The DNA Sequence at Echinomycin Binding Sites Determines the Structural Changes Induced by Drug Binding: NMR Studies of Echinomycin Binding to $[d(ACGTACGT)]_2$ and $[d(TCGATCGA)]_2$†

Dara E. Gilbert and Juli Feigon*

Department of Chemistry and Biochemistry and Molecular Biology Institute, 405 Hilgard Avenue, University of California, Los Angeles, California 90024

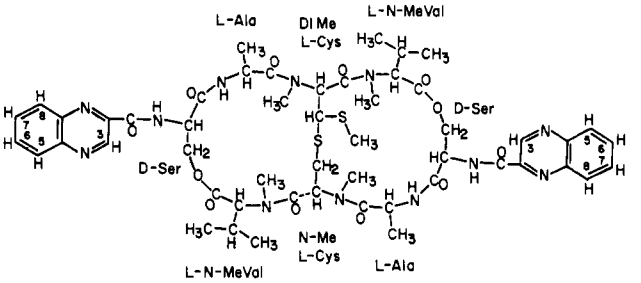
Received August 13, 1990; Revised Manuscript Received November 29, 1990

ABSTRACT: The complexes formed between the cyclic octadepsipeptide antibiotic echinomycin and the two DNA octamers $[d(ACGTACGT)]_2$ and $[d(TCGATCGA)]_2$ have been investigated by using one- and two-dimensional proton NMR spectroscopy techniques. The results obtained for the two complexes are compared to each other, to the crystal structures of related DNA–echinomycin complexes, and to enzymatic and chemical footprinting results. In the saturated complexes, two echinomycin molecules bind to each octamer by bisintercalation of the quinoxaline moieties on either side of each CpG step. Binding of echinomycin to the octamer $[d(ACGTACGT)]_2$ is cooperative so that only the two-drug complex is observed at lower drug–DNA ratios, but binding to $[d(TCGATCGA)]_2$ is not cooperative. At low temperatures, both the internal and terminal A·T base pairs adjacent to the binding site in the $[d(ACGTACGT)]_2$–2 echinomycin complex are Hoogsteen base paired (Gilbert et al., 1989) as observed in related crystal structures. However, as the temperature is raised, the internal A·T Hoogsteen base pairs are destabilized and are observed to be exchanging between the Hoogsteen base-paired and an open (or Watson–Crick base-paired) state. In contrast, in the $[d(TCGATCGA)]_2$–2 echinomycin complex, no A·T Hoogsteen base pairs are observed, the internal A·T base pairs appear to be stabilized by drug binding, and the structure of the complex does not change significantly from 0 to 45 °C. Thus, the structure and stability of the DNA in echinomycin–DNA complexes depends on the sequence at and adjacent to the binding site. While we conclude that no single structural change in the DNA can explain all of the footprinting results, unwinding of the DNA helix in the drug–DNA complexes appears to be an important factor while Hoogsteen base pair formation does not.

Echinomycin is a cyclic octadepsipeptide antibiotic that contains a thioacetal cross bridge and two quinoxaline rings attached to the two D-Ser residues (Chart I). Echinomycin is active against Gram-positive bacteria and displays potent antitumor activity as well (Lee & Waring, 1978; Waring & Wakelin, 1974). The first physical characterizations of the drug's interaction with DNA showed that it bound tightly as a bisintercalator (Waring & Wakelin, 1974). Subsequent footprinting studies from the laboratories of Waring and Dervan showed that the binding site was centered around CpG steps. Results of DNase I and II footprints indicated a maximum binding site size of 6–8 bases (Low et al., 1984), while MPE–Fe(II) footprinting studies showed smaller binding sites of 4 bases with the sequence NCGN, the strongest of which, on the basis of the limited sampling of 4-base sequences in the restriction fragment DNA, were TCGT and ACGT (van Dyke & Dervan, 1984). The footprinting studies with DNase I and II also showed that A·T-rich regions adjacent and distal to the echinomycin binding sites were more readily cleaved than those in the free DNA.

Concurrent with the DNase I and II MPE-Fe(II) footprinting results, the crystal structure of triostin A, a closely related echinomycin analogue, with d(CGTACG) was solved (Wang et al., 1984). A related structure was subsequently

Chart I



reported for echinomycin complexed with d(CGTACG) (Ughetto et al., 1985). Consistent with the strong binding sites determined from the footprinting, two echinomycins were found to bind per duplex, with the quinoxaline rings intercalating (or stacking at the ends of the hexamer) on each side of the CpG steps. The unexpected structural feature of these complexes was that the A·T bases next to the intercalation sites were Hoogsteen base paired with the adenines in the syn conformation. A crystal structure of a triostin A–d(GCGTACGC) complex also showed Hoogsteen base pairs for the A·T as well as terminal G·C bases (Quigley et al., 1986; Wang et al., 1986). This alternative base-pairing scheme was proposed to be the basis of the enhanced sensitivity to DNase I and II of DNA bases adjacent to echinomycin binding sites (Mendel & Dervan, 1987; Wang et al., 1984; Ughetto et al., 1985).

Echinomycin binding to DNA has more recently also been probed by footprinting with the chemical reagent diethyl pyrocarbonate (DEPC). Both Mendel and Dervan (1987) and

†This work was supported by grants from NIH (R01 GM 37254-01) and ONR (Contract No. N00014-88-K-0180) and an NSF Presidential Young Investigator Award with matching funds from AmGen Inc., Sterling Drug Inc., E. I. du Pont de Nemours and Co., Inc., and Monsanto Co. to J.F. and by a USPHS national research service award (GM 07185) and a UCLA dissertation year fellowship to D.E.G.

0006-2960/91/0430-2483$02.50/0

Portugal et al. (1988) found that DNA bases adjacent to echinomycin binding sites were hypersensitive to cleavage by DEPC. However, Mendel and Dervan (1987) concluded that the results were consistent with Hoogsteen base pairing while Portugal et al. (1988) concluded that the results were inconsistent with Hoogsteen base pairs at these sites. Chemical modification with the thymine-specific reagents OsO_4 and $KMnO_4$ showed that regions adjacent to echinomycin binding sites are also hypersensitive to these reagents and that this hypersensitivity is dependent on the sequence adjacent to the binding site (McClean & Waring, 1988; McClean et al., 1989). Experiments with DNA fragments containing adenines modified at N7 so that they are incapable of forming Hoogsteen base pairs are still hypersensitive to OsO_4 (McClean et al., 1989).

The questions remain whether any echinomycin–DNA complexes contain Hoogsteen base pairs in solution and what structures give rise to the footprinting results. NMR studies of complexes of echinomycin with the DNA tetramers [d(TCGA)]$_2$ and [d(ACGT)]$_2$ indicated that Hoogsteen A·T base pairs formed in the [d(ACGT)]$_2$ complex but not in the [d(TCGA)]$_2$ complex (Gao & Patel, 1988). Subsequent studies by Gao and Patel showed that [d(GCGC)]$_2$ complexes formed Hoogsteen base pairs in a pH-dependent manner while [d(CCGG)]$_2$ did not and that the complex [d(AAACGTTT)]$_2$–echinomycin had no Hoogsteen base pairs (Gao & Patel, 1989).

As part of a program to study the sequence-specific binding of echinomycin and related derivatives, we have investigated the binding of echinomycin to the DNA octamers [d(ACGTACGT)]$_2$ and [d(TCGATCGA)]$_2$ as well as the hexamer [d(CGTACG)]$_2$. These longer sequences contain adjacent binding sites and allow us to address certain questions that cannot be addressed in the study of complexes with only a single binding site. In our initial report on the echinomycin complex with [d(ACGTACGT)]$_2$ (Gibert et al., 1989), we showed that two drugs bound per duplex at the CpG steps and that the binding was cooperative. At low temperatures, both the internal and terminal A·T bases in the complex are found to be Hoogsteen base paired. However, as the temperature is raised, only the terminal A·T bases remain stably Hoogsteen base paired, while those in the interior of the duplex appear to be exchanging between a Hoogsteen base-paired and an open (or possibly Watson–Crick) state. It was concluded that the physiologically relevant structure for A·T bases in this binding site in vivo might therefore not be Hoogsteen base pairs.

In this paper we present the complete assignment procedures for the [d(ACGTACGT)]$_2$–2 echinomycin complex, hereafter referred to as complex A, as well as more complete details on the structure of the complex. These results are compared to those obtained and reported here for the DNA–drug complex [d(TCGATCGA)]$_2$–2 echinomycin, hereafter referred to as complex T. For both complexes, bisintercalation at the same CpG binding sites is observed and the sugar pucker of the nucleotides in the binding site is the same. Similar contacts between peptide resonances and the minor-groove markers are observed for both complexes. In contrast to complex A, however, no Hoogsteen base pairs are observed in complex T, and the central A·T base pairs appear to be more stable in complex T than in the free DNA. In addition, binding of echinomycin to [d(TCGATCGA)]$_2$ is not cooperative, so at low drug:DNA ratios both one and two drugs bound per duplex are observed. Thus, the conformation and stability of DNA in echinomycin–DNA complexes depends on the sequence at and adjacent to the binding site. While no single structural change appears adequate to explain all of the footprinting results, the primary determinant of the enhanced sensitivity of DNA bases near echinomycin binding sites to chemical reagents and DNase I and II would appear to be the unwinding of the DNA duplex adjacent to the intercalation sites of the drug.

Materials and Methods

Sample Preparation. The octamer d($A_1C_2G_3T_4A_5C_6G_7T_8$) and hexamer d($C_1G_2T_3A_4C_5G_6$) were synthesized and purified as described previously (van der Marel et al., 1981). Additional d(ACGTACGT) and the octamer d($T_1C_2G_3A_4T_5C_6G_7A_8$) were synthesized on an ABI 381A synthesizer using β-cyanoethyl phosphoramidite chemistry on the 10-μmol scale. The octamers were purified by gel filtration according to the method of Kintinar et al. (1987). The DNA was converted to the sodium form by passage over a Bio-Rad AG 50W-X4 column. The [d(ACGTACGT)]$_2$ sample contained 2 mM DNA and 50 mM NaCl, pH 6.5 (meter reading) with no added buffer in 400 μL. The [d(TCGATCGA)]$_2$ sample contained 1.3 mM DNA and 50 mM NaCl, pH 6.5 (meter reading) with no added buffer in 400 μL. The samples were lyophilized once from water and then twice from 99.96% D_2O, dried a third time in the NMR tube under a stream of N_2 gas, and finally dissolved in 400 μL of 99.996% D_2O. Spectra in H_2O were obtained on the same samples after redrying in the NMR tube under N_2 gas and redissolving in 90% H_2O/10% D_2O. Partially deuterated (at purine H8) DNA samples were prepared as in Gilbert et al. (1989).

The echinomycin used was a gift from the National Cancer Institute or was purchased from Fermical Inc., Detroit, MI. Saturated complexes of two drugs per duplex were formed by adding 2 equiv of echinomycin in methanol to the DNA sample in the NMR tube and slowly drying the sample in the tube under a stream of N_2 gas for 12–18 h. For experiments in D_2O, the dried sample was redissolved in D_2O, redried, and redissolved in 99.996% D_2O. For spectra in H_2O, the dried sample was redissolved in 90% H_2O/10% D_2O. Partially saturated complexes were prepared by adding the desired number of equivalents of echinomycin in methanol and drying similarly. Complete complex formation was monitored by NMR.

NMR Spectroscopy. All NMR experiments were done on a General Electric GN500 (500.119 MHz, ^1H) spectrometer. Chemical shifts were referenced to the chemical shift of water, which had been calibrated relative to 2,2-dimethyl-2-silapentane-5-sulfonate (DSS). Phase-sensitive nuclear Overhauser effect (NOESY) spectra in D_2O were obtained by using the method of States et al. (1982) and the standard pulse sequence (Kumar et al., 1980) with preirradiation of the solvent during the recycle delay. A total of 1024 complex points were collected in t_2 and 200–350 points in t_1. Phase-sensitive NOESY spectra in H_2O were acquired by replacing the last pulse of the standard sequence with 1$\bar{1}$ spin-echo pulse sequence and phase cycling appropriately to suppress the large solvent resonance (Sklenář & Bax, 1987). The carrier was set on the water resonance, and the delay τ was adjusted so that the excitation maximum was centered at the imino resonances. HOHAHA spectra were acquired by using the MLEV 17 mixing scheme and two 2-ms trim pulses for the spin lock (Bax & Davis, 1985). COSY (Aue et al., 1984) and DQF-COSY (Rance et al., 1983) spectra were acquired with standard pulses sequences and phase cycling. All 2D NMR spectra were processed on a VAX 8800 using the Fortran program FTNMR (Hare Research). Detailed descriptions of the acquisition and

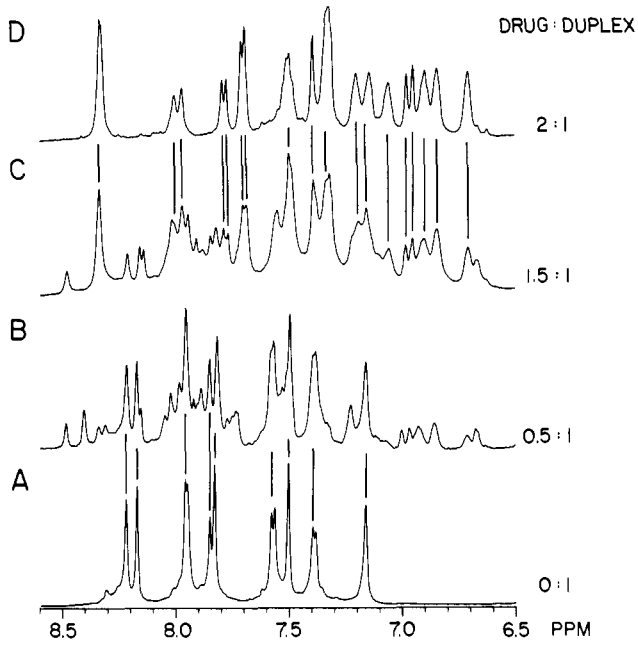


FIGURE 1: Proton NMR spectra of the aromatic region of [d-(TCGATCGA)]$_2$–echinomycin as a function of added drug concentration at 30 °C in D_2O. DNA:echinomycin ratios of (A) 1:0, (B) 1:0.5, (C) 1:1.5, and (D) 1:2 are shown. Solid lines connect the free DNA resonances to the free DNA resonances in the spectrum of the 1:0.5 complex and the resonances of the fully saturated complex to the resonances from the fully saturated complex in the spectrum of the 1:1.5 complex. Echinomycin (0.5 equiv) in methanol was added to the free DNA (1.3 mM duplex, 50 mM NaCl, pH 6.5) in the NMR tube. After each addition, the sample was dried down and redissolved in 99.996% D_2O. The spectra were acquired with a sweep width of 5000 Hz in 8K complex points. Data were line broadened by 3 Hz prior to Fourier transformation.

processing parameters for each experiment are given in the figure captions.

Results

Echinomycin–DNA Complex Formation. Spectra of the aromatic resonances of [d(TCGATCGA)]$_2$ as a function of drug concentration are given in Figure 1. At drug:DNA ratios below 2:1, many more new resonances are observed, in addition to the free DNA resonances, than could be accounted for if only one type of drug–DNA complex were being formed. These correspond to two 1:1 drug-duplex complexes as well as the fully saturated 2:1 drug–duplex complex. As increasing amounts of drug are added, the number of complexes with two drugs bound increases until all the echinomycin binding sites are saturated. At a drug:DNA ratio of 2:1, all of the resonances from the free DNA and the 1:1 complexes have disappeared and only the 2:1 complex resonances are observed.

In NOESY spectra obtained on freshly prepared 2:1 echinomycin–[d(TCGATCGA)]$_2$ complexes, only cross peaks from the fully saturated complex are observed (see Figure 5). However, approximately 2 days after the complex is made, cross peaks from 1:1 echinomycin–[d(TCGATCGA)]$_2$ complexes and free DNA also appear in the spectra and excess echinomycin is observed as a precipitate in the NMR tube. The total percentage of 1:1 complexes and free DNA is approximately 10%. The fully saturated complex can be readily reformed by adding methanol to the sample, redrying it, and redissolving it in aqueous solution. NOESY spectra of the 1:1 echinomycin–[d(TCGATCGA)]$_2$ complex showed exchange cross peaks between the free DNA and resonances corresponding to the 1:1 and 2:1 drug–duplex complexes (supplementary material). However, no exchange cross peaks were

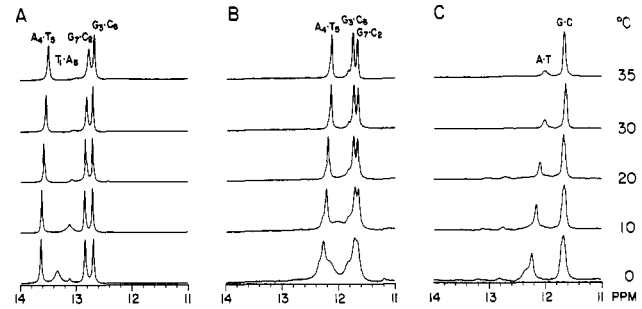


FIGURE 2: Imino proton spectra as a function of temperature. (A) Free [d(TCGATCGA)]$_2$, 1.3 mM duplex, 50 mM NaCl, pH 6.5. (B) [d(TCGATCGA)]$_2$–2 echinomycin, 1.3 mM duplex, 50 mM NaCl, pH 6.5. (C) [d(ACGTACGT)]$_2$–2 echinomycin, 2 mM duplex, 50 mM NaCl, pH 6.5. Assignments of the imino resonances are indicated. Spectra were acquired with a sweep width of 10000 Hz in 8K complex points. Spectra were obtained with the 1$\bar{1}$ spin echo pulse sequence (Sklenář & Bax, 1987) H_2O with the delay τ adjusted to give maximum excitation in the imino region, τ = 60 μs, Δ = 50 μs. All spectra were line broadened by 3 Hz prior to Fourier transformation.

observed between the free DNA and the 2:1 drug–duplex complex.

The results on complex formation between echinomycin and [d(TCGATCGA)]$_2$ are in marked contrast to those observed for [d(ACGTACGT)]$_2$ (Gilbert et al., 1989). In the latter case, at drug:DNA ratios below 2:1, only resonances corresponding to the free DNA or a 2:1 drug–duplex complex were observed. The complexes formed remain stable indefinitely. Under no conditions were exchange cross peaks between free DNA and the echinomycin–DNA complex observed.

Imino Proton Spectra of the Complexes as a Function of Temperature. Figure 2 shows the imino proton spectra of [d(TCGATCGA)]$_2$, [d(TCGATCGA)]$_2$–2 echinomycin, and [d(ACGTACGT)]$_2$–2 echinomycin as a function of temperature. The imino proton resonances shift upfield approximately 1 ppm when the complexes form (Figure 2B,C). This large upfield complexation shift is characteristic of in of intercalative binding (Feigon et al., 1984a). The free [d(TCGATCGA)]$_2$ melts as is expected for duplex DNA. First the terminal T imino resonance broadens and disappears, followed by the G_7 imino resonance, which begins to broaden at about 30 °C. The imino protons in complex T behave similarly (Figure 2b). The terminal resonance disappears first at about 10 °C. At 35 °C the G_7 imino resonance is sharper in the complex than it is in free DNA, indicating that the imino protons are exchanging less rapidly with the solvent in the complex than in the free DNA. The internal T imino resonance remains sharp over the entire temperature range shown. This is in striking contrast to complex A, in which the internal T imino proton resonances broaden and almost disappear by 35 °C (Figure 2C).

NOESY Spectra of the Nonexchangeable Resonances of the Complex [d(ACGTACGT)]$_2$–2 Echinomycin. A NOESY spectrum of complex A is shown in Figure 3. We anticipated that the standard pattern of cross peaks expected for B-DNA would not be observed for this complex, and this is indeed the case. The presence of Hoogsteen base pairs with adenine in the syn conformation [discussed below and in Gilbert et al. (1989)] disrupts the base–H1′,H2′,H2″ connectivities normally used to make sequential assignments in right-handed B-DNA helices, as has previously been observed for Z-DNA structures (Feigon et al., 1984b). The sequential assignments are also disrupted by the intercalation of the quinoxaline rings.

In addition to the intraduplex NOEs seen for the octamer, there are intramolecular contacts for the echinomycin that define the structure of the drug in the complex as well as many intermolecular cross peaks that define the interactions between

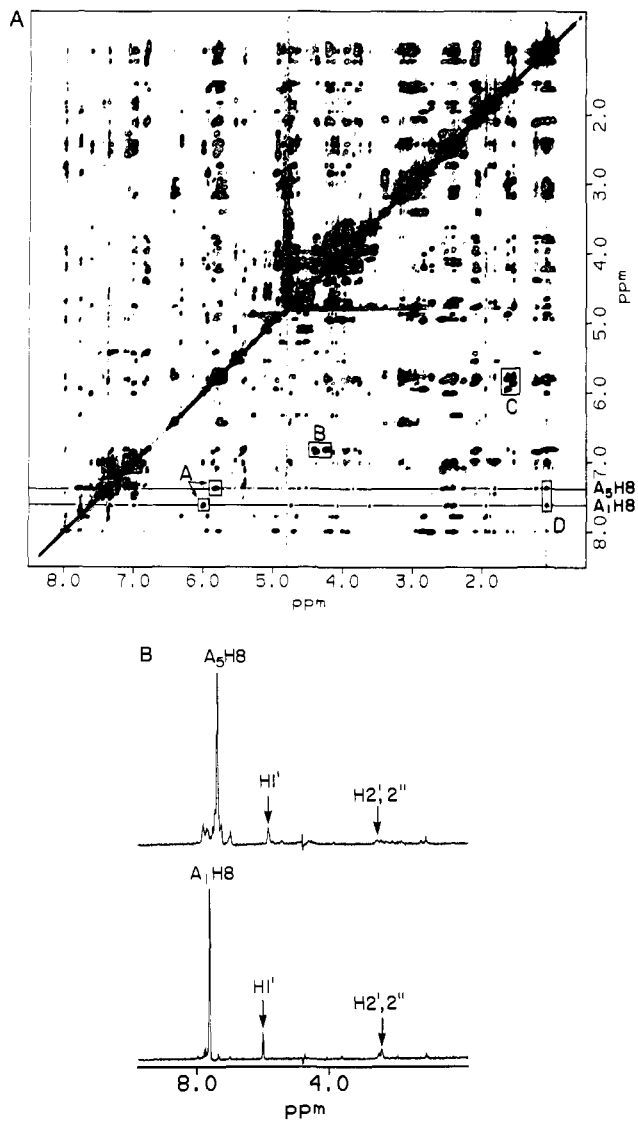


FIGURE 3: (A) NOESY spectrum of d[(ACGTACGT)]$_2$–2 echinomycin in 50 mM NaCl, pH 6.5 in D_2O, 20 °C. The sample contained 2 mM complex. The lines indicate the positions of A_1H8 and A_5H8 in the large spectrum. Several important cross peaks are boxed: (A) the H8–H1′ cross peaks for A_1 and A_5, (B) the H6–H3′ cross peaks for C_2 and C_6, (C) the H1′–AlaCH$_3$ cross peaks for C_2, C_6, G_3, and G_7, and (D) the H8–ValCH$_3$ cross peaks for A_1 and A_5. The spectrum was acquired with a sweep width of 5000 Hz in both dimensions and a mixing time of 200 ms. A total of 2K points in t_2, 32 scans per t_1 block, and 266 t_1 values were acquired. The data in both dimensions were apodized by a skewed sine-bell squared function (skew = 1.5) phase-shifted by 60°. Two hundred sixty-six points were apodized in both dimensions. Data in t_1 were zero-filled to 1K points prior to Fourier transformation. (B) Cross sections of the NOESY spectrum of [d(ACGTACGT)]$_2$–2 echinomycin from Figure 3A, showing the cross peaks of A_1H8 and A_5H8. The H8–H1′ and H8–H2′,2″ cross peaks are indicated.

the DNA and the drug. These are discussed in detail below.

Assignments of the Nonexchangeable Resonances of the DNA in the Complex [d(ACGTACGT)]$_2$–2 Echinomycin. There were several major obstacles to overcome in assigning the DNA resonances in this complex, as indicated above. Due to the disruption of base–H1′,H2′,H2″ sequential connectivities (Feigon et al., 1983; Scheek et al., 1983; Hare et al., 1983), we had to rely on other methods to make the DNA assignments. Echinomycin binds essentially irreversibly to [d(ACGTACGT)]$_2$ on the NMR time scale and therefore, two-dimensional exchange spectroscopy could not be used to assign the DNA resonances in this complex. The protons were initially assigned to proton types. In the HOHAHA spectrum (supplementary material) there are two CH6–CH5 cross peaks seen in the base–H1′ region, and two TH6–TMe cross peaks seen in the base–H2′,2″ region. The purine H8 resonances were identified by the disappearance of their cross peaks in NOESY spectra taken after deuteration (Gilbert et al., 1989). This also allowed us to unambiguously distinguish the H8 and H2 resonances. We note that this is critical in determining the base-pairing scheme, since in standard B-DNA with Watson–Crick base pairs the AH2 resonances are usually identified on the basis of their strong cross peaks to their corresponding imino protons, while for a Hoogsteen base pair the imino would instead have a strong cross peak to AH8.

Once the protons were assigned as to type, it was possible to trace two short sequences of connectivities from TMe to GH8 to CH6 in the NOESY spectrum, but they could not be placed relative to each other. Fortunately the base and H1′ resonances of T_8 and A_1 could be assigned by comparing the NOESY spectrum of the [d(ACGTACGT)]$_2$–2 echinomycin complex to that of the [d(CGTACG)]$_2$–2 echinomycin complex (supplementary material). Two base–H1′ cross peaks appear in the spectrum of complex A that do not appear in the spectrum of the hexamer complex and can therefore be assigned to the terminal bases. Once T_8 and A_1 were identified, G_7 and C_6 could be assigned; C_2 could be assigned by connectivities from A_1H1′ to Q_1H7 to C_2H1′.

After the H1′ resonances were assigned, the rest of the sugar resonances for each base were assigned on the basis of scalar couplings in the HOHAHA and DQF-COSY spectra. We note that the 2′–2″ and 1′–2′ cross peaks for both C_2 and C_6 cannot be seen in the HOHAHA spectrum but are observed in the DQF-COSY spectrum (data not shown). The sugar resonances of A_4 and T_5 are quite broad due to exchange between conformations (discussed below), but they can be partially traced.

Assignment of the Exchangeable Resonances of the DNA in the Complex [d(ACGTACGT)]$_2$–2 Echinomycin. A portion of the NOESY spectrum of complex A in H_2O at 1 °C is shown in Figure 4A. The region shown includes the imino and aromatic resonances and all the cross peaks to them. The most upfield imino resonances have strong cross peaks to C_2 and C_6 amino protons and are assigned as the G imino resonances. The C aminos have a characteristically strong NOESY cross peak between the hydrogen-bonded and non-hydrogen-bonded protons in the aromatic region. The overlap of the two imino resonances is too great to distinguish G_3 and G_7. The downfield imino resonance has a cross peak to the T_4Me resonance, identifying it as the internal A·T imino resonance. The terminal A·T imino resonance at 12.32 ppm has a cross peak to the water resonance only at all temperatures. The chemical shifts of all the assigned DNA resonances in complex A are listed in Table I.

Assignment of the Drug Resonances in the Complex [d(ACGTACGT)]$_2$–2 Echinomycin. Initially the amino acid spin systems were identified in the HOHAHA spectrum (supplementary material). The usual sequential assignment procedure used for proteins and peptides (Wüthrich, 1986) could not be applied here because the amides of Val and Cys are methylated. Once the spin systems were identified, the NMe resonances of Val and Cys were assigned on the basis of NOE cross peaks to the C_αH resonances of Val and Cys. The quinoxaline 5, 6, 7, and 8 protons are easily identified because they are the only protons from either the drug or DNA that are expected to have COSY cross peaks to each other in the aromatic region. In this region of the spectrum, two sets of

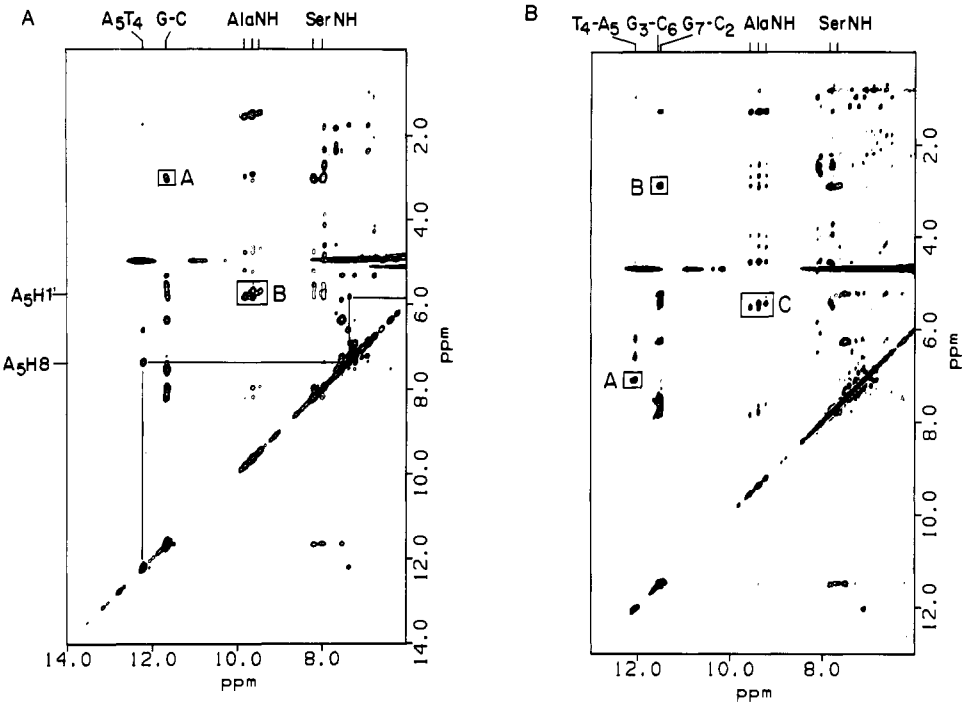


FIGURE 4: (A) Expanded region of the NOESY spectrum of $[d(ACGTACGT)]_2$–2 echinomycin, 50 mM NaCl, pH 6.5, in 90% H_2O/10% D_2O at 1 °C and τ_m = 50 ms, showing the imino and aromatic resonances and their cross peaks. The solid line connects the imino to H8 to H1′ resonances of A_5. The boxed cross peaks are (A) G imino–ValNCH$_3$ and (B) AlaNH–G_3 and G_7 H1′. The spectrum was acquired with a sweep width of 10000 Hz in both dimensions, and 266 t_1 values of 64 scans, 4K points. The final pulse in the NOESY experiment was replaced with the 1$\bar{1}$ spin-echo pulse sequence (Sklenář & Bax, 1987), τ = 60 μs, Δ = 50 μs. Prior to Fourier transformation the FIDs were corrected by using a Gaussian window function with K = 32 and extrapolation with M = 16 (Marion et al., 1989) to remove the residual water signal; the data in both dimensions were apodized by a skewed sine-bell squared function (skew = 1.5) phase-shifted by 60°. Two hundred sixty-six points were apodized in both dimensions, and data in t_1 were zero-filled to 2K points. (B) Expanded region of the NOESY spectrum, τ_m = 50 ms, in 90% H_2O/10% D_2O of the $[d(TCGATCGA)]_2$–2 echinomycin sample described in Figure 2B. The region of the spectrum including the imino and aromatic resonances and their cross peaks is shown. The boxed cross peaks are (A) T_5 imino–A_4H2, (B) G_3 and G_7 imino–ValNCH3, and (C) AlaNH–G_3 and G_7 H1′. The DNA imino and peptide amide protons are indicated. The spectrum was acquired and processed as in (A) above. Each t_1 spectrum was 128 scans, and 245 t_1 values were collected. The data were apodized with a skewed sine-bell squared function (skew = 1.4) phase-shifted by 70°. Two hundred forty-five points were apodized in each dimension, and data in t_1 were zero-filled to 2K.

Table I: Assignments of the DNA Resonances in the Complex $[d(ACGTACGT)_2]$–2 Echinomycin at 20 °C[a]

	H8,H6	H5,Me/H2	H1′	H2′	H2″	H3′	H4′	H5′,5″	amino[b]	imino[b]
A	7.61	7.37	6.00	2.42	2.53	4.77	4.10	3.63		
	7.63									
C	6.84	5.42	5.86	1.23	2.14	4.40	3.78	4.16, 3.98	7.53, 6.36	
G	7.98		5.78	2.74	2.74	4.72	3.43	4.25		11.67
T	6.97	1.83	5.54	1.21	2.11	4.80	4.42	3.75		12.22
A	7.38	7.37	5.83	2.43	2.58	5.13				
C	6.82	5.40	5.79	0.98	2.41	4.25	3.82		7.53, 6.30	
	6.80									
G	7.98		5.95	2.88	2.88	4.99	4.21	4.03		11.64
T	7.77	1.95	6.33	2.37	2.41	4.68	9.18			12.31

[a] Chemical shifts relative to DSS. [b] Chemical shifts at 1 °C.

cross peaks for quinoxaline (Q) protons 5–8 can be traced. The remaining two unassigned echinomycin resonances in the aromatic region have cross peaks to the GH8 protons. These are therefore identified as the two QH3 resonances, since they are the only aromatic resonances, other than the already assigned CH6 resonances, that might have cross peaks to the GH8. Each set of quinoxaline resonances corresponds to a quinoxaline ring at either end of the peptide ring. The ring designated 1 has NOE cross peaks to A_1 and C_2 and ring 2 has cross peaks to A_5 and C_6. Similarly, the peptide resonances were distinguished on the basis of NOEs between the drug and DNA. Amino acids designated 1 have cross peaks to bases 1–4 and those designated 2 have cross peaks to bases 5–8. (See Figure 7 for the numbering scheme.) The complete assignments of the echinomycin resonances in complex A are listed in Table II.

NOESY Spectra of the Complex $[d(TCGATCGA)]_2$–2 Echinomycin. A NOESY spectrum of the complex d-$(TCGATCGA)]_2$–2 echinomycin at 25 °C is shown in Figure 5. Spectra obtained on this complex were generally better resolved than those of complex A. This is at least partially attributable to the fact that only a single stable structure, with no Hoogsteen base pairs, is formed for this DNA sequence (see below and Discussion). As expected, intercalation of the quinoxaline disrupts some of the standard B-DNA sequential connectivities used to make sequential assignments of DNA (Feigon et al., 1983; Hare et al., 1983). It was possible to assign the DNA resonances in this complex by using NOESY and HOHAHA spectra without recourse to deuteration experiments.

In addition to the intraduplex NOEs seen for the DNA, there are intramolecular contacts for the echinomycin that

Table II: Chemical Shifts of the Echinomycin Resonances in the Complex [d(ACGTACGT)]$_2$–2 Echinomycin at 20 °C[a]

		α	β	β′	γ	NH[b]	NCH_3	SCH_3
serine	1	5.28	4.85	4.62		7.98		
	2	5.27	4.88	4.52		8.19		
valine	1	4.89, 4.87	2.44		1.07		3.19	
	2	4.67, 4.65	2.42		1.09		3.16	
cysteine	1	6.45	4.79			2.97		2.09
		6.40	4.76				2.92	2.07
	2	5.77	3.43	2.85			2.93	
alanine	1	4.78	1.56			9.63, 9.46		
	2	4.89	1.64			9.81, 9.63		

		3	5	6	7	8
quinoxaline	1	7.77	7.55	7.39	7.02	7.02
	2	7.45	7.50	7.25	7.08	7.00

[a] Chemical shifts relative to DSS. [b] Chemical shifts at 1 °C.

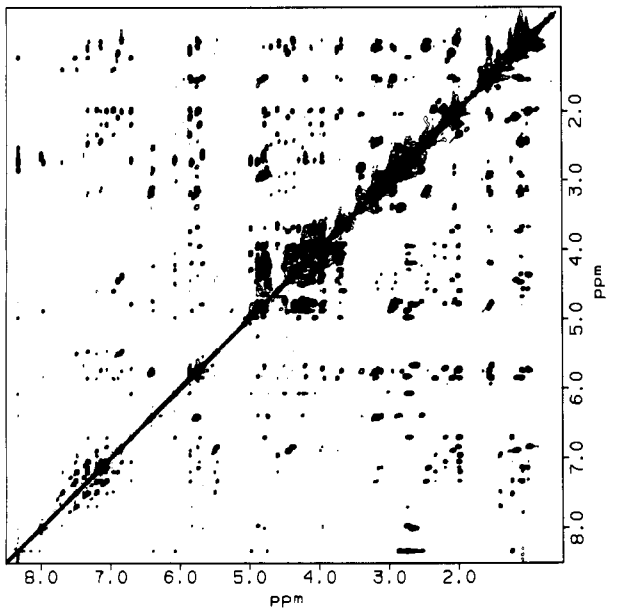


FIGURE 5: NOESY spectrum, τ_m = 100 ms, of [d(TCGATCGA)]$_2$–2 echinomycin in D_2O. The sample was 1.3 mM complex, 50 mM NaCl, pH 6.5, 30 °C. The spectrum was acquired with a sweep width of 5000 Hz in both dimensions. Three hundred eleven t_1 values of 32 scans and 1K complex points were collected. The data in both dimensions were apodized with a skewed sine-bell squared function (skew = 1.3) phase-shifted by 70°. Three hundred eleven points were apodized in each dimension, and spectra in t_1 were zero-filled to 1K prior to Fourier transformation.

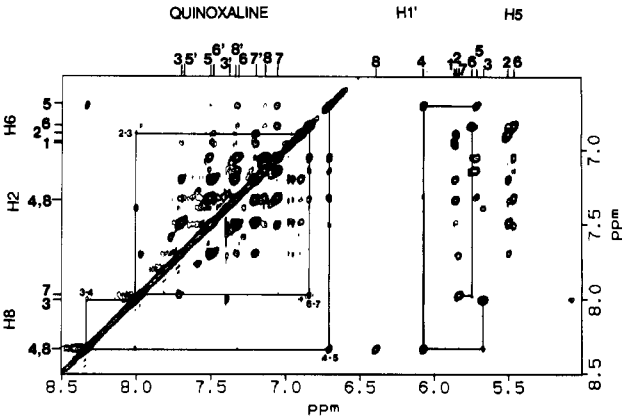


FIGURE 6: Expanded region of the NOESY spectrum of [d-(TCGATCGA)]$_2$–2 echinomycin in D_2O, showing the aromatic resonances and their cross peaks to the H1′ region at 30 °C and τ_m = 300 ms. The sample is the same as that described in Figure 5. The quinoxaline, aromatic DNA, H1′, and H5 resonance assignments are indicated. The sequential connectivities in both regions are shown. In the aromatic region the C_2H6 to G_3H8 to A_4H8 to T_5H6 cross peaks and the C_6H6 to G_7H8 cross peaks are connected by solid lines and the cross peaks are labeled. In the base–H1′ region the 3-4-5 and 6-7 sequential connectivities are drawn. The spectrum was acquired with a sweep width of 5000 Hz in both dimensions. Two hundred thirteen t_1 values of 32 scans and 2K complex points were collected. The data were apodized with a skewed sine-bell squared function (skew = 1.1) phase-shifted by 80°. Data in t_1 were zero-filled to 1K prior to transformation.

define the structure of the drug in the complex and many intermolecular cross peaks that define the interactions between the DNA and the drug. These interactions are discussed in detail below.

Assignment of the Nonexchangeable DNA Resonances in the [d(TCGATCGA)]$_2$–2 Echinomycin Complex. The assignment of the DNA resonances in this complex was much simpler than the assignment of complex A. Figure 6 shows an expansion of the aromatic and base–H1′ regions of a NOESY spectrum of complex T. Sequential base–base (H6,H8) connectivities are observed in the aromatic region for bases C_2–T_5 and C_6–G_7. These assignments are confirmed in the base–H1′ region, where sequential base–H1′ connectivities can be seen for bases 3–5 and 1–7. With bases 3–5 assigned, the rest of the resonances can be placed in the sequence unambiguously. Sequential assignments of bases 1 and 2 can be made through the quinoxaline protons, i.e., T_1H6 and C_2H6 have cross peaks to quinoxaline ring protons. The A_4H2 resonance was identified by NOE cross peaks to the internal A·T imino resonance observed in NOESY spectra in H_2O (see below). The A_8H2 unfortunately overlaps the A_4H2 and several quinoxaline resonances.

Once the H1′ resonances were assigned sequence-specifically, the assignments were extended to the rest of the sugar protons via scalar coupling in the HOHAHA (supplementary material) and DQF-COSY spectra (data not shown). As seen with complex A, the 2′–2″ and 1′–2′ cross peaks for both C_2 and C_6 were only observed in the DQF-COSY spectrum. Complete assignments were obtained for all but some of the H5′,H5″ resonances (Table II).

Further confirmation of some of the assignments was obtained via exchange cross peaks between free DNA, partially saturated complexes, and the fully saturated complexes (supplementary material).

Assignment of the Exchangeable Resonances of the DNA in the Complex [d(TCGATCGA)]$_2$–2 Echinomycin. The exchangeable resonances in this complex were assigned in essentially the same way as those in complex A, except the assignments were made from spectra taken at 10 °C, as indicated in Figure 4B. The internal adenine imino resonance had a strong cross peak to AH2, indicative of Watson–Crick base pairing. A list of the assigned DNA resonances in complex T is shown in Table III.

Table III: Chemical Shifts of the DNA Resonances in the Complex [d(TCGATCGA)]$_2$–2 Echinomycin at 30 °C[a]

	H8/H6	H5/CH3/H2	H1′	H2′	H2″	H3′	H4′	H5′,5″	amino[b]	imino[b]
T_1	6.96	1.42	5.86	2.34	1.99	4.60	3.93	3.71, 3.65		12.04
	6.95	1.36								
C_2	6.90	5.50	5.86	2.08	1.17	4.45	3.75	3.98, 4.10	6.45, 7.69	
G_3	8.00		5.67	2.66	2.63	4.80	4.20			11.74
A_4	8.33	7.31	6.07	2.76	2.71	4.98	4.30	4.06, 3.70		
T_5	6.71	1.24	5.72	2.20	2.00	4.80	3.97			12.24
C_6	6.84	5.46	5.74	2.01	0.99	4.38	3.68	3.96, 4.12	6.45, 7.72	
G_7	7.97		5.83	2.75	2.75	4.08	4.20			11.67
A_8	8.33	7.31	6.39	2.85	2.54	4.76	4.25		6.80	

[a] Chemical shifts relative to DSS. [b] Chemical shifts at 10 °C.

Table IV: Chemical Shifts of the Echinomycin Resonances in the Complex [d(TCGATCGA)]$_2$–2 Echinomycin at 30 °C[a]

serine		α	β	β′	γ	NH[b]	NCH_3	SCH_3
serine	1	4.88	4.36	4.29		8.21, 8.23		
	2	4.88	4.36	4.29		8.16, 8.18		
valine	1	4.50	2.46		1.09, 1.12		3.18	
	1′	4.70	2.49		1.08, 1.08		3.22	
	2	4.37	2.43		1.01, 1.10		3.11	
	2′	4.59	2.45		1.00, 1.06		3.17	
cysteine	1	6.40	4.75				2.95	2.05
	2	5.75	3.39	2.96			2.91	
alanine	1	4.83	1.56			9.57, 9.76		
	2	4.78	1.54			9.57, 9.40		

		3	5	6	7	8
quinoxaline	1	7.71	7.69	7.48	7.20	7.33
	2	7.39	7.51	7.33	7.06	7.14

[a] Chemical shifts relative to DSS. [b] Chemical shifts at 10 °C.

Assignment of the Drug Resonances in the Complex [d-(TCGATCGA)]$_2$–2 Echinomycin. The amino acid spin system were identified by HOHAHA and NOESY experiments. With the exception of the Ser resonances, the assignment procedure for the echinomycin resonances was identical with that discussed for complex A. The spin systems were identified in the HOHAHA spectrum (supplementary material). Sequential assignments of the peptide resonances were, again, made on the basis of DNA–drug contacts. The Ser $C_\alpha H$ protons and the water resonance overlap at 30 °C, making it difficult to identify the Ser spin systems. Therefore, the Ser resonances were assigned on the basis of the NH–CαH and NH–$C_\beta H$ cross peaks seen in a HOHAHA spectrum in water (spectrum not shown). The quinoxaline ring proton resonances were assigned as described above for complex A but were better resolved in this complex. The chemical shifts of the echinomycin resonances in complex T are listed in Table IV.

Structural Features of the DNA in the Complex [d-(ACGTACGT)]$_2$–2 Echinomycin. There are profound changes in the structure of the duplex caused by the binding of echinomycin. The A·T base pairs immediately adjacent to the binding site are Hoogsteen base paired (Gilbert et al., 1989). There are strong cross peaks from both adenine H8 protons to their respective H1′ resonances (Figure 3A, boxes A). The base–H1′ cross peaks are much more intense than the base–H2″ cross peaks characteristic of the syn conformation. The Hoogsteen base pairs A_5 and T_4 in the interior of the duplex are most stable at 1 °C. At 20 °C there is appreciable exchange between the Hoogsteen conformation and another, probably unpaired, conformation (Gilbert et al., 1989). Therefore, at the temperature shown, the base–H1′ cross peak for A_1 is much stronger than that for A_5 (Figure 3B). The strong imino–H8 cross peak for the A_5-T_4 base pair proves that the base pair is in fact Hoogsteen base paired at lower temperatures (Figure 4A). Unfortunately, the conformational exchange between Hoogsteen and open base pairs and the degree of spectral overlap preclude determining a three-dimensional structure of the complex.

Other structural changes are induced by echinomycin binding and these can also be seen in the NOESY spectrum in D_2O (Figure 3). The strong base–H3′ cross peaks seen for C2 and C6 are characteristic of the C3′ endo sugar pucker. In addition, the fine structure of the COSY H1′–H2′ and H1′–H2″ cross peaks are consistent with a C3′ endo sugar pucker. A C3′ endo sugar pucker would widen the minor groove, allowing room for the peptide ring to bind.

The lack of sequential cross peaks in the base–H1′ region shows that the duplex is not a canonical right-handed helix. It is difficult to draw conclusions about the structure of the center of the duplex because the resonances from bases at the center of the duplex are very broad and show few weak NOEs due to exchange between Hoogsteen and unpaired (or Watson–Crick) conformations (Gilbert et al., 1989). The cross peaks that we do see are consistent with the unwinding of the duplex seen in the crystal structures (Wang et al., 1984, 1986; Ughetto et al., 1985; Quigley et al., 1986) and shown by viscometric studies (Wakelin & Waring, 1976).

Binding of echinomycin induces some minor asymmetry in the DNA duplex. The A_1H8 and C_2H6 and C_6H6 resonances are all doubled. Enough of the symmetry is maintained so that it is not possible to assign the two DNA strands separately.

Structure of Echinomycin in the Complex [d-(ACGTACGT)]$_2$–2 Echinomycin. Details of the structure of the echinomycin molecules in the complex can be obtained from the NOESY spectra in D_2O (Figure 3) and H_2O (Figure 4A). There are many strong intramolecular NOEs. The methyl resonances of Ala and Val have strong cross peaks to each other. The NCH_3 resonances of Cys and Val have cross peaks to each other and to the methyl protons of Ala and Val. In fact, all the methyl groups with the exception of the SCH_3 in the thioacetal cross bridge have cross peaks to one another. These methyl groups form a hydrophobic face on one side of the echinomycin molecule. There are strong cross peaks between the NCH_3 and α protons of Cys and Val as well as cross

Table V: DNA–Echinomycin Contacts in the Complex [d(ACGTACGT)]$_2$–2 Echinomycin[a]

	H8/H6	Me/H5/H2	H1′	H2′	H2″	H3′	H4′	H5′,5″ amino	imino
A_1	$Q_1$7 (w) Val_1CH_3 (m)		$Q_1$7,8 (m) Val_1CH_3 (m)	$Q_1$7,8 (s)	$Q_1$8 (s)		Val_1CH_3 (m)		
C_2	$Q_1$7 (m) Ala_1CH_3 (m) $Q_1$5 (m)	$Q_1$7 (m)	$Q_1$7,8,6 (m) Ala_1CH_3 (s) Val_1CH_3 (m)	Q_1,7,8 (s) Ala_1CH_3 (m)	$Q_1$7,8 (s) Ala_1CH_3 (m)		Ala_1CH_3 (m)		
G_3	$Q_2$3 (m)		Ala_1CH_3 (s); Val_2NCH_3 (m); Ala_1NH (s)	$Q_2$3 (m)			Ala_1CH_3 (m)		Val_2NCH_3 (s); SerNH (w); $Q_2$6,5 (w)
T_4	$Ser_1\alpha$ (w)								
A_5	Val_2CH_3 (m)	$Ser_1\beta$ (w); $Ser_1\beta'$ (w); $Ser_1\alpha$ (w)	$Q_2$7,8 (m)						
C_6	$Q_2$7 (m) $Q_2$8 (m) Ala_2CH_3 (m)	Ala_2CH_3 (w)	Ala_2CH_3 (s) $Q_2$7 (m)	Ala_2CH_3 (s) $Q_2$8 (m)	Ala_2CH_3 (s)	Ala_2CH_3 (w)	Ala_2CH_3 (w)		
G_7	$Q_1$3 (M)	Ala_2CH_3 (m); $Cys_2\alpha$ (w)					Ala_2CH_3 (w)		Val_1NCH_3 (w); SerNH (m); $Q_1$5,6,7 (w)
T_8	$Ser_2\alpha$ (w)		$Ser_2\beta$ (w); $Ser_2\beta'$ (w)			Val_1CH_3 (w)			

[a] Cross-peak intensities are listed as strong (s), medium (m), or weak (w).

Chart II

peaks to the α protons of Ala. There are many cross-strand NOEs between opposite sides of the ring. The quinoxaline ring protons 7 and 8 have cross peaks to the Val CH_3. This shows that the hydrophobic face is on the same side of the drug as the quinoxaline rings. Qualitative analysis of the NOEs indicates that the solution structure of the drug in this complex is very similar to the crystal structure of the drug in the complex 2 echinomycin–[d(CGTACG)]$_2$ (Ughetto et al., 1985) and to its structure in other echinomycin–DNA complexes in solution (Gao & Patel, 1988).

Interactions between Echinomycin and DNA in the Complex d(ACGTACGT)]$_2$–2 Echinomycin. The binding site and important DNA–drug interactions can be defined on the basis of the intermolecular NOESY cross peaks seen in the NOESY spectrum shown in Figure 3. A complete list of DNA–drug contacts is given in Table V. The mode of binding is defined by cross peaks between the quinoxaline ring protons and the bases as well as cross peaks between the peptide ring and minor-groove markers on the DNA.

NOE cross peaks are observed between Q_1 and the aromatic resonances of A_1 and C_2. The Q_2 resonances have cross peaks between the quinoxaline rings and the bases show that the rings are stacking between the A_1-C_2 and A_5-C_6 bases at each binding site, as illustrated in Chart II. Weak cross peaks observed between the Q_2H3 proton and the G_3H8 and between Q_1H3 and G_7H8 define the orientation of the quinoxaline rings with the long axis parallel to the base pairs.

The important interactions that define the binding site and sequence specificity are between the peptide ring and the DNA. The majority of contacts are between the peptide ring and minor-groove markers. The DNA–peptide interactions for complex A at 20 °C are illustrated schematically in Figure 7A. The Ala_1 and Ala_2 methyls have strong cross peaks to both the sugars of the C_2 and G_3 and C_6 and G_7, respectively, at each binding site (Figure 3B), which indicates that the methyl group is positioned between the two sugars. The Val_1 methyl groups have cross peaks to the A_1H8, A_1H1′ and A_1H4′, T_8H3′, and C_2H1′ resonances. Val_2 methyl resonances have cross peaks to the A_5H8 and T_4H_6. Many of the methyl groups have cross peaks to minor-groove markers. The hydrophobic face is toward the duplex in the minor groove. Most important in terms of the structure of the DNA, there are cross peaks between the A_1 and A_5 H8 and the Val methyl resonances (Figure 3, box D). The H8 protons are in the minor groove only when the base pairs are Hoogsteen base paired.

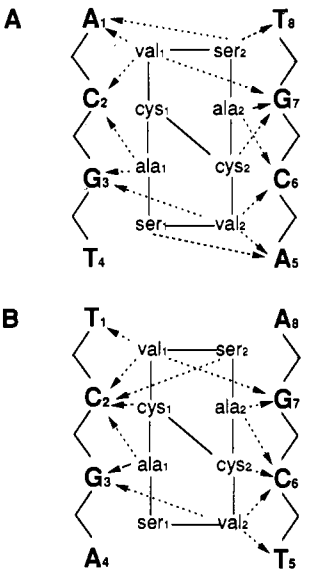


FIGURE 7: Schematic illustrations of the DNA peptide contacts in (A) complex A and (B) complex T. The arrows indicate NOESY cross peaks seen in Figures 3 and 5, respectively, and discussed in the text. The lengths of the arrows are not correlated to the distances between the protons.

The methyl resonances form a hydrophobic face pointing toward the minor groove. Cross peaks between the G·C iminos and Val NCH_3 resonances (Figure 4A, box A) show that the NCH_3 groups point into the center of the minor groove. The Val NMe points down toward the bottom of the groove and the side-chain methyls of Ala and Val point toward the edges of the groove. The Ala amide resonances are shifted downfield 1 ppm relative to the random coil resonance positions, indicative of hydrogen bonding. The Ala amide resonances also have strong cross peaks to the minor-groove marker GH1′ protons (Figure 4A, box B). This puts the Ala NH protons quite close to the N3 of guanine. This set of cross peaks and the downfield shift of the amide resonance is consistent with the pattern of hydrogen bonding seen in the crystal structure of the echinomycin–[d(CGTACG)]$_2$ complex (Ughetto et al., 1985)). The hydrogen bonds between the Ala amide and GN3 have been shown to be important for the sequence specificity of the drug (Wang et al., 1984).

The interactions of the drug with the DNA are not symmetric. The Ala_1 methyl has a stronger cross peak to C_2H6 than Ala_2 methyl has to C_6H6, and Ala_2 is closer to G_7H8 than Ala_1 is to G_3H8. This asymmetry is also seen in the crystal structure. There are two hydrogen bonds between Ala_2 and

Table VI: DNA–Echinomycin Contacts in the Complex $[d(TCGATCGA)]_2$–2 Echinomycin[a]

	H8/H6	H5/H2/Me	H1′	H2″	H2′	H3′	H4′	H5′,5″ amino	imino
T_1	$Q_1$7 (m) $Q_1$6 (m)	$Q_1$6 (m) $Q_1$5 (m)	Val_1NCH_3 (w)	$Q_1$7 (m)	$Q_1$7 (m)				
C_2	$Q_1$7 (m)	$Q_1$7 (m) $Q_1$6 (m)	$Q_1$7 (w); $Q_1$5 (w); Val_1NCH_3 (w) $Val_1\beta$ (w); Ala_1CH_3 (s); Val_1CH_3 (s); Cys_1NCH_3 (w); Val_1CH_3 (m)	Ala_1CH_3 (s)	Ala_1CH_3 (s)	$Ser_1\alpha$ (w)	Ala_1CH_3 (m)		
G_3	$Q_2$3 (w) Ala_1NH (s)		$Q_2$3 (w) Ala_1CH_3 (m); Ala_1NH (s)	Ala_1CH_3 (w) $Q_2$3 (w)	Ala_1CH_3 (w) $Q_2$3 (w)				$Q_2$6 (w) $Q_2$5 (s); Ala_1CH_3 (m); Val_2NCH_3 (m)
A_4									
T_5	$Q_2$7 (m) $Q_2$6 (m)	$Q_2$7 (m) $Q_2$5 (m)	$Q_2$7 (w) $Q_2$6 (m); Val_2CH_3	$Q_2$7 (m) $Q_2$8 (w)					
C_6	$Q_2$7 (m) $Q_2$6 (m)	$Q_2$7 (m) $Q_2$6 (s)	$Q_2$7 (m); Ala_2CH_3 (m) $Q_2$8 (m); Cys_2NCH_3 (w); Val_2NCH_3 (w); Val_2CH_3 (w)	Ala_2CH_3 (m) Val_2CH_3 (w)	Ala_2CH_3 (m) Val_2CH_3 (w)	Ala_2CH_3 (w)	Ala_2CH_3 (w)		
G_7	$Q_1$3 (w)		$Q_1$3 (w); Ala_2NH (m); Ala_2CH_3 (s)	Ala_2CH_3 (m); $Q_1$3 (w)	Ala_2CH_3 (m); $Q_1$3 (w)	Ala_2CH_3 (m); $Q_1$3 (w)			$Q_2$7 (w); $Q_2$6 (w); $Q_2$5 (w); Ala_2CH_3 (s); Val_1NCH_3 (m)
A_8									

[a] Cross-peak intensities are listed as strong (s), medium (m), and weak (w).

G_7 and only one between G_3 and Ala_1. The asymmetry seen in solution is consistent with the crystal structure.

The two drug molecules bind the self-complementary duplex with a 2-fold axis of symmetry. One quinoxaline ring has cross peaks to A_1 and C_2, while the other has cross peaks to A_5 and C_6. Ala_1 methyl has cross peaks to C_2 and G_3, while Ala_2 methyl has cross peaks to C_6 and G_7. The orientation of the two drugs is illustrated in Chart II.

Structural Features of the DNA in the Complex [d-(TCGATCGA)]$_2$–2 Echinomycin. Unlike complex A, complex T has no strong base–H1′ cross peaks. The AH8–H1′ cross peaks of both A_4 and A_8 are similar in intensity to the other base–H1′ cross peaks and less intense than the base–H2′,2″ cross peaks. There are no bases in the syn conformation. There are also no imino to H8 cross peaks in the NOESY spectrum taken in H_2O shown in Figure 4B. There is, instead a strong cross peak from the T_5 imino to A_4H2 (Figure 4B, box A). The absence of both strong H8–H1′ and imino–H8 cross peaks shows that there are no Hoogsteen base pairs in this complex.

Although the A·T base-pairing scheme adjacent to the drug binding is different in the two complexes, the sugar puckers for the bases at the binding site are similar. The intensity of the CH6–H3′ NOESY cross peaks and their COSY fine structure are similar to that seen in complex A, indicating that in this complex the C_2 and C_6 sugars are A-DNA-like (C3′ endo).

The most striking change in the structure of this duplex is a large unwinding. The unwinding can be seen by the presence of several intraduplex cross peaks that are not expected for B-DNA. These include a weak cross peak between A_4H8 and T_5H6 (Figure 6), which is also seen in the NOESY spectrum with a 100-ms mixing time. In "standard" B-DNA this distance is approximately 6.5 Å. The presence of a NOESY cross peak at 100 ms means that this distance is roughly 4.5 Å in the complex. Unwinding of the helix brings these A_4H8 and T_5H6 protons closer together. The H1′ protons are also brought closer together as the helix unwinds. Weak H1′–H1′ cross peaks between A_4H1′ and T_5H1′ can also be seen in the shorter mixing time NOESY. This distance is 5.5–6.5 Å in B-DNA. In the complex the distances are reduced to the order of 4.5 Å.

As seen in complex A, the symmetry of the duplex in complex T is also broken by echinomycin binding. The resonances for T_1H8 and T_1CH_3 are doubled. However, the differences between the two strands are not great enough to make it possible to assign the strands separately.

Structure of Echinomycin in the Complex d-[(TCGATCGA)]$_2$–2 Echinomycin. There are many intramolecular NOEs for echinomycin that can be seen in the NOESY spectrum shown in Figure 5. The pattern of cross peaks is similar to that described for complex A. In addition to the QH8 to Val CH_3 cross peaks seen in complex A, there are also cross peaks of medium intensity between both QH8 resonances and Val NCH_3 not seen in complex A. These orient the quinoxaline ring perpendicularly away from the ring toward the hydrophobic face. The structure of the drug appears to be much the same as is seen in complex A, and it is also similar to the crystal structure.

Interactions between DNA and Echinomycin in the Complex [d(TCGATCGA)]$_2$–2 Echinomycin. As was discussed for complex A, the echinomycin binding site and mode of binding can be defined on the basis of the cross peaks between the aromatic protons of the Ts and Cs and the quinoxaline ring protons QH5–8 and cross peaks between the peptide ring and minor-groove markers observed in Figure 5.

The quinoxaline ring protons have cross peaks to T_1 and C_2 at one end of each binding site and to T_5 and C_6 at the other end of each binding site. The Q_2H3 resonance also has a cross peak to G_3H8 and Q_1H3 has a cross peak to G_7H8, which orient the long axis of the drug parallel to the long axis of the A-T base pair.

The interactions that are most interesting in terms of binding site and sequence specificity are those between the peptide ring and the minor-groove markers on the DNA. These are illustrated schematically in Figure 7B. The interactions are very similar to those seen in complex A. A complete list of DNA–drug contacts is given in Table VI.

The drug does not interact with the two strands of the binding site symmetrically. Ala_1 is closer to C_2 than Ala_2 is to C_6, and Ala_2 is closer to G_7 than Ala_1 is to G_3. The cross peaks between T_1, C_2, and Q_1H5-8 are less intense than those between T_5, C_6, and Q_2H5-8. This may be due to end effects, which allow greater degrees of freedom for the terminal bases.

This is different than complex A, which had stronger cross peaks between the terminal bases and the quinoxaline rings.

The orientation of the two drugs is the same as in complex A. There is a 2-fold axis of symmetry relating the two binding sites. The symmetry is not perfect, as indicated by the fact that the T_1H6 resonance is split and there are four resolvable Val spin systems. Two of the Vals have cross peaks to T_1 and C_2, and two others have cross peaks to T_5 and C_6.

Discussion

Mode of Binding. Preliminary NMR results have been published showing that echinomycin binds $[d(ACGTACGT)]_2$ as a bisintercalator (Gilbert et al., 1989). The imino resonances show large upfield shifts of approximately 1 ppm on complexation, which is characteristic of intercalative binding (Feigon et al., 1984a). As predicted from footprinting and crystallographic results, two drugs bind per duplex at the CpG sites and the binding to the two drugs is almost symmetric. The two quinoxaline rings bisintercalate between the A·T and G·C base pairs at each binding site. Strong NOE cross peaks observed between the quinoxaline ring resonances and base resonances discussed above prove that the quinoxaline rings are intercalating on either side of the CpG steps.

Echinomycin also binds $[d(TCGATCGA)]_2$ as a bisintercalator, with two drugs bound per duplex at the two CpG sites when saturating levels of drug are added. The characteristic large upfield shift of the imino resonances upon complexation is also observed in this complex. Strong cross peaks between the aromatic protons of the Ts and Cs and the quinoxaline ring protons QH5–QH8 (discussed in detail above) also prove that echinomycin binds as a bisintercalator in solution. The orientation of the two drugs relative to the duplex is the same as seen in complex A (see Chart II).

The peptide ring binds in the minor groove. There are a large number of contacts between the peptide ring and minor-groove marker protons, as already discussed. The mode of binding and central CpG binding sites for the two octamers are the same. This is in agreement with the results of Gao and Patel (1988) on echinomycin binding to the DNA tetramers d(ACGT) and d(TCGA). This lends further support to the idea that the hydrogen bonding between Ala and G and the structure of the peptide ring are the determinants of sequence specificity (Wang et al., 1984).

Hoogsteen Base Pair Formation. We have previously shown that the terminal A·T base pairs in complex A are Hoogsteen base paired over a wide temperature range and that internal A·T base pairs are Hoogsteen base paired at low temperature but not at higher temperatures (Gilbert et al., 1989). In contrast, in the complex of $[d(TCGATCGA)]_2$ and echinomycin there are no Hoogsteen base pairs. The terminal and internal A·T base pairs are Watson–Crick base paired at temperatures ranging from 0 to 45 °C.

It is not clear from the NMR results what different structural features might cause the difference in base pairing in the two different complexes. Hoogsteen base pair formation brings the 1′ carbons on the sugars closer together by approximately 2 Å. It has been suggested that the narrower minor groove at the site of the Hoogsteen base pairs results in increased van der Waals contacts between the peptide and the phosphate sugar backbone, thereby stabilizing the complex (Wang et al., 1984). However, this cannot be the predominant reason or we would expect to see Hoogsteen base pairing in both complexes. Gao and Patel (1988) proposed that the stacking interactions between the bases and the quinoxaline ring must contribute to the sequence requirements for Hoogsteen base pair formation and we concur. According to theoretical calculations done by Singh et al. (1986), the stacking between the two A·T base pairs in the complex $[d(CGTACG)]_2$–2 echinomycin was much more favorable if the adenines were in the Hoogsteen conformation. The increase in stabilization came from increased stacking between the two A·T base pairs rather than between the adenine and the quinoxaline ring. Actually, according to their calculations the stacking interactions between the adenines and the rings were approximately the same for both the Watson–Crick and Hoogsteen conformations. This explains the results for the internal base pairs nicely but cannot explain the presence of Hoogsteen base pairs at the ends of the duplex. It is not clear why the experimental and theoretical results are so disparate. It may be that, despite theoretical arguments to the contrary, the double ring of the adenine does stack better over the quinoxaline ring in complex A when the A·T bases are Hoogsteen base paired but that Hoogsteen geometry does not increase the stacking between the T and the quinoxaline ring in complex T enough to overcome the increased electrostatic repulsion caused by bringing the phosphates closer together.

Stability of the DNA–Echinomycin Complexes. The stability of the $[d(ACGTACGT)]_2$–2 echinomycin complex has been discussed in a previous report (Gilbert et al., 1989). It was shown that $[d(ACGTACGT)]_2$ binds two echinomycin molecules per duplex and that the binding is cooperative and essentially irreversible. There is no exchange between free and complexed DNA in exchange experiments done on a partially saturated complex even at mixing times up to 2 s (data not shown).

The titration of echinomycin with $[d(TCGATCGA)]_2$ shown in Figure 1 is in marked contrast to that of $[d(ACGTACGT)]_2$. The binding is not cooperative. There are four distinct sets of cross peaks in the NOESY spectrum of the half-saturated complex, corresponding to free DNA, two different complexes containing one drug per duplex, and the fully saturated complex (supplementary material). There are exchange cross peaks between the free DNA and the complexes with one drug bound and between these and the fully saturated complex. However, there does not appear to be any exchange between the free DNA and the fully saturated complex. The exchange between complexes is slow on the NMR time scale. Exchange cross peaks are seen in NOESY spectra with a 200-ms mixing time. Based on this, a rough estimate of the off rate is $\sim 5\ s^{-1}$.

Not only is complex T kinetically less stable than complex A, it is also thermodynamically less stable. When complex T is first formed, there are only cross peaks from the fully saturated complex in the NOESY spectrum. However, when spectra are taken approximately 2 days after the complex is made, cross peaks from less than fully saturated complexes and free DNA also appear in the spectra. The presence of exchange peaks and the presence of free DNA in a 2:1 drug–duplex complex indicate that the binding constant is not as high and that the dissociation constant is higher for the $[d(TCGATCGA)]_2$ complex than for the $[d(ACGTACGT)]_2$ complex. The increased van der Waals contacts between the peptide ring and the phosphate sugar backbone of the DNA caused by the formation of Hoogsteen base pairs may account for the increased stability of complex A vs complex T.

Stability of Base Pairing within the Duplex. All of the imino proton resonances of complex T remain sharp over the temperature range shown in Figure 2. The G_3–C_2 imino is, in fact, sharper in the complex than in the free DNA. Echinomycin binding raises the melting temperature of the duplex, as expected for intercalative binding (Patel, 1977). The

slow exchange rate for the central A·T imino in complex T is in marked contrast to the results obtained with complex A shown in Figure 2C. The imino resonance from the central A·T base pairs broadens and disappears at much lower temperatures than in complex T, indicating that the central A·T base pairs in complex A are exchanging much more rapidly than those in complex T. Although imino proton exchange can only be rigorously used to determine base-pair lifetimes when the experiments are done in the presence of infinite catalyst concentration, it is reasonable to correlate exchange rates with relative base-pair lifetime in this case because the binding of the drug in complex T does not have the same effect on imino exchange rates as it does in complex A. If the drug is catalyzing exchange more efficiently in complex A, one would also expect that to be the case in complex T.

The fact that the DNA duplex in complex T appears to be stabilized rather than destabilized upon drug binding is supported by the fact that there are no major changes in the NOESY spectra of complex T as a function of temperature up to 45 °C (data not shown), in contrast to the results obtained on complex A in which the resonances from T_4 and A_5 broaden as a function of temperature until at 45 °C they are broadened almost to baseline (Gilbert et al., 1989). Because the DNA duplex begins to melt at higher temperatures, we are unable to unambiguously determine if the T_4·A_5 Hoogsteen base pairs are exchanging between an open or a Watson–Crick base-paired state. It is clear, however, that it is a change in the conformation of the A·T base pairs and not in the mode of drug binding (i.e., the quinoxaline ring swinging out of the duplex), since the neighboring C·G resonances do not broaden to the same extent as the temperature is raised.

The destabilization of the internal A·T base pairs accounts for the difference in the intensities of the base–quinoxaline cross peaks in the two complexes. The cross peaks between Q_2 and A_5 in complex A are less intense than the cross peaks between T_5 and Q_2 in complex T, because A_5 is exchange broadened.

Differences in stacking interactions between the ends and the center of the duplex account for the 0.2 ppm upfield shift of Q_2 relative to Q_1 in complex T. In complex A at 20 °C the center of the duplex is more like the ends, because the central A·T base pairs are exchanging between a Hoogsteen and an unpaired (or Watson–Crick) state. As a result, the chemical shifts of the two quinoxaline rings in complex A are very similar.

Relevance to Footprinting Results. Several laboratories have shown that hypersensitivity of DNA to DEPC and other chemical cleavage reagents in DNA–echinomycin complexes depends upon the sequence adjacent to the CpG binding site (Mendel & Dervan, 1988; Portugal et al., 1988; McClean & Waring, 1988). Adenines immediately 3′ of the CpG binding site are more reactive to DEPC than those 5′ of the binding site, while OsO_4 reacts more strongly with thymines that are 3′ of the binding site.

The structural basis for the hypersensitivity of DNA to these cleavage reagents in DNA–echinomycin complexes has been unclear. The Hoogsteen base pairs seen in solution under some conditions and in the crystal structure cannot be the only structural perturbation responsible for the hypersensitivity, because TCGA sites are also hypersensitive, and we have shown that there are no Hoogsteen base pairs in the complex of echinomycin with $[d(TCGATCGA)]_2$. In addition, work by McClean et al. (1989) shows that oligonucleotides containing modified adenines that are incapable of forming Hoogsteen base pairs are still hypersensitive to OsO_4. In fact, TCGA sites are more reactive to DEPC than ACGT sites. However, a pronounced structural change in the DNA that can be seen in both complexes is the large unwinding of the duplex. This widens the major groove, making the N7 of adenine more accessible to reaction with DEPC. We propose that this is the primary structural change in the DNA duplex on echinomycin binding that results in the hyperreactivity of bases adjacent to the CpG binding sites to cleavage reagents.

Helix unwinding alone does not, however, explain the apparent sequence specificity of the enhanced cleavage by DEPC. We have shown that all the adenines in $[d(ACGTACGT)]_2$ are Hoogsteen base paired at low temperature and that the central base pairs begin to exchange between the Hoogsteen conformation and an unpaired state as the temperature is raised (Gilbert et al., 1989). At higher temperature the central A·T base pairs are destabilized, but they remain Hoogsteen base paired for some fraction of the time. This could account for the sequence specificity of the enhanced cleavage by DEPC. TCGA sites are more susceptible because the N7 of adenine is always available to react with DEPC. Because the adenines in ACGT are sometimes Hoogsteen base paired, the adenine N7 is blocked from reacting with DEPC by the hydrogen bond to TH3. OsO_4, a chemical probe used to detect changes in the structure of the DNA helix by reaction at the C5=C6 double bond of thymine, reacts more strongly with thymines 3′ of the binding site. The combination of base pair destabilization and unwinding at ACGT sites may make the C5=C6 double bond more accessible than simple unwinding of the helix seen at TCGA sites.

Footprinting results showed that echinomycin binds with "all or none" behavior, exhibiting similar degrees of protection over a wide range of drug concentrations (Low et al. 1984). In our studies with short oligonucleotides, it appears that the drug only binds cooperatively to adjacent ACGT binding sites.

Our results are consistent with those obtained by footprinting and suggest a structural basis for the sequence-dependent variation of enhanced cleavage by DEPC and OsO_4.

Comparison to Other NMR Solution Studies. Solution NMR studies of echinomycin complexed to several tetramers and one octamer have been published previously (Gao & Patel, 1988, 1989). The results obtained from the tetramers ACGT and TCGA (Gao & Patel, 1988) and GCGC and CCGG (Gao & Patel, 1989) showed that pur-C-G-pyr binding sites form Hoogsteen base pairs and pyr-C-G-pur binding sites do not (deduced indirectly from observation of a strong AH8–H1′ NOE, indicative of the syn conformation). Preliminary results also showed that the complex of the octamer AAACGTTT and echinomycin did not have any Hoogsteen base pairs (Gao & Patel, 1989).

This work confirms the conclusions of Gao and Patel regarding the sequence specificity of Hoogsteen base pairing in DNA–echinomycin complexes. This work also addresses questions that cannot be answered by studies of DNA molecules containing only one binding site. We show that cooperative binding is sequence dependent. We have also shown that echinomycin binds symmetrically to adjacent binding sites. The destabilization of the A·T base pairs adjacent to echinomycin binding sites within a duplex occurs only when adenine is at the 5′ position of the binding site. It is not possible to compare our results regarding destabilization to those of Gao and Patel (1989) on the octamer AAACGTTT because only preliminary results regarding this complex have been published.

Comparison to the Crystal Structures. The solution structure of the complex A at low temperatures is similar to

the crystal structure of $[d(GCGTACGC)]_2$–2 triostin A solved by Rich and co-workers (Quigley et al., 1986; Wang et al., 1986). The terminal base pairs in both structures are Hoogsteen base paired. The central A·T base pairs in the crystal structure are also Hoogsteen base paired. In solution, the central A·T base pairs of the DNA in complex A are Hoogsteen base paired at lower temperatures. However, at higher temperatures the base pairs are exchanging between the Hoogsteen conformation and an unpaired (or Watson–Crick base-paired) state. Thus, although the A·T base pairs in the crystal structure are in the Hoogsteen conformation, this is only one of the structures that the duplex may form in solution. At physiological temperatures Hoogsteen base pairing of the central region of the duplex is very unstable, suggesting that an open rather than a base-paired state for A·T bases in ACGT binding sites may be the relevant conformation in vivo.

No crystal structures have been solved containing DNA with pyr-C-G-pur binding sites. We anticipate that such a crystal structure will be in agreement with the solution structure we have observed, i.e., there will be no Hoogsteen base pairs and the helix will be unwound.

Acknowledgments

We thank Dr. Gijs A. van der Marel and Dr. Jacques H. van Boom for some of the d(ACGTACGT) oligonucleotide that was used in the early stages of this work.

Supplementary Material Available

Four figures showing HOHAHA spectra of complexes A and T and expanded regions of NOESY spectra of complex A, $[d(CGTACG)]_2$–2 echinomycin complex, and $[d(TCGATCGA)]_2$–1 echinomycin complex (6 pages). Ordering information is given on any current masthead page.

References

Aue, W., Bartholdi, E., & Ernst, R. R. (1976) *J. Chem. Phys. 64*, 2229–2246.

Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355–360.

Feigon, J., Leupin, W., Denny, W. A., & Kearns, D. R. (1983) *Biochemistry 22*, 5943–5951.

Feigon, J., Denny, W. A., Leupin, W., & Kearns, D. R. (1984a) *J. Med. Chem. 27*, 450–465.

Feigon, J., Wang, A. H.-J., van der Marel, G. A., van Boom, J. H., & Rich, A. (1984b) *Nucleic Acids Res. 12*, 1243–1263.

Gao, X., & Patel, D. J. (1988) *Biochemistry 27*, 1744–1751.

Gao, X., & Patel, D. J. (1989) *Q. Rev. Biophys. 22*, 93–138.

Gilbert, D. E., van der Marel, G. A., van Boom, J. H., & Feigon, J. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 3006–3010.

Hare, D. R., Wemmer, D. E., Chou, S. H., Drobny, G., & Reid, B. R. (1983) *J. Mol. Biol. 171*, 319–336.

Lee, S. J., & Waring, M. J. (1978) *Biochemistry 73*, 129–144.

Kintanar, A., Klevit, R. E., & Reid, B. R. (1987) *Nucleic Acids Res. 15*, 5845–5861.

Kumar, A., Ernst, R. R., & Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun. 95*, 1–6.

Low, C. M. L., Drew, H. R., & Waring, M. J. (1984) *Nucleic Acids Res. 12*, 4865–4879.

Marion, D., Ikura, M., & Bax, A. (1989) *J. Magn. Reson. 84*, 425–430.

McClean, M. J., & Waring, M. J. (1988) *J. Mol. Recognition 1*, 138–151.

McClean, M. J., Seela, F., & Waring, M. J. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 9687–9691.

Mendel, D., & Dervan, P. B. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 910–914.

Patel, D. J. (1977) *Biopolymers 16*, 2739–2754.

Portugal, J., Fox, K. R., McLean, M. J., Richenberg, J. L., & Waring, M. J. (1988) *Nucleic Acids Res. 16*, 3655–3670.

Quigley, G. J., Ughetto, G., van der Marel, G. A., van Boom, J. H., Wang, A. H.-J., & Rich, A. (1986) *Science 232*, 1255–1258.

Rance, M., Sørensen, M., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479–485.

Scheek, R. M., Russo, N., Boelens, R., & Kaptein, R. (1983) *J. Am. Chem. Soc. 105*, 2914–2915.

Singh, U. C., Pattabiraman, N., Langridge, R., & Kollman, P. A. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 6402–6406.

Sklenář, V., & Bax, A. (1987) *J. Magn. Reson. 75*, 378–383.

States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286–292.

Ughetto, G., Wang, A. H.-J., Quigley, G. J., van der Marel, G. A., van Boom, J. H., & Rich, A. (1985) *Nucleic Acids Res. 13*, 2305–2323.

van der Marel, G. A., van Boeckel, C. A. A., Wille, G., & van Boom, J. H. (1981) *Tetrahedron Lett. 22*, 3887–3890.

van Dyke, M. M., & Dervan, P. B. (1984) *Science 225*, 1122–1127.

Wakelin, L. P. G., & Waring, M. J. (1976) *Biochem. J. 157*, 721–740.

Wang, A. H.-J., Ughetto, G., Quigley, G. J., Hakoshima, T., van der Marel, G. A., van Boom, J. H., & Rich, A. (1984) *Science* 4225, 1115-1121.

Wang, A. H.-J., Ughetto, G., & Quigley, G. J. (1986) *J. Biomol. Stereodyn. 4*, 319–342.

Waring, M. J., & Wakelin, L. P. G. (1974) *Nature 252*, 653–657.

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