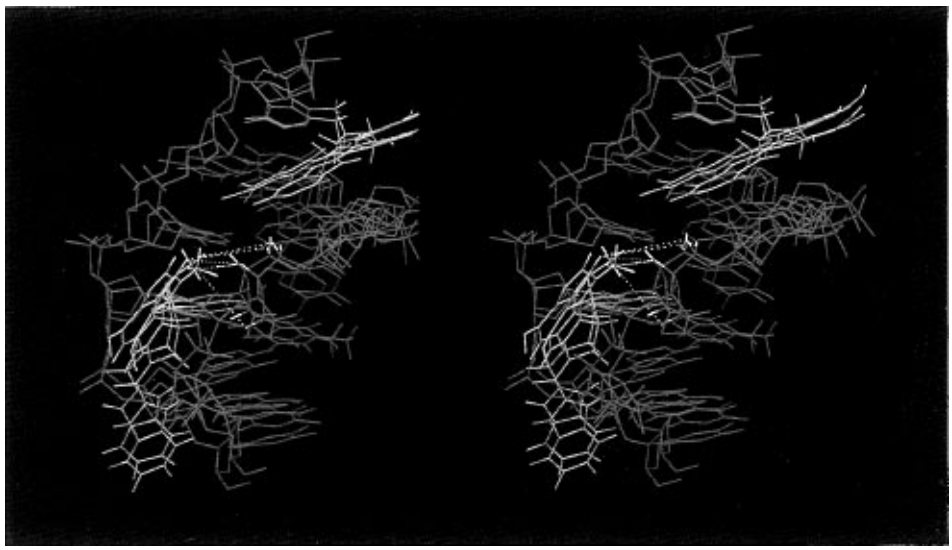

Figure 9. Stereoview showing final 20 ps averaged structures resulting from the *in vacuo* rMD (100 ps) analysis of the 10-mer 5A,8A CPI-I bisadduct. The two overlapping structures correspond to two averaged rMD products derived from A- and B-form starting structures. The B-form-derived structure has its 5-CPI-I methyl and H2 hydrogens and 17T H1′, H4′, H5′, and H5″ hydrogens colored white and connected by lines indicating NOESY connectivity. Other colors are green (adenine), red (thymine), cyan (CPI-I), and magenta (backbone).

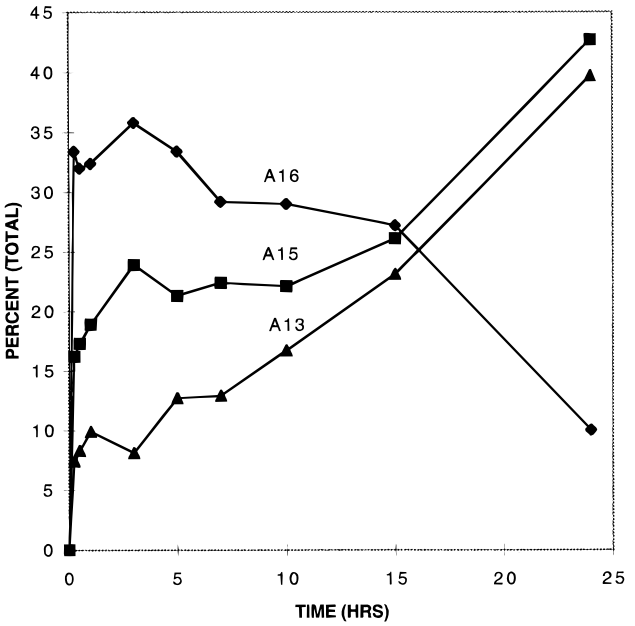

Figure 10. Time-course gel electrophoresis of CPI-I alkylation of AA-21: densitometric analysis of the autoradiogram produced after incubating the 5′-end-labeled DNA (0.40 pmol) with CPI-I (840 pmol) at room temperature for periods of time specified below. A 10 μL sample calf thymus DNA (1 mg/mL) was added to the reaction mixture (20 μL), and the mixture was subjected to thermal cleavage (90 °C, 45 min). The relative concentrations of AA-21 bisadduct thermal cleavage products (vertical axis) are plotted against sampling times (horizontal axis) of the time-course reaction to give three curves: A16 corresponding to cleavage adjacent to the 3′-end adenine, A15 corresponding to cleavage at the 3′-end penultimate adenine, and A13 corresponding to the third adenine from the 3′-end. (For numbering, see Table 1.)

ment of AT-Step Alkylation Appears To Direct the Second Alkylation Step. Comparison of time-course gel electrophoresis results for the A-tract AA-21 sequence to the previously described results for the AT-21 sequence reveals that the A-tract second alkylation not only occurs at the same relative position but, in fact, proceeds faster than the AT-step sequence second alkylation. Consequently, AT-step properties do not appear to enhance reactivity at this site. The results suggest that these two contrasting sequences share a second alkylation site that is the most downstream, nondistorted, targetable adenine not sterically blocked by the already attached drug. Further, the results suggest that position and not conformational enhancement of reactivity is critical to alkylation. Although the targeted second alkylation site may not be an enhanced reaction site, it is arguably the least affected site in terms of the negative impact of bending distortion. The previously mentioned 5′-AAAAA[1]-3′ (+)-CC-1065 monoadduct projects distortion toward a site between four and five base pairs upstream from the covalently modified adenine A[1]. The presence of a comparable 5′-end TA-step junction site in the 5′-TAATTA-3′ CPI-I monoadduct dictates that, of the available adenine second alkylation targets, the AT-step adenine is farthest from the junction site. As no direct measure of monoadduct DNA conformational distortion could be obtained from the bisadduct NMR spectra or from time-course gel electrophoresis, monoadduct duplex behavior was modeled using solvated molecular dynamics analysis (described below).

Molecular Dynamics Analyses of the AT-10 and AA-10 Monoadduct Display Distortion Focused on the 3T·18A Site. MD analysis of the AT-10 8A monoadduct (solvated) yields structures characterized by backbone distortion in the 3T·18A region. In three separate 100 ps analyses, base pairing is disrupted at either 3T·18A or at both 3T·18A and the adjacent G·C base pair (Figure 11). After approximately 60 ps, 18A (P-O5′) and (C5′-C4′) backbone torsional angles begin highly distortive rotations, followed by an increase in the 3T H3 to 18A N1 distance, culminating in base-pair opening (OPN, Figure 11). Opening of the 3T·18A base pair permits an unusual helical twist angle at the 3T·18A site (interbase-pair parameter: TWS, Figure 11) and an appreciable shear angle (intrabase-pair parameter: SHR, Figure 11) expressed by the 18A adenine projecting into the major groove. MD analysis (solvated) of the AA-10 (AA-10 is analogous to AT-10, but an A_5-tract, 5′-TAAAAA-3′, replaces the 5′-TAATTA-3′ sequence) 8A monoadduct produces a result similar to that of the AT-10 monoadduct (unpublished results). Base pairing failed in a similar fashion after approximately 50 ps.

In monoadduct models, the first-attached drug's "long-distance" influence on the reactivity of other target adenines takes the form of distortion in the 3T·18A region (11T−35A in 21-mers), reducing the reactivity of this site. In the MD

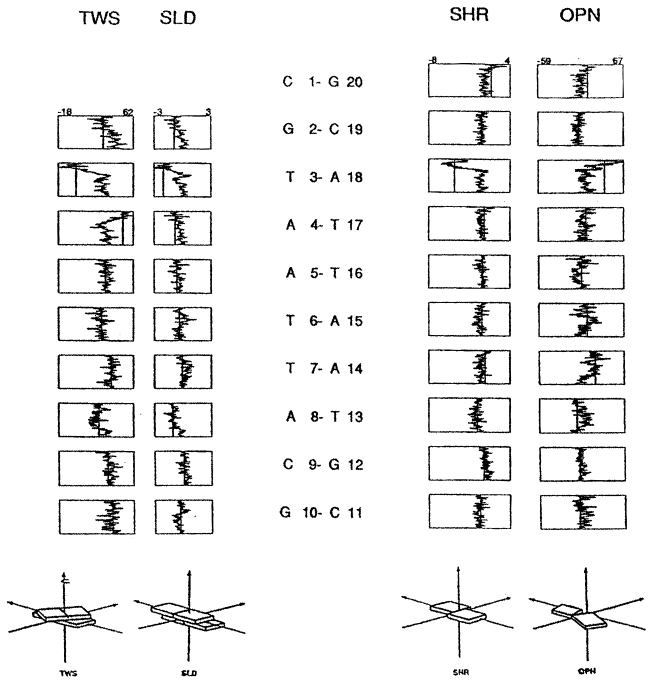

Figure 11. Analysis results (Dials and Windows)[32] from molecular dynamics (solvated, 100 ps) of the 10-mer 8A CPI-I monoadduct showing changes in selected helical and base pairing properties over 100 ps: TWS = twist, SLD = slide, SHR = shear, and OPN = open.

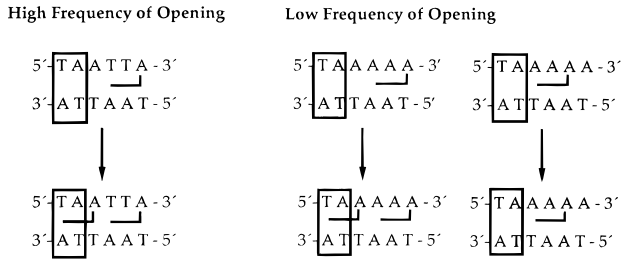

Figure 12. CPI-I monoadduct 5′-TA-step junction sites (within box) of the "high frequency of opening" duplex [5′-TAATTA-3′]$_2$ and the "low frequency of opening" duplex 5′-TAAAAA-3′/5′-TTTTTTA-3′, which are mutually exclusive of the second alkylation site. Targeted adenines outside the junction site that are not sterically blocked by the first attached CPI-I are shown.

simulations of monoadduct behavior, instability in the duplex region unoccupied by drug is released by backbone torsional angle adjustment coupled with 3T·18A base-pair opening in a fashion similar to that predicted by Ramstein and Lavery.[17] The resulting strand two 18A base displacement into the major groove of the modeled system would diminish this adenine's capacity to react with the drug in the minor groove. Consequently, rMD results suggest that the second step of the bisalkylation process will be qualitatively different from the first in terms of target adenine, reaction rate, etc.

This two-step CPI-I bisalkylation process consisting of qualitatively different steps is also reflected in the 5′-TAA^2-TTA^1-3′ (AT-21) and 5′-TAA^{2}AAA^1-3′ (AA-21) gel electrophoresis results (Figure 12). In the first step, rapid drug attachment at A^1 generates a distortional "wave" that is focused on a junction within the 5′-end TA-step. A similar junction site appears to arise during the less-common alkylation of the AA-21 5′-TAAA^{1}A-3′. Neither adenine bordering the junction site (Figure 12) is targeted by the second CPI-I in any of these monoadducts. Instead, for 5′-TAAAAA^1-3′, the one remaining nonsterically blocked adenine three base pairs to the 5′-side of the first alkylation site is targeted for the second alkylation step. For 5′-TAAAA^{1}A-3′, failure to alkylate at the TA junction site together with drug minor groove steric obstruction eliminates all possible second alkylation sites.

Reaction Reversibility and Alkylation Pattern. If CPI-I 5′-TAATTA^1-3′ monoadduct distortion is severe enough to prevent alkylation at the symmetrical strand two A^1 site, the question of the significance of the ASYM and SYMM bisadduct's reverse reaction is moot. However, if the SYMM bisadduct is produced and a reverse reaction occurs at either drug site, reattachment of CPI-I to the resulting 8A monoadduct can occur at several competing strand one and two adenines. In this case, the reverse reaction rates can affect the relative proportion of SYMM and ASYM products following lengthy drug incubation. Reverse reaction of the SYMM structure returns a monoadduct to the pool, which can then react with CPI-I to yield either another SYMM or an ASYM product. If the ASYM product's reverse reaction rate is less favorable than that of SYMM, the relative rates of these reverse reactions could determine which product accumulates. Even if the energy barrier to the reaction of monoadduct to form ASYM is relatively higher than that of SYMM, reverse reaction rates could be responsible for the accumulation of ASYM.

Comparison of the Fate of 5′-TAATTA-3′ and 5′-TAAAAA-3′ upon Bizelesin Cross-Linking and CPI-I Same-Strand Bisalkylation. Within the CPI-I reaction mixtures, rapid monoadduct buildup (0.25 h) is followed by its slow conversion to the ASYM bisadduct (>24 h). The relatively rapid Bizelesin cross-linking reaction contrasts sharply with this slow replacement of monoadduct by the ASYM product. In addition to the different reaction rates and alkylation sites, CPI-I cannot match Bizelesin's steric influence on duplex conformation.[7] While both CPI-I and Bizelesin monoadducts generate distortion, Bizelesin directs this distortion in a way that drives AT-step base-pair opening and Hoogsteen base pairing in the 5′-TAATTA-3′ sequence or, in the case of 5′-TAAAAA-3′, modifies the bending properties of the A-tract structure,[15] allowing for subsequent cross-linkage. Consequently, bending distortion disappears from each of these two Bizelesin-cross-linked sequences. Pre-cross-linkage Bizelesin adduct conformational adjustments dissipate whatever distortional effects result from the drug binding and initial adenine attachment. These adjustments place the second 5′-PuNTTA-3′ alkylation target in the correct position in relation to the remaining Bizelesin reactive cyclopropyl moiety. The status of the 5′-TAATTA-3′ sequence as Bizelesin's preferred cross-linking target is attributable to this sequence's capacity to readily respond to the drug with gross conformational adjustments at the AT-step, ultimately providing a suitable target that can be trapped by cross-linkers.

Comparison of CPI-I/Bizelesin−DNA Interactions with Interactions Involving Other Agents That Target A·T-Rich DNA Regions. Our results suggest that covalent modification of the 3′-end adenine of either the AA-21 or AT-21 target sequence generates backbone distortion that in both instances is focused on a junction site five base pairs upstream from the alkylation site. While CPI-I attachment appears to retain a distorted junction site, Bizelesin redirects distortional stress toward the AT-step region and dissipates it through base-pair opening, adenine base rotation, and Hoogsteen base pairing. Results from the AT-step and A-tract sequence studies suggest that CPI-I can neither redirect the distortional stress from the junction site nor dissipate it. Consequently, while Bizelesin unbends the cross-linked region, the CPI-I alkylation products

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

remain bent. The localization of a conformationally distorted junction site at the TA-step resembles TA-step behavior observed in other contexts. Analysis of DNA base dynamics[18] discloses that large-amplitude base dynamics is associated with all TA-steps when a thymine precedes or an adenine follows the TA-step. The results further imply that the adenine residue following the TA-step is more influential than the residue preceding the TA-step. The TA-step adenine experiences large-amplitude (20−50°) slow motion (10 ms to 1 μs), a sharp discontinuity in the attached sugar conformation, and an unusual average orientation.[19] A large number of unusual structural features are also observed, including high negative propeller twist, large rise, negative buckle, and large opening factors, which may explain the flexibility of the TA-step.[18] This TA-step conformational flexibility has been proposed as an important component of transcription initiation in the Pribnow box (5′-TTAACT-3′).[20] These studies of innate TA-step behavior suggest that CPI alkylation-related bending distortion would exacerbate the TA-step conformational anomalies and potentially disrupt the proper orientation of second alkylation reactants at the TA-step reaction site.

The most interesting parallel with CPI-I adduct behavior involves the minor and major groove interactions of homeodomain factors that specifically target 5′-TAATTA-3′ and related sequences. Recent homeodomain binding specificity studies[21,22] have shown that DNA minor−major groove interactions are more complex than would be described by simple amino acid side chain−base pair additive interactions. The observed "nonadditivity" was attributed to energetic coupling of different interactions between homeodomain and DNA components. This coupling is explained by entropic considerations associated with the mutual stabilization of functional group interactions. One such case involves the unexpectedly poor binding affinity of the wild-type engrailed homeodomain with the 5′-TAATCC binding site (versus the consensus 5′-TAATTA-3′ binding site). A proposed explanation is that the presence of C·G base pairs at position 5 and/or 6 causes conformational changes that weaken interactions made by homeodomain residues at other base pairs.[21] The CPI-I/Bizelesin results provide a relatively simplified example of the sequence-specific conformational linkage between apparently isolated interactions (i.e., alkylation events) involving DNA and its binding partners.

Implications for Template-Directed Design of DNA−DNA Crosslinkers. A bisadduct of the pyrrolo[1,4]benzodiazepine (P[1,4]B) tomaymycin was used to guide the template-directed design of the DNA−DNA cross-linker DSB-120.[23−25] In both the P[1,4]B study and the present study, the predicted bisadduct reaction sites were based on sequence specificity findings from monoalkylation reactions. However, CPI-I results show that alteration of the duplex conformation by monoadduct formation diminishes the predictability of a second reaction's sequence specificity, based on the same criteria used to predict monoadduct formation.

While distortive drugs such as CPI-I fail as models for template-based cross-linker design, use of other classes of drugs (e.g., the P[1,4]Bs) that minimally affect duplex conformation increases the feasibility of the template-directed design approach. Although a template-directed design based on CPI-I bisadducts does not lead to a Bizelesin-like tail-to-tail interstrand cross-linker design, the bisadduct provides a possible head-to-tail design for a tetranucleotide intrastrand cross-linker.

Conclusions

Comparison of the CPI-I and Bizelesin reactions with 5′-TAA^{2}TTA^1-3′ demonstrates that Bizelesin's capacity to alkylate both strand one and strand two A^1 stems from its pre-cross-linkage steric rearrangement of the duplex conformation. In part because of bound-CPI-I's inability to "dictate" a suitable monoadduct conformation, strand two A^1 remains inaccessible to the second CPI-I drug, and a second site, strand one A^2, is targeted for alkylation. In purely comparative terms, absence of the ureadiyl linker that joins two CPI-I molecules into a single entity (i.e., Bizelesin) precludes the production of a detectable quantity of symmetrical bisadduct at strand one and strand two A^1.

The monoadduct intermediate in Bizelesin cross-linkage of either 5′-TAATTA-3′ or five-adenine A-tract generates distortional stress, which is presumably propagated through the DNA sequence in much the same fashion as the CPI-I monoadduct. The Bizelesin monoadduct differs from the CPI-I monoadduct in that the former can refocus the distortion within the cross-linkable region, rearrange the duplex conformation, and trap out nonbent cross-linked entities. Contrasting the behaviors of the AT-step and A-tract sequences in response to CPI-I and Bizelesin alkylation helps to elucidate the range of DNA conformational response to external stimuli in AT-rich regions. These conformational probes into the behavior of AT-rich DNA regions provide insight that may prove useful in understanding the role of DNA in the DNA sequence-specific protein recognition and binding process.

Experimental Section

Chemicals. CPI-I 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. Electrophoretic reagents [acrylamide, TEMED, ammonium persulfate, and bis(acrylamide)] were purchased from Bio-Rad. [γ-^{32}P]ATP was purchased from ICN, and X-ray film, intensifying screens, and developing chemicals were purchased from Kodak.

NMR-Scale Oligonucleotide Preparation and Purification. The self-complementary 10-mer d(CGTAATTACG)$_2$ was synthesized on a 10-μmol scale by using automated phosphoramidite chemistry[26] on an Applied Biosystems automated DNA synthesizer (model 381A). The oligomers were deprotected with saturated ammonium hydroxide at 55 °C overnight. Solvent was evaporated at room temperature. The oligomers were purified on a Machery-Nagel Nucleogen-DEAE 60-7 HPLC column with an increasing gradient from 15 mM sodium phosphate to 1 M NaCl and 15 mM sodium phosphate in 20% acetonitrile/aqueous buffer, pH 6.8. The purified single-strand 10-mers were then desalted on four C_{18} Sep Pak cartridges (Waters), and the solvent was evaporated at room temperature.

NMR-Scale Adduct Preparation and Purification. The CPI-I bisadduct was prepared by mixing 5 mg of CPI-I in 0.2 mL of DMF solution with 30 mg of purified 10-mer in 0.75 mL of buffer containing 20 mM sodium phosphate and 200 mM sodium chloride. The reaction mixture was stirred for 4 days in the dark (5 °C), lyophilized to dryness overnight, desalted, and separated from excess drug on C_{18} Sep-Pak

(18) McAteer, K.; Ellis, P. D.; Kennedy, M. A. *Nucleic Acids Res.* **1995**, *23*, 3962−3966.
(19) Kennedy, M. A.; Nuutero, S. T.; Davis, J. T.; Drobny, G. P.; Reid, B. R. *Biochemistry* **1993**, *32*, 8022−8035.
(20) Lefevre, J.-F.; Lane, A. N.; Jardetzky, O. *Biochemistry* **1988**, *27*, 1086−1094.
(21) Ades, S. E.; Sauer, R. T. *Biochemistry* **1994**, *33*, 9187−9194.
(22) Ades, S. E.; Sauer, R. T. *Biochemistry* **1995**, *34*, 14601−14608.
(23) Wang, J.-J.; Hill, G. C.; Hurley, L. H. *J. Med. Chem.* **1992**, *35*, 2995−3002.
(24) Bose, D. S.; Thompson, A. S.; Ching, J.; Hartley, J. A.; Berardini, M. D.; Jenkins, T. C.; Neidle, S.; Hurley, L. H.; Thurston, D. E. *J. Am. Chem. Soc.* **1992**, *114*, 4939−4941.
(25) Jenkins, T. C.; Hurley, L. H.; Neidle, S.; Thurston, D. E. *J. Med. Chem.* **1994**, *37*, 4529−4937.
(26) Gait, M. J., Ed. *Oligonucleotide Synthesis−A Practical Approach*; IRL: Oxford, England, 1984.

cartridges (Waters). The relative amount of bisadduct was monitored by one-dimensional NMR. The above-described reaction had to be repeated four times in order to produce a product that was approximately 95% bisadduct.

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 General Electric GN-500 FT NMR spectrometer (bisadduct) and Bruker AMX 500 FT NMR spectrometer (unmodified duplex, 10-mer). Proton chemical shifts 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). NOE difference spectra were obtained at 5 °C in 90% H_2O/10% D_2O. A 1−3−3−1 pulse sequence was used to suppress the H_2O signal. Signals were irradiated for 1 s with a total recycle delay of 2.4 s.

Phase-sensitive two-dimensional NOESY spectra were obtained for two mixing times, 200 and 400 ms. All spectra were acquired with 16 scans at each of 1024 t_1 values, with 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 45°) 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[27,28] with a 2.5 s pulse repetition time, a sweep width of 24.396 ppm, and a 90° pulse width of 28.75 ms.

Restrained Molecular Dynamics. Restrained molecular dynamics for the bisadduct incorporated 212 NOESY (200 ms) cross-peak derived constraints that were based on strong (ca. 2.00−2.75 Å), medium (ca. 2.75−3.75 Å), weak (ca. 3.75−4.50 Å), and very weak (ca. 4.50−5.00 Å) classifications of cross-peak intensity.[29] The AMBER force field pseudoenergy terms for the interproton distances have the form of flatwells (between r_2 and r_3) with parabolic sides (between r_1 and r_4) within a distance defined by the boundaries of cross-peak classifications: strong ($r_1 = 0.00$, $r_2 = 2.00$, $r_3 = 2.75$, $r_4 = 3.75$), medium ($r_1 = 0.00$, $r_2 = 2.75$, $r_3 = 3.75$, $r_4 = 4.75$), weak ($r_1 = 0.00$, $r_2 = 3.75$, $r_3 = 4.50$, $r_4 = 5.50$), and very weak ($r_1 = 0.00$, $r_2 = 4.50$, $r_3 = 5.00$, $r_4 = 6.00$). A total of 192 DNA−DNA and 20 drug−DNA distance restraints were applied during the rMD trajectories. Flat-angle and distance restraints for base-pair hydrogen-bonding interactions were also applied throughout the duplex region.

For the RMSD study (*in vacuo*), the 10-mer AT-10 was expanded to a 14-mer, 5′-CGCGTAATTACGCG-3′, with retention of CPI-I adducts and distance restraints at corresponding positions. 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 like 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 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 15 ps isothermal phase (300 K) of the rMD analysis, an average structure was generated using CARNAL (AMBER 4.1)[30] for each of the A- and B-form starting structures. RMSD analysis was conducted (CARNAL) within and between the two families of A- and B-form starting structures and the average rMD products (Table 4).

Interproton distances were incorporated into the rMD calculations of the solvated system. rMD calculations were performed using the SANDER module of AMBER 4.0 on an SGI 4D35G personal Iris workstation.[31] The AMBER restraints were applied as described above for the RMSD study. Helical and base-pairing properties were calculated using the Curves program of Dials and Windows.[32]

Heat-Induced Strand Breakage Assay of Drug-Modified DNA. Twenty-one-base-pair oligonucleotides (Table 1) were synthesized on an automated DNA synthesizer (Applied Biosystems 381A) by the phosphoramidite method.[26] The oligomers were deprotected with saturated ammonium hydroxide at 55 °C overnight. The ammonium hydroxide solution was evaporated under vacuum. The DNA strand containing the CPI-I alkylation site was 5′-end-labeled, hybridized with the unlabeled complementary strand, and purified. Individual DNA strands were hybridized, 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, and extracted with annealing buffer [10 mM Tris−HCl (pH 7.6) and 100 mM NaCl].

Time-Course Study of CPI-I-Modified DNA Duplexes. Aliquots (0.40 pmol) of purified duplexes in 10 mM Tris−HCl and 10 mM NaCl (pH 7.6) were modified with 10 μL of 840 pmol CPI-I at room temperature for periods of 5 min, 15 min, 30 min, 1 h, 3 h, 5 h, 1 day, and 2 days. Calf thymus DNA (10 μg) was then added to the reaction mixture, and unreacted drug molecules were removed by phenol/$CHCl_3$ extraction followed by ethanol precipitation. DNA pellets were dissolved in double distilled water and subjected to thermal cleavage (100 °C, 45 min) to induce strand breakage at drug modification sites. After thermal treatment, samples were lyophilized, and the DNA pellets were redissolved in alkaline sequencing dye solution (80% formamide, 10 mM NaOH) and applied to 20% denaturing sequencing gel electrophoresis. After electrophoresis, the gels were exposed to X-ray films.[33] Purine- and pyrimidine-specific sequencing reactions and an adenine-specific sequencing reaction were carried out.[34]

Concentration Dependency of CPI-I DNA Alkylation on the Drug. The reaction mixtures (20 μL) consisted of 10 mM NaCl, 10 mM Tris−HCl (pH 7.6), 0.81 pmol of DNA, and different amounts of CPI-I (0, 0.84, 8.4, 84, and 840 pmol). The mixture was incubated at room temperature for 1 day. Postreaction workup was the same as in the time-course study of CPI-I-modified DNA duplexes described above.

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 U72779 (CPI-I) and Bizelesin. We thank David Bishop for proofreading, editing, and preparing the paper.

Supporting Information Available: Text and figures describing and showing NOESY data, bis-CPI-I−DNA adduct exchangeables ^{1}H NMR data, ^{1}H NOEs, and gel electrophoresis of CPI-I alkylation of AA-21 (16 pages). Ordering information is given on any current masthead page.

JA960017A

(27) Sklenar, V.; Bax, A. *J. Magn. Reson.* **1987**, *74*, 469−479.
(28) Blake, P. R.; Summers, M. F. *J. Mag. Reson.* **1990**, *86*, 622−625.
(29) Lin, C. H.; Hill, G. C.; Hurley, L. H. *Chem. Res. Toxicol.* **1992**, *5*, 167−182.
(30) 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 (UCSF)*; University of California: San Francisco, 1995.
(31) Pearlman, D. A.; Case, D. A.; Caldwell, J.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. *AMBER 4.0 (UCSF)*; University of California: San Francisco, 1991.
(32) Ravishanker, G.; Swaminathan, S.; Beveridge, D. L.; Lavery, R.; Sklenar, H. *J. Biomol. Struct. Dyn.* **1989**, *6*, 669.
(33) Reynolds, V. L.; Molineaux, I. J.; Kaplan, D. J.; Swenson, D. H.; Hurley, L. H. *Biochemistry* **1985**, *24*, 6228−6237.
(34) Maxam, A. M.; Gilbert, W. *Methods Enzymol.* **1980**, *65*, 499−560.