Sequence Specificity of Quinoxaline Antibiotics. 1. Solution Structure of a 1:1 Complex between Triostin A and $[d(GACGTC)]_2$ and Comparison with the Solution Structure of the [*N*-MeCys^3,*N*-MeCys^7]TANDEM–$[d(GATATC)]_2$ Complex[†,‡]

Kenneth J. Addess and Juli Feigon*

Department of Chemistry and Biochemistry and Molecular Biology Institute, University of California, Los Angeles, California 90024-1569

Received March 31, 1994; Revised Manuscript Received July 15, 1994[⊗]

ABSTRACT: Triostin A, a naturally occurring quinoxaline antibiotic that contains *N*-methyl groups on the valine and cysteine residues, binds sequence specifically to DNA at NCGN sites. [*N*-MeCys^3,*N*-MeCys^7]-TANDEM (CysMeTANDEM), a synthetic quinoxaline antibiotic, differs in its chemical structure from triostin A only at the valine residues, which contain no *N*-methyl substituents. CysMeTANDEM has a sequence specificity different from triostin A, binding specifically to DNA at NTAN sites. To understand the factors that determine the sequence specificity of these quinoxaline antibiotics, the solution structure of a 1:1 complex of triostin A with the DNA hexamer $[d(GACGTC)]_2$ has been determined using NMR-derived distance and dihedral angle restraints. The solution structure of the triostin A–$[d(GACGTC)]_2$ complex is compared directly to the solution structure of a 1:1 complex of CysMeTANDEM with $[d(GATATC)]_2$ and is also compared to the crystal structure of 2:1 complex of triostin A with $[d(CGTACG)]_2$. Triostin A binds to $[d(GACGTC)]_2$ as a bis-intercalator around the CpG step, and the peptide ring of the drug binds in the minor groove of the DNA. The central C·G base pairs of the complex are underwound with an average helical twist angle of ~–9.0° and buckle inward by about 25°. There are intermolecular hydrogen bonds between each of the Ala NH and the GN3 protons of the CpG binding site. Similar structural features are observed in the solution structure of the CysMeTANDEM–$[d(GATATC)]_2$ complex. However, in the structure of the triostin A–$[d(GACGTC)]_2$ complex, two intermolecular hydrogen bonds between each of the Ala CO oxygens of the drug and the 2-amino protons of guanine are observed. These hydrogen bonds do not form in the CysMeTANDEM–DNA complex. Instead, CysMeTANDEM contains two intramolecular hydrogen bonds between the Ala CO atoms and the Val amide protons, making the Ala CO atoms unavailable to form two intermolecular hydrogen bonds. The role of these intermolecular hydrogen bonds in the CpG specificity of triostin A is discussed.

Triostin A is a quinoxaline antibiotic that was first isolated from *Streptomyces* strain s-2-210L (Yoshida & Katagiri, 1969; Katagiri *et al.*, 1974). Triostin A contains two planar aromatic quinoxaline rings that are covalently attached to a cyclic octadepsipeptide ring (Chart 1). The two L-Cys[1] residues form a disulfide bond. Both L-Cys and L-Val residues contain methylated amide nitrogens; the two D-Ser residues are amide bonded to the two quinoxaline rings. Echinomycin, another well-known naturally occurring quinoxaline antibiotic, has a chemical structure similar to triostin A, except that echinomycin contains a thioacetal bridge instead of a disulfide bridge. [*N*-MeCys^3,*N*-MeCys^7]TANDEM (hereafter referred to as CysMeTANDEM) is a synthetic derivative of triostin A that differs in its chemical structure only at the two valine residues, which contain no *N*-methyl substituents (Chart 1). DNase I footprinting studies showed that triostin A and echinomycin bind sequence specifically to double-stranded DNA at NCGN sites (Low *et al.*, 1984; van Dyke & Dervan, 1984). In contrast, NMR studies of the binding of Cys-MeTANDEM to the DNA octamers $[d(GGATATCC)]_2$[2] and $[d(GGTTAACC)]_2$ established that CysMeTANDEM binds sequence specifically to DNA at NTAN sites (Addess *et al.*, 1992). This TpA specificity has been corroborated by footprinting studies of CysMeTANDEM on a series of DNA fragments containing alternating A-T sequences (Waterloh *et al.*, 1992; Lavesa *et al.*, 1993).

Rich and co-workers solved the crystal structure of the 2:1 drug–DNA complex of triostin A bound to the DNA hexamer $[d(CGTACG)]_2$ (Wang *et al.*, 1984; Ughetto *et al.*, 1985). The crystal structure showed that two molecules of triostin A bind sequence specifically in the minor groove of the two CpG steps by bis-intercalation with the quinoxaline rings bracketing the two CpG binding sites. The two A·T base pairs between the quinoxaline rings of the two drugs are Hoogsteen base-paired in the complex. Bis-intercalative binding by triostin A to each CpG binding site causes the DNA to unwind to a helical twist angle of about –10° and the

† This work was supported by a grant from the NIH (R01 GM 37254) and by an NSF Presidential Young Investigator Award with matching funds from AmGen Inc., Monsanto Co., and Sterling Winthrop Drug Inc. to J.F.

‡ Coordinates of all final structures have been deposited at the Protein Data Bank, Brookhaven National Laboratory, Upton, NY 11973 (filename 185D).

* Address correspondence to this author at the Department of Chemistry and Biochemistry.

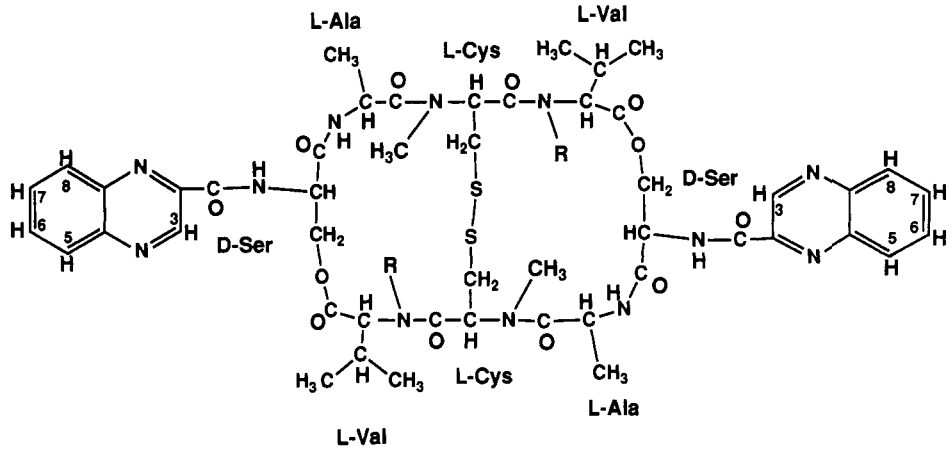
⊗ Abstract published in *Advance ACS Abstracts,* September 15, 1994.

[1] Abbreviations: 2D NMR, two-dimensional nuclear magnetic resonance spectroscopy; CysMeTANDEM, [*N*-MeCys^3,*N*-MeCys^7]-TANDEM; Cys, cysteine; HOHAHA, homonuclear Hartmann–Hahn spectroscopy; NOESY, nuclear Overhauser effect spectroscopy; P.COSY, purged COSY; Ser, serine; Val, valine; Ala, alanine; RMSD, root mean square difference.

[2] Here and for subsequent sequences, the central two bases of the sequence-specific binding site of the bound drug are underlined.

0006-2960/94/0433-12386$04.50/0

Chart 1: Chemical Structures of CpG-Specific Triostin A (R = CH_3), and TpA-Specific [*N*-MeCys^3,*N*-MeCys^7]TANDEM (R = H)



base pairs to buckle inward by about 20° (Ughetto *et al.*, 1985; Wang *et al.*, 1984, 1986). On the basis of the crystal structure of the 2:1 drug–DNA complex, they proposed that the structural basis of the sequence-specific binding of triostin A to NCGN sites was the formation of three intermolecular hydrogen bonds, two between each of the two Ala NH and guanine N3 and one between an Ala CO and a guanine 2-amino group (Ughetto *et al.*, 1985; Wang *et al.*, 1984, 1986). The second possible intermolecular hydrogen bond between the symmetry-related Ala CO and guanine 2-amino group was not observed in the crystal structure.

The enzymatic and chemical footprinting and crystallographic studies of these drugs raised questions about what determines that triostin A binds specifically to CpG steps and CysMeTANDEM binds specifically to TpA steps. To investigate the structural basis for these differences, we solved the three-dimensional solution structure of CysMeTANDEM bound to the DNA hexamer [d(GATATC)]$_2$ (Addess *et al.*, 1993). The structural features of the DNA in the CysMeTANDEM–[d(GATATC)]$_2$ complex are similar to those of the DNA in the crystal structure of the triostin A–[d(CGTACG)]$_2$ complex. The two central A·T base pairs of the complex are underwound with an average helical twist angle of about −10° and buckle inward by about 20°. The structure of the CysMeTANDEM–[d(GATATC)]$_2$ complex also contains two intermolecular hydrogen bonds between each of the Ala NH protons of the drug and the AN3 atoms of the TpA binding site, analogous to the two intermolecular hydrogen bonds between the Ala NH and GN3 of each CpG binding site in the crystal structure. However, the NMR structure of the peptide ring of CysMeTANDEM in the complex differs from that of triostin A in the crystal structure due to the formation of two intramolecular hydrogen bonds between the Val NH protons and the Ala carbonyls. These hydrogen bonds were also observed in the crystal structure of TANDEM alone (Viswamitra *et al.*, 1981; Hossain *et al.*, 1982). These intramolecular hydrogen bonds cannot form in triostin A because of the *N*-methyl substituents attached to both Val amides.

Although the crystal structure of the triostin A complex shows only one intermolecular hydrogen bond between an Ala CO and a guanine 2-amino group, recent footprinting studies indicate that echinomycin cannot bind tightly to DNA in which inosines (I) are substituted for guanine (G) on either strand of DNA (Marchand *et al.*, 1992). This suggests that two hydrogen bonds between each of the two Ala carbonyls and the two guanine 2-amino groups are critical to the sequence-specific binding of triostin A to DNA in solution and indicates that the crystal structure does not completely explain the structural basis of the CpG specificity of triostin A in solution.

As part of a program to determine the structural basis of the sequence specificity of these quinoxaline antibiotics, we have solved the three-dimensional structure of a 1:1 complex of triostin A bound to the DNA hexamer [d(GACGTC)]$_2$ using two-dimensional ^{1}H NMR derived distance and dihedral angle restraints. Details of the structure are presented and compared to the solution structure of the analogous CysMeTANDEM–[d(GATATC)]$_2$ complex and the crystal structure of the 2:1 complex of triostin A with [d(CGTACG)]$_2$.

MATERIALS AND METHODS

Sample Preparation. The hexamer d($G_1A_2C_3G_4T_5C_6$) was synthesized on an ABI 381A synthesizer using β-cyanoethyl phosphoramidite chemistry on a 10-μmol scale and purified by gel filtration as previously described (Addess *et al.*, 1992). NMR samples of the free DNA were prepared as previously described (Addess *et al.*, 1992).

Triostin A was a gift from Dr. M. Shin, Shionogi Research Laboratories, Osaka, Japan. The quinoxaline antibiotic was assessed to be pure by one-dimensional proton NMR spectroscopy. The triostin A–[d(GACGTC)]$_2$ complex was formed by adding triostin A dissolved in acetonitrile directly to the free DNA solution in the NMR tube. The concentration of the triostin A stock solution was calculated on the basis of the UV absorbance using an extinction coefficient of 11 500 M^{-1} cm^{-1} at 325 nm (Low *et al.*, 1984). The acetonitrile/water mixture was then evaporated under a stream of N_2 gas, and the sample was redissolved in D_2O or 90% H_2O/10% D_2O. Triostin A was added until a saturated complex formed; i.e., excess triostin A precipitated out of solution. Complex formation was assayed by one- and two-dimensional NMR. The final conditions of the NMR sample were ~2 mM DNA and triostin A and 150 mM NaCl, pH 6.6 (meter reading).

NMR Spectroscopy. All NMR experiments were done at 500 MHz on a General Electric GN500 spectrometer. Phase-sensitive nuclear Overhauser effect (NOESY) spectra in D_2O were acquired using the hypercomplex method (States *et al.*, 1982) with low-power preirradiation of the HDO peak during the recycle delay (Kumar *et al.*, 1980). The NOESY spectrum used in the assignments of triostin A–[d(GACGTC)]$_2$ was obtained with a mixing time of 150 ms at 25 °C. NOESY spectra used in the structure determination were obtained with mixing times of 75 and 100 ms at 25 °C. Phase-sensitive NOESY spectra in H_2O were obtained by replacing the last 90° pulse with a $1\bar{1}$ spin–echo pulse sequence and the

appropriate phase cycle to suppress the large water resonance (Sklenář & Bax, 1987). The carrier frequency was centered at the water resonance, and the delay τ was adjusted so that the excitation maximum was at ~12 ppm. HOHAHA spectra were acquired using the MLEV17 mixing sequence and 1.5-ms trim pulses for the spin lock (Bax & Davis, 1985). ROESY spectra were acquired with a spin-lock field of 5000 Hz (Bothner-by *et al.*, 1984). P.COSY spectra were acquired with a flip angle mixing pulse of 90° (Marion & Bax, 1988). All 2D NMR spectra were processed on a Personal Iris 4D25 using FELIX 1.1 (Hare Research). NOESY spectra were baseline flattened with a first-order polynomial in t_2 and with a second-order polynomial in t_1. Detailed descriptions of other acquisition and processing parameters are given in the figure captions.

Structure Determination. NOE cross-peak volumes of the nonexchangeable proton resonances at the two different mixing times (75 and 100 ms) were obtained by cross-peak integration using the FELIX 2.05 software (Hare Research, Woodinville, WA). NOE cross-peak volumes of exchangeable proton resonances were obtained by cross-peak integration of a NOESY spectrum of the sample in H_2O acquired with τ_m = 50 ms and a $1\bar{1}$ spin–echo observe pulse. Calculating the volumes involve summing all the points of the transformed spectra within the limits of a rectangle that surrounds the cross peak. These volumes were converted into distances by the isolated two-spin approximation, using the CH5–CH6 distance of 2.44 Å as a calibration. The distances obtained for cross peaks on both sides of the diagonal of both 75- and 100-ms mixing time NOESY spectra were averaged. The criteria used for the upper and lower bounds for distances between the nonexchangeable protons were the same as those for the distances between the nonexchangeable protons in the calculation of the three-dimensional structure of the Cys-MeTANDEM–[d(GATATC)]2 complex (Addess *et al.*, 1993). The lower bounds for distances were set to 1.9 Å, while the upper bounds that were used depended on the distance between the two protons, as follows: $d \leq 3.0$, upper bound $1.1d$; $3.0 < d \leq 3.5$, upper bound $1.2d$; $3.5 < d \leq 4.5$, upper bound $1.3d$; and $4.5 < d$, upper bound $1.4d$, where d symbolizes the interproton distance. Because exchange between DNA resonances of the free hexamer and the complex was observed in both D_2O NOESY spectra, thereby influencing the accuracy of the NOE cross-peak volumes and the calculated distances, an additional 0.5 Å was added to the upper bounds of the nonexchangeable interproton distance restraints. For distances involving exchangeable resonances, the upper bound was increased by 20% to account for a possible decrease in the NOE cross-peak volume due to exchange with H_2O. The dihedral bond angle restraints were obtained by calculating the proton–proton torsional angles, ϕ_{HH} ± 30%, of each deoxyribose sugar on the basis of whether the sugar adopted mostly an N-type geometry (P = 9°) or S-type geometry (P = 162°), keeping the sugar pucker amplitude, τ_m, at a constant value of 40° (Rinkel & Altona, 1987). Sugars were determined to be either N-type or S-type on the basis of the P.COSY spectrum shown in Figure 5.

A total of 366 distance restraints, representing 2 × 183 asymmetric distance restraints, were used in the structure calculations. Thirty-seven of these 183 asymmetric distance restraints were between the drug and the DNA. Like the CysMeTANDEM–[d(GATATC)]$_2$ complex, the triostin A–[d(GACGTC)]$_2$ complex has a dyad symmetry in solution; however, no symmetry restraints were imposed during the refinement. Ten initial structures were generated by embedding distances into Cartesian coordinates using the distance geometry algorithm (Crippen & Havel, 1988) with the substructure embedding routine in X-PLOR 3.1 (Brünger, 1992; Kuszewski *et al.*, 1992). In the substructure embedding routine, 184 of the 516 atoms of the drug–DNA complex were embedded into Cartesian space. Five of these ten structures, called TD1–TD5, were refined by using the standard energy minimization, restrained molecular dynamics, and simulated annealing routines in X-PLOR 3.1 (Brünger, 1992). The protocol used to refine these structures has been described in detail previously (Addess *et al.*, 1993). All structures were visualized using Insight 2.2 (Biosym Technologies, San Diego, CA). Helical parameters of the complex were calculated with the program NEWHELIX91 (obtained from R. E. Dickerson).

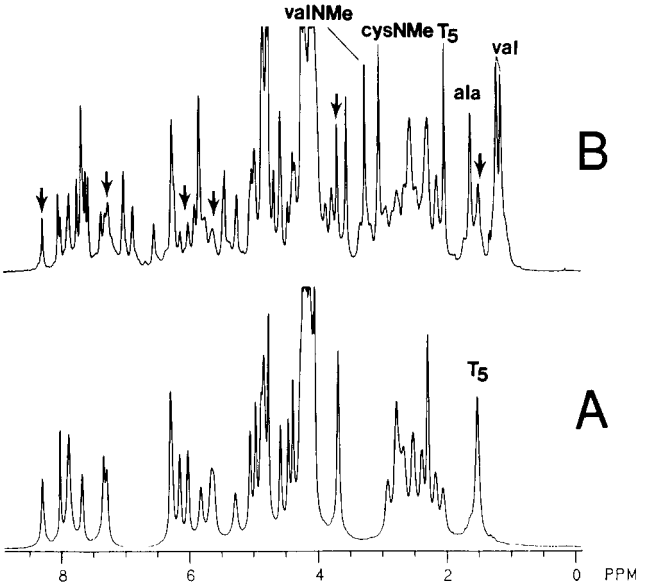


FIGURE 1: One-dimensional ^{1}H NMR spectra of (A) [d(GACGTC)]$_2$ and (B) the 1:1 complex of triostin A + [d(GACGTC)]$_2$ in D_2O at 25 °C. Free DNA peaks in (B) are indicated by arrows. Samples are 2 mM DNA duplex and 150 mM NaCl, pH 6.5. Assignments of the thymine methyls, Ala methyl, Val methyl, Val NMe, and Cys NMe are indicated. The spectra were acquired with a sweep width of 5000 Hz in 2K complex points and line broadened by 3 Hz prior to Fourier transformations.

RESULTS

Complex Formation. The one-dimensional (1D) proton NMR spectra in D_2O of [d(GACGTC)]$_2$ containing no triostin A and 1 equiv of triostin A per DNA duplex are shown in panels A and B of Figure 1, respectively. Addition of triostin A to the DNA duplex causes the characteristic shift in the DNA resonances observed in the formation of the CysMe-TANDEM–[d(GATATC)]$_2$ complex (Addess *et al.*, 1993). A change in the chemical shift of the T_5 methyl resonance and the appearance of the Cys N-methyl, Ala methyl, Val methyl, and Val N-methyl resonances of the drug, which are labeled in Figure 1B, indicate that triostin A binds sequence specifically to the CpG step of [d(GACGTC)]$_2$. However, in contrast to the titration of the [d(GATATC)]$_2$ duplex with CysMe-TANDEM, titration of [d(GACGTC)]$_2$ with triostin A does not result in a saturated complex at ≥1:1 ratio of drug to DNA. This is indicated by the presence of resonances at the chemical shifts of the free DNA in the spectrum (Figure 1B). The integrated intensity of the T_5Me resonance peak of the complex compared to that of the T_5Me resonance of "free DNA" indicates that only about 70% of the available CpG binding sites are occupied by triostin A. This indicates that the binding of triostin A to [d(GACGTC)]$_2$ is weaker than the binding of CysMeTANDEM to [d(GATATC)]$_2$.

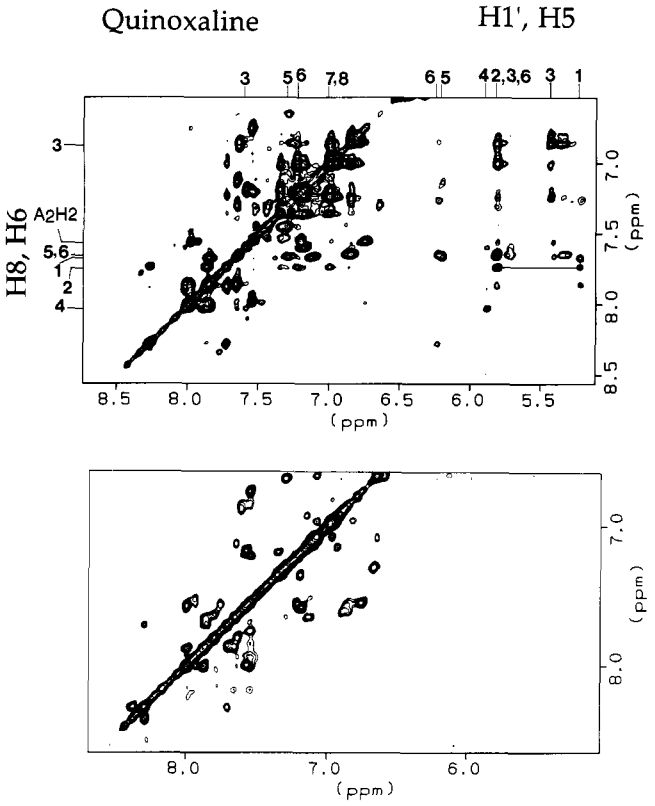


FIGURE 2: (A, top) Expanded region of the NOESY spectrum (τ_m = 150 ms) of the triostin A–[d(GACGTC)]$_2$ complex in D_2O at 25 °C, showing the aromatic and the base–H1′ region. The sample is the same as in Figure 1. The quinoxaline, aromatic, H5, and H1′ resonances are indicated. The base–H1′ sequential connectivity between bases 1 and 2 is indicated by a solid line. The spectrum was acquired with a sweep width of 5000 Hz in both dimensions, 300 t_1 values of 64 scans, and 2K complex points. A total of 300 points were apodized in both dimensions with a sine squared bell phase shifted by 60°. Data in t_1 were zero filled to 1K prior to Fourier transformation. (B, bottom) ROESY of (A) at 25 °C. The spectrum was acquired with a spin-lock field of 5 kHz for 100 ms, a sweep width of 5000 Hz in both dimensions, 230 t_1 values of 64 scans, and 2K complex points. A total of 230 points were apodized in both dimensions with a sine squared bell phase shifted by 60°. Data in t_1 were zero filled to 1K prior to Fourier transformation.

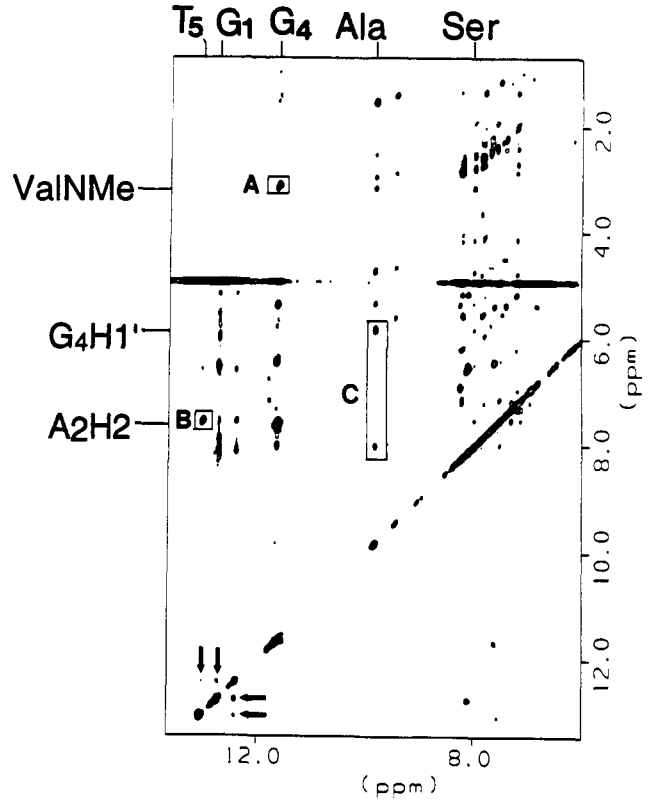


FIGURE 3: Expanded region of the H_2O NOESY spectrum of the 1:1 triostin A–[d(GACGTC)]$_2$ complex at 5 °C, showing the imino, amino, and aromatic resonances and their cross peaks. Assignments of the iminos, Ser NH, Ala NH, A_5H2, G_5H1', and CysNMe are indicated. The boxed cross peaks are (A) G_4 imino to CysNMe, (B) T_5 imino to A_2H2, and (C) Ala NH to G_4H1' and to Ser NH. Some exchange cross peaks to a small amount of another drug–DNA complex are indicated by arrows. The spectrum was acquired with a sweep width of 10 000 Hz in both dimensions, 256 t_1 values of 64 scans, and 4K points. Before processing, the residual water signal was removed from the FIDs by a Gaussian window function with K = 32 and an extrapolation with M = 16 (Marion *et al.*, 1989).

Chemical Exchange in the Triostin A–[d(GACGTC)]$_2$ Complex. Figure 2A shows the expanded region of a NOESY spectrum (τ_m = 150 ms) of the triostin A–[d(GACGTC)]$_2$ complex. Many of the cross peaks shown in this region of the spectrum are due to chemical exchange between the resonances of the "free" DNA and the DNA resonances of the complex. In addition, we observe exchange cross peaks between the resonances of triostin A in the complex and some resonance intensities for triostin A that is not intercalated. The exchange cross peaks can be identified in the ROESY spectrum shown in Figure 2B. The integrated intensity of the diagonal peak of the Cys αH resonance of the nonintercalated triostin A in the NOESY spectrum indicates that the concentration of the triostin A not associated with the intercalated complex in the D_2O solution containing about 2 mM duplex DNA hexamer is about 600 μM. This is about 100 times greater than the saturating concentration of free triostin A in aqueous solution. For nonintercalated triostin A to remain in solution at a concentration of 600 μM, it must be bound to the DNA in a nonspecific, nonintercalative manner. The chemical shifts of the DNA resonances of this nonspecific complex are almost identical to the chemical shifts of the free DNA. The chemical exchange cross peaks in Figure 2 reveal that the specific and nonspecific, nonintercalative complexes of triostin A with [d(GACGTC)]$_2$ are in slow exchange on the NMR time scale. This result is quite different from the NOESY spectrum of a half-saturated complex of CysMeTANDEM–[d(GATA-TC)]$_2$, which contains no chemical exchange cross peaks between the free DNA and the resonances of the specific drug–DNA complex (Addess *et al.*, 1993) and shows no evidence for nonspecific drug binding.

In the case where slow exchange is observed, the exchange rate between the specific and nonspecific complexes is much smaller than the smallest chemical shift difference between the resonances of the specific complex and the resonances of the nonspecific complex, or $k_{ex} \ll 2\pi|\nu_s - \nu_{ns}|$. The smallest chemical shift difference between the specific and nonspecific triostin A–[d(GACGTC)]$_2$ complexes is about 10 Hz, which occurs for the Cys αH resonances. Therefore, k_{ex} between the specific and the nonspecific complexes must be $\ll 60$ s^{-1}.

Assignment of the Resonances in the Triostin A–[d(GACG-TC)]$_2$ Complex. Even though the NOESY spectrum of the triostin A–[d(GACGTC)]$_2$ complex has many exchange cross peaks and is not as well resolved as the NOESY spectrum of the CysMeTANDEM–[d(GATATC)]$_2$ complex (Addess *et al.*, 1993), the NOE cross peaks are sufficiently resolved to assign a majority of these cross peaks. Assignment of the nonexchangeable and exchangeable DNA resonances of the 1:1 complex was made using information provided by the D_2O NOESY (Figure 2A), H_2O NOESY (Figure 3), P.COSY (Figure 4), and HOHAHA spectra (not shown).

The assignments of the DNA aromatic and H1′ protons are indicated in the expanded region of the NOESY spectrum

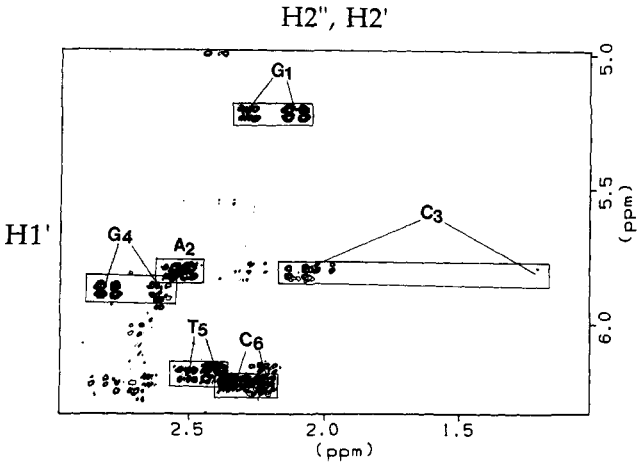


FIGURE 4: P.COSY spectrum of the 1:1 triostin A–[d(GACGTC)]$_2$ complex showing the region containing the H1′–H2′ and H1′–H2″ cross peaks at 25 °C. Cross peaks for each sugar are boxed and labeled. The spectrum was acquired with a sweep width of 5000 Hz in both dimensions, 16 scans per t_1 value, 4K complex points, and 850 t_1 values, respectively. A reference spectrum of one t_1 value was acquired with a sweep width of 5000 Hz, 256 scans, and 8K complex points. The reference spectrum was subtracted from each two-dimensional data set. A total of 850 points, respectively, were apodized with a sine bell phase shifted by 45°. The data were zero filled to 2K points in t_1 prior to Fourier transformation.

shown in Figure 2A. The intercalation of the DNA bases by the quinoxaline ring interrupts the base–H1′ sequential connectivities that are observed in the NOESY spectrum of the free hexamer (data not shown). However, a base–H1′ connectivity between G_1 and A_2 is observed, which is indicated by a line in Figure 2A. An almost identical connectivity from G_1 to A_2 has been observed in the NOESY spectrum of the CysMeTANDEM–[d(GATATC)]$_2$ complex (Addess *et al.*, 1993). The T_5 and C_6 residues of the triostin A-DNA complex also resonate at frequencies similar to the T_5 and C_6 residues of the CysMeTANDEM–DNA complex. Both the H2′ and H2″ protons of C_3 and G_4 of the complex also resonate at frequencies similar to the H2′ and H2″ protons of the T_3 and A_4 residues of the CysMeTANDEM complex (Addess *et al.*, 1993). C_3H2′ and C_3H2″ are shifted upfield by approximately 1 ppm relative to these resonances in the free hexamer, while the G_4H2″ proton resonates at higher field than the G_4H2′ proton (see Figure 4). Starting from these assignments, we identified all the nonexchangeable protons except for a few H5′ and H5″ resonances.

Figure 3 shows the NOESY spectrum of the triostin A–[d(GACGTC)]$_2$ complex in H_2O. The region shown includes an expansion of the imino and aromatic regions of the spectrum. Signals of the DNA resonances from both the free hexamer and the complex are present in the H_2O NOESY spectrum as well as in the D_2O NOESY spectrum. The exchange cross peaks between the imino resonances of the free hexamer and those of the complex are indicated by arrows. The G_4 imino of the complex can be identified easily by its NOE to the Val *N*-methyl resonances of triostin A (Figure 3, box A). This NOE defines an important structural difference between the CysMeTANDEM complex and the triostin A complex (see below). The G_4 imino has a chemical shift of ~11.5 ppm, which is about 1 ppm upfield from where G_4 resonates in the free hexamer. This upfield shift is characteristic of bis-intercalative binding of the DNA by the drug (Feigon *et al.*, 1984). The chemical shifts of the G_1 imino and T_5 imino, the latter of which is identified on the basis of its cross peak to the A_2H2 proton (Figure 3, box B), are very similar to these corresponding imino protons in the CysMeTANDEM–[d(GATATC)]$_2$ complex. The C aminos were identified on the basis of their cross peaks to the G iminos.

The chemical shifts of the four peptide spin systems of triostin A were identified in a HOHAHA spectrum and a P.COSY spectrum. Assignment of the quinoxaline ring protons is indicated in the expanded region of the NOESY spectrum shown in Figure 2A. With the exception of the Val *N*-methyl groups, the NOE cross-peak patterns of the triostin A–[d(GACGTC)]$_2$ complex are very similar to those of the CysMeTANDEM–[d(GATATC)]$_2$ complex. The proton chemical shifts of the triostin A–[d(GACGTC)]$_2$ complex are listed in Table 1.

Conformation of the Deoxyriboses. An expanded region of the P.COSY spectrum of the triostin A–[d(GACGTC)]$_2$ complex showing the H1′ to H2′,H2″ region is displayed in Figure 4. Many of the extra cross peaks seen in this spectrum are H1′ to H2′,H2″ cross peaks of the "free" DNA. The patterns displayed by G_1 and T_5 and the extent of overlap of the H1′ to H2′,H2″ cross peaks of the A_2 and C_6 are almost identical to those of the corresponding nucleotides in the CysMeTANDEM–[d(GATATC)]$_2$ complex, indicating that the conformations of these sugars are identical in both complexes (Addess *et al.*, 1993). The cross-peak pattern of the C_3 sugar is also similar to the pattern of the T_3 sugar in the CysMeTANDEM–[d(GATATC)]$_2$ complex. These results indicate that the deoxyribose sugars of the corresponding nucleotides in both complexes are essentially the same. In the structure of the CysMeTANDEM–[d(GATATC)]$_2$ complex, the values for pseudorotation phase angle *P* for the G_1, A_2, G_4, and T_5 sugars are close to S-type (near C2′ endo),

Table 1: Assignments of the Triostin A–[d(GACGTC)$_2$] Complex

	[d(GACGTC)]$_2$ chemical shifts (ppm)[a]									
	H8/H6	Me/H5 H2	H1′	H2′	H2″	H3′	H4′	H5′,H5″	imino[b]	amino[b]
G_1	7.69		5.23	2.12	2.29	5.00	3.52	3.98	12.74	
A_2	7.75		5.83	2.54	2.54	4.96	4.19			
C_3	6.87	5.44	5.82	1.19	2.11	4.37	3.74	3.98, 4.19		6.48, 7.63
G_4	8.04		5.89	2.82	2.53	4.81	4.32		11.65	
T_5	7.68	2.00	6.21	2.43	2.51	5.01	4.19	4.01, 4.09	13.03	
C_6	7.66	5.82	6.25	2.27	2.27	4.55	4.02			6.58, 8.11

	triostin A chemical shifts (ppm)[a]						
	α	β	β′	γ	γ′	NH	NMe
serine	5.43	4.57	4.83				
valine	4.85	2.55		1.12	1.19		3.24
cysteine	6.53	3.31	3.15				3.03
alanine	4.79	1.59				9.85	
quinoxaline	7.59 (3)[c]	7.36 (5)	7.24 (6)	6.89 (7)	6.89 (8)		

[a] Chemical shifts relative to DDS. [b] Chemical shifts at 10 °C. [c] Positions in parentheses.

Table 2: X-PLOR Energies (kcal mol^{-1}) for the Five Solution Structures of the Triostin A–[d(GACGTC)]$_2$ Complex

	TD1	TD2	TD3	TD4	TD5
Distance Geometry Structures					
E_{total}[a]	173	213	224	293	366
E_{NOE}	28	32	26	44	57
Restrained Molecular Dynamics and Energy-Minimized Refined Structures					
E_{total}	−220	−220	−225	−216	−215
E_{NOE}	9	11	11	11	10

[a] E_{total} includes potential energy functions for the Lennard-Jones potential energy function (E_{VDW}), NOE distance restraints (E_{NOE}), covalent bond lengths (E_{BOND}), electrostatic interactions (E_{ELEC}), covalent bond angles (E_{ANGL}), dihedral angle restraints (E_{CDIH}), dihedral angle energies (E_{DIH}), and improper dihedral angles (chirality and planarity) (E_{IMPR}). The formulas for these potential energy functions are given in Brünger (1992).

Table 3: Atomic RMS Difference (Å) between Energy-Refined Structures of Triostin A–[d(GACGTC)]$_2$[a,b]

structure	TD1	TD2	TD3	TD4	TD5
TD1	1	1.12	1.26	1.50	1.18
TD2		2	1.45	1.50	1.40
TD3			3	1.47	1.48
TD4				4	1.39
TD5					5

[a] Calculated for heavy atoms of fission A and the four internal residues of [d(GACGTC)]$_2$. [b] Average: 1.37 ● 0.13 Å.

while those for the T_3 and C_6 sugars are close to N-type (near C3′ endo) (Addess *et al.*, 1993). Thus, all the sugars in the structure of the triostin A–[d(GACGTC)]$_2$ complex are predominantly S-type (near C2′ endo), except for the C_3 and C_6 sugars, which are N-type (near C3′ endo). The geometry of the C_3 sugar is also similar to the C sugar of the CpG step in echinomycin–DNA complexes in solution (Gao & Patel, 1988, 1989; Gilbert & Feigon, 1991, 1992). For G_4, the H2″ resonance is shifted upfield from the H2′ resonances, similar to the A_4 of the CysMeTANDEM complex, which indicates that G_4 sugar and the A_4 sugar are in similar chemical environments.

Structure Calculations. Five of the ten generated starting structures had no distance violations greater than 0.5 Å. These were refined with energy minimization and restrained molecular dynamics. The total potential energy values and the NOE energy values of the five starting and final refined structures are shown in Table 2. The starting structures TD1–TD5 agree very well with the experimental distance restraints. Visually, all starting structures are underwound right-handed DNA helices with triostin A bound in the minor groove of the CpG step as a bis-intercalator (not shown). Refinement caused only a slight improvement in the NOE potential energy term but caused a major improvement in the total potential energy term. This relieved all the covalent, dihedral bond angle violations as well as flattening the bases and the quinoxaline rings. In each of the final structures, no more than 6 of the 366 distances had violations of greater than or equal to 0.2 Å and no final structure contained a violation greater than 0.4 Å.

Figure 5 shows stereoviews of five superimposed energy-refined structures, TD1–TD5, into the major groove and from the side of the complex. RMSDs calculated using all the heavy atoms of the drug and the four internal residues of the DNA between all possible pairs of structures TD1–TD5 are listed in Table 3. Because the 1:1 complex of triostin A with [d(GACGTC)]$_2$ is not as stable as the complex formed between CysMeTANDEM and [d(GATATC)]$_2$ (Figure 2) and is in slow exchange with its nonspecific, nonintercalative counterpart (Figure 3A,B), the refined structure cannot be as precisely determined as the structure of the CysMeTANDEM–[d(GATATC)]$_2$ complex. Nevertheless, the most important parts of the structure for this work, the triostin A and the internal C_3·G_4 pairs, are fairly well defined. The average pairwise RMSD for triostin A and the two internal G·C residues excluding the phosphate backbone is 0.98 ± 0.17 Å. The high definition of this part of the structure is important because it clearly reveals some key similarities and differences between the structure of the triostin A–[d(GACGTC)]$_2$ complex and the structure of the CysMeTANDEM–[d(GATATC)]$_2$ complex. These are discussed below.

DISCUSSION

Kinetics of Binding of Triostin A to [d(GACGTC)]$_2$. The NOESY spectrum of the triostin A–[d(GACGTC)]$_2$ complex (Figure 3A) contains exchange cross peaks that are not observed in the NOESY spectrum of the CysMeTANDEM–[d(GATATC)]$_2$ complex (Addess *et al.*, 1993). Since triostin A is essentially insoluble in aqueous solution, the exchange cross peaks in the NOESY spectrum of the triostin A–[d(GACGTC)]$_2$ complex must be between the intercalated, specific complex and a nonspecific, nonintercalated drug–DNA complex. This suggests that binding of triostin A to a DNA sequence containing a CpG step involves a two-step kinetic mechanism according to eq 1. In the first step, which

$$\text{drug} + \text{DNA} \xrightarrow{k_1} \text{nonspecific drug–DNA complex} \underset{K_d}{\overset{K_a}{\rightleftarrows}} \text{intercalated drug–DNA complex} \quad (1)$$

$$K_a = k_{on}/k_{off} \quad (2)$$

is irreversible in aqueous solution (since any uncomplexed triostin A will precipitate out), the drug and the DNA associate to form a nonspecific complex. NMR evidence for nonspecific binding of a quinoxaline antibiotic to oligonucleotides that do not contain sequence-specific binding sites has been observed in 1:1 mixtures of CysMeTANDEM with [d(GGAATTCC)]$_2$ and [d(CGCGATCGCG)]$_2$ (Addess *et al.*, 1992; Powers *et al.*, 1989). Specific complex formation occurs in the second step of the mechanism. The reversibility of this second step, which is observed from the slow exchange between the specific and nonspecific complexes in the NOESY spectra of triostin A–[d(GACGTC)]$_2$ (Figure 3A), depends on the difference in free energy between the specific complex and the nonspecific complex, as reflected by the association constant K_a. Because the specific (bis-intercalated) complex formed by [d(GACGTC)]$_2$ with triostin A is not as stable as the specific complex formed by [d(GATATC)]$_2$ with CysMeTANDEM, the K_a value for the triostin A–[d(GACGTC)]$_2$ complex, determined from [intercalated drug–DNA complex]/[nonspecific complex], (2.3) is lower than the K_a value for CysMeTANDEM bound to [d(GATATC)]$_2$ (≫2.3). The reasons for these different K_a values are discussed in more detail in the following paper (Addess & Feigon, 1994).

Structure of the Triostin A–[d(GACGTC)]$_2$ Complex. As a result of chemical exchange and weaker binding, the structure of the triostin A–[d(GACGTC)]$_2$ complex was calculated with fewer and less precise restraints than the three dimensional structure of the CysMeTANDEM–[d(GATATC)]$_2$ complex (Addess *et al.*, 1993). Thus, the average pairwise root mean square deviation (RMSD) calculated for the drug and the four internal base pairs of the five energy-refined structures

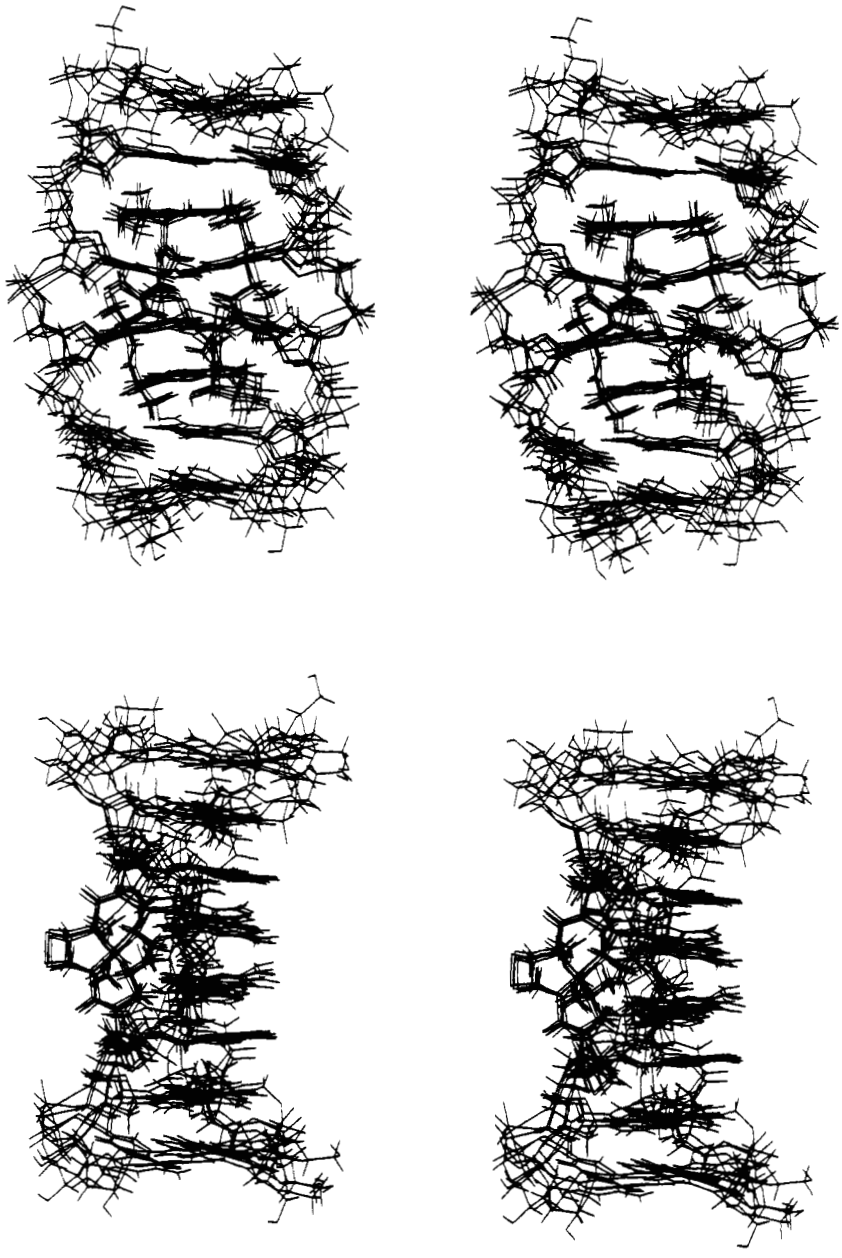


FIGURE 5: Stereoviews into the major groove (top) and of the side (bottom) of five superimposed refined structures of the triostin A–[d(GACGTC)]$_2$ complex (TD1–TD5). The structures were obtained by metric matrix distance geometry, molecular dynamics, and minimization using the program X-PLOR.

of the CpG-specific complex (1.37 Å) was 0.47 Å higher than that of the most precise structures for the TpA-specific CysMeTANDEM complex (0.90 Å). The biggest difference in the precision between the CysMeTANDEM complex and the triostin A complex is at the ends of the DNA, which were much better defined in the CysMeTANDEM complex. Nevertheless, the structure of the triostin A–[d(GACGTC)]$_2$ complex is of high enough precision to be compared directly with the structure of the CysMeTANDEM–[d(GATATC)]$_2$ complex. This enabled us to define some important similarities and differences between the two structures.

The view into the major groove and of the side of the triostin A–[d(GACGTC)]$_2$ complex of one of the five energy refined structures is shown in Figure 6. The view into the major groove of the complex shows that the structure of the complex contains an approximate 2-fold axis of symmetry located in the middle of the central CpG base pairs, perpendicular to the plane of the figure. Since symmetry was not imposed during the refinement, there is a slight variation in the two symmetric halves of the molecule. The average pairwise RMSD between the two symmetric halves of the five final structures of the triostin A–[d(GACGTC)]$_2$ complexes is about 1.0 Å. The variations in the two strands and the two symmetrical halves of the triostin A complex represent essentially a doubling of the number of structures shown. The view into the major groove also shows that the central C·G base pairs have an inward buckle. The view of the side of the complex shows that the global conformation of the DNA is an underwound right-handed B-DNA helix with the most severe local unwinding occurring between the two central CpG base pairs of the complex. The central C_3pG_4 step has an average helical twist angle of $-8.2 \pm 1.7°$. The average helical twist between the $A_2{\cdot}T_5$ base pairs and the $C_3{\cdot}G_4$ base pairs of the five structures is $-31.1 \pm 4.0°$. The average inward buckle of the central C·G base pairs is $25.5 \pm 3.6°$. Table 4 lists the average values of the pseudorotation phase angle P for the six deoxyribose sugars from each strand of the DNA. G_4 and T_5 adopt C2′ endo sugar puckers, G_1 is O4′ endo, G_4 is C1′ exo (near C2′ endo), and C_3 and C_6 are C3′ endo.

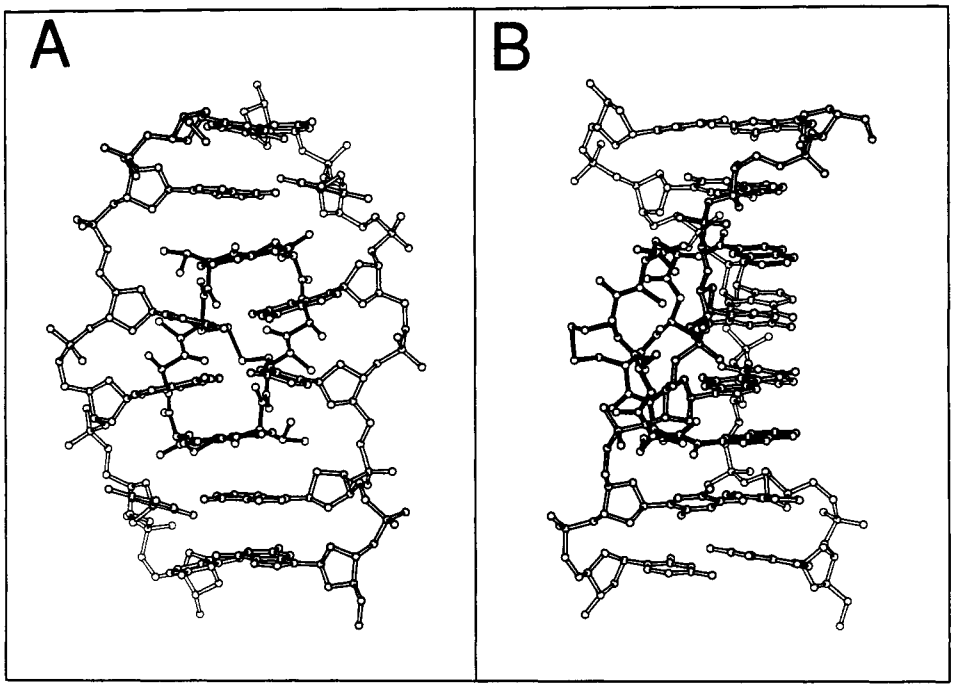


FIGURE 6: One of the five energy-refined structures of the triostin A–[d(GACGTC)]$_2$ complex that was generated by metric matrix distance geometry. The view is (A) into the major groove and (B) of the side of the complex.

Table 4: Sugar Conformations of Triostin A–[d(GACGTC)]$_2$

	pseudorotation phase angle P (deg)	puckering amplitude τ_m (deg)	puckering mode
G_1	99 ± 38	46 ± 3	O4′ endo
A_2	125 ± 69	42 ± 2	C1′ exo
C_3	17 ± 11	46 ± 4	C3′ endo
G_4	147 ± 17	44 ± 3	C2′ endo
T_5	162 ± 17	47 ± 3	C2′ endo
C_6	2 ± 4	46 ± 1	C3′ endo

As can be seen in Figure 6A, the two quinoxaline rings intercalate between the symmetry-related A_2·T_5 and C_3·G_4 base pairs, both of which are separated by ~5.0 Å to accommodate the intercalated quinoxaline rings. Figure 7 illustrates the stacking arrangements of the bases and quinoxaline rings of the triostin A complex as viewed down the helical axis. The quinoxaline rings stack very well with both A_2 bases of the A_2·T_5 base pairs while having virtually no stacking with the T_5 bases (Figure 7B). Both the C_3 and G_4 bases of the central CpG base pair stack over the quinoxaline rings, as seen in Figure 7C. There are also strong stacking interactions between the two C·G base pairs of the central C_3pG_4 steps of the complex (Figure 7D).

The side view of the complex (Figure 6B) shows that the peptide ring of triostin A is positioned in the minor groove of the DNA. Figure 8 shows a view into the minor groove triostin A bound to the central C_3pG_4 step of the complex. Van der Waals interactions occur between the Ala Me of the drug and the C_3C1', C_3C2', and G_4O4' of the DNA. The Val *N*-methyl groups of the drug face directly into the minor groove of the DNA. As a result, the triostin A-DNA complex contains van der Waals interactions between the *N*-methyl group of valine and both the 2-amino groups of G_4 and the O2 of C_3. Intermolecular hydrogen bonds are observed between the Ala NH and G_4N3 and between the Ala CO and G_4NH2. The average calculated distances between the Ala CO and G_4-NH2 group from the five refined structures is 2.2 ± 0.1 Å and between the Ala NH and G_4N3 is 2.0 ± 0.1 Å.

Comparison with the Crystal Structure of the 2:1 Triostin A–[d(CGTACG)]$_2$ Complex. The solution structure of the triostin A–[d(GACGTC)]$_2$ complex and the crystal structure

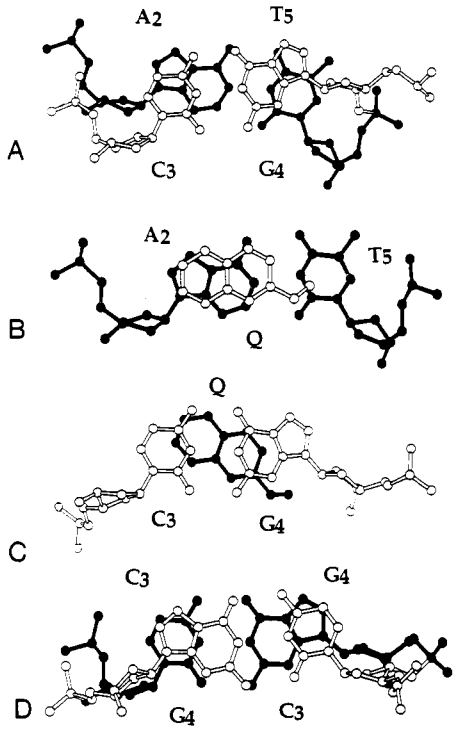


FIGURE 7: Stacking arrangements of the bases and the quinoxaline rings as viewed down the helical axis. Shown are stacking between (A) A_2·T_5 and C_3·G_4 (with the quinoxaline ring deleted), (B) A_2·T_5 and the quinoxaline ring, (C) the quinoxaline ring and C_3·G_4, and (D) the two C_3·G_4 base pairs.

of the 2:1 triostin A–[d(CGTACG)]$_2$ complex are similar in many ways, even though the solution structure contains a single CpG binding site and the crystal structure contains two CpG binding sites (Wang *et al.*, 1984, 1986; Ughetto *et al.*, 1985). In both the crystal and solution structures, the helical twist angles at the CpG steps are about −10° ; the C·G base pairs of the CpG binding site buckle inward in a range of 20–25°. The pattern of stacking interactions observed between the quinoxaline rings of triostin A and the G·C base pairs of the CpG binding site is also very similar in both the crystal structure and the solution structure. In both cases, the

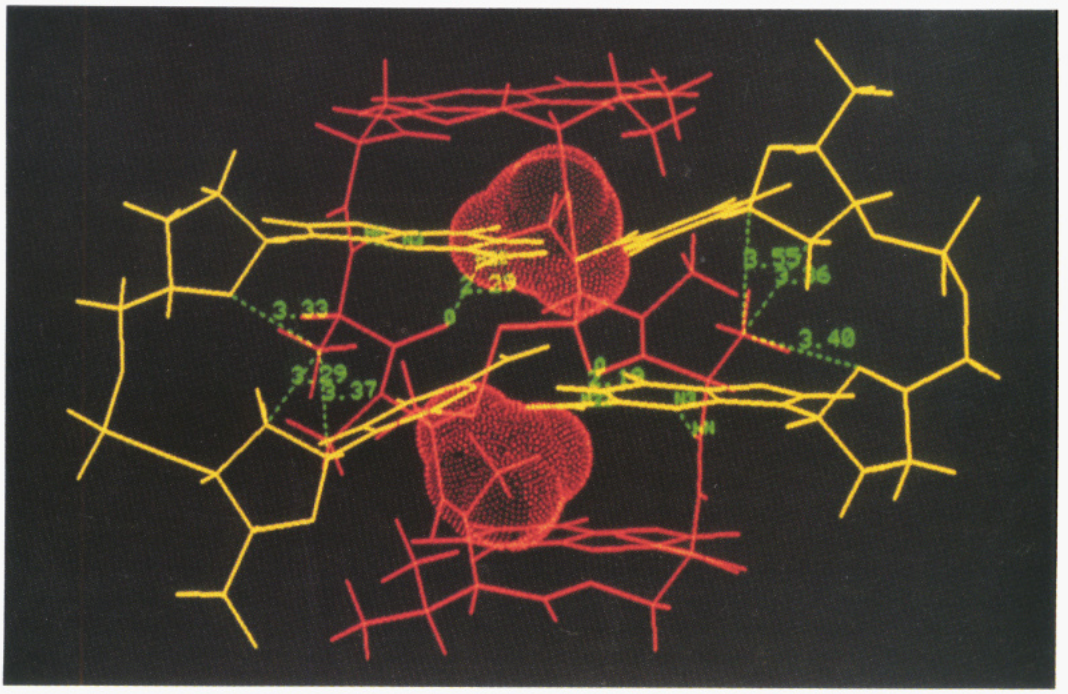


FIGURE 8: Structure of triostin A bound to the central CpG step. Triostin A is illustrated in red and the DNA is illustrated in yellow. The Val *N*-methyl groups are drawn with van der Waals surfacing. Van der Waals interactions between the Ala Me and the C_3C2', C_3C1', and G_4O4' are indicated by green dashed lines. Intermolecular hydrogen bonds between the Ala CO and G_4NH2 and between the Ala NH and G_4N3 are indicated with green dashed lines.

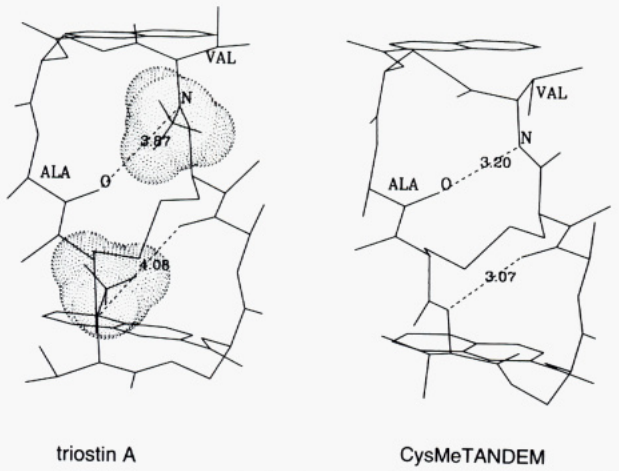


FIGURE 9: Structures of triostin A and CysMeTANDEM. The distances between each of the two Ala carbonyls and Val nitrogens of both drugs are indicated. The two Val *N*-methyl groups of triostin A are highlighted with van der Waals surfacing.

Table 5: Selected Dihedral Angles of the Peptide Backbone

residue	dihedral angle	triostin A (mean NMR)	CysMeTANDEM (mean NMR)[a]
Ser_1	N–C–CA–CB	−143	−139
	N–C–CA–N	−21	−16
	C–CA–CB–OG	74	63
Cys_1	C–CA–N–C	−117	−121
	N–C–CA–N	72	80
Ala_1	C–CA–N–C	−89	−81
	N–C–CA–N	159	157
Val_1	C–CA–N–C	−123	−129
	OG–C–CA–N	59	−148
Ser_2	N–C–CA–CB	−142	−140
	N–C–CA–N	−20	−17
	C–CA–CB–OG	77	63
Cys_2	C–CA–N–C	−111	−124
	N–C–CA–N	74	86
Ala_2	C–CA–N–C	−88	−89
	N–C–CA–N	159	157
Val_2	C–CA–N–C	−124	−136
	OG–C–CA–N	64	−127

[a] From Addess *et al.* (1993).

quinoxaline rings overlap equally with the base pairs of the CpG step, as displayed in Figure 7C. The interactions between the atoms of the peptide ring and the atoms of the minor groove of the CpG binding site are also very similar in both structures. Strong van der Waals interactions occur between the Ala methyl groups and the deoxyribose sugars of the CpG binding site. The distances between the Ala methyls and the deoxyribose cytosine C1′, cytosine C2′, and guanine O4′ average 3.6 ± 0.2, 3.6 ± 0.5, and 3.0 ± 0.1 Å in the crystal structure and 3.5 ± 0.1, 3.2 ± 0.1, and 3.3 ± 0.2 Å in the solution structure. Intermolecular hydrogen bonds between the Ala NH of triostin A and GN3 of the CpG binding site are observed in both the crystal and solution structures.

There are, however, several differences between drug binding sites in the crystal structure and solution structure. In the crystal structure, the sugar conformations of each of the C_1 and G_6 residues, which form one of the two G·C base pairs in each of the CpG binding sites, are O4′ exo and C3′ endo, respectively. These are different from the sugar puckers of C_3 and G_4 (Table 2) in the solution structure. However, the sugar puckers of each of the C_5 and G_2 residues, which form the second C·G base pair of the binding site, are C3′ endo and C2′ endo, respectively, similar to the sugar pucker of the C_3 and G_4 residues. This suggests that the puckering mode of the C_1 and G_6 residues of the crystal structure could be due to 3′- and 5′- end effects or to crystal packing. The central T·A base pairs between the two drugs of the crystal structure are Hoogsteen base paired with the central adenine residues adopting *syn* glycosidic torsional angles, while in the solution structure, the two A_2·T_5 base pairs which flank the single CpG binding site are Watson–Crick base paired.

In the crystal structure of triostin A bound to [d(CG-TACG)]$_2$ only one of the two possible intermolecular hydrogen bonds between the Ala CO and the 2-amino groups of guanine is observed per CpG binding site, indicating that the two Val *N*-methyl groups are positioned differently in the two sym-

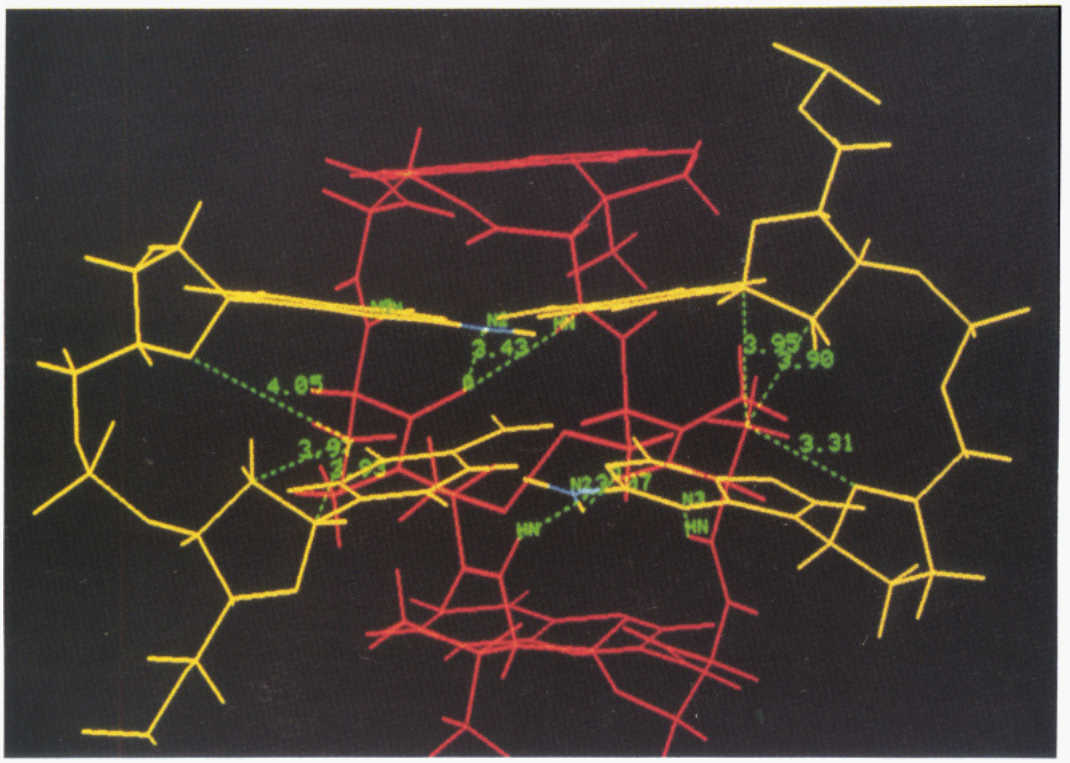


FIGURE 10: Model of CysMeTANDEM bound to the central CpG step of [d(GACGTC)]$_2$. CysMeTANDEM is illustrated in red and the DNA is illustrated in yellow. The 2-amino nitrogen of guanine is illustrated in blue. Intermolecular hydrogen bonds between the Ala NH protons and GN3 nitrogens and intramolecular hydrogen bonds between the Val NH protons and Ala CO groups are drawn with green dashed lines. Distances between the Ala CO groups and GNH2 groups are indicated in green. Distances between the Ala methyl groups and selected deoxyribose atoms are indicated in green.

metrical halves of the complex (Wang *et al.*, 1984, 1986; Ughetto *et al.*, 1985). In the solution structure, we observe both intermolecular hydrogen bonds in each of the five refined structures. The two Val *N*-methyl groups are symmetrical to each other in the structure of the complex. Recently, it was shown by DNase I footprinting that echinomycin cannot bind tightly to DNA in which either the Watson or the Crick strand contains inosines in place of guanines (Marchand *et al.*, 1992). This suggests that both hydrogen bonds are needed for bis-intercalative binding of triostin A to CpG. The critical nature of these hydrogen bonds to the sequence-specific binding to triostin A to CpG is discussed in greater detail in the following paper (Addess & Feigon, 1994).

Comparison with the Solution Structure of the CysMeTANDEM–[d(GATATC)]$_2$ Complex. Overall comparison between the structures of the two complexes reveals that many of the structural features of these complexes are similar (Addess *et al.*, 1993). The two most important changes in the structure of the DNA in the CysMeTANDEM–[d(GATATC)]$_2$ complex are unwinding of the helix and buckling of the base pairs flanking the site of intercalation. Also, the deoxyribose sugars in the CysMeTANDEM–[d(GATATC)]$_2$ complex adopt sugar puckers near C2′ endo with the exception of the T_3 sugars and the C_6 sugars, which adopt sugar puckers near C3′ endo. Similar features are observed in the structure of the DNA in the triostin A–[d(GACGTC)]$_2$ complex. The degree of unwinding and buckling of the central CpG step is very similar to that of the central TpA step of the CysMeTANDEM complex. The relative positions of the quinoxaline rings of triostin A and the C_3pG_4 base pairs (Figure 7) are also similar to the positions between the quinoxaline rings and the T_3pA_4 step of the CysMeTANDEM–[d(GATATC)]$_2$ complex. All the sugars in the triostin A–[d(GACGTC)]$_2$ are predominantly S-type (near C2′ endo) except the C_3 and C_6 sugars, which are closer to N-type conformations (near C3′ endo) (Table 4). The two hydrogen bonds between the symmetry related Ala NH protons and the G_4N3 nitrogens in the structure of the triostin A–[d(GACGTC)]$_2$ complex are analogous to the two intermolecular hydrogen bonds between the Ala NH and A_4N3 of the TpA step of the CysMeTANDEM–[d(GATATC)]$_2$ complex.

The principal difference between the two complexes is that the structure of triostin A bound to [d(GACGTC)]$_2$ is different from the structure of CysMeTANDEM bound to [d(GATATC)]$_2$. Figure 9 shows the structure of triostin A alongside of the structure of CysMeTANDEM from the CysMeTANDEM–[d(GATATC)]$_2$ complex. The two Val *N*-methyl groups of triostin A are drawn with van der Waals molecular surfaces. Table 5 lists selected dihedral angles of the peptide rings of the two different drugs. CysMeTANDEM contains two intramolecular hydrogen bonds between each Val NH and Ala CO (Addess *et al.*, 1993). These intramolecular hydrogen bonds cannot form in triostin A because of the *N*-methyl substituents attached to both Val amides. If the conformation of triostin A bound to DNA were the same as the conformation of CysMeTANDEM bound to DNA, the distance between the Ala carbonyl and the Val *N*-methyl would be 1.8 Å, which would result in unfavorable van der Waals interactions. This explains why the Val *N*-methyl groups of the drug face directly into the minor groove of the DNA and the two dihedral angles defined by the serine side-chain oxygen, the valine carbonyl, the valine Cα, and the valine nitrogen (OG–C–CA–N) are different from these dihedral angles in the structure of the CysMeTANDEM–DNA complex (Table 5). As a result, triostin A has a wider peptide ring than CysMeTANDEM. The distance between the oxygen of the Ala CO and the Val nitrogen is about 1 Å longer in triostin A (~4.0 Å) than in CysMeTANDEM (~3.1 Å), and the distance from the Cα carbon of one Cys_1 residue to the Cα carbon of Cys_2 is about 4.2 Å in triostin A and about 3.9 Å

in CysMeTANDEM. These distances are about the same as the corresponding distances in the crystal structure of triostin A alone (Sheldrick *et al.*, 1982) and bound to [d(CGTACG)]$_2$ (Ughetto *et al.*, 1985; Wang *et al.*, 1984, 1986) and the crystal structure of TANDEM alone (Viswamitra *et al.*, 1981; Hossain *et al.*, 1982). This indicates that the conformation of the peptide ring of the drug does not change when it binds to DNA.

Figure 10 shows a view into the minor groove of a model of CysMeTANDEM bound to the central CpG base pairs of [d(GACGTC)]$_2$. As a result of intramolecular hydrogen bonds between the Val NH and Ala CO, the distance between a 2-amino group of the guanine base and the alanine carbonyl of CysMeTANDEM is about 3.2 Å, which is about 1.0 Å farther than the distance between the G_4NH2 and Ala CO of triostin A (Figure 8). Therefore, formation of intermolecular hydrogen bonds between the Ala CO and the G_4NH2 would require breaking the *intra*molecular hydrogen bonds between the Ala CO and the Val NH that form in the structure of CysMeTANDEM; breaking these *intra*molecular hydrogen bonds to form *inter*molecular hydrogen bonds would be very energetically unfavorable.

This explains why CysMeTANDEM does not bind to CpG steps, but it does not explain why the drug binds so strongly to TpA steps. One possibility is that the stacking interactions between the quinoxaline rings and the TpA step are favorable enough so that CysMeTANDEM can bind in a bis-intercalative manner around a TpA step by forming only two intermolecular hydrogen bonds. Triostin A binds less tightly to a CpG step than CysMeTANDEM does to a TpA step. This suggests that intercalation of the quinoxaline rings around a CpG step is not as stable energetically, and in order to stabilize the complex between triostin A and [d(GACGTC)]$_2$, four intermolecular hydrogen bonds are required. To understand how the relationship between hydrogen bonding and the stacking interactions in both complexes determine the difference in the sequence specificity between triostin A and CysMeTANDEM, we have examined the binding of both antibiotics to DNA containing a CpI step. This is presented in the following paper (Addess & Feigon, 1994).

ACKNOWLEDGMENT

K.J.A. thanks Karl Koshlap and Shiva Malek for synthesizing the DNA used in these studies, Peter Schultze and Ed Wang for productive discussions, and Thorsten Dieckmann for comments on the manuscript.

REFERENCES

Addess, K. J., & Feigon, J. (1994) *Biochemistry* (following paper in this issue).
Addess, K. J., Gilbert, D. E., Olsen, R. K., & Feigon, J. (1992) *Biochemistry 31,* 339–350.
Addess, K. J., Sinsheimer, J. S., & Feigon, J. (1993) *Biochemistry 32,* 2498–2508.
Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65,* 355–360.
Bothner-by, A. A., Stephen, R. L., Lee, J., Warren, C. D., & Jeanloz, R. W. (1984) *J. Am. Chem. Soc. 106,* 811–813.
Brünger, A. T. (1992) *X-PLOR (Version 3.1) Manual,* Yale University Press, New Haven, CT, and London.
Crippen, G., & Havel, T. F. (1988) *Distance Geometry and Molecular Conformation,* Research Studies Press, Taunton, Somerset, England.
Feigon, J., Leupin, W., Denny, W. A., & Kearns, D. R. (1984) *J. Med. Chem. 27,* 450–465.
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., & Feigon, J. (1991) *Biochemistry 30,* 2483–2494.
Gilbert, D. E., & Feigon, J. (1992) *Nucleic Acids Res. 20,* 2411–2420.
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.
Hossain, M. B., van der Helm, D., Olsen, R. K., Jones, P. G., Sheldrick, G. M., Egert, E., Kennard, O., Waring, M. J., & Viswamitra, M. A. (1982) *J. Am. Chem. Soc. 104,* 4301–4308.
Katagiri, K., Yoshida, T., & Sato, K. (1974) in *Antibiotics III. Mechanism of Action of Antimicrobial and Antitumour Agents* (Corcoran, J., & Hahn, F. E., Eds.) pp 234–251, Springer-Verlag, Berlin, Heidelberg, and New York.
Kumar, A., Ernst, R. R., & Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun. 95,* 1–6.
Kuszewski, J., Nilges, M., & Brünger, A. T. (1992) *J. Biomol. NMR 2,* 33–56.
Lavesa, M., Olsen, R. K., & Fox, K. R (1993) *Biochem J., 289,* 605–607.
Low, C. M. L., Olsen, R. K., & Waring, M. J. (1984) *FEBS Lett. 176,* 414–419.
Marchand, C., Bailly, C., McLean, M., Moroney, S. E., & Waring, M. J. (1992) *Nucleic Acids Res. 20,* 5601–5606.
Marion, D., & Bax, A. (1988) *J. Magn. Reson. 80,* 528–533.
Marion, D., Ikura, M., & Bax, A. (1989) *J. Magn. Reson. 84,* 425–430.
Olsen, R. K., Ramasamy, K., Bhat, K. L., Low, C. M. L., & Waring, M. J. (1986) *J. Am. Chem. Soc. 108,* 6032–6036.
Powers, R., Olsen, R. K., & Gorenstein, D. G. (1989) *J. Biomol. Struct. Dyn. 7,* 515–553.
Rinkel, L. J., & Altona, C. (1987) *J. Biomol. Struct. Dyn. 4,* 621–649.
Sheldrick, G. M., Guy, J. J., Kennard, O., Rivera, V., & Waring, M. J. (1984) *J. Chem. Soc., Perkin Trans. 2,* 1601–1605.
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 Dyke, M. M., & Dervan, P. B. (1984) *Science 225,* 1122–1127.
Viswamitra, M. A., Kennard, O., Cruse, W. B. T., Egert, E., Sheldrick, G. M., Jones, P. G., Waring, M. J., Wakelin, L. P. G., & Olsen, R. K. (1981) *Nature 289,* 817–819.
Wang, A. H.-J., Ughetto, G., Quigley, G. J., Hakoshima, T., van der Marel, G. A., van Boom, J. H., & Rich, A. (1984) *Science 225,* 1115–1121.
Wang, A. H.-J., Ughetto, G., Quigley, G. J., & Rich, A. (1986) *J. Biomol. Struct. Dyn. 4,* 319–342.
Waring, M. J., & Wakelin, L. P. G. (1974) *Nature 252,* 653–657.
Waterloh, K., Olsen, R. K., & Fox, K. R. (1992) *Biochemistry 31,* 6246–6253.
Yoshida, T., & Katagiri, K. (1969) *Biochemistry 8,* 2645–2651.