NMR Characterization of a Heterocomplex Formed by Distamycin and Its Analog 2-ImD with d(CGCAAGTTGGC):d(GCCAACTTGCG): Preference for the 1:1:1 2-ImD:Dst:DNA Complex over the 2:1 2-ImD:DNA and the 2:1 Dst:DNA Complexes

Bernhard H. Geierstanger,[†] Tammy J. Dwyer,[‡,§] Yadagiri Bathini,[∥,⊥] J. William Lown,[∥] and David E. Wemmer*[,‡]

Contribution from the Department of Chemistry and the Graduate Group in Biophysics, University of California, Berkeley, California 94720, and the Department of Chemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2G2

Received December 15, 1992

Abstract: The minor-groove binder distamycin (Dst), its pyrrole–imidazole–pyrrole analog, 2-imidazole–distamycin (2-ImD), and the oligonucleotide d(CGCAAGTTGGC):d(GCCAACTTGCG) were found to form a 1:1:1 2-ImD-Dst:DNA complex. As characterized by 2D NOE spectroscopy combined with molecular modeling, one 2-ImD and one Dst molecule bind simultaneously in a head-to-tail orientation contacting the minor groove of the central AAGTT:AACTT site. The 2-ImD ligand lies along the AAGTT strand with the imidazole nitrogen of the ligand specifically interacting with the guanine amino group. The distamycin ligand lies along the AACTT strand. The molecular structure of the 1:1:1 2-ImD:Dst:DNA complex is very similar to those of the complexes formed between two 2-ImD molecules (2:1 2-ImD:DNA complex) or two distamycin molecules (2:1 Dst:DNA complex) and the same oligonucleotide duplex. Competition titrations confirm that the 1:1:1 2-ImD:Dst:DNA complex forms preferentially over the 2:1 2-ImD:DNA complex, as well as over the newly discovered 2:1 Dst:DNA complex. These results indicate that (i) the hydrogen-bond-accepting imidazole nitrogen of the 2-ImD ligand in the 1:1:1 2-ImD:Dst:DNA complex is responsible for the strand-specific recognition of a GC base pair, (ii) the availability of a single hydrogen bond acceptor on the ligand molecule per guanine amino group enhanced both specificity and affinity of DNA binding, and (iii) different distamycin-like ligands can be combined in the 2:1 binding motif to expand the range of DNA sequences that can be specifically recognized through the minor groove.

Introduction

Peptide-linked polypyrroles, such as the naturally occurring antibiotics netrospin and distamycin (Dst) (**1**), bind noncovalently to the minor groove of DNA, preferentially to AT-rich DNA sequences.[1] It is presumed that binding to GC-containing sequences is unfavorable because of steric interference between the pyrrole H3 protons of the ligands and the guanine amino group at the floor of the minor groove.[2] DNA binding ligands of this class have been modified in an attempt to vary their sequence specificity,[3,4] fostering their use in cancer therapy and as tools in molecular biology. One strategy, the replacement of one pyrrole ring by an imidazole ring to form specific ligand:DNA complexes with GC-containing sequences,[2,3] has recently proven successful in two cases:[4–6] The imidazole–pyrrole–pyrrole ring system 1-methylimidazole-2-carboxamide–netropsin (2-ImN) (**2**), designed by Wade and Dervan, was shown to bind specifically to the five-base-pair site TGACT:AGTCA.[4,5] Concurrently, Wemmer and Lown designed the distamycin analog 2-imidazole–distamycin (2-ImD) (**3**), in which the imidazole nitrogen of this pyrrole–imidazole–pyrrole ring system targets the guanine amino group in the AAGTT:AACTT binding site.[6] Structural analysis of these ligand:DNA complexes by NMR combined with molecular modeling[5a,6] led us to expand upon the current design criteria for sequence-specific minor groove ligands:

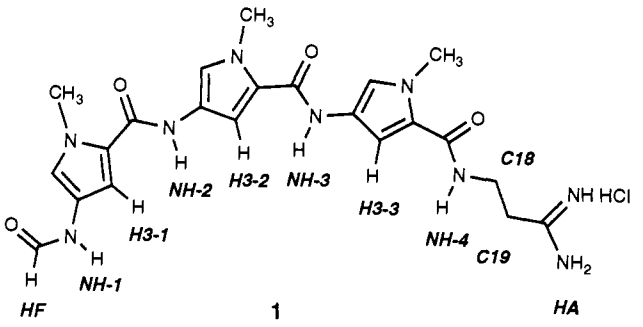
(*i*) *Binding in the 2:1 mode optimizes contacts to the DNA groove*: Both the 2-ImN[5a] and the 2-ImD[6] complexes with DNA revealed that these molecules bind with high cooperativity in a 2:1 ligand:DNA mode similar to the distamycin:DNA complexes characterized previously.[7] In each complex two ligand molecules stack side-by-side in an antiparallel fashion and fill the minor groove without introducing major distortions in the DNA. The 2:1 ligand:DNA binding mode optimizes hydrogen bonding, van der Waals contacts, and electrostatic interactions between the cationic ligands and the minor groove of DNA.[5a,6,7]

(*ii*) *Hydrogen bonds to the imidazole nitrogens on the ligand allow the recognition of guanine amino groups*: In the 2:1 2-ImN:

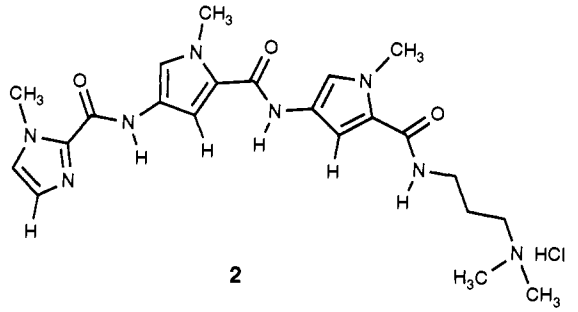
* Address correspondence to this author at the Department of Chemistry, University of California, Berkeley, CA 94720.
† Graduate Group in Biophysics, University of California.
‡ Department of Chemistry, University of California.
§ Current Address: College of Arts & Sciences, California State University, San Marcos, San Marcos, CA 92096.
∥ Department of Chemistry, University of Alberta.
⊥ Current address: SynPhar Laboratories Inc., 24 Taiho Alberta Center, Edmonton, AB, Canada T6E 5V2.

(1) Zimmer, C.; Wähnert, U. *Prog. Biophys. Mol. Biol.* **1986**, *47*, 31.
(2) Kopka, M. L.; Yoon, C.; Goodsell, D.; Pjura, P.; Dickerson, R. E. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 1376.
(3) Lown, J. W.; Krowicki, K.; Bhat, U. G.; Skorobogaty, A.; Ward, B.; Dabrowiak, J. C. *Biochemistry* **1986**, *25*, 7406. Kissinger, K.; Krowicki, K.; Dabrowiak, J. C.; Lown, J. W. *Biochemistry* **1987**, *26*, 5590. Lee, M.; Chang, D. K.; Hartley, J. A.; Pon, R. T.; Krowicki, K.; Lown, J. W. *Biochemistry* **1988**, *27*, 445. Burckhardt, G.; Luck, G.; Zimmer, C.; Storl, J.; Krowicki, K.; Lown, J. W. *Biochim. Biophys. Acta* **1989**, *1009*, 11. Lee, M.; Krowicki, K.; Shea, R.; Lown, J. W.; Pon, R. T. *J. Molecular Recognition* **1989**, *2*, 84.
(4) Wade, W. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1987**, *109*, 1574.
(5) (a) Mrksich, M.; Wade, W. S.; Dwyer, T. J.; Geierstanger, B. H.; Wemmer, D. E.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 7586. (b) Wade, W. S.; Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 8783.
(6) Dwyer, T. J.; Geierstanger, B. H.; Bathini, Y.; Lown, J. W.; Wemmer, D. E. *J. Am. Chem. Soc.* **1992**, *114*, 5911.
(7) (a) Pelton, J. G.; Wemmer, D. E. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 5723. (b) Pelton, J. G.; Wemmer, D. E. *J. Am. Chem. Soc.* **1990**, *112*, 1393.

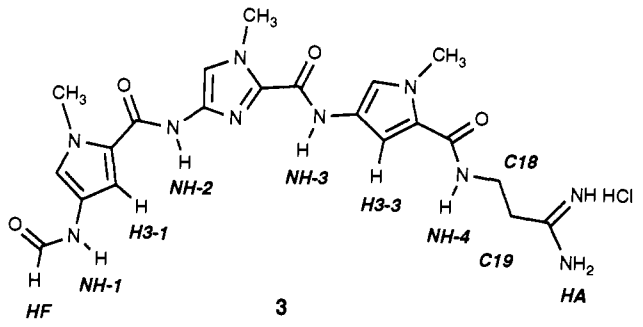
0002-7863/93/1515-4474$04.00/0



1



2



3

DNA complex, the imidazole nitrogen of each ligand recognizes a single guanine amino group of the target sequence TGACT:AGTCA.[5a] Evidence for the formation of hydrogen bonds between the imidazole nitrogens and the guanine amino protons has been provided indirectly by NOE crosspeaks between the guanine amino protons and ligand protons spatially close to the imidazole nitrogen. A similar complex is formed by the pyrrole–imidazole–pyrrole ring system 2-ImD. The two-ligand binding site of 2-ImD is the minor groove of the AAGTT:AACTT sequence, as confirmed by intermolecular ligand–DNA NOEs and molecular modeling.[6] The imidazole nitrogens of the two 2-ImD ligands are in close proximity to the amino group of the central guanine. Since 2-ImD was found not to bind specifically to the site AAATT:AATTT, a good 2:1 distamycin:DNA binding site, the ligand imidazole nitrogens proximal to the amino group of guanine are responsible for the recognition of the GC-containing sequence. In contrast to the 2:1 2-ImN:DNA complex, we did not find NOE peaks between any ligand protons and the guanine amino protons in the 2:1 2-ImD:DNA complex.

Crosspeaks to guanine amino protons are usually not observed in NOESY spectra due to line broadening caused by rotation about the C–N bond and proton exchange with solvent.[8,9] The presence of such NOE peaks in the 2:1 2-ImN:DNA complex indicates that rotation of the guanine amino protons is slowed significantly due to strong interactions with each ligand imidazole nitrogen. While, in the case of the 2:1 2-ImN:DNA complex with TGACT:AGTCA, only one hydrogen bond acceptor (imidazole nitrogen) is located in the vicinity of each guanine amino group, a single guanine amino group can interact with two imidazole nitrogens in the 2:1 2-ImD:DNA complex with AAGTT:AACTT. This can possibly result in weaker hydrogen bonding and decreased binding affinity and specificity and would explain the absence of NOE peaks to the guanine amino group in the 2:1 2-ImD complex.

These observations led us to predict that the affinity of distamycin-like ligands for specific DNA sequences could be enhanced by choosing pairs of ligand molecules such that only a single imidazole ring per pair would interact with each guanine amino group. For the AAGTT:AACTT binding site this can be achieved by combining one distamycin molecule and one pyrrole–imidazole–pyrrole ligand, 2-ImD, in a heteromeric 2:1 ligand:DNA complex. In the present work we examine the binding of distamycin, and mixtures of distamycin and its imidazole derivative 2-ImD, to this binding site and show that, indeed, a 1:1:1 2-ImD:Dst:DNA complex is optimal at this site. In addition an unexpected 2:1 complex of distamycin with this DNA sequence is characterized.

Experimental Section

Synthesis of 2-ImD (2) and Oligonucleotides. The distamycin analog 2-ImD (**2**) was synthesized as described previously.[6] Distamycin was purchased from Sigma and used without further purification. The oligomers d(CGCAAGTTGGC) and d(GCCAACTTGCG) were synthesized and purified as reported previously.[7a]

Sample Preparation. NMR samples were prepared by dissolving the undecamer oligonucleotide duplex in 0.25 mL of 20 mM sodium phosphate buffer (pH 7.0) and then lyophilizing to dryness. For experiments carried out in D_2O, the solid was lyophilized twice from 99.9% D_2O (Cambridge Isotope Laboratories) and finally redissolved in 0.5 mL of 99.96% D_2O (Cambridge Isotope Laboratories). For experiments in H_2O, the solid was redissolved in a 90% H_2O/10% D_2O mixture to a final volume of 0.5 mL.

Two stock solutions of 2-ImD·HCl were prepared by dissolving 1.93 and 1.88 mg of the ligand each in 100 μL of 99.96% D_2O. The concentrations of the stock solutions were determined to be 19.3 and 22.6 mM, respectively, by UV absorbance at 304 nm ($\epsilon = 3.6 \times 10^4$ M^{-1} cm^{-1}). Similarly, a stock solution of distamycin of 19.2 mM was prepared by dissolving 1 mg in 100 μL of 99.96% D_2O ($\epsilon = 3.4 \times 10^4$ M^{-1} cm^{-1} at 304 nm). Stock solutions of the ligands were stored at −70 °C. Extinction coefficients for d(CGCAAGTTGGC) and d(GCCAACTTGCG) were calculated[10] to be 1.03×10^5 and 9.93×10^4 M^{-1} cm^{-1}, respectively. The concentrations of the double-stranded DNA samples were determined to be 1 mM by UV absorbance at 260 nm and 80 °C.

1D NMR Titration. 1D NMR titration spectra were acquired at 600 MHz on a Bruker AMX-600 spectrometer. 2-ImD·HCl and distamycin were titrated into the NMR sample containing the DNA in approximately 0.25 mol equiv per addition unless indicated otherwise. Due to experimental uncertainties in the concentrations of the ligand stock solutions, ligand to DNA ratios are approximated from the intensity ratios of NMR resonances. 1D spectra in D_2O were acquired at 25 °C using 8192 complex points, 128 scans, and a spectral width of 6024 Hz. A presaturation pulse was applied during the 2.0-s recycle delay to suppress the residual HDO resonance. For 1D spectra in H_2O, the spectral width was 13 514 Hz using 128 scans and a 1:1 jump and return sequence for solvent suppression.[11]

2D NOESY Spectra. All 2D NMR spectra were acquired at 600 MHz on a Bruker AMX-600 spectrometer. NOESY spectra in D_2O were acquired at 25 °C using the standard TPPI phase cycle.[12] The spectra were collected with 1024 complex points in t_2 using a spectral width of 6024 Hz and a mixing time of 200 ms. 481–493 t_1 experiments were recorded and zero-filled to 1 K. For each t_1 value, 48 or 80 scans were signal averaged using a recycle delay of 2 s. A presaturation pulse was applied during the recycle and mixing periods to suppress the residual HDO resonance.

NOESY spectra in water were acquired at 25 °C, replacing the last 90° pulse by a 1:1 jump and return sequence[9] to suppress the solvent

(8) Boelens, R.; Scheek, R. M.; Dijkstra, K.; Kaptein, R. *J. Magn. Reson.* **1985**, *62*, 378.
(9) Skelenár, V.; Brooks, B. R.; Zon, G.; Bax, A. *FEBS Lett.* **1987**, *216*, 249.
(10) Warshaw, M.; Cantor, C. *Biopolymers* **1970**, *9*, 1079.
(11) Plateau, P.; Guéron, M. *J. Am. Chem. Soc.* **1982**, *104*, 7310.
(12) Drobny, G.; Pines, A.; Sinton, S.; Weitekamp, D. P.; Wemmer, D. E. *Faraday Symp. Chem. Soc.* **1979**, *13*, 49.

signal, using the pulse sequence: delay $90°_x$–t_1–$90°_x$–τ_{mix}–$90°_x$–Δ^{11}–$90°_\phi$–t_2. Phase-sensitive detection was accomplished using TPPI. The spectra were collected into 2048 complex points in t_2 using a spectral width of 13 514 Hz and mixing times of 100 and 200 ms. 526–645 t_1 experiments with 64 or 80 scans were recorded and zero-filled to 2 K. The delay period Δ^{11} was calibrated to give optimum excitation in the imino region of the ^{1}H spectrum (11–13 ppm) with a single null at the water resonance.

The data were processed with FTNMR (Hare Research) on a Vax 4000-300 computer or FELIX (Hare Research) on a Silicon Graphics IRIS/4D workstation. The 2D NOESY data were apodized with a skewed sine bell function in both dimensions (800 or 1600 points, phase 60°, skew 0.5 in t_2; 481–645 points, phase 60°, skew 0.7 in t_1). The first row of the data matrix was multiplied by 0.5 prior to Fourier transformation in t_1 to suppress t_1 ridges.

Distance Restraints. Intermolecular distance restraints were generated from the volume integrals of the crosspeaks in the H_2O NOESY spectra acquired at mixing times of 100 and 200 ms. Since spin diffusion contributes more significantly at longer mixing times, we calculated distance restrains only for crosspeaks that were present in the 100-ms NOESY. Volume integrals were measured for each mixing time using FELIX. The crosspeak volumes were scaled according to $V_{corr} = V_{obs}/\sin(\Delta\omega\Delta^{11})$, where V_{corr} is the corrected crosspeak volume, V_{obs} is the measured crosspeak volume, Δ^{11} is the delay adjusted to optimize excitation of the region of interest, and $\Delta\omega$ is the difference in frequency between the null and the crosspeak of interest in ω_2. The crosspeak volumes were classified semiquantitatively into three categories: strong (less than 2.5 Å), medium (between 2.5 and 3.7 Å), or weak (greater than 3.7 Å) relative to the volume integrals of the cytosine H5–H6 crosspeak volumes at each mixing time. In all, for modeling the 1:1:1 2-ImD:Dst:DNA complex, 21 intermolecular ligand–DNA restraints, 2 intermolecular ligand–ligand restraints, and 12 intramolecular restraints of the ligands were used. For the 2:1 Dst:DNA homocomplex, the respective numbers of restraints were 24, 2, and 13. Listings of the intermolecular ligand–DNA and ligand–ligand restraints and the achieved distances are available as supplemenary material.

Structure Refinement. The starting structure for the minimization of the 2:1 2-ImD:DNA complex with d(CGCAAGTTGGC):d(GCCAACTTGCG)[6] was used as the initial model for the 1:1:1 2-ImD:Dst:DNA complex and the 2:1 Dst:DNA complex. This starting structure had been constructed using the Biopolymer module of Insight II (Biosym) from standard B-form DNA.[6] The ligands from the energy-minimized NMR structure of the 2:1 Dst:DNA complex with d(CGCAAATTGGC):d(GCCAATTTGCG)[7] had been docked manually into the minor-groove binding pocket on the Silicon Graphics workstation. To model the two complexes in this work, the two 2-ImD ligand molecules were transformed into one 2-ImD and one distamycin molecule or into two distamycin molecules, respectively, with the help of the Builder module of Insight II. Energy minimizations were performed using the Discover module of Insight II. Hydrogen bonds for standard Watson–Crick base pairing were included as NOE restraints using a force constant of 100 (kcal/mole)/Å^2 while a force constant of 25 (kcal/mole)/Å^2 was used for the experimentally derived NOE restraints. The cutoff distance for nonbonded interactions was set at 12 Å with a switching distance of 2 Å. A distance-dependent dielectric of the form $\epsilon = R$ was used to account for solvent effects. The energy of the complexes was initially minimized using 100 steps of a steepest descents algorithm, and a final root mean square derivative of <0.001 (kcal/mole)/Å^2 was achieved using 20 000 steps of conjugate gradient minimization.

Results

Titration of the AAGTT:AACTT Site with 2-ImD and Distamycin. The results of the titration of the duplex d(CGCAAGTTGGC):d(GCCAACTTGCG) with 2-ImD and distamycin (Dst) to approximately equimolar amounts are shown in Figure 1. The individual aromatic base protons of the central three-base-pair site in the free duplex are indicated by stars. These were assigned in the NOESY spectrum of the free duplex by standard sequential methods.[13] As 2-ImD is added to the DNA sample, new resonances appear which are indicative of the formation of a 2:1 2-ImD:DNA complex (filled squares in Figure 1b), as described previously.[6] Upon addition of 0.3 equiv of distamycin to the 0.7:0:1 2-ImD:Dst:DNA sample, another set of resonances

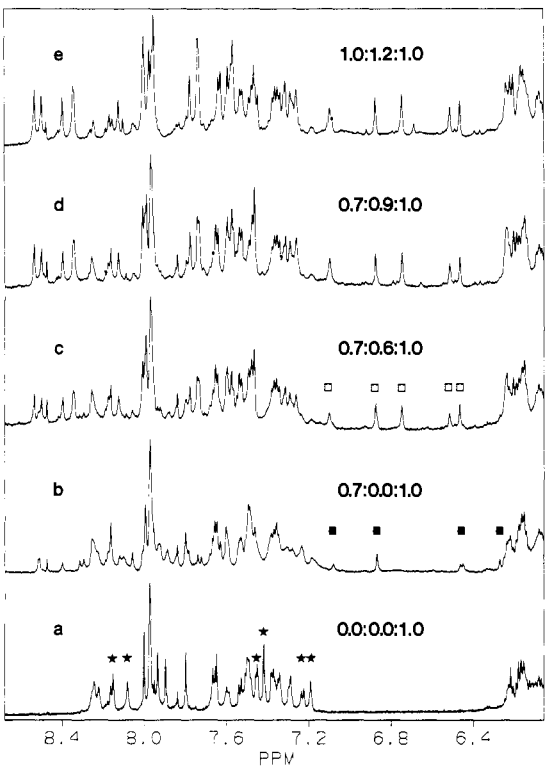


Figure 1. Titration of d(CGCAAGTTGGC):d(GCCAACTTGCG) with 2-ImD and distamycin (Dst) at 25 °C: (a and b) addition of 2-ImD; (c and d) addition of distamycin; (e) addition of distamycin and 2-ImD to correct for undertitration. The approximate 2-ImD:Dst:DNA ratio, based on the NMR signal intensities, is indicated for each spectrum. Stars denote the aromatic proton resonances of uncomplexed DNA for the central three base pairs of the ligand binding site. Filled squares indicate proton resonances of the 2:1 2-ImD:DNA complex, while proton resonances of the 1:1:1 2-ImD:Dst:DNA complex are shown as open squares.

appears. Further addition of distamycin increases the intensities of this new set of peaks (indicated by open squares in Figure 1c) while the intensities of the 2:1 2-ImD:DNA complex and of free DNA decrease. At a 2-ImD:Dst:DNA ratio of 1:1.2:1 (Figure 1e), this new set of peaks represents the predominant ligand:DNA complex. These observations are consistent with the formation of a 1:1:1 2-ImD:Dst:DNA complex. A comparison of the peak intensities of complexed to uncomplexed DNA as a function of concentration of added ligand confirms that this single complex has a 2-ImD:Dst:DNA stoichiometry of 1:1:1. The sharpness of the ligand and DNA signals throughout the titration indicates that the ligands are in slow exchange[9] with a k_{off} of less than a few per second at ambient temperature.

Titration of the AAGTT:AACTT Site with Distamycin. The results of the titration of the AAGTT:AACTT duplex with distamycin are shown in Figure 2a–c. The individual aromatic base protons of the central three-base-pair site in the free duplex are indicated by stars. Upon the addition of 0.3 molar equiv of distamycin, the number of resonances doubles. At a 2-ImD:Dst:DNA stoichiometry of 0:1:1, the intensity ratio of complexed DNA to free DNA is 1:1, indicating that half of the DNA exists as complex. At a stoichiometry of 0:2:1, only one set of complexed DNA and ligand resonances (indicated by filled circles in Figure 2c) is present, confirming the exclusive formation of a slowly exchanging 2:1 Dst:DNA complex. This 2:1 Dst:DNA complex exhibits a higher degree of cooperativity than distamycin binding to AATTT:AATTT,[7] in which two 1:1 binding modes are evident up to a ligand:DNA ratio of approximately 0.75:1 and the 2:1 binding mode becomes predominant only at higher ratios. With AAGTT:AACTT, even at a Dst:DNA ratio of 0.3:1, a 1:1 complex of distamycin is not detectable.

(13) Hare, D. R.; Wemmer, D. E.; Chou, S.-H.; Drobny, G.; Reid, B. R. *J. Mol. Biol.* **1983**, *171*, 319.

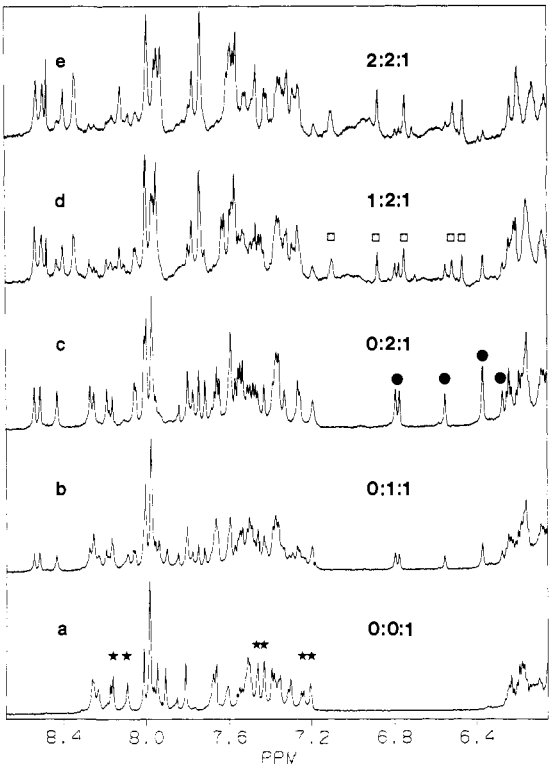


Figure 2. Titration of d(CGCAAGTTGGC):d(GCCAACTTGCG) with distamycin (Dst) and 2-ImD at 25 °C: (**a**–**c**) addition of distamycin; (**d** and **e**) addition of 2-ImD. Filled circles indicate proton resonances of the 2:1 Dst:DNA complex, while open squares denote the 1:1:1 2-ImD:Dst:DNA complex. The approximate 2-ImD:Dst:DNA ratio is indicated for each spectrum.

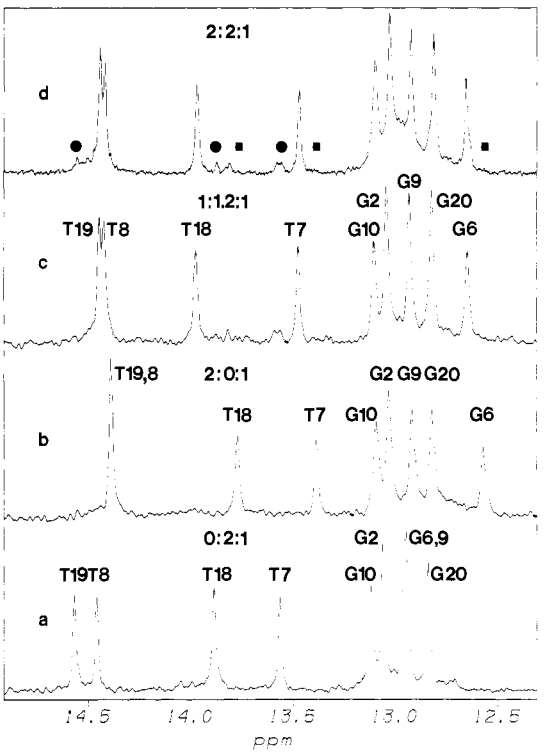


Figure 3. Imino proton region of the spectra of the three ligand complexes with d(CGCAAGTTGGC):d(GCCAACTTGCG) at 25 °C in 90% H_2O/10% D_2O: (**a**) 2:1 Dst:DNA complex; (**b**) 2:1 2-ImD:DNA complex; (**c**) 1:1:1 2-ImD:Dst:DNA complex (endpoint of titration in Figure 1); (**d**) 1:1:1 2-ImD:Dst:DNA complex (endpoint of titration in Figure 2). The approximate 2-ImD:Dst:DNA ratio is indicated for each spectrum. The imino proton resonances are labeled. The terminal imino protons G22 and G12 are not observed due to proton exchange with solvent. Filled circles denote (from left) the positions of T19, T18, and T7 imino protons of the 2:1 Dst:DNA complex, while the filled squares show the positions of T18, T7, and G6 imino resonances of the 2:1 2-ImD:DNA complex. The first blocks of 2D H_2O NOESY spectra of each complex were zero-filled to 8192 data points for the spectra in Figure 3a–c. The spectrum in Figure 3d was acquired with 8192 data points.

Competition Experiments. In order to test relative binding affinities of the different complexes, 2-ImD was added to the AAGTT:AACTT sample already containing 2 equiv of distamycin. The results are shown in Figure 2d and e. Prior to the addition of 2-ImD, the sample is purely a 2:1 Dst:DNA complex (Figure 2c). Upon addition of 0.25 molar equiv of 2-ImD, a set of resonances corresponding to the 1:1:1 2-ImD:Dst:DNA complex appears. Subsequent additions of 2-ImD increase the resonance intensities of the 1:1:1 2-ImD:Dst:DNA complex (indicated as open squares in Figure 2d) while the 2:1 Dst:DNA complex signals decrease. At a 2-ImD:Dst:DNA ratio of approximately 2:2:1 (Figure 2e), the 1:1:1 2-ImD:Dst:DNA complex is dominant, preferentially formed relative to both the 2:1 Dst:DNA and the 2:1 2-ImD:DNA complexes.

Figure 3 shows a comparison of the imino proton resonances for the three observed ligand:DNA complexes. The guanine G6 imino resonance line is shifted upfield by about 0.3 ppm in the 1:1:1 2-ImD:Dst:DNA complex (Figure 3c,d) and by 0.4 ppm in the 2:1 2-ImD:DNA complex (Figure 3b) relative to the 2:1 Dst:DNA complex (Figure 3a). This suggests, in combination with the results of molecular modeling, a strong interaction between the 2-ImD imidazole nitrogens and the guanine G6 amino group in the center of the binding site in each complex. Figure 3d shows the imino proton region of an AAGTT:AACTT sample titrated with approximately 2 equiv of 2-ImD and 2 equiv of distamycin. The resolvable imino peak positions for the 2:1 2-ImD:DNA complex are indicated with filled squares, while those for the 2:1 Dst:DNA complex are shown with filled circles. In the presence of an excess of both ligands, the 1:1:1 2-ImD:Dst:DNA complex is by far the dominant form in solution. Only small amounts of the 2:1 Dst:DNA complex, but not of the 2:1 2-ImD:DNA complex, can be detected. According to the integrated peak intensities, the 2:1 Dst:DNA complex accounts for only 7–8% (±2%) of the complexes present. The relative complex populations are not measurably affected as more 2-ImD is added (data not shown). This confirms that the 1:1:1 2-ImD:Dst:DNA complex exhibits at least a 10-fold higher binding affinity for AAGTT:AACTT than either 2:1 ligand:DNA complex.

Signal Assignments. 2D NOESY spectra of the 1:1:1 2-ImD:Dst:DNA complex in D_2O were acquired at an approximate 2-ImD:Dst:DNA ratio of 1:1:1. Each aromatic base proton (H8 of purines, H6 of pyrimidines) was assigned through its connectivity to the C1′H of its own sugar and the C1′H of its 5′ neighbor. Assignment of the imino protons in the H_2O NOESY by standard methods[14] facilitated the assignment of the adenine H2 protons in the complex. Ligand amide protons and pyrrole H3 protons were assigned by intramolecular connectivities among these protons in the NOESY spectrum of the complexes in H_2O. Ligand methylene protons were assigned via intermolecular contacts to the formyl proton (HF) on the opposite ligand of the 1:1:1 2-ImD:Dst:DNA complex (C19 protons), as well as to A4 H2 and A15 H2 (C18 and C19 of Dst and 2-ImD, respectively). Assignment of the ligand methylene protons was confirmed by NOE crosspeaks to NH-4 and amidinium protons of the ligands. The chemical shift assignments for the deoxyribose ring C1′ protons and aromatic protons of the DNA in the free duplex and in the 1:1:1 2-ImD:Dst:DNA complex are shown in Table I. The chemical shift assignments for the bound ligand protons are given in Table II.

The 2D NOESY spectra of the 2:1 Dst:DNA complex in D_2O and H_2O were assigned similarly (spectra were acquired at 2 mol

(14) Wüthrich, K. *NMR of Proteins and Nucleic Acids*; J. Wiley & Sons: New York, 1986.

Table I. Chemical Shift Assignments of the d(CGCAAGTTGGC):d(GCCAACTTGCG) Duplex, Free and in the 1:1:1 2-ImD:Dst:DNA Complex[a]

	H6/H8			H1′		
	free duplex	1:1:1 complex	Δδ	free duplex	1:1:1 complex	Δδ
			Strand 1			
C1	7.65	7.63	−0.02	5.80	5.78	−0.02
G2	7.96	7.99	+0.03	5.90	5.89	−0.01
C3	7.35	7.36	+0.01	5.49	5.83	+0.34
A4	8.25	8.53	+0.28	5.93	5.55	−0.38
A5	8.08	8.34	+0.26	6.02	5.72	−0.30
G6	7.45	7.72	+0.27	5.82	5.14	−0.68
T7	7.18	7.08	−0.10	6.02	5.69	−0.33
T8	7.28	7.24	−0.04	5.85	5.22	−0.63
G9	7.89	7.71	−0.18	5.65	5.61	−0.04
G10	7.79	7.73	−0.06	6.01	5.92	−0.09
C11	7.48	7.44	−0.04	6.21	6.21	0.00
			Strand 2			
G2	7.99	7.97	−0.02	6.02	6.00	−0.02
C13	7.50	7.51	+0.01	6.08	6.07	−0.01
C14	7.48	7.47	−0.01	5.32	5.78	+0.46
A15	8.24	8.49	+0.25	5.95	5.55	−0.40
A16	8.15	8.39	+0.24	6.11	6.13	+0.02
C17	7.23	7.25	+0.02	5.71	5.45	−0.26
T18	7.41	7.30	−0.11	6.03	5.58	−0.45
T19	7.33	7.28	−0.05	5.85	5.28	−0.57
G20	7.93	7.77	−0.16	5.85	5.80	−0.05
C21	7.37	7.33	−0.04	5.81	5.70	−0.11
G22	7.96	7.94	−0.02	6.17	6.15	−0.02

[a] Chemical Shifts are given in ppm with the residual HDO signal referenced to 4.80 ppm (25 °C).

Table II. Chemical Shift Assignments of the 2-ImD and the Distamycin (Dst) Ligand in the 1:1:1 2-ImD:Dst:DNA Complex with d(CGCAAGTTGGC):d(GCCAACTTGCG)

proton	2-ImD	Dst	proton	2-ImD	Dst
HF	7.98	8.11	NH-4	7.98	7.75
NH-1	10.05	10.33	C(18)H[a]	2.48	2.42
H3-1	5.78	6.18		3.71	3.71
NH-2	9.59	8.83	C(19)H[a]	1.25	1.23
H3-2	n.a.[c]	6.45		2.24	2.20
NH-3	10.63	9.71	HA[a,b]	9.37	9.33
H3-3	6.50	6.23		8.25	8.25

[a] Not stereospecifically assigned. [b] HA = amidinium protons. [c] Not applicable.

equiv of added ligand). The chemical shift assignments for the 2:1 Dst:DNA complex will be reported elsewhere.[15]

Intermolecular Contacts. The NOESY spectrum of the 1:1:1 2-ImD:Dst:DNA complex in H_2O is shown in Figure 4. The spectrum contains numerous intermolecular contacts that permit the placement of the ligands on the DNA. The ligand–ligand and ligand–DNA contacts are summarized in Table III. Strong NOE crosspeaks between the H3-1 and H3-3 pyrrole protons of the 2-ImD ligand and the H2 proton of A5 and A16, respectively, and between the H3-1 and H3-3 protons of the distamycin ligand and A16 H2 and A5 H2, respectively, define the orientations of the ligands. Further, strong NOEs observed between the formyl protons (HF) of each ligand and the C19 methylene protons on the opposite ligand confirm the antiparallel arrangement. Additional contacts between ligand amide protons and DNA protons along the minor groove of the central five-base-pair binding site confirm the placement of the ligands. Most intermolecular contacts are shown diagrammatically in Figure 5. With the exception of the NOEs between A16 C2H and NH-3 of the 2ImD ligand and NH-1 of the Dst ligand, the contacts are similar to those observed in the 2:1 2-ImD:DNA complex.[6]

The analysis of the D_2O and H_2O spectra of the 2:1 Dst:DNA complex (not shown) reveals similar contacts (Table IV, Figure 6). Stronger contacts in the 2:1 Dst:DNA complex between NH-1 of Dst 1 and NH-3 of Dst 2 with A5 C2H are indicative of minor ligand rearrangements with respect to the 1:1:1 2-ImD:Dst:DNA complex (both NOE contacts were present in the 200-ms mixing time NOESY, but absent in the 100-ms NOESY). The orientation and general placement of the ligands in the 2:1 Dst:DNA complex, the 1:1:1 2-ImD:Dst:DNA complex, and the 2:1 2-ImD:DNA complex[6] are, however, identical.

Molecular Modeling. Semiquantitative modeling of the 1:1:1 2-ImD:Dst:DNA complex and the 2:1 Dst:DNA complex were carried out using the Insight II/Discover model building and simulation package. All ligand–DNA and ligand–ligand contacts listed in Tables III and IV (except for ligand–DNA contacts involving amidinium protons and methylene protons) plus intramolecular ligand contacts were used as restraints. Listings of the intermolecular ligand–DNA and ligand–ligand restraints and the achieved distances in both models are available as supplementary material.

The structure of the 1:1:1 2-ImD:Dst:DNA complex obtained from the energy minimization procedure described above is presented in Figure 7. The ligands fit snugly into the minor groove similar to models of distamycin complexes with AT-rich DNA oligomers[7] and the 2:1 2-ImD:DNA complex.[6] The positively charged tails of the ligands lie deep in the minor groove while the van der Waals surface of the ring systems line the wall of the minor groove. The imidazole nitrogen of the 2-ImD ligand points toward the NH_2 group of G6. For the 2-ImD molecule, hydrogen bonds are assigned by Insight II between NH-3 and T7 O2, but not between N1 of the imidazole and the NH_2 group of G6. The distance between N1 and the nitrogen of the imino group is on the order of 3.5 Å, slightly longer than the maximum distance for a hydrogen bond. Two hydrogen bonds between the distamycin molecule and DNA are observed (NH-3 and T18 O2 and NH-4 and T19 O2). The distamycin molecule does not sterically interfere with the NH_2 group of G6. According to molecular modeling, the stacking of the ligand molecules with respect to each other is similar to the staggered arrangement found in the 2:1 complexes formed by 2-ImD with AAGTT:AACTT[6] and distamycin with AAATT:AATTT and AAATTT:AAATTT.[7]

A similar description applies to the minimized model of the 2:1 Dst:DNA complex with AAGTT:AACTT. A comparison of the minimized structures of the 1:1:1 2-ImD:Dst:DNA complex, the 2:1 2-ImD:DNA complex,[6] and the 2:1 Dst:DNA complex indicates slight differences in the ligand positions with respect to the guanine G6 amino group. Given the experimental limitations, however, at the present resolution the structures of all three complexes are essentially identical.

Discussion

Previous NMR and modeling studies on the 2-ImN[5a] and the 2-ImD[6] complexes with DNA suggested that binding and specificity of distamycin analogs to GC-containing DNA sequences could be enhanced by using a single imidazole nitrogen per guanine amino group. Formation of a 1:1:1 2-ImD:Dst:DNA complex indicates that this idea is correct. The 1:1:1 2-ImD:Dst:DNA complex consists of one 2-ImD molecule and one distamycin molecule stacked side-by-side in the minor groove of the AAGTT:AACTT binding site. The imidazole ring of the 2-ImD ligand specifically recognizes the amino group of the central guanine and determines the specificity of the complex.

Competition experiments (Figures 2c–e and 3c–d) clearly establish that the 1:1:1 2-ImD:Dst:DNA complex forms in preference to a previously characterized 2:1 2-ImD:DNA complex[6] and a 2:1 Dst:DNA complex identified in a control titration of the AAGTT:AACTT duplex with distamycin (Figure 2c). When 2-ImD is added to a AAGTT:AACTT sample complexed to saturation with 2 equiv of distamycin, the 1:1:1 2-ImD:Dst:DNA complex forms stoichiometrically (Figure 2c–e). At a 2-ImD:Dst:DNA ratio of approximately 2:2:1, the peak intensities of the three ligand:DNA complexes indicate the relative populations,

(15) Leheny, R. A.; *et al.* Unpublished results.

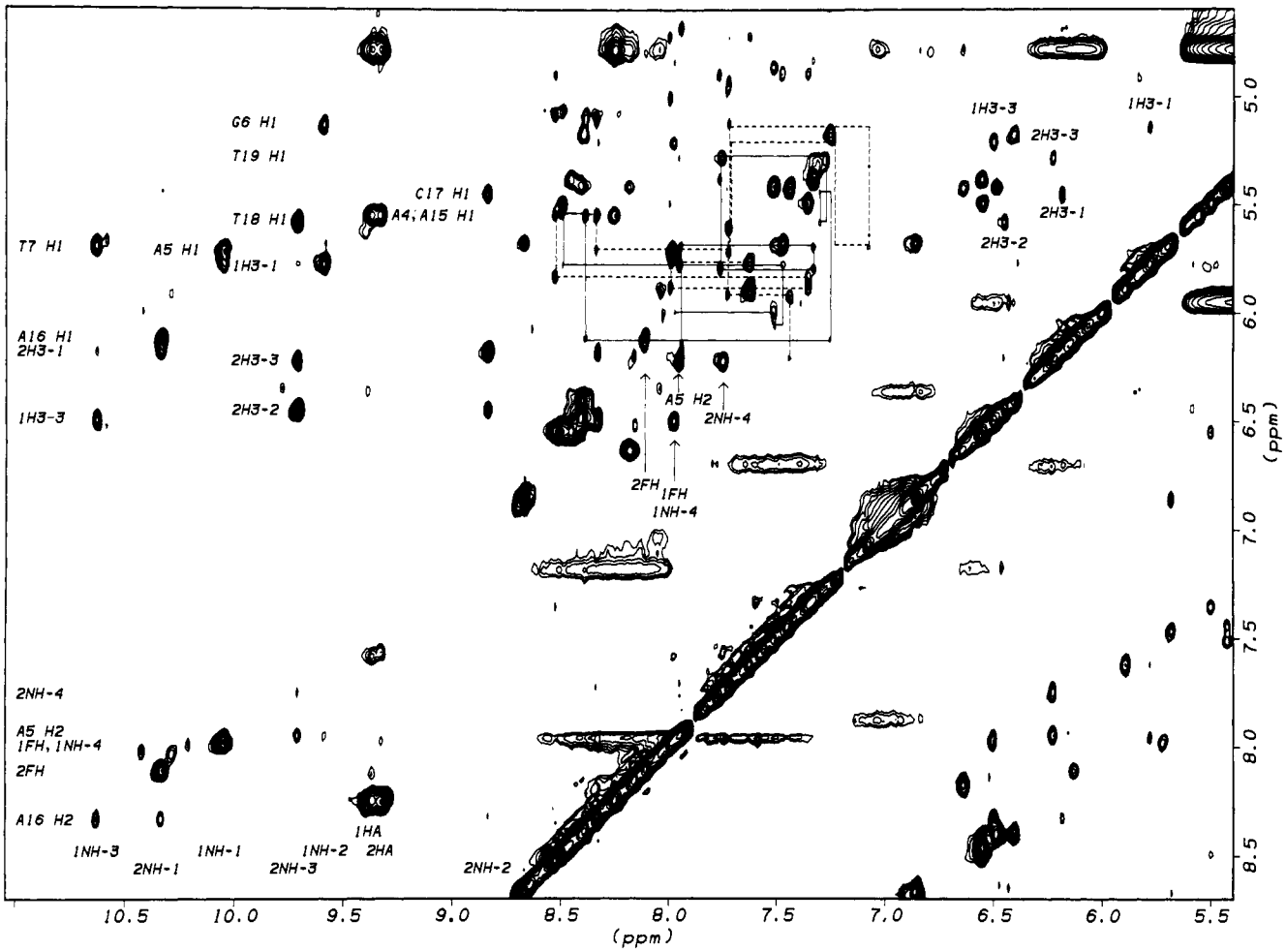


Figure 4. Expansion of the aromatic and amide region of a NOESY spectrum of the 1:1:1 complex of 2-ImD and distamycin (Dst) with d(CGCAAGTTGGC):d(GCCAACTTGCG) (in 90% H_2O/10% D_2O, 25 °C, τ_{mix} = 200 ms). Sequential aromatic to C1′H connectivities for the AAGTT strand are denoted by dashed lines, while for the AACTT strand they are shown as solid lines. Ligand–ligand and ligand–DNA crosspeaks are labeled according to structures **1** and **3**. Protons of the 2-ImD ligand and the distamycin ligand are marked with 1 and 2, respectively, in front of the denotation. Labels below or above crosspeaks denote the chemical shift along the ω_2 (horizontal) axis, while labels to the left or right of peaks indicate the chemical shift along the ω_1 (vertical) axis.

Table III. Ligand–DNA and Ligand–Ligand Intermolecular Contacts for the 1:1:1 2-ImD:Dst:DNA Complex with d(CGCAAGTTGGC):d(GCCAACTTGCG)[a]

2-ImD	DNA	Dst
	Ligand–DNA	
	A4 C1′H	HA
	A4 C2H	HA, C(18)H, C(19)H
	T19 C1′H	NH-4, H3-3
NH-1, HF	A5 C1′H	
H3-1	A5 C2H	H3-3
	T18 C1′H	NH-3, H3-2
	C17 C1′H	NH-2, H3-1
NH-2, H3-1	G6 C1′H	
NH-3	T7 C1′H	
	A16 C1′H	HF, NH-1
NH-3, H3-3	A16 C2H	NH-1, H3-1
NH-4, H3-3	T8 C1′H	
HA	A15 C1′H	
HA, C(18)H, C(19)H	A15 C2H	
	Ligand–Ligand	
C(19)H		HF
HF		C(19)H

[a] Identified in the H_2O NOESY acquired at 100-ms mixing time.

and hence the relative binding constants. With both ligands present in excess, the 1:1:1 2-ImD:Dst:DNA complex is by far most populated, with no 2:1 2-ImD:DNA complex and little 2:1 Dst:DNA complex detectable (Figure 3d). From these data we estimate that the relative binding constant of the 1:1:1 2-ImD:Dst:DNA complex exceeds that of the 2:1 Dst:DNA complex by a factor of about 15. No estimate can be given for the relative binding affinities for the 2:1 Dst:DNA and the 2:1 2-ImD:DNA complexes. On the basis of the detection of the 2:1 Dst:DNA complex but not of the 2:1 2-ImD:DNA complex in the presence of excess ligand, the binding constant for the 2:1 Dst:DNA complex is somewhat higher than that for the 2:1 2-ImD:DNA complex.

A concern not addressed by the NMR titration studies is the absolute binding affinities in these complexes. Recent calorimetric studies have shown that the binding affinity of 2-ImD at the AAGTT:AACTT site is at least as high as that of distamycin at the AAATT:AATTT site. This suggests that the absolute binding constants of the 1:1:1 2-ImD:Dst:DNA complex and the 2:1 Dst:DNA complex with AAGTT:AACTT are comparable to or higher than that of distamycin at the AAATT:AATTT site. Details of the calorimetric studies will be presented elsewhere.[16]

The binding site of the 1:1:1 2-ImD:Dst:DNA complex with AAGTT:AACTT is clearly defined by the intermolecular ligand–DNA NOE contacts summarized in Table III and Figure 5. The pyrrole–imidazole–pyrrole ring system of the 2-ImD ligand is oriented with its formyl group on the 5′ end of the AAGTT strand and extends in the minor groove from A5 to T8. The pyrrole–pyrrole–pyrrole ring system of distamycin spans the same A5-G6-T7 sequence as 2-ImD although with the formyl group pointing to the 5′ end of the AACTT strand. Intermolecular contacts between the methylene protons on C19 of each ligand and the formyl proton on the opposite ligand confirm the relative

(16) Marky, L.; Rentzeperis, D.; Geierstanger, B.; Dwyer, T.; Wemmer, D. In preparation.

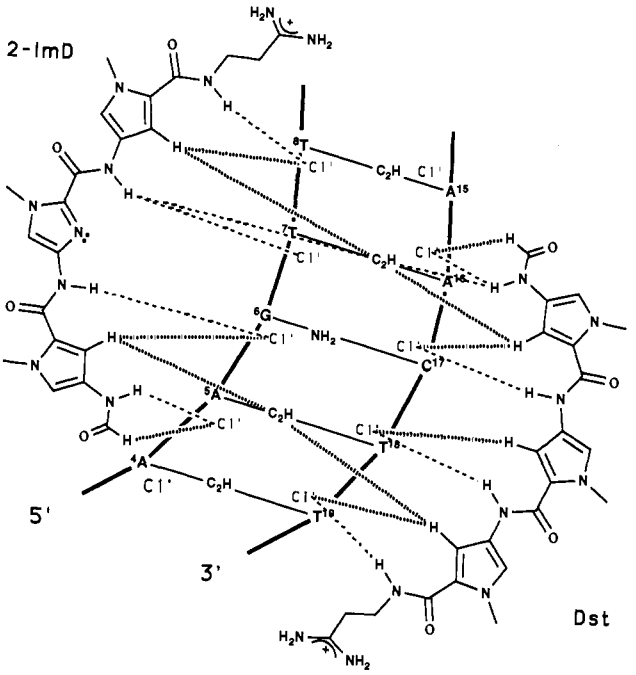


Figure 5. Intermolecular contacts between the 2-ImD and the distamycin (Dst) ligand of the 1:1:1 2-ImD:Dst:DNA complex with d(CGCAAGTTGGC):d(GCCAACTTGCG). NOE contacts between amide protons of the ligands and DNA protons are indicated as dashed lines, while contacts between other ligand protons and DNA are shown as stippled lines. Ligand–DNA contacts involving the amidinium and methylene protons as well as ligand–ligand contacts are not included.

Table IV. Ligand–DNA and Ligand–Ligand Intermolecular Contacts for the 2:1 Dst:DNA Complex with d(CGCAAGTTGGC):d(GCCAACTTGCG)[a]

Dst 1	DNA	Dst 2
	Ligand–DNA	
	A4 C1′H	HA
	A4 C2H	HA, C(18)H, C(19)H
	T19 C1′H	NH-4, H3-3
NH-1, HF	A5 C1′H	
NH-1, H3-1A5 C2H	NH-3, H3-3	
	T18 C1′H	NH-3, H3-2
	C17 C1′H	NH-2, H3-1
NH-2, H3-1	G6 C1′H	
NH-3, H3-2T7 C1′H		
	A16 C1′H	NH-1, HF
NH-3, H3-3	A16 C2H	NH-1, H3-1
NH-4, H3-3	T8 C1′H	
HA	A15 C1′H	
HA, C(18)H, C(19)H	A15 C2H	
	Ligand–Ligand	
C(19)H		HF
HF		C(19)H

[a] Identified in the H_2O NOESY acquired at 100-ms mixing time.

orientations of the ligands. The two ligands occupy the minor groove of the central portion of the AAGTT:AACTT oligomer similar to the case of the 2:1 2-ImD:DNA complex characterized previously.[6] Importantly, no exchange of position was observed between the 2-ImD and the distamycin ligand. Interactions between the imidazole nitrogen of 2-ImD and the amino group guanine G6, therefore, determine the ligand arrangement in the complex. The guanine G6 imino resonance line shifts upfield by about 0.3 ppm in the 1:1:1 2-ImD:Dst:DNA complex relative to the 2:1 Dst:DNA complex (Figure 3) (compared to 0.4 ppm in the 2:1 2-ImD:DNA complex). While all other guanine imino resonances are within 0.03 ppm in the three complexes, the shift of the guanine G6 imino line points to strong interactions between the 2-ImD imidazole nitrogens and the guanine G6 amino group in the center of the binding site. Because there is no ligand hydrogen bond acceptor close to the G6 amino group in the 2:1

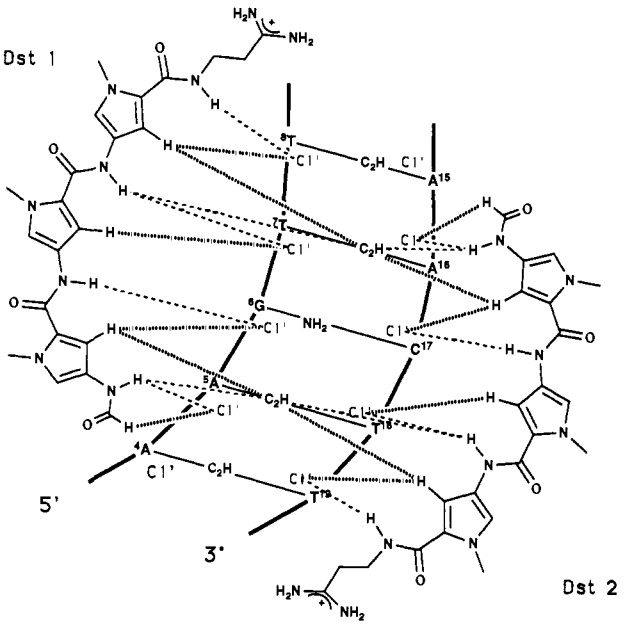


Figure 6. Intermolecular contacts between distamycin (Dst) ligands of the 2:1 Dst:DNA complex with d(CGCAAGTTGGC):d(GCCAACTTGCG). NOE contacts between amide protons of the ligands and DNA protons are indicated as dashed lines, while contacts between other ligand protons and DNA are shown as stippled lines. Ligand–DNA contacts involving the amidinium and methylene protons as well as ligand–ligand contacts are not included.

Dst:DNA complex and, according to molecular modeling, the ligands in all three complexes appear to be positioned similarly, we suggest that the characteristic upfield shift of the G6 imino resonance is not caused by ring current effects.

Interestingly, NOE contacts between any ligand protons and the amino protons of G6 could not be observed in the 1:1:1 2-ImD:Dst:DNA complex. On the basis of the modeling results one would expect NOEs to H3-2 and NH-3 of the 2-ImD ligand. Due to line broadening caused by rotation of the amino group, guanine amino protons are usually not observed in NMR spectra of uncomplexed oligomers.[8,9] NMR characterization of a similar 2:1 complex with DNA formed by the imidazole–pyrrole–pyrrole ring system 2-ImN revealed the presence of NOEs between the guanine amino groups and the H4 proton adjacent to the imidazole nitrogen.[5a] This was interpreted as evidence for the formation of a hydrogen bond between the guanine amino protons and the imidazole nitrogen responsible for sequence specificity of this complex. In contrast, similar NOEs involving guanine G6 amino protons were not observed in the 2:1 2-ImD:DNA complex with AAGTT:AACTT although interactions of the nitrogen of the central imidazole ring with the guanine amino protons also determine specific binding.[6] There are two possible explanations for these observations. First, since the imidazole ring of 2-ImN is at the end of the molecule, this ligand might be able to adopt a position optimal for hydrogen bond formation more easily, resulting in a significant slowing of the rotation of the guanine amino group. Alternatively, in the 2:1 2-ImD:DNA complex, two imidazole nitrogens are in the vicinity of the guanine amino group, possibly "competing" for hydrogen bond formation. In the 2:1 2-ImN:DNA complex,[5a] each imidazole nitrogen of the two ligands bonds with a separate guanine amino group. On the basis of this last argument, we had expected to observe NOEs to the guanine G6 amino protons in the 1:1:1 2-ImD:Dst:DNA complex in which binding affinity and specificity is enhanced as only one single hydrogen bond acceptor per guanine amino group is positioned for complexation. The absence of NOE peaks to the guanine G6 amino protons even in the 1:1:1 2-ImD:Dst:DNA complex is probably due to the exact position of the imidazole nitrogen with respect to the guanine amino protons. However, the molecular model of the 1:1:1 2-ImD:Dst:DNA complex

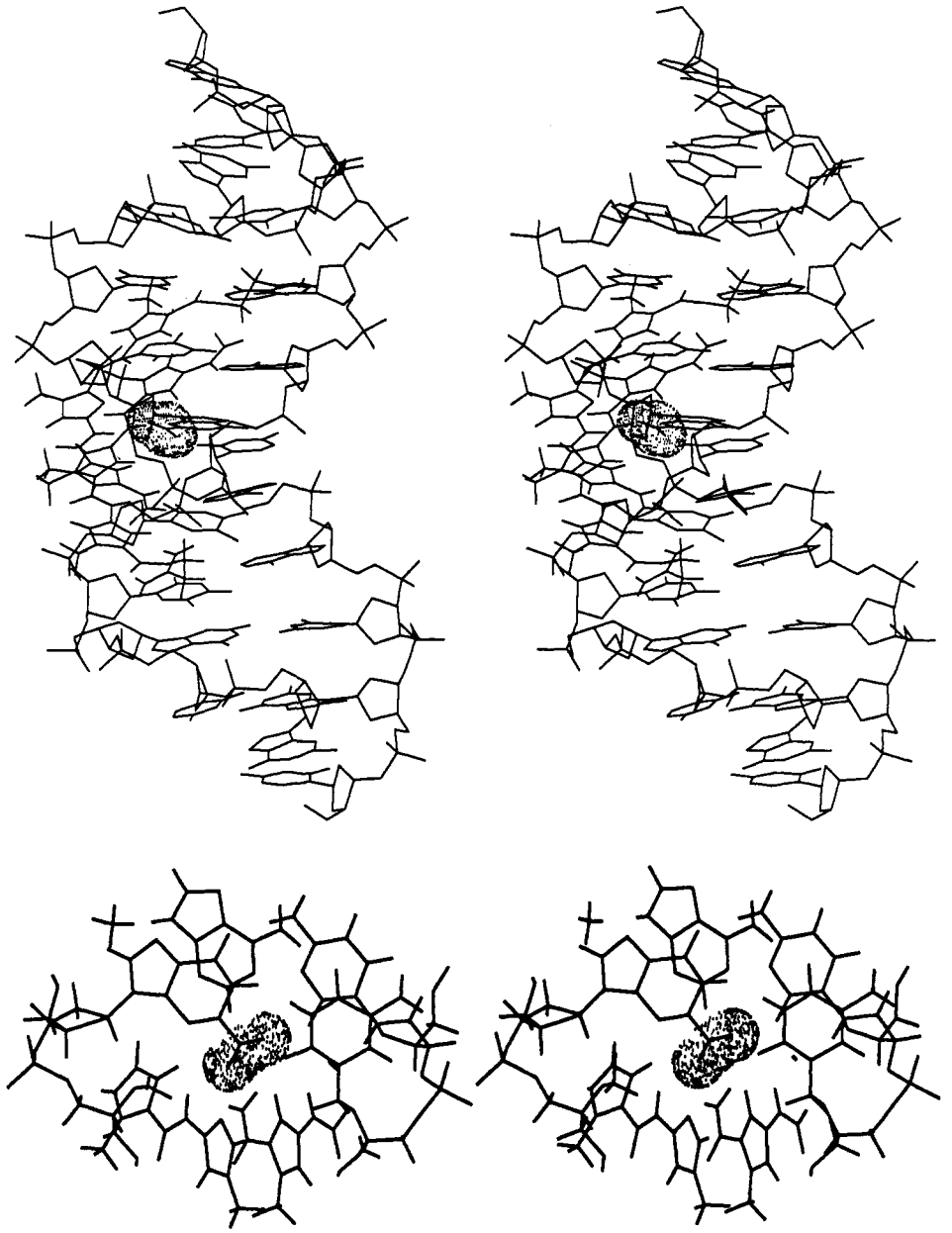


Figure 7. Stereo diagram of the molecular model of the 1:1:1 2-ImD:Dst:DNA complex with d(CGCAAGTTGGC):d(GCCAACTTGCG) obtained by energy minimization with semiquantitative distance restraints from NOESY spectra (see text). (a, top) Side view. For clarity, hydrogen atoms are omitted from the DNA model but not from the ligand molecules. (b, bottom) View approximately along the helical axis. The central G6-C17 base pair is shown above the A5-T18 base pair. All atoms are displayed. The 2-ImD molecule lies to the left of the Dst ligand. In both views the amino group of the guanine base in the 5′-AAGTT-3′ recognition site is highlighted as a van der Waals surface. Note that the imidazole nitrogen of the 2-ImD ligand points toward the amino group of the guanine base.

(Figure 7) is not of sufficiently high resolution to confirm this explanation conclusively.

Prior to this work, distamycin binding to GC-containing sequences was thought to be disfavored due to steric interference between the H3 protons on the pyrrole ring of distamycin and the guanine amino group at the bottom of the minor groove.[2] However, we observe that distamycin forms a tight, slowly exchanging 2:1 complex with AAGTT:AACTT (Figure 2). The complex is very similar to the distamycin complexes seen with AAATT:AATTT[7a] but forms with higher cooperativity. On the basis of intermolecular NOE contacts (Table IV) and molecular modeling, the placement and orientation of the ligands are virtually identical to those found in the 2:1 2-ImD:DNA[6] and the 1:1:1 2-ImD:Dst:DNA complexes with AAGTT:AACTT. In fact, preliminary modeling studies indicate that the minor groove of the AAGTT:AACTT binding site is sufficiently spacious to allow the positioning of the distamycin ligands in a way that steric clashes between the ligand H3 protons and the guanine amino group can be avoided. A more detailed analysis, however, must be postponed until the structure of the 2:1 Dst:DNA complex is determined at higher resolution.[15] Furthermore, the binding of 2 equiv of distamycin to a TACGTA:TACGTA site[17] was reported recently. NMR results[18] confirm that distamycin binding to TACGTA:TACGTA occurs as a 2:1 Dst:DNA complex similar to the one observed with AAGTT:AACTT. The binding affinities of distamycin to GC-containing sequences, however, must be measured accurately to allow a comparison with the conventional all-AT distamycin binding sites and to evaluate the significance of these observations.

In summary, the formation of a heterocomplex of 2-ImD and distamycin with AAGTT:AACTT is highly favored over the formation of either 2:1 2-ImD:DNA or 2:1 Dst:DNA complex. Since both the binding site and the ligand complex are symmetric with respect to the central G-C base pair and the 2-ImD ligand lies along the AAGTT strand but not the AACTT strand, the

(17) Capobianco, M. L.; Colonna, F. P.; Forni, A.; Garbesi, A.; Iotti, S.; Moretti, I.; Samori, B.; Tondelli, L. *Nucleic Acids Res.* **1991**, *19*, 1695.
(18) Geierstanger, B. H. Unpublished results.

1:1:1 2-ImD:Dst:DNA complex is capable of distinguishing a G-C from a C-G base pair in the minor groove. This clearly points to the importance of specific interactions between the hydrogen-bond-accepting imidazole nitrogen of 2-ImD and the guanine amino protons at the bottom of the minor groove. Placing a single hydrogen bond acceptor per guanine amino group of the binding site has a dramatic effect on the binding constant of the pyrrole–imidazole–pyrrole ring system 2-ImD and, therefore, provides a new guideline in the design of sequence-specific minor-groove binding molecules. Furthermore, this study suggests that different distamycin-like ligands can be combined using the 2:1 binding mode as a general platform, to expand the range of DNA sequences which can be specifically recognized.

Acknowledgment. This work was supported by NIH Grant GM-43129 (to D.E.W.) and by grants (to J.W.L.) from the Natural Sciences and Engineering Research Council of Canada and through instrumentation grants from the U.S. Department of Energy (DE FG05-86ER75281) and the National Science Foundation (DMB 86-09305 and BBS 87-20134).

Supplementary Material Available: Listings of the intermolecular ligand–DNA and ligand–ligand restraints and the achieved distances (2 pages). Ordering information is given on any current masthead page.