Sequence Specificity of Quinoxaline Antibiotics. 2. NMR Studies of the Binding of [*N*-MeCys^3,*N*-MeCys^7]TANDEM and Triostin A to DNA Containing a CpI Step†

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: The binding of CysMeTANDEM and triostin A to DNA containing a CpI step has been studied by one- and two-dimensional ^{1}H NMR spectroscopy. CysMeTANDEM binds sequence specifically to CpI steps as well as TpA steps as a bis-intercalator with the peptide backbone in the minor groove of the DNA. Only nonspecific, nonintercalative binding is observed between triostin A and DNA containing a CpI step. Comparison of the CysMeTANDEM–[d(GGACITCC)]$_2$ complex to the CysMeTANDEM–[d(GGATATCC)]$_2$ complex indicates that the structures of both complexes are very similar. However, CysMeTANDEM binds less tightly to [d(GGACITCC)]$_2$ than to [d(GGATATCC)]$_2$. The NMR evidence presented provides molecular insight into the role of stacking interactions and hydrogen bonding between the drug and the DNA in the sequence-specific binding of CysMeTANDEM to TpA sites and of triostin A to CpG sites.

The quinoxaline antibiotics are a family of drugs that include the naturally occurring triostin A and the synthetic derivative [*N*-MeCys^3,*N*-MeCys^7]TANDEM. Both drugs are composed of a cyclic octadepsipeptide ring that is covalently attached to two quinoxaline rings by amide linkages to two D-Ser[1] residues (Chart 1). [*N*-MeCys^3,*N*-MeCys^7]TANDEM (hereafter referred to as CysMeTANDEM) differs from triostin A in its chemical structure only at the two valine residues, which contain no *N*-methyl substituents (Chart 1). DNase I footprinting, X-ray crystallographic, and NMR studies of these quinoxaline antibiotics bound to a variety of DNA sequences have established that both triostin A and echinomycin, another naturally occurring quinoxaline antibiotic, bind sequence specifically to DNA at NCGN sites and that CysMeTANDEM binds sequence specifically at NTAN sites (Addess *et al.,* 1992, 1993; Gao & Patel, 1988, 1989; Gilbert & Feigon, 1991, 1992; Gilbert *et al.*, 1989; Lavesa *et al.*, 1993; Low *et al.*, 1984, 1986; Ughetto *et al.*, 1985; Wang *et al.*, 1984, 1986; Waterloh *et al.*, 1992).

This has raised questions about what structural factors cause the difference in sequence specificity between triostin A and CysMeTANDEM. In a previous study, we investigated these differences by solving the solution structure of the triostin A–[d(GACGTC)]$_2$[2] complex and comparing it with the solution structure of the analogous CysMeTANDEM–[d(GATATC)]$_2$ complex (Addess *et al.*, 1993; Addess & Feigon, 1994). Both complexes have very similar DNA structures and contain two intermolecular hydrogen bonds between the Ala NH and the purine N3 of the YpR binding site. The main difference between the two complexes is in the structure of the drug. The solution structure of CysMeTANDEM contains intramolecular hydrogen bonds between the Val NH and Ala CO. As a result, the Ala carbonyls of CysMeTANDEM are not available to form intermolecular hydrogen bonds with the DNA. By contrast, the Ala carbonyls of triostin A form two additional intermolecular hydrogen bonds with the guanine 2-amino protons of the DNA in the solution structure of the triostin A–[d(GACGTC)]$_2$ complex in solution. Recently, it was shown by DNase I footprinting that echinomycin loses its ability to bind sequence specifically to DNA in which inosines (I) are substituted for guanines (G) on either or both strands of DNA. This suggests that both intermolecular hydrogen bonds are critical to the CpG recognition of echinomycin and (presumably) triostin A (Marchand *et al.*, 1992).

Why these additional intermolecular hydrogen bonds are needed for the sequence-specific binding of triostin A to CpG is not completely understood. Triostin A does not bind as tightly to a CpG step as CysMeTANDEM does to a TpA step (Addess & Feigon, 1994). This suggests that the quinoxaline rings stack less favorably with a CpG step than with a TpA step and that, in order for the quinoxaline rings of triostin A to bis-intercalate around a CpG step, the peptide ring needs to form four intermolecular hydrogen bonds with the DNA, while for the quinoxaline rings of CysMeTANDEM to bis-intercalate around a TpA step, the peptide ring needs to form only two intermolecular hydrogen bonds with the DNA.

In order to determine how the differences in stacking interactions and hydrogen bonding between both complexes explain the difference in sequence specificity of these two drugs, we have studied the binding of CysMeTANDEM and triostin A to DNA containing a CpI site using one- and two-dimensional NMR spectroscopy. Since the minor groove of

† 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.

* 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; COSY, correlation spectroscopy; Cys, cysteine; CpI complex, CysMeTANDEM–[d(GGACITCC)]$_2$; TpA complex, CysMeTANDEM–[d(GGATATCC)]$_2$; HOHAHA, homonuclear Hartmann–Hahn spectroscopy; *N*-MeCys, *N*-methylcysteine; NOESY, nuclear Overhauser effect spectroscopy; P.E.COSY, primitive exclusive COSY; P.COSY, purged COSY; Ser, serine; Val, valine.

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

0006-2960/94/0433-12397$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)

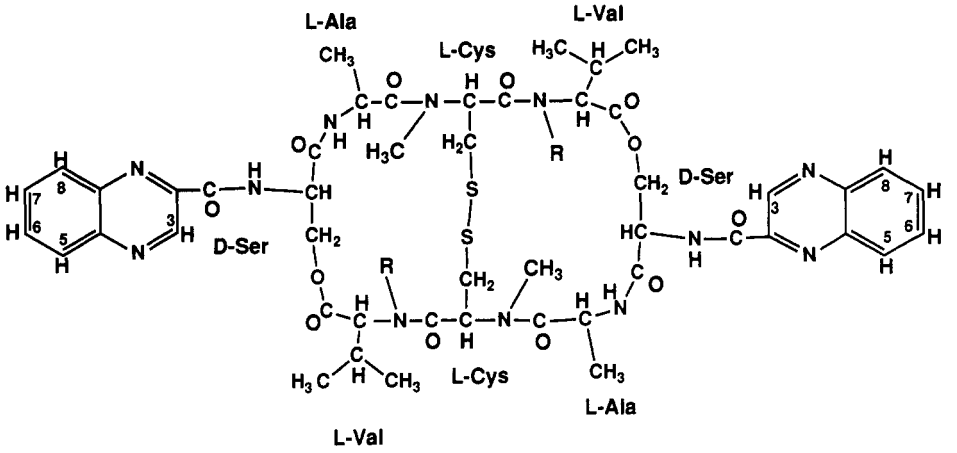
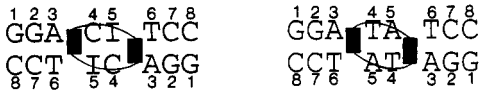


Chart 2: Sequence and Numbering System of the DNA Octamers [d(GGACITCC)]$_2$ and [d(GGATATCC)]$_2$ with the Orientation of the Bound CysMeTANDEM Shown Schematically



a CpI site resembles the minor groove of a TpA site, removal of the guanine exocyclic 2-amino groups from the minor groove of a CpG site not only may eliminate a sequence-specific binding site for triostin A but create a potential new binding site for CysMeTANDEM. We present NMR evidence that shows that CysMeTANDEM binds sequence specifically to DNA containing a CpI step as shown in Chart 2 and that triostin A binds to DNA containing a CpI step in a nonspecific, nonintercalative manner only. The structure of the 1:1 complex of CysMeTANDEM with [d(GGACITCC)]$_2$ is very similar to the solution structure of the complex between the drug and the DNA octamer [d(GGATATCC)]$_2$. However, CysMeTANDEM does not bind as tightly to CpI as it does to TpA. In light of the NMR data on the binding of CysMeTANDEM and triostin A to [d(GGACITTCC)]$_2$, the molecular basis for the sequence specificity of triostin A and CysMeTANDEM is discussed.

MATERIALS AND METHODS

Sample Preparation. d($G_1G_2A_3C_4I_5T_6C_7C_8$) was synthesized on an ABI 381A synthesizer using β-cyanoethyl phosphoramidite chemistry on a 10-μmol scale and purified by gel filtration by the method of Kintanar *et al.* (1987). 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. CysMeTANDEM was a gift from Dr. Richard K. Olsen, Utah State University. Both quinoxaline antibiotics were assessed to be pure by one-dimensional proton NMR and used without further purification. The CysMeTANDEM–[d(GGACITCC)]$_2$ complex was prepared in the same way as the CysMeTANDEM–[d(GGATATCC)]$_2$ complex (Addess *et al.*, 1992).

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). 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.

RESULTS

Binding of CysMeTANDEM to [d(GGACITCC)]$_2$. The minor groove of the CpI dinucleotide sequence resembles the minor groove of the TpA dinucleotide sequence. Since the peptide ring of CysMeTANDEM interacts with the minor groove, it should bind sequence specifically to DNA at a CpI step. To test this, we investigated the binding of CysMeTANDEM to the DNA octamer [d(GGACITCC)]$_2$. Figure 1 shows the one-dimensional ^{1}H NMR spectra of [d(GGACITCC)]$_2$ containing no drug (panel A) and a 1:1 mixture of drug and DNA duplex (panel B). The addition of 1 equiv of CysMeTANDEM per DNA duplex causes the appearance of the Cys NMe, Val NMe, Ala methyl, Val methyl, and a second set of T_6 methyl, I_5H8, and A_3H8 resonances, which are labeled in Figure 1B. This indicates that CysMeTANDEM binds sequence specifically to [d(GGACITCC)]$_2$. However, the addition of 1 equiv of drug results in only partial saturation of the DNA duplex. This is indicated by the simultaneous presence of a second set of DNA resonances in the NOESY spectrum of the 1:1 mixture of drug and DNA. This second set of resonances has identical chemical shifts to those of the free DNA resonances shown in Figure 1A. These are labeled with asterisks in Figure 1B. The integrated intensity of the T_6 resonance of the 1:1 drug–DNA complex compared with the that of the T_6* resonance indicates that about 50% complex formation between the drug and the DNA occurs when 1 equiv of [d(GGACITCC)]$_2$ is mixed with 1 equiv of CysMeTANDEM. This is in contrast to the addition of one equivalent of CysMeTANDEM to [d(GGATA-

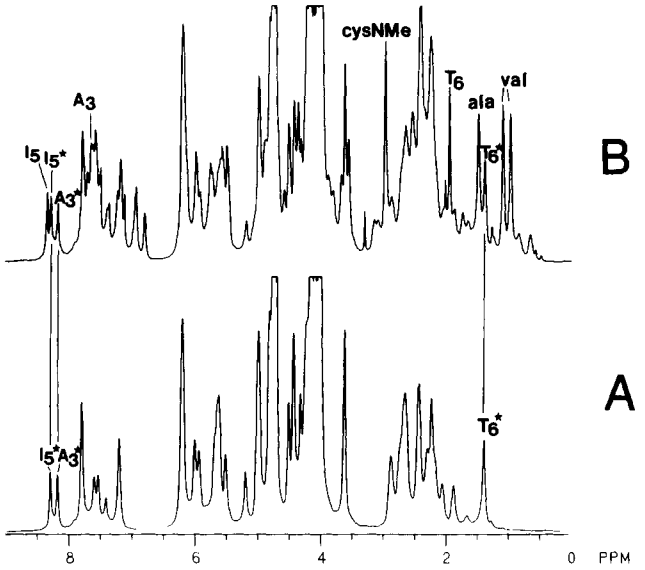


FIGURE 1: One-dimensional ^{1}H NMR spectra of (A) [d(GGACITCC)]$_2$ and (B) a 1:1 mixture of CysMeTANDEM + [d(GGACITCC)]$_2$ in D_2O at 25 °C. Samples are 2 mM DNA duplex and 150 mM NaCl, pH 6.5. Solid lines connect free DNA resonances with resonances that have identical chemical shifts in the 1:1 mixture. These resonances are labeled with asterisks. Assignments of the I_5-H8, T_6 methyl, Ala methyl, Cys *N*-methyl, and Val methyl resonances are indicated. The spectra were acquired with a sweep width of 5000 Hz in 4K complex points and were line broadened by 3 Hz prior to Fourier transformation.

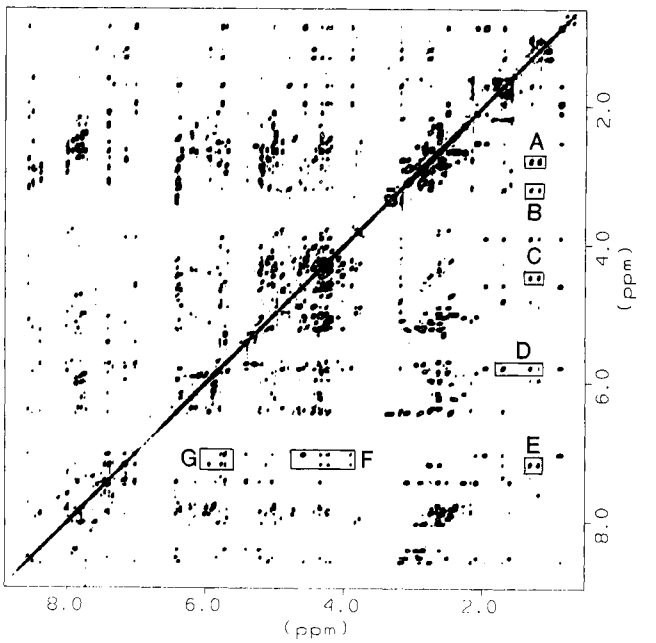


FIGURE 2: NOESY spectrum of a 1:1 mixture of CysMeTANDEM and [d(GGACITCC)]$_2$ at 25 °C. Sample conditions are given in Figure 1. Several important cross peaks are boxed: Val CH_3 to Val βH (A), Val CH_3 to Val NMe (B), Val CH_3 to Val αH (C), Ala CH_3 to C_4H1' and Val CH_3 to C_4H1' (D), QH7, QH8 to Val CH_3 (E), C_4H6 and QH7, QH8 to C_4H3', C_4H4', C_4H5' and C_4H5'' (F), C_4-H6 to C_4H1', C_4H5 and QH7, QH8 to A_3H1', C_4H1', and C_4H5 (G). A total of 400 t_1 blocks of 2K points and 32 scans were acquired. The data in both dimensions were apodized by a sine-bell squared function phase-shifted by 50°. A total of 400 points were apodized in both dimensions. Data in t_1 were zero filled to 1K points prior to Fourier transformation.

TCC)]$_2$, which results in 100% complex formation. This indicates that CysMeTANDEM bind less tightly to [d(GGACITCC)]$_2$ than to [d(GGATATCC)]$_2$.

Figure 2 shows the NOESY spectrum (τ_m = 150 ms) of a 1:1 mixture of CysMeTANDEM with [d(GGACITCC)]$_2$ at 25 °C. NOE cross peaks corresponding to resonances of the 1:1 complex and to resonances of the free DNA are observed. The spectrum is very similar to the NOESY spectrum of a mixture of [d(GGATATCC)]$_2$ with CysMeTANDEM at less than a 1:1 ratio of drug to DNA. In addition, the NOESY spectrum of CysMeTANDEM with [d(GGACITCC)]$_2$ contains exchange cross peaks between the resonances of the specific complex and the resonances of the nonintercalated DNA octamer. Similar exchange is observed between a bis-intercalative, specific complex and a nonspecific, nonintercalative complex of triostin A with [d(GACGTC)]$_2$ (Addess & Feigon, 1994). Since resonances of both specific and nonspecific complexes are observed, the exchange rate is slow on the NMR time scale and, thus, $k_{ex} \ll 2\pi|\nu_s - \nu_{ns}|$. The smallest chemical shift difference between the specific and nonspecific species in Figure 2 involves the I_5H8, which is shown in Figure 4. This difference is about 10 Hz; therefore, k_{ex} is much smaller than 60 s^{-1}.

In Figure 2, the NOE cross peaks of the 1:1 CysMeTANDEM–[d(GGACITCC)]$_2$ complex (hereafter referred to as the CpI complex) are similar to corresponding cross peaks in the NOESY spectrum of the CysMeTANDEM–[d(GGATATCC)]$_2$ complex (hereafter referred to as the TpA complex). There are cross peaks from the quinoxaline H7, H8 and the C_4H6 protons to the C_4H3', C_4H4', C_4H5', and C_4H5'' protons (box F) and to the A_3H1' and C_4H1' protons (box G). These cross peaks are almost identical to those involving the quinoxaline, A_3 and T_4 protons in the NOESY spectrum of the TpA complex. This indicates that CysMeTANDEM binds sequence specifically to the CpI step of [d(GGACITCC)]$_2$ as a bis-intercalator and that the structures of the CpI complex and the TpA complex are similar to each other.

Binding of Triostin A to [d(GGACITCC)]$_2$. To investigate how triostin A binds to [d(GGACITCC)]$_2$, we added 1 equiv of triostin A in 50% (v/v) methanol/acetonitrile to an aqueous solution containing the DNA duplex. During drying of this sample under nitrogen gas, some of the triostin A precipitated out of solution, but much of it remained in aqueous solution. This can be seen in the NOESY spectrum of the mixture of triostin A with [d(GGACITCC)]$_2$, shown in Figure 3. The addition of triostin A to [d(GGACITCC)]$_2$ results in the appearance of additional cross peaks from the peptide spin systems of the drug. For example, cross peaks from the Val methyl side-chain protons to the Val βH (box A), Val *N*-methyl (box B), and Val αH protons (box C) are observed. On the basis of the integrated intensities of the triostin A cross peaks, the concentration of the drug in the NMR sample containing 2 mM DNA duplex is about 500 μM. Since the saturating concentration of triostin A in aqueous solution is about 6 μM, this indicates that about one in four molecules of DNA duplex is bound by triostin A. The binding of triostin A to the DNA octamer is also indicated by weak base to base cross peaks from the C_4H6 to both the A_3H8 and I_5H8 protons (Figure 3, box D). These cross peaks are not observed in the NOESY of the free octamer (data not shown). This indicates that the binding of triostin A to [d(GGACITCC)]$_2$ causes small changes in the structure of the DNA.

However, unlike CysMeTANDEM, triostin A does not bind to [d(GGACITCC)]$_2$ in a sequence-specific manner at the CpI step. In Figure 2, the boxes labeled D, F, and G contain cross peaks between the C_4 resonances and the CysMeTANDEM resonances that result from sequence-specific

Table 1: Assignments of the [*N*-MeCys^3,*N*-MeCys^7]TANDEM–[d(GGACITCC)]$_2$ Complex at 25 °C

	[d(GGACITCC)]$_2$ chemical shifts (ppm)[a]									
	H8/H6	Me/H5 H2	H1′	H2′	H2″	H3′	H4′	H5′,H5″	imino[b]	amino[b]
G_1	7.95		5.69	2.50	2.65	4.90	4.34		12.68	
G_2	7.83		5.31	2.57	2.57	5.00	4.41	4.12, 4.23	12.87	
A_3	7.90	7.61	5.95	2.62	2.76	5.13	4.42	4.24, 4.35		
C_4	7.04	5.72	5.78	0.88	1.96	4.59	3.91	4.34, 4.23		
I_5	8.59	7.35	6.40	3.18	2.77	5.10	4.58	4.20, 4.26	14.14	
T_6	7.87	2.17	6.36	2.64	2.74	5.22	4.44		12.99	
C_7	7.81	5.99	6.22	2.44	2.64	4.73	4.37			7.13, 8.20
C_8	7.86	6.01	6.44	2.48	2.48	4.75	4.24			6.88, 8.29

	CysMeTANDEM chemical shifts (ppm)[a]						
	α	β	β′	γ	γ′	NH[b]	NMe
serine	5.81	4.65	4.91			8.30	
valine	4.48	2.79		1.20	1.32	8.62	
cysteine	6.47	3.34	3.42				3.19
alanine	4.82	1.71				9.20	
quinoxaline	7.74 (3)[c]	7.48 (5)	7.42 (6)	7.02 (7)	7.02 (8)		

[a] Chemical shifts relative to DSS. [b] Chemical shifts at 1 °C. [c] Positions in parentheses.

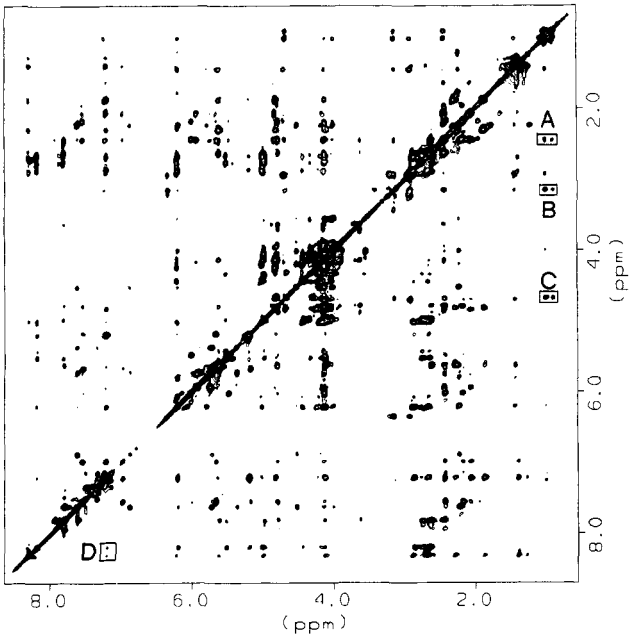


FIGURE 3: NOESY spectrum of a 1:1 mixture of triostin A and [d(GGACITTCC)]$_2$. The sample contains ~2 mM DNA duplex and 150 mM NaCl, at pH 6.5. The following cross peaks are boxed: Val CH_3 to Val βH (A), Val CH_3 to Val NMe (B), Val CH_3 to Val αH (C), and A_3H8 to C_4H6 and C_4H6 to A_5H8 (D). A total of 350 t_1 blocks of 2K points and 32 scans were acquired. The data in both dimensions were apodized by a sine-bell squared function phase-shifted by 60°. Data in t_1 were zero filled to 1K points prior to Fourier transformation.

binding of CysMeTANDEM at the CpI step of [d(GGACITTCC)]$_2$. In addition, the box labeled E in Figure 2 contains NOE cross peaks between the Val methyls and the quinoxaline H8. These cross peaks indicate that the quinoxaline rings of CysMeTANDEM are intercalated between the bases of the DNA in the CpI complex. These cross peaks are not observed in the NOESY spectrum of triostin A bound to [d(GGACITCC)]$_2$ (Figure 3). Since the NOESY spectrum of triostin A bound to [d(GGACITCC)]$_2$ does not contain specific drug–DNA NOE cross peaks, we cannot determine the nature of the interactions between triostin A and the DNA octamer. Therefore, we can only conclude that triostin A binds to [d(GGACITCC)]$_2$ in a nonspecific, nonintercalative manner.

Assignments of the CysMeTANDEM–[d(GGACITCC)]$_2$ Complex. Assignments of the DNA and CysMeTANDEM resonances of the CpI complex and the TpA complex are very similar (Addess *et al.*, 1992). This indicates that many of the structural features observed in the TpA complex are also present in the CpI complex. Assignments of the CpI complex are given in Table 1. An expanded region of the NOESY spectrum of the CpI complex is shown in Figure 4. The assignment of the DNA aromatic, H1′, and the drug quinoxaline protons of the complex are indicated. Exchange cross peaks involving the A_3H8 and I_5H8 resonances of both specific and nonspecific complexes are labeled with arrows. In Figure 4, there is a base to base NOE between the C_4H6 and the A_5H8 of the CpI complex, which is not observed in the NOESY spectrum of the free octamer (data not shown). A corresponding NOE cross peak between the T_4H6 and A_5H8 is observed in the NOESY spectrum of the TpA complex. The H2′ and H2″ resonances of the C_4 and I_5 nucleotides of the CpI complex show unusual chemical shifts like those of the T_4 and A_5 nucleotides of the TpA complex. The C_4H2″ and C_4H2′ resonances are upfield shifted by ~1 ppm from these resonances in the free octamer. The I_5H2′ proton has a larger chemical shift than the I_5H2″ proton, which is the reverse of what is observed in the free octamer (data not shown).

An expanded region of the H_2O NOESY spectrum of the CpI complex at 1 °C is shown in Figure 5. Exchange cross peaks between the specific complex and the nonspecific complex are labeled with arrows. Imino resonances from both the specific complex and the free DNA are shown. The chemical shifts of the G_1, G_2, and T_6 imino resonances are similar in both the CpI complex and the TpA complex. The chemical shift of the I_5 imino of the complex occurs about 1 ppm upfield from the chemical shift of the I_5 imino proton of the free DNA (labeled with an asterisk). This upfield shift is characteristic of bis-intercalative binding of the DNA by the drug (Feigon *et al.*, 1984). The Ala NH protons have cross peaks to the I_5H2 and I_5H1′ (Figure 5, box A), similar to the cross peaks from Ala NH to both the A_5H2 and A_5H1′ of the TpA complex. These results indicate that Ala NH protons are hydrogen bonded to I_5N3 atoms, analogous to the intermolecular hydrogen bonds between the Ala NH protons and A_5N3 atoms of the TpA complex. This indicates that CysMeTANDEM binds in the minor groove of the CpI step of [d(GGACITCC)]$_2$. The Val NH has a strong cross peak to the Cys αH proton (Figure 5, box B) but does not have a cross peak to the I_5H2. This indicates that CysMeTANDEM

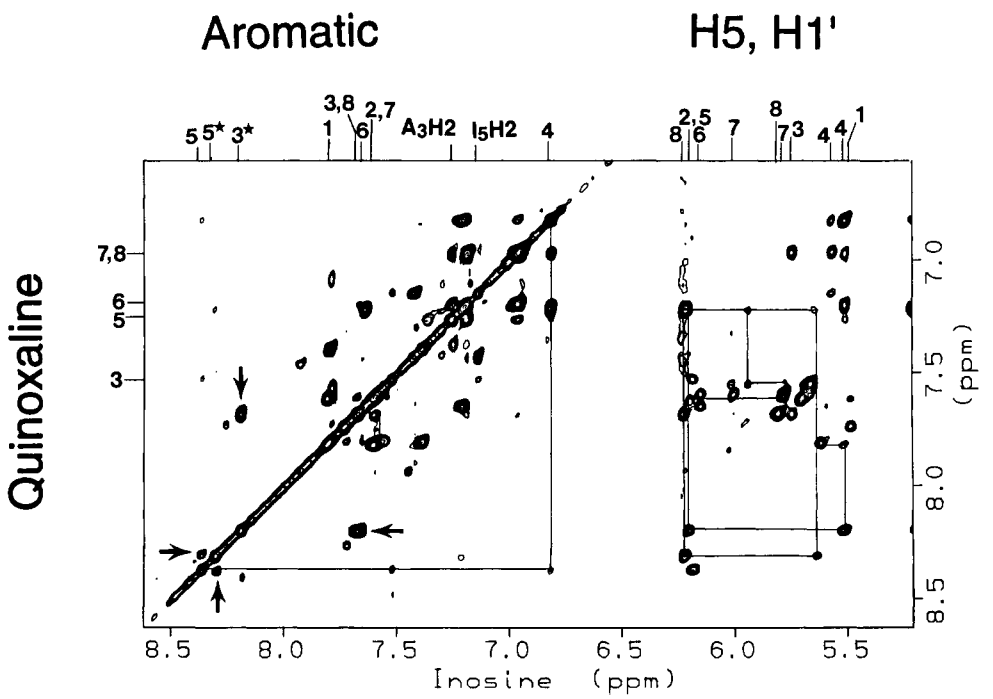


FIGURE 4: Expanded region of the NOESY spectrum (τ_m = 150 ms) of the 1:1 CysMeTANDEM–[d(GGACITCC)]$_2$ complex showing the aromatic and base–H1′ region. Assignments of the quinoxaline, aromatic, H5, and H1′ resonances are indicated. Assignments of the A_3 and I_5 of the free octamer are indicated, and the exchange cross peaks between the A_3 of the free DNA and the complex and between the I_5 of the free octamer and the complex are labeled with arrows. The base–H1′ sequential connectivities for the free [d(GGACITCC)]$_2$ are indicated by solid lines.

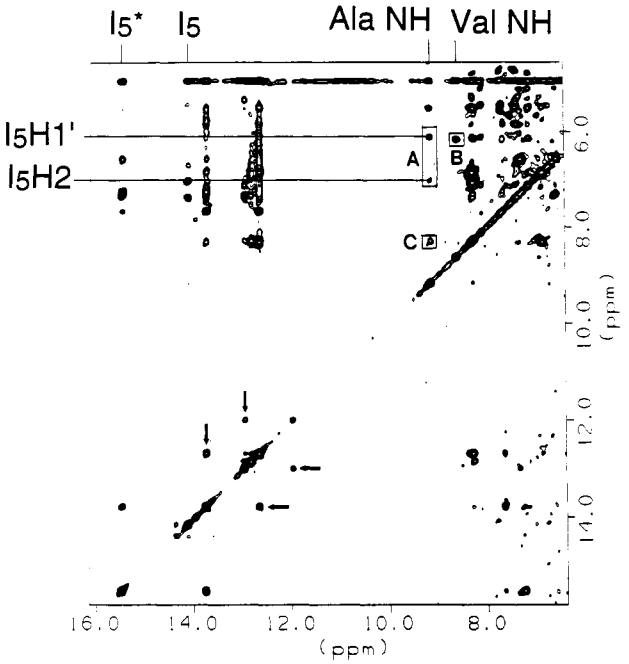


FIGURE 5: Expanded region of the NOESY spectrum in 90% H_2O/10% D_2O of the 1:1 mixture of CysMeTANDEM with [d(GGACITCC)]$_2$ at 1 °C, τ_m = 50 ms. Sample conditions are the same as in Figure 1. Assignments of the I_5 iminos from the complex and the free DNA (labeled with a star), Val NH, Ala NH, I_5H2, and I_5H1′ are indicated. The boxed cross peaks are (A) Ala NH to I_5H2 and to I_5H1′ and (B) Val NH to Cys αH. Some exchange cross peaks to a nonspecific drug–DNA complex are indicated with arrows. The spectrum was acquired with a sweep width of 10 000 Hz in both dimensions, 256 t_1 values of 64 scans, and 2K 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). A total of 256 points were apodized in both dimensions by a sine-bell squared function phase-shifted by 60°.

bound to [d(GGACITCC)]$_2$ contains intramolecular hydrogen bonds between the Val NH protons and Ala CO oxygens

Sugar Conformations in the CysMeTANDEM–[d(GGACITCC)]$_2$ Complex. An expanded region of the P.COSY spectrum of the CpI complex displaying H1′–H2′ and H1′–H2″ cross peaks is shown in Figure 6. The additional cross peaks observed in this region of the spectrum are H1′–H2′ and H1′–H2″ cross peaks of the free DNA. The cross peaks of the 1:1 complex are boxed and labeled. The cross-peak patterns for the G_1, G_2, A_3, T_6, C_7, and C_8 deoxyribose sugars of the 1:1 complex are almost identical to those in the P.COSY of the TpA complex (Addess *et al.*, 1992). The cross-peak patterns for the I_5 sugar are very similar to those of the A_5 sugar, with the I_5H2″ proton resonating upfield from the I_5-H2′ proton. The cross-peak patterns for the C_4 sugar, which are very similar to the patterns of the T_4 sugar of complex A, are very different from the patterns of all the other sugars in the 1:1 complex. For the P.COSY spectrum of the CysMeTANDEM–[d(GGATATCC)]$_2$ complex, all of the sugars except T_4 have similar H1′–H2′ and H1′–H2″ cross-peak patterns characteristic of a predominantly S-type sugar pucker; the T_4 sugar has a cross-peak pattern characteristic of an N-type sugar pucker (Addess *et al.*, 1992). This indicates that all the deoxyribose sugars of the CpI complex, except C_4, are predominantly S-type (near C2′ endo sugar pucker), while the C_4 sugar is predominantly N-type (near C3′ endo sugar pucker).

DISCUSSION

Binding of CysMeTANDEM and Triostin A to CpI. Waring and co-workers showed recently that sequence-specific binding of echinomycin to DNA is abolished completely when an inosine is substituted for guanine but that it binds specifically to a TpD site, where D is a 2,6-diaminopurine (Marchand *et al.*, 1992; Bailly *et al.*, 1993). These results led us to hypothesize that a CpI step should be a specific binding site for CysMeTANDEM. To test this hypothesis, we studied the binding of CysMeTANDEM and triostin A with the DNA octamer [d(GGACITCC)]$_2$, which contains a single CpI step.

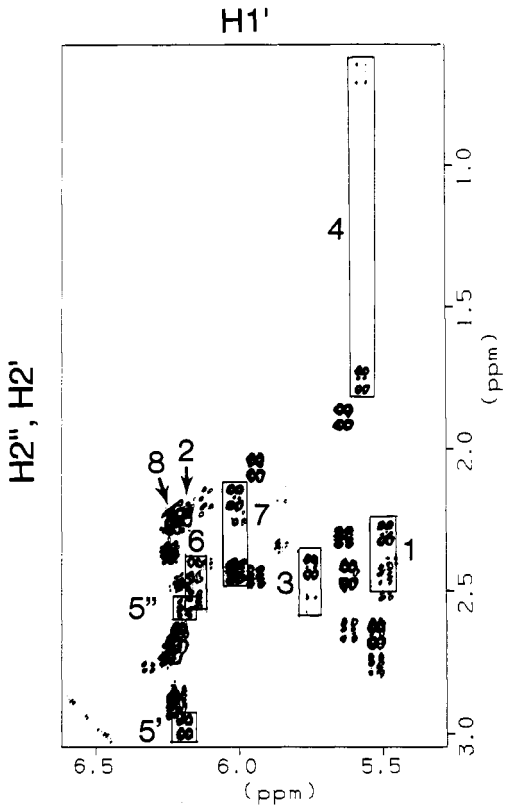


FIGURE 6: Expanded region of the P.COSY spectrum of the CysMeTANDEM–[d(GGACITCC)]$_2$ complex showing the region containing the H1′–H2′ and H1′–H2″ cross peaks at 25 °C. Cross peaks for each sugar of the 1:1 complex are boxed and labeled. The spectrum was acquired with a sweep width of 5000 Hz in both dimensions, 16 scans per t_1 value, 2K complex points, and 800 t_1 values. For each spectrum, a reference spectrum of one t_1 value was acquired with a sweep width of 5000 Hz, 256 scans, and 4K complex points. The reference spectrum was subtracted from each two-dimensional data set. A total of 800 points 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.

The data presented here indicate that CysMeTANDEM binds sequence specifically to a CpI step as well as a TpA step by bis-intercalation of the quinoxaline rings around the YpR base pairs with the peptide ring binding in the minor groove. The data also indicate that triostin A binds to DNA containing a CpI step only in a nonspecific, nonintercalative manner. A representation of the sequence-specific binding of CysMeTANDEM to [d(GGACITCC)]$_2$ is shown in Chart 2.

However, CysMeTANDEM does not bind as tightly to a CpI site as to a TpA site. The addition of 1 equiv of CysMeTANDEM to [d(GGACITCC)]$_2$ results in only 50% formation of the 1:1 drug–DNA complex while the addition of the same equivalent of drug to [d(GGATATCC)]$_2$ results in 100% complex formation (Addess *et al.*, 1992). In addition, we observe chemical exchange between the bis-intercalative CpI complex and its nonspecific, nonintercalative counterpart. Similar results were observed for the binding of triostin A at a CpG site (Addess & Feigon, 1994). In contrast, binding of CysMeTANDEM to the TpA step of either [d(GGATATCC)]$_2$ or [d(GATATC)]$_2$ is essentially irreversible. This indicates that the K_a value, defined in the kinetic scheme of the previous paper as [intercalated drug–DNA complex]/[nonspecific complex], is much lower for the CpI complex than for the TpA complex.

Structure of the DNA in the CpI Complex. A stereoview into the major groove of a model of CysMeTANDEM bound to CpI is shown in Figure 7. This model is based on the solution structure of the CysMeTANDEM–[d(GATATC)]$_2$ complex (Addess *et al.*, 1993) and is shown here to visually illustrate important structural features of a specific complex between CysMeTANDEM and CpI steps. The structures of the DNA in both the CpI complex and the TpA complex are similar. The central C_4·I_5 base pairs of the complex have a helical twist angle of about −10° and an inward buckle of about 20°, similar to the values for the unwinding and buckling angles of the central A_4·T_5 base pairs of the TpA complex (Addess *et al.*, 1992). The C_4 sugar is mainly N-type (near C3′ endo) and the I_5 sugar is mainly S-type (near C3′ endo). The conformation of these sugars is similar to those of the analogous A_4 and T_5 sugars in the TpA complex.

CysMeTANDEM–DNA Interactions in the CpI Complex. The peptide ring of the drug binds in the minor groove of the CpI binding site of the CpI complex, and there are two intermolecular hydrogen bonds between the symmetry-related Ala NH protons and I_5N3 atoms (Figure 7), analogous to the hydrogen bonds between the Ala NH protons and A_5N3 atoms in the structure of the TpA complex. In both the CpI complex and the TpA complex, CysMeTANDEM has intramolecular hydrogen bonds between the Val NH and the Ala CO. The drug–DNA NOE cross peaks are also very similar in both complexes, which indicates that the peptide–DNA interactions are similar. There are strong van der Waals interactions between the Ala methyl groups of the drug and the C_4C1′, C_4C2′, and I_5O4′ atoms of the DNA (Figure 7). In the structure of the CysMeTANDEM–[d(GATATC)]$_2$ complex, there are strong van der Waals interactions between the Ala methyl groups and the TC1′, TC2′, and AO4′ of the TpA binding site (Addess *et al.*, 1993).

NOE cross peaks between the quinoxaline H8, H7, and H6 protons and the C_4H6 and C_4H1′ protons and between the quinoxaline H6 and H5 and the A_3H8 and A_3H1′ protons indicate that the quinoxaline rings intercalate between the A_3·T_6 base pairs and the C_4·I_5 base pairs of the CpI complex (Chart 2). This indicates that the stacking of the quinoxaline rings with the bases of the $A_3pC_4pI_5pT_6$ step in the CpI complex is similar to the stacking of the quinoxaline rings with the bases of the $A_3pT_4pA_5pT_6$ step in the TpA complex.

Role of Intermolecular Hydrogen Bonds between Ala NH and the Purine N3 in the YpR Sequence Specificity of Quinoxaline Antibiotics. In all quinoxaline antibiotic–DNA complexes, there are intermolecular hydrogen bonds between the Ala NH protons and the purine N3 nitrogens. These hydrogen bonds anchor the drug in the minor groove of the DNA in the structures of the triostin A–[d(GACGTC)]$_2$ complex and of the CysMeTANDEM–[d(GATATC)]$_2$ complex, causing the stacking between the quinoxaline rings and the DNA bases in both complexes to be similar (Addess *et al.*, 1993; Addess & Feigon, 1994). The similar stacking in both complexes is why the structure of the DNA in the triostin A–[d(GACGTC)]$_2$ complex is so similar to the structure of the DNA in the CysMeTANDEM–[d(GATATC)]$_2$ complex. Waring and Wakelin (1974) showed that quinoxaline rings alone do not intercalate into DNA. Furthermore, substitution of lactic acid for Ala in TANDEM, which removes the Ala NH groups, results in a loss of protection of the Tyr_1 DNA restriction fragment (Olsen *et al.*, 1986). This, combined with the structural evidence of hydrogen bonding in both triostin A and CysMeTANDEM complexes, suggests strongly that these hydrogen bonds are the critical element in the sequence-specific binding of these quinoxaline antibiotics to YpR sites of the DNA. These hydrogen bonds distinguish the sequence-specific binding of this class of drugs

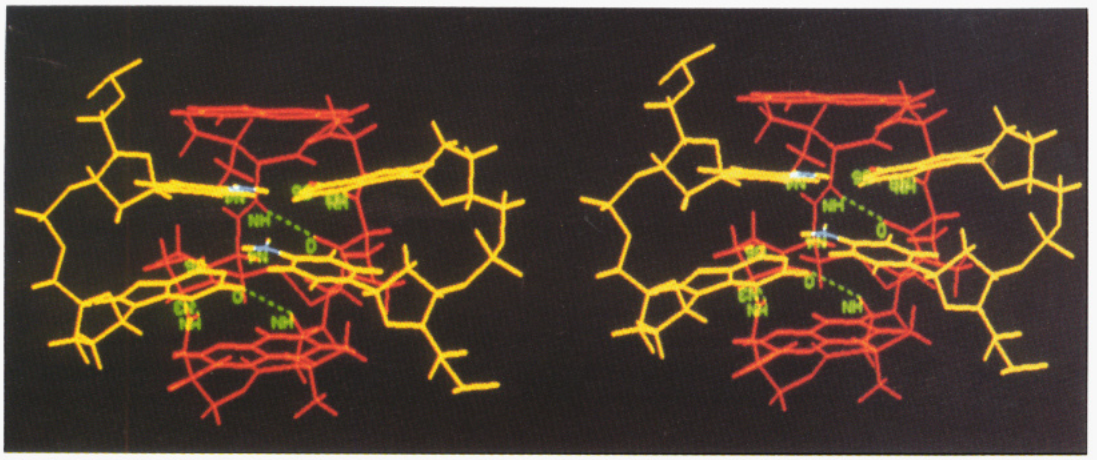


FIGURE 7: Stereoview of the model structure of CysMeTANDEM bound to CpI. CysMeTANDEM is drawn in red and the C·I base pairs are drawn in yellow. Hydrogen bonds are drawn with green dashes. The N4 atom of cytosine is drawn in blue and the O6 atom of inosine is drawn in red.

from that of other DNA-intercalating drugs such as actinomycin, which binds sequence specifically to GpC (RpY) sites in the minor groove by forming two intermolecular hydrogen bonds between the threonine NH and the N3 of guanine (Liu *et al.*, 1991; Kamitori & Takusagawa, 1992; Brown *et al.*, 1994).

Role of Intermolecular Hydrogen Bonds between Ala CO and GNH_2 in the Sequence Specificity of Triostin A. In the triostin A–$[d(GACGTC)]_2$ complex, the intermolecular hydrogen bonds between the Ala NH and GN3 position the carbonyls of the alanine residues close to the 2-amino groups of guanine, thereby allowing the formation of a second set of intermolecular hydrogen bonds between the Ala carbonyl and the guanine 2-amino group (Addess & Feigon, 1994). This second set of hydrogen bonds has been correlated to the CpG specificity of triostin A, but their formation has not been proved directly to be critical to the binding of triostin A or echinomycin to CpG steps. We found that CpI steps, which do not contain exocyclic 2-amino groups like CpG steps, are not sequence-specific binding sites for triostin A but are for CysMeTANDEM. Furthermore, substitution of an inosine for a guanine on either the Watson or the Crick strand of DNA, which removes one of the two exocyclic 2-amino groups, abolishes the sequence-specific binding of echinomycin (Marchand *et al.*, 1992). Therefore, the two intermolecular hydrogen bonds between the Ala CO and G_4NH2 in the solution structure of the triostin A–$[d(GACGTC)]_2$ complex (Addess & Feigon, 1994), together with the two intermolecular hydrogen bonds between the Ala NH and GN3, appear to be required for sequence-specific binding of triostin A and echinomycin to CpG steps.

Role of Stacking in the Sequence Specificity of Triostin A and CysMeTANDEM. Since the minor groove of the CpI site is very similar to the minor groove of a TpA site and all the peptide–DNA interactions of both complexes are in the minor groove, the difference in the binding constant K_a between the two complexes must be due to a difference in the stacking energies between the quinoxaline rings and the base pairs of the YpR binding site. Therefore, the stacking interactions of the two quinoxaline rings with the binding site bases must more favorable for bis-intercalation around TpA than around CpI.

Triostin A binds less tightly to its sequence-specific CpG binding site than CysMeTANDEM does to its sequence-specific TpA binding site. This is also consistent with the stacking interactions for the bis-intercalation of the quinoxaline rings around CpG being less favorable than those for the bis-intercalation of the quinoxaline rings around TpA. Theoretical calculations by Gallego *et al.* (1993) indicate that the difference in stacking interactions is due to dipole–dipole interactions between the quinoxaline rings and the YpR base pairs. They argue that the net electronic dipole moment of a G·C base pair is high (about 5 D) and is parallel to the dipole moment of the quinoxaline ring when the G·C base pair and the quinoxaline are stacked in the manner observed in the structure of the triostin A–$[d(GACGTC)]_2$ complex, which results in unfavorable dipole–dipole interactions. Because the magnitude of the dipole moment for an A·T base pair (about 2 D) is lower than the dipole moment of the G·C base pair and is almost perpendicular to the dipole moment of a quinoxaline ring, the dipole–dipole interaction between the quinoxaline ring and the A·T base pair of the central TpA step is more favorable when they are stacked in a similar manner (Gallego *et al.*, 1993).

If bis-intercalation of the quinoxaline rings around a CpG step is energetically unstable relative to TpA, how does triostin A form a stable complex with $[d(GACGTC)]_2$? In the triostin A–DNA bis-intercalation complex, the two Ala carbonyls of triostin A form intermolecular hydrogen bonds with GNH2 protons. Stabilization is also provided by van der Waals interactions between the Val *N*-methyls and both the 2-amino groups of guanine and the O2 of cytosine (discussed below).

In contrast, the Ala carbonyls of CysMeTANDEM are positioned too far away to hydrogen bond with the GNH2 protons of the CpG binding site due to the narrower peptide ring from the intramolecular hydrogen bonds between the Ala carbonyls and the Val NH protons. Furthermore, CysMeTANDEM contains no Val *N*-methyl groups. As a result, CysMeTANDEM cannot form a stable complex with $[d(GACGTC)]_2$. Sequence-specific binding to TpA does not require the stabilizing forces of additional intermolecular hydrogen bonds since the dipole–dipole interactions between the quinoxaline rings and the DNA bases are favorable. This is why CysMeTANDEM forms a stable bis-intercalative complex with $[d(GATATC)]_2$.

Role of N-Methylation in the Sequence-Specific Binding of Triostin A. Since intermolecular hydrogen bonds between the Ala CO and a purine 2-amino group are not needed for sequence-specific binding at TpA steps, and the quinoxaline rings stack better with TpA bases than with CpG bases, one might expect triostin A to bind tightly to both TpA and CpG, but this is not the case. Though there is evidence to suggest

that triostin A could bind weakly to alternating A-T sequences of DNA (Lee & Waring, 1978), it has never been shown that triostin A can bind in a bis-intercalative manner around a TpA step. In the structure of the triostin A–$[d(GACGTC)]_2$ complex, the Val *N*-methyl groups point directly into the minor groove of the DNA, where they have favorable van der Waals interactions with the guanine 2-amino groups and the cytosine O2 atoms. Removal of the 2-amino groups prevents triostin A from binding in a bis-intercalative manner around a TpA step. The reason for this is not known but could be due to unfavorable dipole–dipole interactions between the Val *N*-methyl groups and the O2 carbonyl in the minor groove of CpI and TpA. These unfavorable interactions could be minimized by the presence of 2-amino groups in the minor groove of a YpR binding site. This could be further studied by examining the sequence specificity of a quinoxaline antibiotic derivative that has oxygens substituted for the *N*-methyl groups of the valine residues.

SUMMARY

In the previous study, we presented the three-dimensional solution structure of the triostin A–$[d(GACGTC)]_2$ complex and compared it with the solution structure of the CysMeTANDEM–$[d(GATATC)]_2$ complex. The solution structures of the two complexes are very similar to each other. The main difference between the two complexes occurs in the stability of the complexes and the structure of the bound drug in each complex. The difference in stability between these two complexes is caused by a difference in their stacking energies. N-Methylation of the Val NH causes triostin A to have a slightly different structure from CysMeTANDEM. This causes a difference in the hydrogen-bonding patterns in the two complexes. In this paper, we show how differences in stacking interactions, hydrogen bonding, and van der Waals interactions in the triostin A and CysMeTANDEM complexes explain why triostin A binds sequence specifically to CpG sites and why CysMeTANDEM binds sequence specifically to TpA sites. These studies indicate that very small changes in the chemical structure of a ligand can completely alter the sequence specificity while having little effect on the overall structure of the complex.

ACKNOWLEDGMENT

K.J.A. thanks Peter Schultze and Ed Wang for productive discussions and Thorsten Dieckmann for comments on the manuscript.

REFERENCES

Addess, K. J., & Feigon, J. (1994) *Biochemistry* (preceding 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.
Bailly, C., Marchand, C., & Waring, M. J. (1993) *J. Am. Chem. Soc. 115,* 3784–3785.
Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65,* 355–360.
Brown, D. R., Kurz, M., Kearns, D. R., & Hsu, V. (1994) *Biochemistry 33,* 651–664.
Gallego, J., Ortiz, A. R., & Gago, F. (1993) *J. Med. Chem. 36,* 1548–1561.
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.
Kamitori, S., & Takusagawa, F. (1992) *J. Mol. Biol. 225,* 445–456.
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.
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.
Lavesa, M., Olsen, R. K., & Fox, K. R (1993) *Biochem J. 289,* 605–607.
Lee, J. S., & Waring, M. J. (1978) *Biochem. J. 173,* 115–128.
Liu, X., Chen, H., & Patel, D. J. (1991) *J. Biomol. NMR 1,* 323–347.
Low, C. M. L., Olsen, R. K., & Waring, M. J. (1984) *FEBS Lett. 176,* 414–419.
Low, C. M. L., Fox, K. R., Olsen R. K., & Waring, M. J. (1986) *Nucleic Acids Res. 14,* 2015–2033.
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.
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.
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. (1979) in *Antibiotics* (Hahn, F. E., Ed.) Vol. 5, Part 2, pp 173–194, Springer-Verlag, Berlin, Heidelberg, and New York.
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.