Structure of a Covalent DNA Minor Groove Adduct with a Pyrrolobenzodiazepine Dimer: Evidence for Sequence-Specific Interstrand Cross-Linking

Terence C. Jenkins,*,† Laurence H. Hurley,‡ Stephen Neidle,† and David E. Thurston§

Cancer Research Campaign Biomolecular Structure Unit, The Institute of Cancer Research, Sutton, Surrey SM2 5NG, U.K., Drug Dynamics Institute, College of Pharmacy, The University of Texas at Austin, Austin, Texas 78712-1074, and School of Pharmacy and Biomedical Sciences, University of Portsmouth, Portsmouth, Hants PO1 2DZ, U.K.

Received July 8, 1994®

The structure of the interstrand cross-linked adduct formed between a C8–C8′-linked pyrrolobenzodiazepine (PBD) dimer (DSB-120; 1,1′-(propane-1,3-diyldioxy)bis[(11a*S*)-7-methoxy-1,2,3,11a-tetrahydro-5*H*-pyrrolo[2,1-*c*][1,4]benzodiazepin-5-one]) and a self-complementary d(CICGATCICG)$_2$ duplex has been determined from high-field 1D- and 2D-NMR data using a simulated annealing procedure. The refined structure supports earlier observations from solution NMR experiments and indicates that the covalently bound molecule spans six DNA base pairs in the minor groove, forming a symmetric cross-link between the spatially separated internal guanines and with active recognition of an embedded 5′-GATC bonding site. This result confirms that template-directed approaches are useful for the design of linked DNA-interactive PBD dimers with viable DNA cross-linking potential. Further, *head-to-head* connection of the PBD moieties results in an overall retention of 5′-GA bonding site preference for each alkylating PBD subunit. Structural analysis indicates that cross-link formation results in a localized perturbation of the DNA duplex, attributable in part to a mutual reduction in dynamic mobility or "covalent clamping" within the Gua4–Cyt7 base tract. However, ligand-induced distortion is confined to the Cyt7 and Ino8 residues on each strand. The Gua(N2)–Gua(N2) cross-link is stabilized by two directed H-bonds from the formed aminal residues to N3 acceptor atoms of adenine bases on the 3′-side of each covalently modified guanine. Evidence for sequence-specific cross-linking with DSB-120 is provided by extended modeling studies which suggest that recognition of the favored d(•GATC•) motif is dominated by van der Waals steric factors, although electrostatic and H-bonded interaction terms also play a key role. This conclusion supports recent covalent footprinting studies revealing that this PBD dimer shows a selectivity for embedded base sequences of the type 5′-(pu/py)GATC(py/pu).

Introduction

Pyrrolo[2,1-*c*][1,4]benzodiazepines (PBDs; see Figure 1) are of considerable current interest due to their ability to recognize and subsequently form covalent bonds to specific base sequences of double-stranded DNA. Such monofunctional alkylating compounds have potential therapeutic application in cancer treatment and as selective anti-infective agents.[1,2]

The PBD antitumor antibiotics are produced by various *Streptomyces* species, with family members (Figure 1) including tomaymycin, anthramycin, sibiromycin, neothramycins A and B, and DC-81 (**1**).[3] The cytotoxic and antitumor activities of monomeric PBDs are attributed to their ability to form covalent DNA adducts via an acid-labile aminal bond to the electrophilic imine C11-position involving the exocyclic N2 of a guanine base in the minor groove of duplex DNA (Figure 2A).[3,4] The C11a*S* chirality provides a right-handed molecular twist, when viewed from the C-ring toward the A-ring, which enables the PBD to mirror the curvature of B-form DNA and maintain isohelical contact with the walls and floor of the minor groove. The structure of an anthramycin-DNA adduct has been studied using indirect methods.[4] Molecular modeling, solution NMR, fluorimetry, and DNA footprinting experiments indicate that PBD monomers are generally sequence-specific with a binding site preference for 5′-puGpu (particularly 5′-AG or 5′-GA) sequences.[3b,5] A correlation between DNA sequence specificity, relative induced distortion of DNA, and the kinetics of covalent binding has been reported.[6] PBDs have been shown to inhibit DNA processing enzymes, with activity in accord with structure-activity predictions,[4] although the role of DNA adduct formation is not understood. An assay system based upon inhibition of *Bam*HI restriction endonuclease has revealed a correlation between DNA binding and biological activity, with sibiromycin representing the most reactive PBD agent examined.[7]

The cytotoxic potency of bifunctional alkylating agents is probably due to the production of *interstrand* DNA cross-links that are difficult for the cell to repair.[8] DNA-reactive molecules of current clinical application, such as the nitrogen mustards or mitomycins, generally act as G/C-specific DNA cross-linking agents, although A/T-specific cross-linkers such as bizelesin, a (+)-CC-1065 analogue, have recently been selected for clinical trials.[9] However, such agents generally show selectivity for only relatively short DNA stretches,[10] indicating a need for new agents capable of recognizing and bonding to more extended and/or mixed sequence tracts of duplex DNA to uniquely define an individual gene target and, hence, effect biological specificity.

* Address correspondence to this author. Tel.: (44)-81-643-8901 ext 4255. FAX: (44)-81-770-7893.
† The Institute of Cancer Research.
‡ University of Texas at Austin.
§ University of Portsmouth.
® Abstract published in *Advance ACS Abstracts*, November 15, 1994.

0022-2623/94/1837-4529$04.50/0

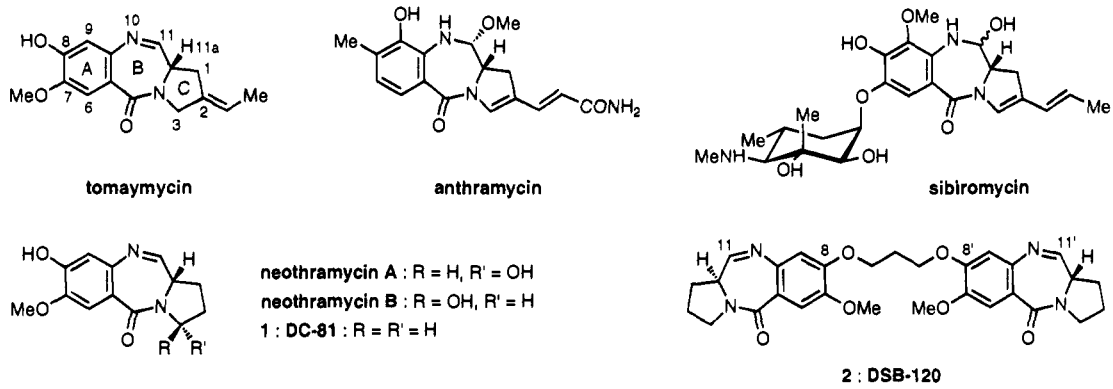

Figure 1. Structures of the PBD monomers, including DC-81 (**1**) and the C8–C8′-linked dimer DSB-120 (**2**).

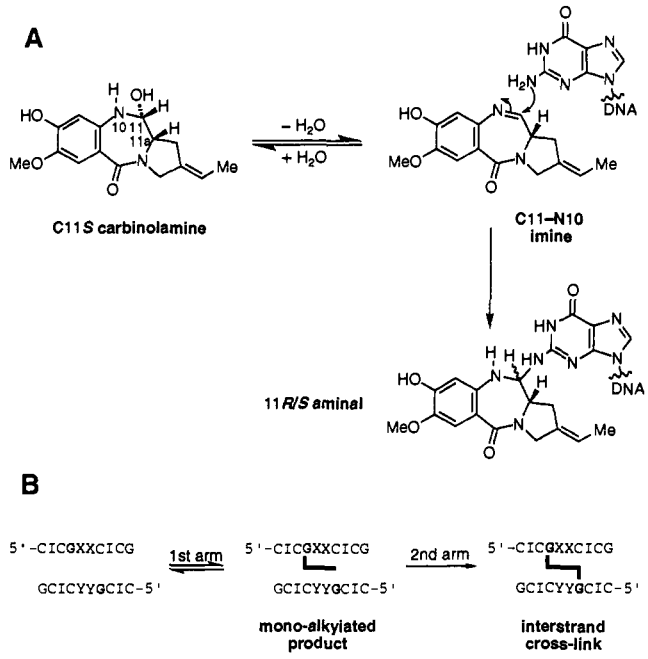

Figure 2. (A) Proposed reaction of tomaymycin hydrate with the exocyclic 2-amino group of guanine in DNA. (B) Two-arm induction of interstrand DNA cross-links.

We recently reported the synthesis of a highly efficient DNA interstrand cross-linking agent, **2** (DSB-120; 1,1′-(propane-1,3-diyldioxy)bis[(11a*S*)-7-methoxy-1,2,3,11a-tetrahydro-5*H*-pyrrolo[2,1-*c*][1,4]benzodiazepin-5-one]), based on the PBD ring system.[11] The PBD dimer **2** consists of two DC-81 (**1**) subunits joined *head-to-head* through the A-ring C8-positions by a flexible 1,3-propanediyldioxy linkage and represents an example of DNA template-directed drug design.[12] Gel electrophoresis and thermal denaturation studies with naked DNA samples have shown that **2** has a high affinity for DNA and forms irreversible interstrand DNA cross-links.[11,13] Further, this compound is the most cytotoxic agent of a series of α,ω-diether homologues *in vitro*, and it has been firmly established that the DNA cross-linking ability of the PBD dimers correlates with cytotoxicity and antitumor activity.[11b] The cellular pharmacology of this series of agents has recently been examined in a range of human tumor cell lines, revealing (i) rapid and highly efficient formation of interstrand DNA cross-links with no evidence of single-strand breaks, (ii) poorly repaired cross-links, and (iii) drug-induced arrest at the G_2/M phase of the cell cycle typical for cross-linking agents.[14] However, cellular sensitivity is determined by the levels of glutathione and p170 glycoprotein, suggesting that resistance may reflect partial inactivation of the drug by GSH binding.

The present study details the structure of the covalently cross-linked adduct formed between **2** and the self-complementary d(CICGATCICG) duplex (Figure 2B), obtained using a NOE distance-restrained simulated annealing procedure with data derived from high-field solution 2D-NMR experiments.[15] Features of the covalent binding behavior of **2** with this decanucleotide duplex have been compared with those of the equivalent 2:1 tomaymycin-DNA adduct using NMR techniques. This preliminary study inferred that tethering of two such PBD units through the A-ring results in conformational stress of the DNA upon cross-link formation, together with a significantly altered presentation of the pyrrole C-ring to the minor groove when compared to the PBD monomer.[15] Structural consequences from such effects are important for the design of cross-linkers using a DNA template-directed strategy, particularly where either tolerance or recognition of further DNA bases is required in the vicinity of the modified guanines.

In order to establish the conformational effects of cross-linking by a C8-linked PBD dimer, we have extended our earlier qualitative NMR studies[15] to determine the structure of the covalent **2**–d(CICGATCICG)$_2$ adduct in aqueous solution. Analysis of this structure provides a quantitative description of the binding-induced DNA helical perturbation and highlights the key molecular features likely to be essential for the design of this class of agent. In particular, this structure reveals that the covalently bound ligand spans six DNA base pairs (bp) in the minor groove, forming a 4-bp symmetric *interstrand* cross-link between the spatially separated internal guanines.

Further, cross-link formation with this agent involves active recognition of the 5′-GATC sequence through hydrogen bonding directed to the spanned adenine bases. This important finding is supported by extended molecular mechanics/dynamics (MM/MD) modeling studies with 10-mer duplexes, d(CICGXXCICG)·d(CICGYYCICG), containing alternative core base sequences. Energy calculations reveal that each PBD subunit in **2** retains a covalent bonding preference for a 5′-GA site, leading to sequence-specific recognition and cross-linking of the d(·GATC·) motif. Searches of the EMBL databank show that such base sequences appear rather infrequently, although certain oncogene sequences (e.g., *raf* and *int*-2) may provide viable gene targets due to a high-frequency occurrence of preferred bonding sites.[16]

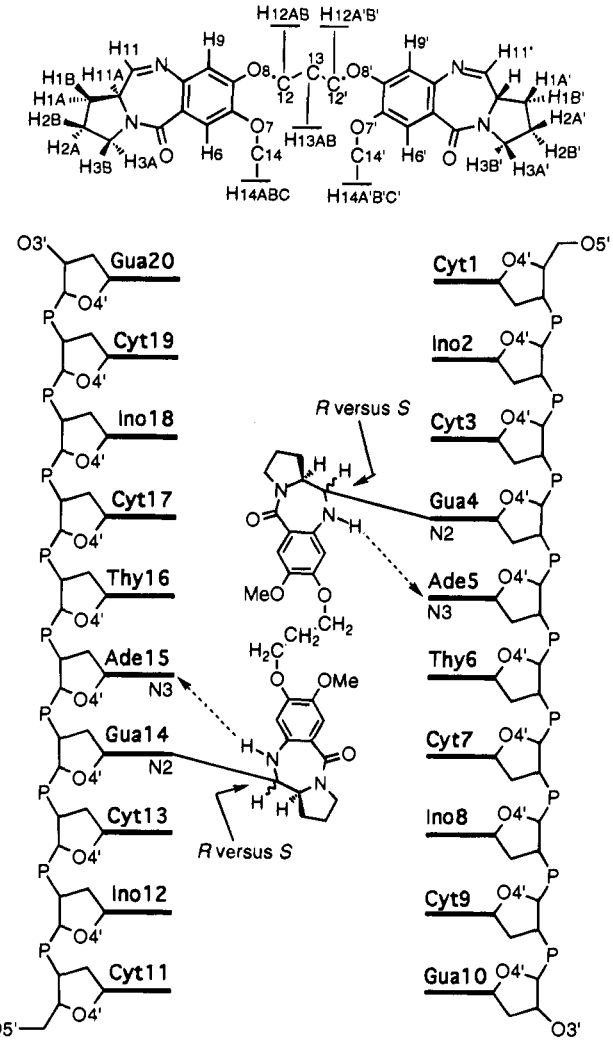

Figure 3. Schematic model of the 1,4-interstrand cross-link between DSB-120 and d(CICGATCICG)$_2$ showing the numbering scheme for the DNA and the ligand. P denotes a 3′-OP-(O_2^-)OCH$_2$-5′ phosphodiester linkage.

The present structural observations provide a structural basis for conclusions from earlier biophysical studies and reinforce findings from cellular pharmacology experiments.

Results and Discussion

Structure Refinement of the Cross-Linked d(CICGATCICG)$_2$ Adduct. The two-step mechanism for interstrand cross-link formation proposed on the basis of derived NMR data is shown schematically in Figure 2B.[15] The numbering scheme used for the electrophilic dimer ligand **2** and the self-complementary 10-mer duplex is shown in Figure 3. Initial molecular models for the ligand, DNA duplex, and interstrand cross-linked adduct were generated as described in the Experimental Section. Models with 11*S* + 11′*S*, 11*R* + 11′*S*, 11*S* + 11′*R* and 11*R* + 11′*R* chirality at the B-ring C11/11′-positions were constructed to determine the effect of stereochemistry upon structure refinement (see later).

The cross-linked DNA minor groove adduct was examined using a simulated isothermal dynamic annealing procedure. The four-stage MM/MD–MD/MM protocol adopted, using X-PLOR,[17] has been used for analogous NMR and MD studies with reversible, noncovalent minor groove-binding ligands.[18] Thermal equilibration of the system was achieved within ≤5–10 ps, as judged from analysis of the velocity distribution and component energy terms. In the absence of counterions and explicit solvent molecules, a postheating acquisition period of 30 ps proved sufficient to ensure satisfactory convergence of the rms-averaged atomic coordinates. Indeed, structures obtained after more prolonged simulations for 50 and 100 ps showed little significant difference, although data acquired during 110-ps simulation (i.e., 10-ps heating and simulated annealing for 100 ps) were used to determine the ultimate structure. In contrast to the dynamic stability of this system, protracted 'heating' of a native 12-mer duplex has been reported to induce a transition from B- to A-type DNA during simulation periods >30 ps, although less divergent behavior may result if time-averaged distance restraints are employed.[19a,b] A recent study has shown that extended MD is essential for DNA duplexes to ensure conformational sampling.[19c]

Target NOE-derived interproton distance restraints (Table 2 from ref 15) were applied in a symmetric manner, assuming equivalence for both DNA strands (i.e., Cyt1≡Cyt11, Ino2≡Ino12, etc.) and each PBD moiety in the adduct, with exact magnetic degeneracy for symmetrically disposed protons. No attempt was made to assign intermolecular DNA-ligand NOE contacts at this stage as the self-complementary nature of the DNA prevented distinction of each strand. Similarly, the centrosymmetric symmetry of both the ligand molecule and the complex leads to a possible 4-fold degeneracy for all observed NOEs. Thus, for example, the strong NOE from H9/H9′ ligand protons to Ade5(H2)[15] was initially treated as four distinct NOEs, i.e., H9–A5/15(H2) and H9′–A5/15(H2). Inappropriate or out-of-range NOEs were removed during subsequent stages of refinement, until all NOEs could be distinguished unambiguously. Error ranges assigned to the NOEs were uniformly scaled for target distances representing the relative NOE intensities, with 2.75 Å (strong, 1.8–3.15 Å), 3.5 Å (medium, 1.8–4.1 Å), 4.5 Å (weak, 1.8–5.3 Å), and 5.0 Å (very weak, 1.8–6 Å), respectively. Thus, maximum r_{app} distances were constrained to remain within 15–20% of the assigned target values.[18a,b] Soft glycosidic torsion constraints were applied equally to corresponding bases in each DNA strand (see the Experimental Section).

Structures were initially examined using only short-range NOE restraints ≤3.2 Å involving the intramolecular DNA and ligand components and the intermolecular ligand-DNA contacts. Further NOE restraints were applied or removed in a stepwise manner as the structure converged, until all experimental NOEs[15] were accommodated. Long-range NOEs (≥4.1 Å) were incorporated only during later stages of refinement. NOE degeneracy (see above) was ultimately achieved and retained only for protons in the –O(CH$_2$)$_3$O– linker, involving contacts with Ade5–Cyt7 and Ade15–Cyt17 base segments.

Stereochemical effects at the C11 covalent linkage site of PBD antibiotics (see Figure 3) have been studied using NMR and biophysical techniques[4,5,20] and have recently been reviewed.[3b,c] We have examined the consequences of a change in stereochemistry at the DNA-ligand link. Attempted refinement with *R* stereochemistry at either C11 ligand position resulted in

destabilization of the cross-linked DNA adduct, leading to disruption of base pairing associated with the covalently modified guanine residue. Such major ligand-induced distortion causes the complementary cytosine (i.e., Cyt7/17) to uncouple from the stacked helix and become displaced from the helical axis into the major groove, with a concomitant, energetically unfavorable distortion of the local DNA backbone. In the case of the C11*R* + C11′*R* ligand complex, induced perturbation in the adduct results in a severe loss of helical integrity, with propagated distortion effects and reduced base stacking. In contrast, stable adducts retaining general B-type DNA integrity were formed *only* with the symmetric C11*S* + C11′*S* ligand molecule, and it proved possible to accommodate all observed NOEs without significant violation (i.e., within ±0.5 Å). This result is in accord with NMR studies of this adduct, which indicate exclusive *S* stereoisomers for the two C11 chiral centers.[15] Equivalent C11*S* geometry has recently been reported for the X-ray crystal structure of a 2:1 anthramycin-DNA complex, where attempted crystallographic refinement with the less favorable C11*R* chirality leads to similar distortive effects.[21]

In the case of the favored C11*S* + C11′*S* cross-linked adduct, a total of 288 interproton distance restraints (i.e., 202 DNA–DNA, 30 ligand–ligand, and 56 intermolecular DNA-ligand contacts), together with 20 DNA χ-angle constraints, were used to generate the ultimate structure. The interproton DNA-ligand distances used for structure refinement accounted for all intermolecular NOE contacts determined from the solution NMR behavior[15] of the cross-linked adduct (Table 4 in the supplementary material). These NOEs represent contact of reporter protons positioned in the floor and walls of the DNA minor groove tract with the A- and B-ring protons presented by the concave face of the PBD dimer (i.e., H6/6′, H9/9′, H11/11′, and H11A/11A′). Additional strong NOEs are provided by close base contacts involving the $-O(CH_2)_3O-$ linker moiety, particularly for the H12/H12′ protons (see Figure 3), consistent with effective penetration of the minor groove by this portion of the molecule. Final NOE violations, following NOE-restrained dynamic annealing for 100 ps, were within ±0.07 Å (rms = 0.008) of the target interproton distances prior to final conjugate gradient relaxation and <±0.13 Å (rms = 0.019) for all 288 NOEs in the fully refined adduct. The dynamic progress of the adduct system during the sampled annealing period is shown in Figure 4. The low-energy structure for the refined C11*S* + C11′*S* symmetrically cross-linked d(CICGATCICG)$_2$-**2** adduct is shown in Figure 5, where the covalently modified guanine residues are necessarily on opposite strands of the DNA duplex.

DNA Structure. The duplex clearly retains a general B-like DNA conformation (Figure 5A). Alternative refinement of the adduct with an A-DNA duplex initial geometry resulted in a closely similar structure with no significant difference (NOE violations <±0.17 Å, final rms = 0.024), providing qualitative support for the NOE-driven dynamic annealing procedure. Helix analysis of the final duplex (Table 1) shows an average helical rise of 3.3 ± 0.4 Å and a mean helical rotation of 40 ± 6° (9.1 bp/turn). These values compare favorably with data for DNA-ligand complexes refined using an equivalent procedure,[18a,b] suggesting that covalent modification of the duplex induces no significant alteration in overall helical stacking.

Figure 4. Dynamic progress of the cross-linked DNA–ligand adduct, sampled at 5-ps intervals during 100 ps, showing the molecular mobility (i.e., excursion) of the component residues.

Examination of the structure reveals that the two covalently modified bases, Gua4 and Gua14, are stacked well into the helix and retain good H-bonded geometry with their complementary cytosines. Indeed, there appears to be little drug-induced perturbation of the exocyclic Gua(N2) amino group, and base stacking within the core Gua4–Cyt7 base tract is not appreciably disrupted. These observations reinforce conclusions inferred from the earlier NMR study.[15]

Helical parameters for the host DNA duplex are collected in Table 1, showing that perturbation of base pairing is largely confined to the core AT-tract, manifest primarily as increased buckle terms. The base pair parameters are nevertheless in accord with values expected for the B-type DNA family. In agreement with the reported NMR behavior,[15] most sugar puckers are near C2′-*endo*. However, Table 1 shows that base steps Ino2pCyt3 and Cyt3pGua4, together with the corresponding Cyt7pIno8 and Ino8pCyt9 steps at the 3′-end of the cross-link site, have decreased rise parameters, indicating localized helical compression that may contribute to the unusual NMR behavior inferred for the Ino8/18 protons.[15] Further, Ino8/18 has an unusually low glycosidic torsion angle of −74°. In support of a localized perturbation effect, bases Cyt7/17 flanking the modified guanines to each 3′-side are also characterized by furanose pseudorotations of 176° (high C2′-*endo* or near C3′-*exo*), compared to a mean value of 163 ± 8° for the duplex overall, and increased α- and δ-phosphate backbone torsion angles. However, such drug-induced perturbation effects are confined to the Cyt7–Ino8 and Cyt17–Ino18 base tracts and not extensively propagated through the helix.

Figure 6A shows the minor groove width of the duplex determined from the interstrand P···P and H4′···H5′ separations, indicating that the groove is effectively widened by **2** in the bonding region upon adduct formation. Similar effects have been reported for noncovalent groove-binding ligands.[22,23] Analysis of the dynamic behavior of the DNA duplex from the rms-averaged atomic coordinates taken from snapshots

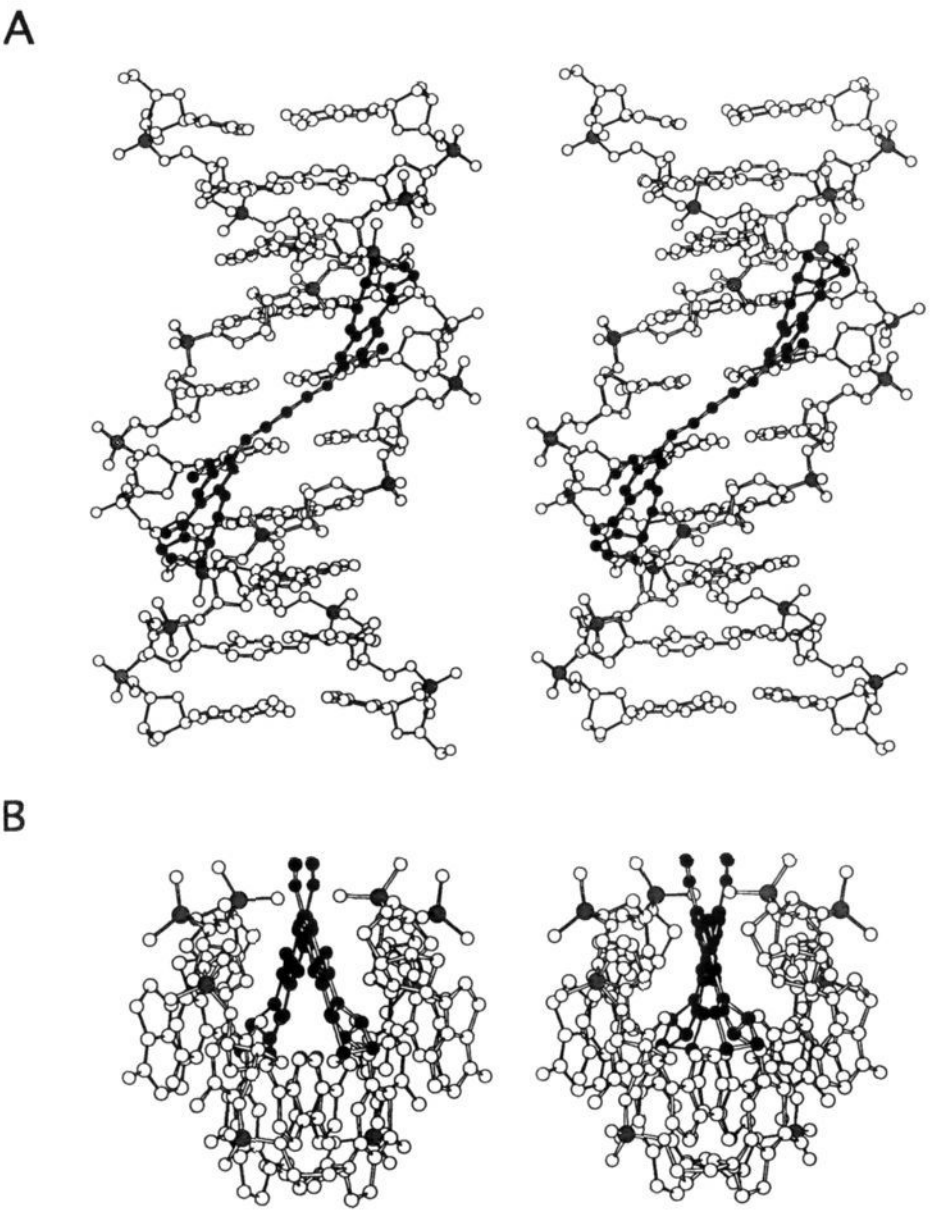

Figure 5. Stereoviews of the cross-linked 10-mer duplex following X-PLOR refinement. (A) View of the entire adduct looking into the DNA minor groove. (B) Close-up view of the Gua4−Cyt7 core region of the DNA showing the orientation of the DSB-120 molecule with respect to the walls of the minor groove. All H-atoms have been removed for clarity. The ligand is highlighted, and the DNA backbone phosphorus atoms are shown stippled.

Table 1. Helical Parameters for the Cross-Linked d(CICGATCICG)$_2$−Ligand Adduct[a]

base pair	tip (°)	incline (°)	buckle (°)	propeller twist (°)	base step	roll (°)	twist (°)	slide (Å)	rise (Å)
Cyt1·Gua20	−4	−3	19	−9	Cyt1−Ino2	10	44	−0.8	3.68
Ino2·Cyt19	9	−2	15	−12	Ino2−Cyt3	−7	38	−1.1	2.98
Cyt3·Ino18	1	3	8	−5	Cyt3−Gua4	−5	41	−1.0	3.17
Gua4·Cyt17	−5	0	−2	−3	Gua4−Ade5	6	40	−2.1	3.41
Ade5−Thy16	2	−3	12	−6	Ade5−Thy6	−6	37	−1.7	3.64
Thy6·Ade15	−2	−3	−12	−6	Thy6−Cyt7	6	40	−2.1	3.41
Cyt7·Gua14	5	0	2	−3	Cyt7−Ino8	−5	41	−1.0	3.17
Ino8·Cyt13	−1	3	−8	−5	Ino8−Cyt9	−7	38	−1.1	2.98
Cyt9·Ino12	−9	−2	−15	−12	Cyt9−Gua10	10	44	−0.8	3.68
Gua10·Cyt11	4	−3	−19	−9					

[a] Calculated using the NEWHEL92 program.

accumulated during the MD annealing period (Figures 4 and 6B) indicates that cross-link formation restricts the mobility of atoms associated with bases within the bonding site. Thus, the mobilities of residues within the Gua4−Cyt7 base tract are effectively "stiffened" by a ~3-fold factor relative to the exposed strand regions. The motion of the ligand molecule is similarly clamped by a ~1.5-fold factor, although atoms positioned deep within the minor groove suffer more restricted movement.

It is likely that such effects will lead to increased time-averaged exposure of the exocyclic 2-amino protons of

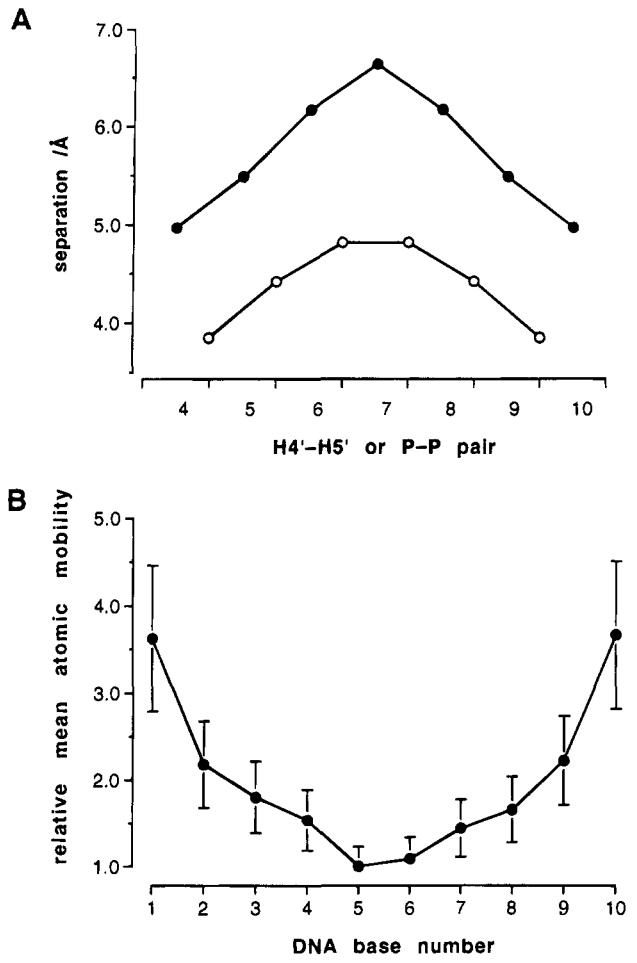

Figure 6. (A) Plot of P···P [$i - (i + 15)$ less 5.8 Å, ●] and mean H4′···H5′ [$i - (i + 16)$, ○] interstrand distances for the refined DNA–drug adduct, where i is the residue on the first strand. (B) Relative atomic excursion of the DNA nucleotides during MD for 100 ps with sampling at 0.2-ps intervals. Data represent the rms mean ± SD for all atoms in each residue, averaged for the two strands, and are normalized to the least-mobile bases (Ade5/15; relative mobility = 1.0).

Gua4/14, thereby facilitating solvent access and leading to faster exchange, in accord with earlier NMR observations.[15] Qualitative support for a drug-induced alteration of mean DNA flexibility is provided by comparative studies of electrophoretic mobility for native DNA and derived adducts formed with both PBD monomers and dimers (J. A. Hartley, unpublished results). However, the consequences of such localized molecular stiffening upon enzymatic recognition and DNA repair processes are uncertain. Nevertheless, the poor inherent repair observed for intracellular cross-links[14] induced by **2** suggests that this lesion is not repaired by a DNA repair mechanism that is scanning for DNA distortion.

DNA–Ligand Interactions. The ligand alignment determined for **2** in the cross-linked adduct (Figure 5) is in qualitative agreement with that inferred from earlier 2D-NMR experiments.[15] The methylene protons in the linker function are positioned close to the floor of the minor groove, and the groove penetration of the pyrrole C-ring is shallow. In contrast, the aromatic phenyl rings are aligned parallel to the walls of the minor groove and the H6/6′ protons are in close contact with the groove floor.

The two phenyl A-rings of **2** are oriented by 34° with respect to one another (Figure 5B) in the DNA adduct. This value is similar to the 30° twist reported for the ligand in the crystal structure of the d(CGCGAATTCGCG)$_2$-propamidine complex, where the ligand contains an equivalent 1,3-diphenoxypropane moiety.[23a] Further, the oxygen–oxygen (O8···O8′) distance of 4.8 Å in the propanediyldioxy bridge determined here for the bound **2** molecule compares well with that of 4.7 Å in the propamidine complex[23a] and a separation of 4.72 Å calculated for a fully antiperiplanar model $-O(CH_2)_3O-$ diether. Thus, the linker in **2** adopts a low-energy, extended *all-trans* conformation in the adduct (Figure 5A). The central C8–O(CH$_2$)$_3$O–C8′ linkage in **2** has torsion angles of (C7–C8–O8–C12) 175°, (C8–O8–C12–C13) −164°, and (O8–C12–C13–C12′) −174°; these angles resemble those determined for analogous DNA-propamidine and related complexes.[23] Similarly, the 7-methoxy groups are rotated from the plane of each A-ring, such that the C8–C7–O7–C14 torsion angle is −177°.

The C11···C11′ separation for the two covalent linkage sites in the final cross-linked structure is 12.5 Å, which corresponds favorably with interstrand Gua4(N2)···Gua14(N2) and Gua4(H22)···Gua14(H22) distances of 11.5 and 12.5 Å, respectively, in native B-type DNA. Equivalent separations for A-form DNA containing an embedded d(·GATC·)$_2$ cross-link site are 13.9 and 15.4 Å, respectively. This correspondence indicates that the linker function in **2** has the optimum separation to effect *interstrand* cross-linking of a 4-bp bonding site within a normal B-type DNA structure and disagrees with earlier suggestions for optimal linker length from modeling studies.[12,24]

The shape complementarity achieved between **2** and the host DNA in the adduct results in the formation of hydrogen bonds between the N10–H10 donor groups and the N3 acceptor atoms of adenine bases positioned on the 3′-side of the alkylated guanines (Figure 3). Thus, two directed H-bonds are formed involving the spanned bases in the cross-link site, at a H10···Ade5/15(N3) separation of 1.88 Å and with a subtended N10–H10···Ade5/15(N3) angle of 142°. It is tempting to suggest that the formation of such H-bonded interaction with the 3′-adjacent base is implicated in the known 5′-GA bonding preference of DNA-binding PBDs,[3-6] particularly as this arrangement would stabilize the formed covalent adduct. An equivalent H-bonded geometry has recently been suggested for a 2:1 anthramycin–DNA covalent adduct.[21] Preliminary modeling studies with a sp^3-hybridized, N10/10′-diprotonated adduct, suggested as a possible alternative species for covalently bound PBDs,[15] indicate that cross-link stabilization is considerably reduced compared to the neutral aminal species (Figure 2A), largely as a result of steric clash and disruption of the favorable H-bonded geometry.

Sequence Selectivity of Interstrand DNA Cross-Linking. The possible sequence-specific interstrand DNA cross-linking of **2** was examined using an equivalent MM–MD–MM strategy, without applied NOE restraints but using glycosidic constraints, for 1:1 covalent adducts with the 10-mer duplex variants d(CICGXXCICG)·d(CICGYYCICG), where X·Y represents an embedded complementary base pair. These sequences, generated by base mutation of the parent d(CICGATCICG)$_2$ duplex, were selected to assess the

Table 2. Interaction Energies (kcal/mol) of Covalent Complexes between DSB-120 and d(CICGXXCICG)·d(CICGYYCICG) Duplexes

XX(YY) core sequence	E_{bond} net bonding energy[a]	intermolecular energy terms		
		ΔE_{vdW}[b]	ΔE_{elec}[b]	$\Delta E_{H\text{-}bond}$[d]
AT	−73.1	−72.8	−8.5	−4.9
AA(TT)	−62.9	−65.2	−3.6	−3.0
TA	−59.6	−62.0	−1.4	−0.5
CG	−56.4	−57.5	0.3	0.7
GG(CC)	−50.1	−55.7	7.2	1.9
GC	−41.8	−50.3	13.4	3.4

[a] Net bonding enthalpy calculated using $E_{total(complex)} - [E_{total(DSB\text{-}120)}]$, following X-PLOR minimization. [b] van der Waals or steric contribution. [c] Electrostatic contribution. [d] Hydrogen-bonded energy component.

influence of the spanned core bases upon the energetics of cross-link formation. Models were evaluated for the PBD dimer **2** with C11*S*,C11'*S* stereochemistry, as the incorporation of each *R* chiral C11 center results in ~15−20 kcal/mol poorer interaction energy due to induced helical disruption (see above).

Table 2 details the bonding enthalpies and component intermolecular energy terms computed for each of the six 10-mers following dynamic annealing for 20 ps at 300 K. The interaction energies for DNA−**2** cross-linking show a clear ranking order, with XX = YY = AT representing the most favored bonding sequence. In marked contrast, XX = YY = GC provides the least favorable site for covalent binding. These results are supported by recent covalent footprinting studies of cross-linking using naked DNA samples and plasmid constructs with oligonucleotide inserts of defined sequence (J. A. Hartley and T. C. Jenkins, unpublished results). Such electrophoretic studies confirm a bonding preference for embedded 5'-(pu/py)GATC(py/pu) duplex sequences with a strict requirement for a 4-bp cross-link site. In agreement with the predicted specificity, sequences containing spanned CpG, GpC, or GpG bases are disfavored for cross-linking and provide low-frequency binding sites.

Analysis of the component energy terms (Table 2) reveals that the predicted rank order is largely determined by nonbonded van der Waals steric factors reflecting induced DNA and/or ligand perturbation effects. Interestingly, the electrostatic and H-bonded contributions to the bonding enthalpy, ΔE_{elec} and $\Delta E_{H\text{-}bond}$, respectively, both indicate the same rank order for interaction, with the least-favored adducts leading to endothermic destabilization. Examination of the rms-averaged structures following energy minimization shows that the XX = AT and AA(TT) adducts retain the geometry determined for the d(CICGATCICG)$_2$−**2** solution NMR structure, with H-bonded contact(s) involving the 3'-flanking adenine bases (see Figure 3). Such an arrangement is less favorable in the case of the cross-linked XX = TA and CG adducts, whereas this geometry is disrupted in the XX = GC and GG(CC) structures due to clash involving steric contact between the spanned guanine bases and the $-O(CH_2)_3O-$ linker moiety in the PBD dimer.

Conclusions

The structure determined for the d(CICGATCICG)$_2$−**2** adduct using data from a 2D-NMR solution study confirms and extends our earlier finding[11a] that this PBD dimer forms a symmetric interstrand cross-link with double-stranded DNA involving a 4-bp bonding site but spanning six DNA base pairs overall. Interstrand *minor groove* cross-linking by **2** results in retention of general DNA structural integrity, with perturbation confined mainly to the Cyt7 and Ino8 nucleotides on each strand. Such ligand-induced distortion effects are not propagated through the helix. The overall lack of structural perturbation suggests that such covalent lesions may be less easily recognized by DNA repair nucleases, especially compared to the serious helical distortion induced by cross-linked major groove adducts.

The present structure indicates that the conformation of PBD dimer **2** is ideal for bifunctional covalent interaction with guanines disposed on opposite strands, with a bonding preference for the 5'-GATC sequence due to induction of post alkylation hydrogen bonds involving acceptor bases within the spanned site. Further, the structure confirms evidence from molecular design and solution NMR studies that two suitably linked PBD monomer subunits can be accommodated within a host B-DNA duplex to effect formation of a cross-link. Our observations of binding site size and sequence specificity support conclusions accumulated from both biophysical studies and cellular pharmacology experiments. On this basis we suggest that this structure provides a firm foundation for the template-directed approach[11,12] and, hence, the design of improved or tailored cytotoxic agents with recognition and bonding preferences for heterogeneous DNA sequences.

Experimental Section

Structure Refinement and Energy Minimization. Initial coordinates for the d(CICGATCICG) 10-mer duplex and the mutated DNA sequences (see text) were generated for an idealized B-DNA conformation, selected on the basis of the reported NMR data.[15] Interactive molecular modeling (GEMINI 2.01 package) and structure refinement were performed using a Silicon Graphics Indigo workstation.

Models for subsequent structure refinement and energy minimization were constructed using coordinates taken from the crystal structures of PBD monomers.[25] The two strands of the self-complementary duplex were treated as equivalent for modeling purposes. Initial symmetric models for the cross-linked adduct were generated by docking the ligand within the minor groove with (i) 1:1 stoichiometry, (ii) the concave surface of the molecule presented by the imine moieties facing the convex groove floor in a close isohelical fit, (iii) the C11/C11' atoms of the PBD dimer (Figure 1) positioned adjacent to the exocyclic N2 atoms of Gua4 and Gua14, respectively, at an appropriate C11···Gua(N2) covalent separation of 1.45 Å, and (iv) a fully extended conformation for the 1,3-propanedioxy linkage. A rigid-body refinement procedure was used to align the PBD residues with the walls of the minor groove and to remove unfavorable atomic contacts.

The energy was minimized at the all-atom level by using the X-PLOR 3.1 program.[17] Rationalized molecular electrostatic potentials (MEPs) were used for all DNA bases; MEPs for the 2'-deoxyinosine (Ino) and 2'-deoxy-N^2-methylguanosine residues were determined from the semiempirical MNDO wave functions for the corresponding 3',5'-[$CH_3OP(O_2^-)$]-disubstituted 2'-deoxynucleosides using a published procedure.[26] Such fragments were constructed using idealized B-DNA coordinates, with a methyl group replacing the exocyclic hydrogen (N2)−H22 not involved in Watson−Crick base pairing for the N2-alkylated guanosine. Additional force field parameters for inosine and N^2-methylguanosine were interpolated from the default X-PLOR values for the purine nucleotides.

Atom-centered MEPs were similarly calculated for the isolated PBD dimer molecule and the neutral 11,11'-diamine adduct (i.e., C11/11'-NH_2), generated by formal addition of NH_3 to the C11−N10 imines of DSB-120, for both *all-R*, *all-S*, or

mixed stereochemistry conformations. Force field parameters for the ligand were interpolated from related studies from these laboratories.[11,18]

The stepwise protocol for structure refinement used (i) 400 steps of conjugate gradient (Powell mechanics) minimization to remove initial bad contacts, (ii) molecular dynamics with heating from 10 to 300 K during 10 ps (0.2-fs time step), (iii) 100-ps MD (1-fs time step) with coupling to a heat bath at 300 ± 5 K (friction coefficient = 100 ps^{-1}) and sampling at 0.2-ps intervals, and finally (iv) unrestrained Powell MM minimization of the rms-averaged snapshots to an ultimate rms gradient of ≤0.10 kcal/mol Å. Soft glycosidic torsion angle constraints (10 kcal/mol rad^2) were applied throughout the refinement for all nucleotides, with χ(pyrimidine/purine) values of −115/−110 ± 10°, respectively. The SHAKE algorithm was used in the dynamic annealing steps to maintain bond lengths. Planar restraints were used to maintain planarity for each DNA base, but explicit restraints were not required to maintain a Watson−Crick base-paired geometry. No attempt was made to influence refinements by restraining either internal or terminal base pairs.

The effects of solvent and counterions were simulated[27] by using a distance-dependent dielectric constant with $\epsilon = cr_{ij}$, with c = 1 for the Verlet MD in steps ii and iii and $c = 4$ for the Powell minimization steps i and iv. Nonbonded energy terms were included up to 11.5 Å, with switching between 9.5 and 10.5 Å, and a factor of 0.4 was used to damp 1,4-electrostatic interactions. Hydrogen-bonded interactions were switched on for heavy-atom donor to acceptor distances between 5.5 and 6.5 Å, and terms up to 7.5 Å were included.

Acknowledgment. This research was supported by the Cancer Research Campaign of the U.K. (to T.C.J., S.N., and D.E.T.) and grants from the Welch Foundation and Public Health Service (CA-49751 to L.A.H.). We are grateful to Drs. J. A. Hartley (UCL Medical School, London) and J. O. Trent (ICR) for valuable discussions.

Supplementary Material Available: Five tables giving MEP charges for the nucleotides, helical parameters for the adduct, intermolecular NOE distances used for structure refinement, and final atomic coordinates for the cross-linked adduct (24 pages). Ordering information is given on any current masthead page.

References

(1) (a) Hurley, L. H.; Boyd, F. L. DNA as a Target for Drug Action. *TIPS* **1988**, *9*, 402−407. (b) Hurley, L. H. DNA and Associated Targets for Drug Design. *J. Med. Chem.* **1989**, *32*, 2027−2033.

(2) (a) Dervan, P. B. Design of Sequence-Specific DNA-Binding Molecules. *Science* **1989**, *232*, 464−471. (b) Thurston, D. E.; Thompson, A. S. The Molecular Recognition of DNA. *Chem. Br.* **1990**, *26*, 767−772.

(3) (a) Remers, W. A. In *The Chemistry of Antitumor Antibiotics*; Wiley: New York, 1988; Vol. 2, pp 28−92. (b) Thurston, D. E. Advances in the Study of Pyrrolo[2,1-*c*][1,4]benzodiazepine (PBD) Antibiotics. In *Molecular Aspects of Anticancer Drug−DNA Interactions*; Neidle, S., Waring, M. J., Eds.; Macmillan Press: London, 1993; pp 54−88. (c) Mountzouris, J. A.; Hurley, L. H. Sequence Specificity of the Pyrrolo(1,4)benzodiazepines. In *Advances in DNA Sequence-Selective Agents*; Padwa, A., Ed.; JAI Press: London, 1992; Vol. 1, pp 263−292. (d) Thurston, D. E.; Hurley, L. H. A Rational Basis for Development of Antitumor Agents in the Pyrrolo[1,4]benzodiazepine Group. *Drugs Future* **1983**, *8*, 957−971. (e) Hurley, L. H.; Thurston, D. E. Pyrrolo-[1,4]benzodiazepine Antitumor Antibiotics: Chemistry, Interaction with DNA, and Biological Implications. *Pharm. Res.* **1984**, *2*, 52−59. (f) Thurston, D. E.; Bose, D. S. Synthesis of DNA-Interactive Pyrrolo[2,1-*c*][1,4]benzodiazepines. *Chem. Rev.* **1994**, *94*, 433−465.

(4) (a) Hurley, L. H.; Petrusek, R. L. Proposed Structure of the Anthramycin−DNA Adduct. *Nature* **1979**, *282*, 529−531. (b) Petrusek, R. L.; Anderson, G. L.; Garner, T. F.; Fannin, Q. L.; Kaplan, D. J.; Zimmer, S. G.; Hurley, L. H. Pyrrolo[1,4]-benzodiazepine Antibiotics. Proposed Structures and Characteristics of the In Vitro Deoxyribonucleic Acid Adducts of Anthramycin, Tomaymycin, Sibiromycin and Neothramycins A and B. *Biochemistry* **1981**, *20*, 1111−1119. (c) Petrusek, R. L.; Uhlenhopp, E. L.; Duteau, N.; Hurley, L. H. Reaction of Anthramycin with DNA. *J. Biol. Chem.* **1982**, *257*, 6207−6216.

(5) (a) Hertzberg, R. P.; Hecht, S. M.; Reynolds, V. L.; Molineux, I. J.; Hurley, L. H. DNA Sequence Specificity of the Pyrrolo[1,4]-benzodiazepine Antitumor Antibiotics. Methidiumpropyl−EDTA−Iron(II) Footprinting Analysis of DNA Binding Sites for Anthramycin and Related Drugs. *Biochemistry* **1986**, *25*, 1249−1258. (b) Hurley, L. H.; Reck, T.; Thurston, D. E.; Langley, D. R.; Holden, K. G.; Hertzberg, R. P.; Hoover, J. R. E.; Gallagher, G., Jr.; Faucette, L. F.; Mong, S.-M.; Johnson, R. K. Pyrrolo[1,4]-benzodiazepine Antitumor Antibiotics: Relationship of DNA Alkylation and Sequence Specificity to the Biological Activity of Natural and Synthetic Compounds. *Chem. Res. Toxicol.* **1988**, *1*, 258−268. (c) Boyd, F. L.; Stewart, D.; Remers, W. A.; Barkley, M. D.; Hurley, L. H. Characterization of a Unique Tomaymycin−d(CICGAATTCICG)$_2$ Adduct Containing Two Drug Molecules per Duplex by NMR, Fluorescence, and Molecular Modeling Studies. *Biochemistry* **1990**, *29*, 2387−2403.

(6) Kizu, R.; Draves, P. H.; Hurley, L. H. Correlation of DNA Sequence Specificity of Anthramycin and Tomaymycin with Reaction Kinetics and Bending of DNA. *Biochemistry* **1993**, *32*, 8712−8722.

(7) Puvvada, M. S.; Hartley, J. A.; Jenkins, T. C.; Thurston, D. E. A Quantitative Assay to Measure the Relative DNA-Binding Affinity of Pyrrolo[2,1-*c*][1,4]benzodiazepine (PBD) Anti-Tumor Antibiotics Based on the Inhibition of Restriction Endonuclease *Bam*HI. *Nucleic Acids Res.* **1993**, *21*, 3671−3675.

(8) (a) Kohn, K. W. DNA Damage in Mammalian Cells. *Bioscience* **1981**, *31*, 593−597. (b) Kohn, K. W. DNA Cross-Linking Agents. In *Development of Target-Oriented Anticancer Agents*; Chen, Y.-C., Goz, B., Minkoff, M., Eds; Raven Press: New York, 1983; pp 181−188.

(9) Sun, D.; Hurley, L. H. Analysis of the Monoalkylation and Cross-Linking Sequence Specificity of Bizelesin, a Bifunctional Alkylation Agent Related to (+)-CC-1065. *J. Am. Chem. Soc.* **1993**, *115*, 5925−5933.

(10) (a) Hartley, J. A.; Bingham, J. P.; Souhami, R. L. DNA Sequence Selectivity of Guanine−N7 Alkylation by Nitrogen Mustards is Preserved in Intact Cells. *Nucleic Acids Res.* **1992**, *20*, 3175−3178. (b) Lee, M.; Rhodes, A. L.; Wyatt, M. D.; Forrow, S.; Hartley, J. A. Design, Synthesis, and Biological Evaluation of DNA Sequence and Minor Groove Selective Alkylating Agents. *Anti-Cancer Drug Des.* **1993**, *8*, 173−192.

(11) (a) Bose, D. S.; Thompson, A. S.; Ching, J.; Hartley, J. A.; Berardini, M.; Jenkins, T. C.; Neidle, S.; Hurley, L. H.; Thurston, D. E. Rational Design of a Highly Efficient Non-Reversible DNA Interstrand Cross-Linking Agent Based on the Pyrrolobenzodiazepine Ring System. *J. Am. Chem. Soc.* **1992**, *114*, 4939−4941. (b) Bose, D. S.; Thompson, A. S.; Smellie, M.; Berardini, M. D.; Hartley, J. A.; Jenkins, T. C.; Neidle, S.; Thurston, D. E. Effect of Linker Length on DNA-Binding Affinity, Cross-Linking Efficiency and Cytotoxicity of C8-Linked Pyrrolobenzodiazepine Dimers. *J. Chem. Soc., Chem. Commun.* **1992**, 1518−1520. (c) Jenkins, T. C.; Neidle, S.; Thurston, D. E. Development of Anthramycin-Based Sequence-Selective DNA Cross-Linking Agents. In *Chemistry of Heterocyclic Compounds*; Stibor, I., Ed.; Prague Institute of Chemical Technology: Prague, 1993; pp 173−179.

(12) Wang, J.-J.; Hill, G. C.; Hurley, L. H. Template-Directed Design of a DNA−DNA Cross-Linker Based upon a Bis-Tomaymycin−Duplex Adduct. *J. Med. Chem.* **1992**, *35*, 2995−3002.

(13) Hartley, J. A.; Berardini, M. D.; Souhami, R. L. An Agarose Gel Method for the Determination of DNA Interstrand Cross-Linking Applicable to the Measurement of the Rate of Total and Second-Arm Cross-Link Reactions. *Anal. Biochem.* **1991**, *193*, 131−134.

(14) Smellie, M.; Kelland, L. R.; Thurston, D. E.; Souhami, R. L.; Hartley, J. A. Cellular Pharmacology of Novel C8-Linked Anthramycin-Based Sequence-Selective DNA Minor Groove Cross-Linking Agents. *Br. J. Cancer* **1994**, *70*, 48−53.

(15) Mountzouris, J. A.; Wang, J.-J.; Thurston, D. E.; Hurley, L. H. Comparison of a DSB-120 DNA Interstrand Cross-Linked Adduct with the Corresponding Bis-tomaymycin Adduct: An Example of a Successful Template-Directed Approach to Drug Design Based upon the Monoalkylating Compound Tomaymycin. *J. Med. Chem.* **1994**, *37*, 3132−3140.

(16) Neidle, S.; Puvvada, M. S.; Thurston, D. E. The Relevance of Drug DNA Sequence Specificity to Anti-Tumour Activity. *Eur. J. Cancer* **1994**, *30A*, 567−568.

(17) Brünger, A. T. *X-PLOR version 3.1: A System for X-Ray Crystallography and NMR*; Yale University Press: New Haven, 1992.

(18) (a) Lane, A. N.; Jenkins, T. C.; Brown, T.; Neidle, S. Interaction of Berenil with the *Eco*RI Dodecamer d(CGCGAATTCGCG)$_2$ in Solution Studied by NMR. *Biochemistry* **1991**, *30*, 1372−1385. (b) Jenkins, T. C.; Lane, A. N.; Neidle, S.; Brown, D. G. NMR and Molecular Modeling Studies of the Interaction of Berenil and Pentamidine with d(CGCAAATTTGCG)$_2$. *Eur. J. Biochem.* **1993**, *213*, 1175−1184. (c) Jenkins, T. C.; Parrick, J.; Porssa, M. DNA-Binding Properties of Nitroarene Oligopeptides Designed as Hypoxia-Selective Agents. *Anti-Cancer Drug Des.* **1994**, *9*, 477−493.

(19) (a) Boehncke, K.; Nonella, M.; Schulten, K.; Wang, A. H.-J. Molecular Dynamics Investigation of the Interaction Between DNA and Distamycin. *Biochemistry* **1991**, *30*, 5465–5475. (b) Pearlman, D. A.; Kollman, P. A. Are Time-Averaged Restraints Necessary for Nuclear Magnetic Resonance Refinement? A Model Study for DNA. *J. Mol. Biol.* **1991**, *220*, 457–479. (c) McConnell, K. J.; Nirmala, R.; Young, M. A.; Ravishanker, G.; Beveridge, D. L. A Nanosecond Molecular Dynamics Trajectory for a B-DNA Double Helix: Evidence for Substates. *J. Am. Chem. Soc.* **1994**, *116*, 4461–4462.

(20) (a) Krugh, T. R.; Graves, D. E.; Stone, M. P. Two-dimensional NMR Studies on the Anthramycin–d(ATGCAT)$_2$ adduct. *Biochemistry* **1989**, *28*, 9988–9994. (b) Boyd, F. L.; Cheatham, S. F.; Remers, W.; Hill, G. C.; Hurley, L. H. Characterization of the Structure of the Anthramycin–d(ATGCAT)$_2$ Adduct by NMR and Molecular Modeling Studies. Determination of the Stereochemistry at the Covalent Linkage Site, Orientation in the Minor Groove, and Effect of Drug Binding on Local DNA Structure. *J. Am. Chem. Soc.* **1990**, *112*, 3279–3289.

(21) (a) Kopka, M. L.; Goodsell, D. S.; Grzeskowiak, K.; Cascio, D.; Dickerson, R. E. Crystal Structure of Anthramycin/C–C–A–A–C–G–T–T–G–G: Covalent Drug Binding Within the Minor Groove. Presented at Workshop on DNA-Drug Interactions, Instituto Juan March, Madrid, November 15–17, 1993. (b) Kopka, M. L.; Goodsell, D. S.; Grzeskowiak, K.; Baikalov, I.; Cascio, D.; Dickerson, R. E. Crystal Structure of a Covalent DNA–Drug Adduct: Anthramycin Bound to C–C–A–A–C–G–T–T–G–G, and a Molecular Explanation of Specificity. *Biochemistry* **1994**, in press.

(22) Neidle, S. Minor-Groove Width and Accessibility in B-DNA Drug and Protein Complexes. *FEBS Lett.* **1992**, *298*, 97–99.

(23) (a) Nunn, C. M.; Jenkins, T. C.; Neidle, S. Crystal Structure of d(CGCGAATTCGCG) Complexed with Propamidine, a Short-Chain Homologue of the Drug Pentamidine. *Biochemistry* **1993**, *32*, 13838–13843. (b) Greenidge, P. A.; Jenkins, T. C.; Neidle, S. DNA Minor Groove Recognition Properties of Pentamidine and its Analogues: A Molecular Modeling Study. *Mol. Pharmacol.* **1993**, *43*, 982–988.

(24) Walker, W. L.; Kopka, M. L.; Filipowsky, M. E.; Dickerson, R. E.; Goodsell, D. S. Design of B-DNA Cross-Linking and Sequence-Reading Molecules. *Biochemistry* **1994**, in press.

(25) (a) Arora, S. K. Structural Investigations of the Mode of Action of Drugs. II. Molecular Structure of Anthramycin Methyl Ether Monohydrate. *Acta Crystallogr.* **1979**, *B35*, 2945–2948. (b) Arora, S. K. Structure of Tomaymycin, a DNA-Binding Antitumor Antibiotic. *J. Antibiot.* **1981**, *34*, 462–464.

(26) Lane, A. N.; Jenkins, T. C.; Brown, D.; Brown, T. NMR Determination of the Solution Conformation and Dynamics of the A·G Mismatch in the d(CGCAAATTGGCG)$_2$ Dodecamer. *Biochem. J.* **1991**, *279*, 269–281.

(27) Orozco, M.; Laughton, C. A.; Herzyk, P.; Neidle, S. Molecular Mechanics Modelling of Drug–DNA Structures: The Effects of Differing Dielectric Treatment on Helix Parameters and Comparison with a Fully Solvated Structural Model. *J. Biomol. Struct. Dyn.* **1990**, *8*, 359–373.