Structural and Dynamic Characterization of the Heterodimeric and Homodimeric Complexes of Distamycin and 1-Methylimidazole-2-carboxamide–Netropsin Bound to the Minor Groove of DNA†

Bernhard H. Geierstanger,‡ Jens Peter Jacobsen,§,‖ Milan Mrksich,⊥ Peter B. Dervan,*,⊥ and David E. Wemmer*,§

Graduate Group in Biophysics and Department of Chemistry, University of California, Berkeley, California 94720, and Arnold and Mabel Beckman Laboratories of Chemical Synthesis, California Institute of Technology, Pasadena, California 91125

Received October 20, 1993; Revised Manuscript Received December 14, 1993®

ABSTRACT: NMR spectroscopy combined with molecular modeling was used to characterize a heterodimeric complex with Dst and 2-ImN bound in the minor groove of d(GCCTAACAAGG)·d(CCTTGTTAGGC) (1:1:1 2-ImN·Dst·DNA complex). The imidazole–pyrrole–pyrrole ligand 2-ImN spans 5′-GTTA-3′ of the TAACA·TGTTA binding site with the imidazole nitrogen specifically recognizing the guanine amino group. The Dst ligand lies along the 5′-AACA-3′ sequence and complements the 2-ImN ligand in the formation of the antiparallel side-by-side heterodimeric complex. Titrations of the same site with Dst or 2-ImN alone yield homodimeric complexes (2:1 ligand·DNA) of lower stability than the 1:1:1 2-ImN·Dst·DNA complex. Dst and 2-ImN binding to d(CGCAAACTGGC)·d(GCCAGTTTGCG) was also investigated. The 1:1:1 2-ImN·Dst·DNA complex is again the most stable complex with the AAACT·AGTTT site and is similar to the TAACA·TGTTA complex. No monomeric binding of either 2-ImN or Dst was observed to either site.

The natural products netropsin and distamycin A (Dst) (**1**, Chart 1) are di- and tripeptides that bind in the minor groove of DNA at sites of four or five successive AT base pairs, respectively (Schultz et al., 1982; Schultz & Dervan, 1984; Taylor et al., 1984; Zimmer & Wähnert, 1986; Dervan, 1986). Investigations of 1:1 peptide·DNA complexes by X-ray crystallography and NMR spectroscopy reveal that hydrogen bonding, van der Waals contacts, and electrostatics all contribute to the binding affinity and specificity (Kopka et al., 1985a,b; Coll et al., 1987; Klevit et al., 1986; Pelton & Wemmer, 1988). These 1:1 peptide·DNA complexes proved useful in designing oligo(pyrrolecarboxamides) for recognition of longer AT tracts of DNA (Schultz & Dervan, 1983; Youngquist & Dervan, 1985a,b). Recent two-dimensional NMR studies have demonstrated that distamycin is also capable of binding in the minor groove of five base pair AT-rich DNA sites as a 2:1 peptide·DNA complex (Pelton & Wemmer, 1989; 1990). This dimeric motif, wherein two peptides bind side-by-side and antiparallel in the minor groove, has broadened the structural repertoire for design considerations of sequence-specific minor groove binding molecules.

Recognition of Mixed AT and GC Sequences. Footprinting and affinity cleaving experiments have revealed that the designed peptide 1-methylimidazole-2-carboxamide–netropsin (2-ImN) (**2**) specifically binds 5′-TGTCA-3′ sequences with unbiased orientation preferences, inconsistent with a 1:1 complex of peptide with DNA (Wade & Dervan, 1987; Wade et al., 1992). Direct characterization of this complex with5′-TGACT-3′ by NMR demonstrates that the peptides bind as a side-by-side antiparallel dimer, very similar in structure to the dimeric complexes of distamycin with DNA (Mrksich et al., 1992). In the 2:1 2-ImN·DNA complex the imidazole nitrogen of each ligand targets a guanine amino group. The observation that each peptide makes specific hydrogen bonds to a single strand in the minor groove suggests that a *heterodimer* consisting of the peptides distamycin and 2-ImN might specifically bind 5′-(A,T)G(A,T)$_3$-3′ sequences. Footprinting and affinity cleaving experiments have demonstrated that distamycin and 2-ImN simultaneously bind a 5′-TGTTA-3′ sequence with opposite orientations, consistent with a side-by-side heterodimeric complex (Mrksich & Dervan, 1993a). We describe here direct structural and dynamic studies by NMR spectroscopy on the heterodimeric and homodimeric complexes formed upon binding of Dst and 2-ImN to the oligonucleotide duplexes d(GCCTAACAAGG)·d(CCTTGTTAGGC) and d(CGCAAACTGGC)·d(GCCAGTTTGCG). These data are relevant to a recent NMR study of a different heterodimeric complex wherein Dst and a distamycin analog containing a central imidazole (2-ImD) (**3**) bind a 5′-AAGTT-3′ sequence (Geierstanger et al., 1993).

MATERIALS AND METHODS

Synthesis of Ligands and Oligonucleotides. 2-ImN was synthesized as described previously (Wade et al., 1992). Dst was purchased from Serva and used without further purification. The oligomers d(GCCTAACAAGG), d(CCTTGTTAGGC), d(CGCAAACTGGC), and d(GCCAGTTTGCG) were synthesized and purified as reported previously (Pelton & Wemmer, 1990).

Sample Preparation. NMR samples were prepared by dissolving the undecamer oligonucleotide duplexes in 0.5 mL of 10 mM sodium phosphate buffer (pH 7.0) and then lyophilizing to dryness. For experiments carried out in D_2O

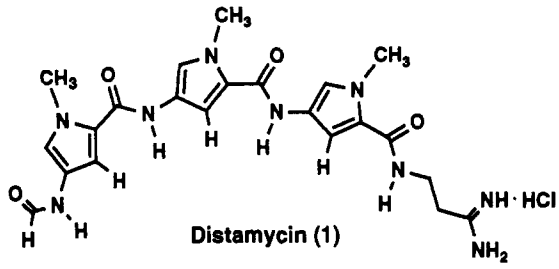
†We are grateful to the National Institutes of Health (GM-27681 to P.B.D. and GM-43129 to D.E.W.) and the National Foundation for Cancer Research for research support and for a National Institutes of Health Research Service Award to M.M. and to the U.S. Department of Energy (DE FG05-86ER75281) and the Natonal Science Foundation (DMB 86-09305 and BBS 87-20134) for instrumentation grants.

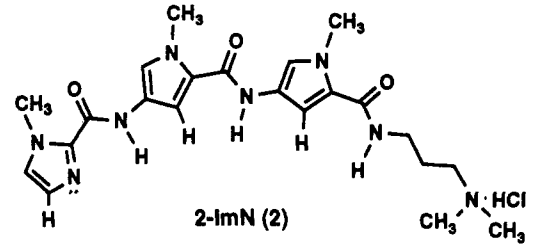
* To whom correspondence should be addressed.
‡ Graduate Group in Biophysics, University of California, Berkeley.
§ Department of Chemistry, University of California, Berkeley.
‖ On leave from the Department of Chemistry, Odense University, 5230 Odense M, Denmark.
⊥ California Institute of Technology.

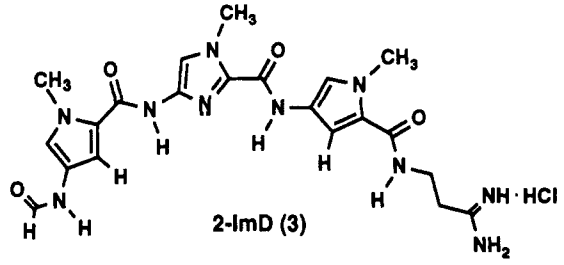
® Abstract published in *Advance ACS Abstracts,* February 15, 1994.

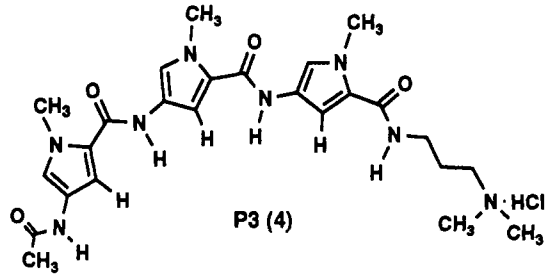
0006-2960/94/0433-3055$04.50/0

Chart 1









the solid was redissolved in 0.5 mL of 99.96% D_2O (Cambridge Isotope Laboratories), while for experiments in H_2O a 90% H_2O/10% D_2O mixture (0.5 mL) was used. Solid 2-ImN·HCl was dissolved in 99.96% D_2O yielding stock solutions of 21.9 and 28.6 mM as determined by UV absorbance at 304 nm ($\epsilon = 3.5 \times 10^4$ M^{-1} cm^{-1}). Similarly, stock solutions of Dst in 99.96% D_2O of 18.6 and 18.8 mM were prepared ($\epsilon = 3.4 \times 10^4$ M^{-1} cm^{-1} at 304 nm). The ligand stock solutions were stored at −70 °C. Extinction coefficients for d(GCCTAACAAGG), d(CCTTGTTAGGC), d(CGCAAACTGGC), and d(GCCAGTTTGCG) were calculated to be 1.11×10^5, 9.79×10^4, 1.02×10^5, and 9.99×10^4 M^{-1} cm^{-1}, respectively (Warshaw & Cantor, 1970). The concentrations of the all double-stranded DNA samples were determined to be 1 mM by UV absorbance at 260 nm and 80 °C.

NMR Experiments and Signal Assignments. NMR experiments were performed at 600 MHz on a Bruker AMX-600 spectrometer unless indicated otherwise. 2-ImN·HCl and Dst were titrated into the NMR sample containing duplex DNA in approximately 0.25 mol equiv per addition. 1D spectra in D_2O were acquired with 4096 complex points over a spectral width of 6024 Hz. The signal was averaged over 128 scans, and a presaturation pulse was applied to suppress the residual HDO resonance. 2D NOESY spectra in D_2O were collected with 1024 complex points in t_2 using a spectral width of 6024 Hz and a mixing time of 200 ms. A total of 483–512 t_1 experiments were recorded and zero-filled to 1K. For each t_1 value 48–64 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. 2D NOESY spectra in water were acquired at 25 °C, replacing the last 90° pulse by a 1–1 jump and return sequence (Sklenàr et al., 1987) to suppress the solvent signal as described previously (Mrksich et al., 1992). The spectra were collected into 2048 complex points in t_2 using a spectral width of 13514 Hz and mixing times of 100 and 200 ms. A total of 533 or 553 t_1 experiments with 48 or 64 scans were recorded and zero-filled to 2K. All 2D NOESY spectra were acquired using TPPI. The data were processed with FELIX (Hare Research) on a Silicon Graphics IRIS/4D workstation. Skewed sine bell functions were used for apodization of the free induction decays. 2D NOESY spectra of the ligand·DNA complex in D_2O and H_2O enabled the assignment of the DNA resonances using standard sequential methods (Hare et al., 1983; Wüthrich, 1986). Ligand resonances were assigned as previously described (Mrksich et al., 1992; Dwyer et al., 1992; Pelton & Wemmer, 1989; Pelton & Wemmer, 1990).

Distance Restraints. A model of the 1:1:1 2-ImN·Dst·DNA heterocomplex with d(GCCTAACAAGG)·d(CCTTGTTAGGC) was obtained based on the NMR data. Intermolecular distance restraints were generated from the volume integrals of the cross peaks in the H_2O NOESY spectrum acquired at a mixing time of 100 ms as described previously (Mrksich et al., 1992; Dwyer et al., 1992). The cross peak volumes were classified semiquantitatively into three categories: strong (1.8–2.5 Å), medium (2.5–3.7 Å), or weak (3.7–4.2 Å) relative to the volume integrals of the cytosine H5–H6 cross peak volumes. In all, 31 intermolecular ligand–DNA restraints, 1 intermolecular ligand–ligand restraint, and 13 intramolecular restraints of the ligands were used (listings of the intermolecular ligand–DNA and ligand–ligand restraints are available as supplementary material).

Structure Refinement. The model of the d(GCCTAACAAGG)·d(CCTTGTTAGGC) duplex was constructed using the Biopolymer module of Insight II (Biosym) from standard B-form DNA. Coordinates for the ligand molecules were obtained from the NMR derived model of the 2:1 Dst·DNA complex with d(CGCAAGTTGGC)·d(GCCAACTTGCG) (Geierstanger et al., 1993) and the 2:1 2-ImN·DNA complex with d(GCATGACTCGG)·d(CCGAGTCATGC) (Mrksich et al., 1992). One Dst and one 2-ImN molecule were roughly arranged in the head-to-tail orientation found in previous 2:1 ligand·DNA complexes and manually docked into the minor groove with the help of the Builder module of Insight II on the Silicon Graphics workstation. Energy minimizations were performed using the Discover module of Insight II (employing the AMBER force field). Hydrogen bonds for standard Watson–Crick base pairing were included as NOE restraints using a force constant of 200 (kcal/mol)/Å^2 while a force constant of 25 (kcal/mol)/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 rms derivative of <0.001 (kcal/mol)/Å^2 was achieved in fewer than 20 000 steps of conjugate gradient minimization.

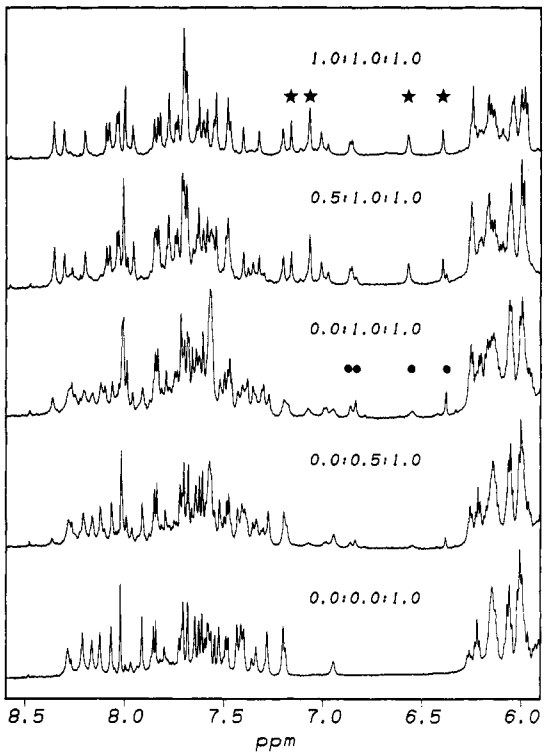


FIGURE 1: Aromatic region of ^{1}H NMR spectra acquired at several points in a titration of d(GCCTAACAAGG)·d(CCTTGTTAGGC) with Dst and 2-ImN at 25 °C (in D_2O). The molar 2-ImN·Dst·DNA ratios are indicated for spectrum a (bottom) through e (top). Ligand resonances of the 2:1 Dst·DNA complex are indicated by dots while that of the 1:1:1 2-ImN·Dst·DNA complex are shown as stars.

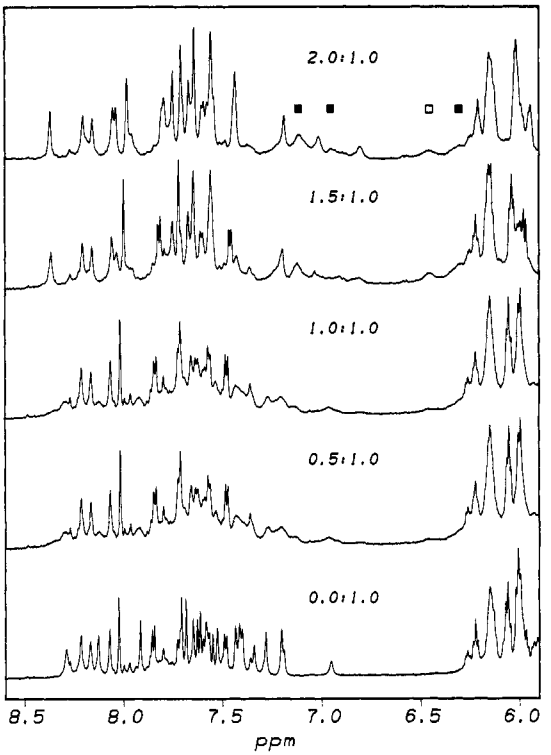


FIGURE 2: Aromatic region of ^{1}H NMR spectra acquired at several points in a titration of d(GCCTAACAAGG)·d(CCTTGTTAGGC) with 2-ImN at 25 °C (in D_2O). The molar 2-ImN·DNA ratios are indicated for each spectrum. Ligand resonances of the 2:1 2-ImN·DNA complex are indicated by squares. The ligand resonance (2H3-2) used for line width analysis at different temperatures is indicated as open square.

RESULTS

Characterization of the Complexes Formed by 2-ImN and Dst with d(GCCTAACAAGG)·d(CCTTGTTAGGC)

NMR Titrations. The 1D NMR spectra recorded at 25 °C during the titrations of the duplex d(GCCTAACAAGG)·d(CCTTGTTAGGC) with Dst and 2-ImN are shown in Figures 1, 2, and 3. The NMR spectra of the DNA duplex titrated with both Dst and 2-ImN show clearly the existence of well-defined complexes at all stoichiometries (Figure 1). Up to a molar 2-ImN·Dst·DNA ratio of 0:1:1 broad lines corresponding to an exchanging 2:1 Dst·DNA homocomplex appear (Figure 1a–c). Addition of 2-ImN results in a new set of sharp resonances while the intensities of the 2:1 Dst·DNA complex as well as those of free DNA duplex are reduced (Figure 1d,e). Only one complex is observed at the end of the titration consistent with the exclusive formation of a 1:1:1 2-ImN·Dst·DNA heterocomplex. The narrow width of the lines from the heterocomplex throughout the second half of the titration indicates that the ligands exchange slowly on the NMR time scale.

Resonances of the ligands and the DNA binding sites in the titration with solely 2-ImN (Figure 2) or Dst (Figure 3) are far broader than the corresponding lines in the 1:1:1 2-ImN·Dst·DNA heterocomplex. In both cases the line broadening is due to chemical exchange of the ligands between different sites as discussed in a later section. On the basis of 2D NOESY spectra in D_2O, both 2-ImN and Dst form 2:1 ligand·DNA homocomplexes with the TAACA·TGTTA site. Throughout all titrations, no evidence of any 1:1 ligand·DNA complex was observed.

Competition Experiments. Both samples containing 2:1 ligand·DNA homocomplexes were further titrated with the absent ligand. For each sample resonances from the heterocomplex and free excess ligand emerged when both ligands were present in a concentration at least equal that of DNA.

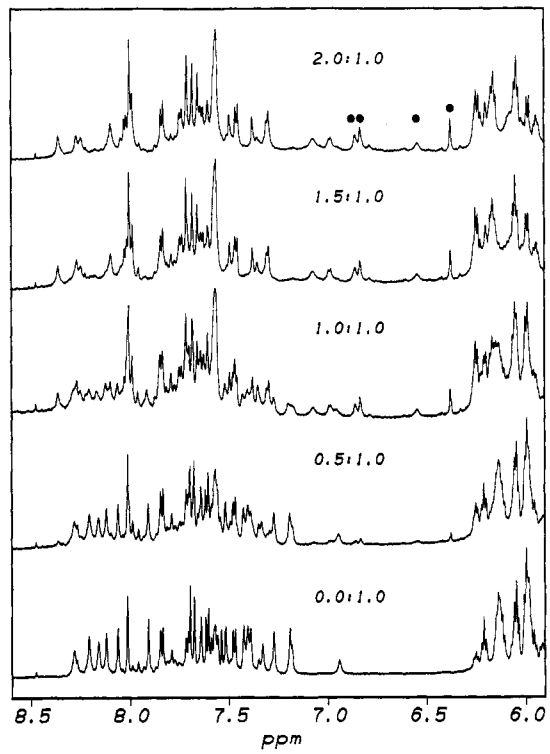


FIGURE 3: Aromatic region of ^{1}H NMR spectra acquired at several points in a titration of d(GCCTAACAAGG)·d(CCTTGTTAGGC) with Dst at 25 °C (in D_2O). The molar Dst·DNA ratios are indicated for each spectrum. Ligand resonances of the 2:1 Dst·DNA complex are indicated by dots.

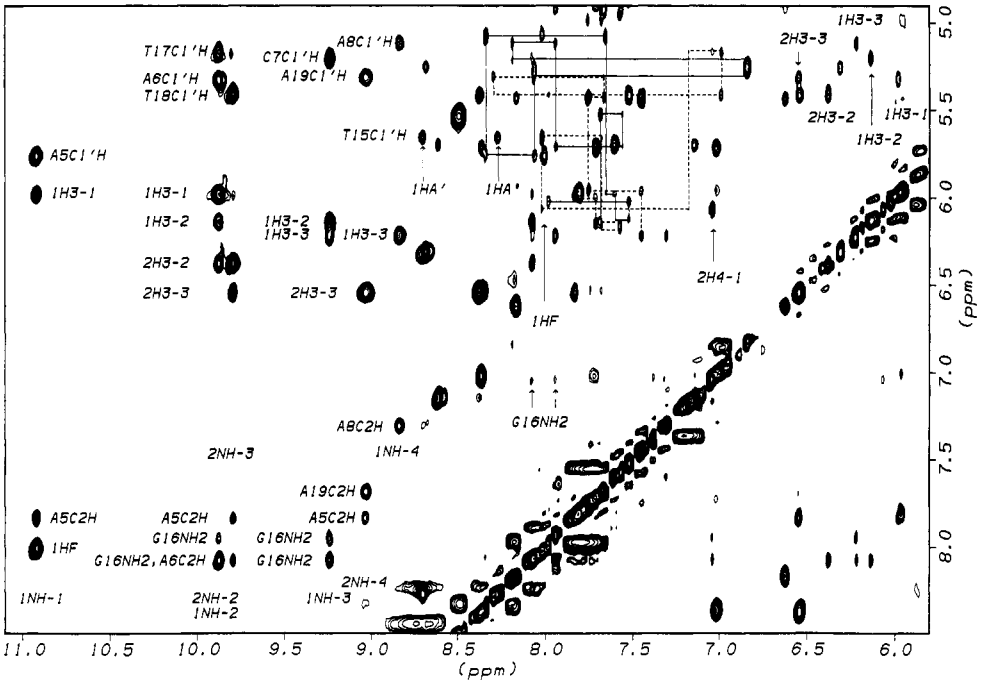


FIGURE 4: Expansion of the aromatic and amide region of a NOESY spectrum of the 1:1:1 complex of 2-ImN and Dst with d(GCCTAACAAGG)·d(CCTTGTTAGGC) (in 90% H_2O/10% D_2O, 25 °C, τ_{mix} = 200 ms). Sequential aromatic to C1'H connectivities for the 5'-TAACAA-3' strand are shown as solid lines; those for the 5'-TTGTTA-3' strand are shown as dashed lines. Labels below or above a cross peaks denote the chemical shift along the ω_2 (horizontal) axis, while labels to the left or right of a peak indicate the chemical shift along the ω_1 (vertical) axis. Protons of Dst (ligand 1) and 2-ImN (ligand 2) are labeled according to Figure 5 with the ligand number in front of the denotions. HA denotes the amidinium protons of Dst.

At a 2-ImN·Dst·DNA ratio of approximately 2:2:1, the 1:1:1 2-ImN·Dst·DNA heterocomplex is the dominant complex in solution (data not shown).

Characterization of the 1:1:1 2-ImN·Dst·DNA Heterocomplex. Part of a NOESY spectrum of the 1:1:1 2-ImN·Dst·DNA heterocomplex in H_2O is shown in Figure 4. The chemical shifts of DNA H1' and H6/H8 protons and of the ligand resonances are listed in Tables 1 and 2. Numerous intermolecular contacts allow the accurate placement of the ligands in the minor groove of the TAACA·TGTTA duplex (Table 3 and Figure 5). The NOESY cross peaks clearly define the two ligand molecules as being stacked on top of each other in an antiparallel fashion. Cross peaks between DNA protons and the pyrrole protons H3-2 and H3-3 as well as the ligand NH protons show that the imidazole–pyrrole–pyrrole ring system of 2-ImN spans the 5'-G16-T17-T18-A19-3' sequence. This is further supported by cross peaks between the imidazole H4-1 proton and C1'H and amino protons of G16. Likewise, the pyrrole–pyrrole–pyrrole ring system of Dst spans the 5'-A5-A6-C7-A8-3' sequence (Figure 6a). The antiparallel arrangement of the ligands is confirmed by ligand to ligand NOE cross peaks.

Characterization of the 2:1 2-ImN·DNA Homocomplex. The lines of the 2:1 2-ImN·DNA complex (Figure 2) sharpen as the temperature is lowered to 8 °C (not shown). 2D NOESY spectra in D_2O (at 8 °C) of the sample titrated with two equivalents of 2-ImN allowed the characterization of the 2:1 2-ImN·DNA complex. Intermolecular contacts and ligand arrangement are virtually identical to the 1:1:1 2-ImN·Dst·DNA heterocomplex. The line broadening observed at higher temperature is caused by the two 2-ImN ligands exchanging positions (Figure 6c) as inferred from exchange cross peaks. The rate of the exchange between the two 2-ImN molecules at the binding site was estimated from a series of 1D NMR spectra at different temperatures in the range between 7 and 25 °C. The contribution to line broadening from chemical exchange at the slow end of intermediate exchange is $\Delta\nu_{1/2} = (\tau_a{*}\pi)^{-1}$ (Harris, 1986) ($\nu_{1/2}$ is the line width at half-maximum, and τ_a is the lifetime of the complex). The exchange rate $k_r = \tau_a^{-1}$ obtained accordingly was 70 s^{-1} (±20 s^{-1}) at 25 °C. The width of the line from a 2-ImN pyrrole proton (H3-2; indicated by an open square in Figure 2) was used to calculate the rate constant at each temperature.

Table 1: Chemical Shift Assignments of the d(GCCTAACAAGG)·d(CCTTGTTAGGC) Duplex, Free and in the 1:1:1 2-ImN·Dst·DNA Complex[a]

	H6/H8			H1'		
	free duplex	1:1:1 complex	Δδ	free duplex	1:1:1 complex	Δδ
strand 1						
G1	8.01	7.99	−0.02	6.05	6.04	−0.01
C2	7.56	7.54	−0.02	6.14	6.13	−0.01
C3	7.58	7.62	+0.04	5.96	5.99	+0.03
T4	7.40	7.68	+0.28	5.61	5.09	−0.52
A5	8.28	8.35	+0.07	5.91	5.78	−0.13
A6	8.12	8.07	−0.05	6.06	5.34	−0.72
C7	7.18	6.85	−0.33	5.38	5.23	−0.15
A8	8.15	8.19	+0.04	5.77	5.13	−0.64
A9	8.06	7.95	−0.11	5.89	5.72	−0.17
G10	7.60	7.58	−0.02	5.60	5.54	−0.06
G11	7.80	7.70	−0.10	6.13	6.14	−0.01
strand 2						
C12	7.83	7.82	−0.01	6.05	6.03	−0.02
C13	7.71	7.73	+0.02	6.12	6.16	+0.04
T14	7.51	7.60	+0.09	6.13	6.19	+0.06
T15	7.42	7.70	+0.28	5.86	5.67	−0.19
G16	7.90	8.04	+0.14	5.99	6.09	+0.10
T17	7.27	7.19	−0.08	5.98	5.18	−0.80
T18	7.39	7.00	−0.39	5.69	5.42	−0.27
A19	8.20	8.30	−0.10	5.99	5.33	−0.66
G20	7.64	7.68	−0.04	5.59	5.44	−0.15
G21	7.67	7.77	−0.10	5.98	5.97	−0.01
C22	7.47	7.46	−0.01	6.21	6.23	+0.02

[a] Chemical shifts are given in ppm (±0.01 ppm) with the residual HDO signal referenced to 4.80 ppm (25 °C).

Table 2: Chemical Shift Assignments of the 2-ImN and the Distamycin (Dst) Ligand in the 1:1:1 2-ImN·Dst·DNA Complex with d(GCCTAACAAGG)·d(CCTTGTTAGGC)[a]

proton	2-ImN	Dst
H4-1	7.06	na
HF	na	8.02
NH-1	na	10.92
H3-1	na	5.99
NH-2	9.88	9.88
H3-2	6.39	6.16
NH-3	9.80	9.24
H3-3	6.56	6.24
NH-4	9.03	8.84
C(18)H[b]	1.89	2.53
	3.34	3.69
C(19)H[b]	1.13	1.66
	1.64	2.49
C(20)H[b]	2.62	na
	2.91	na
HA[b,c]	na	8.70
	na	8.28

[a] Chemical shifts are given in ppm (±0.01 ppm) with the residual HDO signal referenced to 4.80 ppm (25 °C). na = not applicable. [b] Not stereospecifically assigned. [c] HA = amidinium protons.

Table 3: Ligand–DNA and Ligand–Ligand Intermolecular Contacts for the 1:1:1 2-ImN·Dst·DNA Complex with d(GCCTAACAAGG)·d(CCTTGTTAGGC)[a]

2-ImN[b]	DNA	Dst[b]
	Ligand–DNA	
	T15 C1′H	HA′, HA″
	A8 C2H	NH-4, H3-3, C19 H1, C19 H2, C18 H1
	A8 C1′H	NH-4, H3-3
H4-1	G16 C1′H	
H4-1, NH-2	G16 NH_2	NH-3, H3-3
	C7 C1′H	NH-3, H3-2
NH-2	T17 C1′H	
NH-3, H3-2	A6 C2H	NH-2, H3-2, H3-1
	A6 C1′H	NH-2, H3-1
NH-3, H3-2	T18 C1′H	
NH-3, H3-3, NH-4	A5 C2H	NH-1, H3-1
	A5 C1′H	HF, NH-1
NH-4, H3-3	A19 C1′H	
NH-4, C19 H1, C18 H1	A19 C2H	
	Ligand–Ligand	
C19 H1, C19 H2		HF
H4-1		C19 H1, C19 H2

[a] Identified in the H_2O NOESY acquired at 100-ms mixing time. [b] C18, C19, and HA = amidinium protons not stereospecifically assigned.

The rates were then used to calculate the Arrhenius activation energy = $-R\,\partial \ln k_r/\partial T^{-1}$ (R = 8.314 J mol^{-1} K^{-1} is the gas constant, and T is the absolute temperature in Kelvin). The obtained value of 57 kJ/mol (±10 kJ/mol) is of the same magnitude as the free energy of binding of individual ligands of this type (Wade et al., 1993; L. Marky, D. Rentzeperis, B. Geierstanger, T. Dwyer, and D. Wemmer, to be submitted). Although the exchange of positions therefore probably involves free ligand as intermediate, no signals from free ligands are detected.

Characterization of the 2:1 Dst·DNA Homocomplex. The lines of the DNA sample titrated with only Dst (Figure 3) become broader at lower temperature but sharpen when the temperature is increased to 35 °C (not shown). 2D NOESY spectra in D_2O of the DNA titrated with two equivalents of Dst were obtained at both 25 and 35 °C. The cross peaks are distinctly narrower at the higher temperature in accordance with the fast exchange situation (Figure 6b). Intermolecular contacts between the two drug molecules clearly demonstrate an antiparallel side-by-side 2:1 ligand·DNA binding mode.

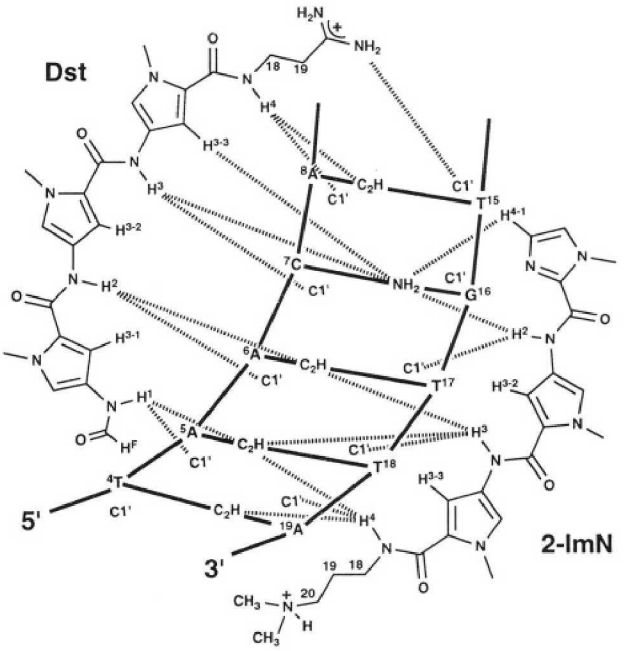


FIGURE 5: Schematic of selected intermolecular NOEs between Dst or 2-ImN and d(GCCTAACAAGG)·d(CCTTGTTAGGC) in the 1:1:1 2-ImN·Dst·DNA complex.

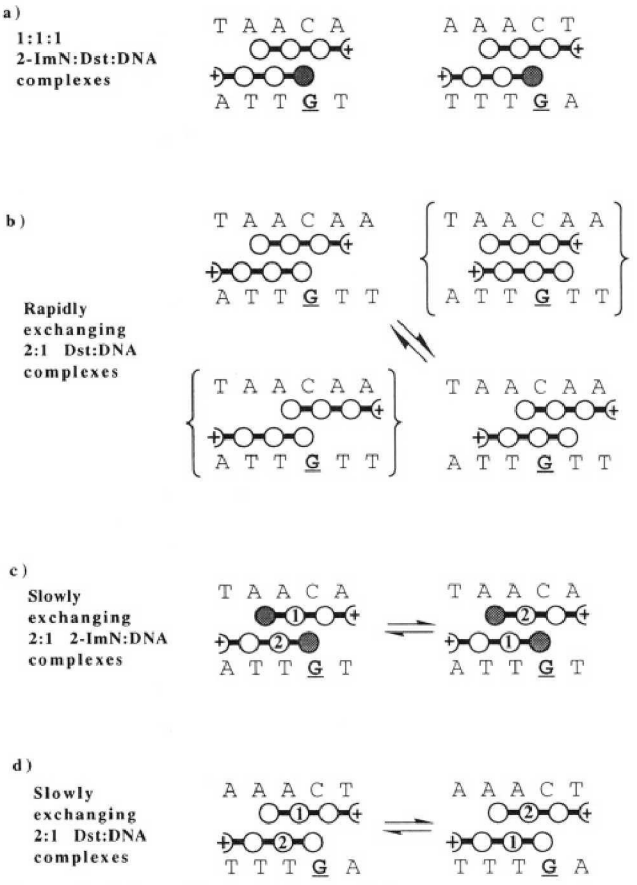


FIGURE 6: Schematic representation of the structure and exchange dynamics of the characterized ligand·DNA complexes. The shaded circle represents the imidazole ring of 2-ImN.

There is no indication in the NOESY spectra of any exchange between the two Dst molecules as in the case of the 2:1 2-ImN·DNA homocomplex. However, cross peaks between the H3 protons on the two Dst molecules and the H1′ of the nucleotides in the binding site show evidence of a sliding of the Dst molecules in the six base TAACAA·TTGTTA binding site (Figure 6b). The intensity of the cross peaks indicates

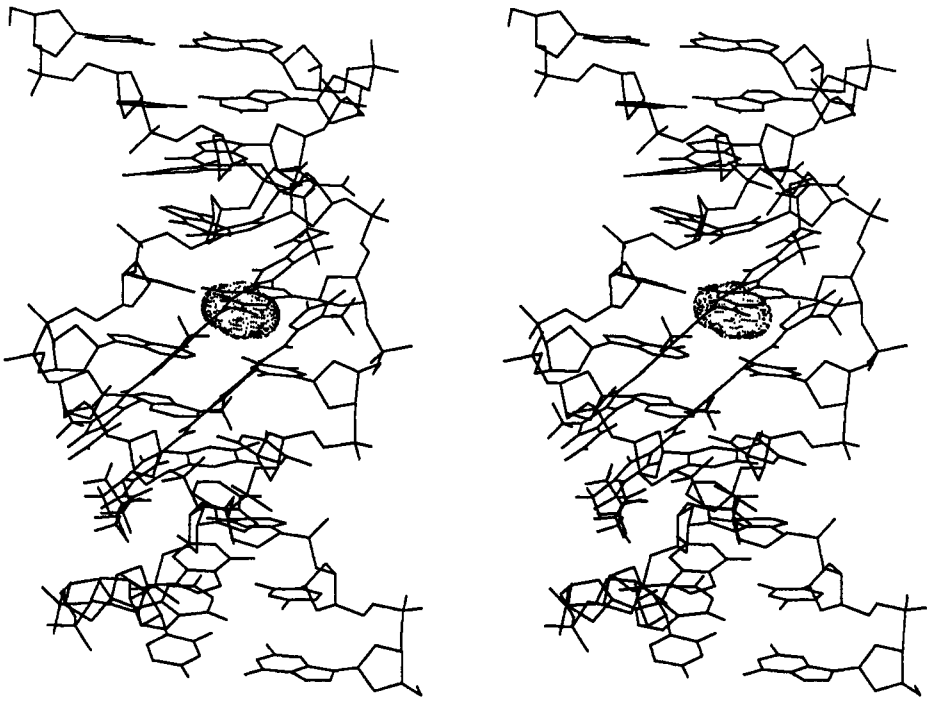


FIGURE 7: Molecular model of the 1:1:1 2-ImN·Dst·DNA complex with d(GCCTAACAAGG)·d(CCTTGTTAGGC) obtained by energy minimization using semiquantitative distance restraints derived from NOESY data. For clarity, hydrogen atoms are omitted for the DNA model but not for the ligand molecules. The guanine amino group recognized by the imidazole nitrogen of 2-ImN is highlighted as van der Waals surface.

that 2:1 ligand·DNA binding to the TAACA·TGTTA subsite is the preferred one, but fast exchange to the minor form binding in the AACAA·TTGTT site is evident. It is not possible from the data to determine whether the fast exchange of the Dst molecules occurs in a correlated sliding between two sites or whether the Dst molecules slide independently between four binding modes. However, the intensities of the observed cross peaks indicate that the major form has a side-by-side position of the two drug molecules equivalent to that of the 1:1:1 2-ImN·Dst·DNA heterocomplex (Figure 6a).

Molecular Modeling of the 1:1:1 2-ImN·Dst·DNA Heterocomplex. A total of 45 intramolecular ligand–ligand and intermolecular ligand–DNA and ligand–ligand distance restraints derived from NOE data were used to obtain the energy minimized model of the 1:1:1 2-ImN·Dst·DNA heterocomplex with TAACA·TGTTA (Figure 7). The antiparallel side-by-side arrangement of the ligand molecules is very similar to the 2:1 2-ImN·DNA complex with TGACT·AGTCA (Mrksich et al., 1992) and other previously characterized 2:1 ligand·DNA complexes (Pelton & Wemmer, 1989, 1990; Fagan & Wemmer, 1990; Dwyer et al., 1992; Geierstanger et al., 1993). The two ligands completely fill the minor groove, stacked on top of each other with the positively charged tail groups pointing in opposite directions. Hydrogen bonds were assigned by the Insight program between the Dst amide protons NH-1, NH-2, NH-3, and NH-4 with A5 N3, A6 N3, C7 O2, and A8 N3, respectively. Although the amidinium protons are likely involved in hydrogen bonds, ambiguities in specifying acceptors remain at the current accuracy of the model. Similar to Dst, the amide protons of 2-ImN complement the pattern of hydrogen bond acceptors on the DNA: NH-2, NH-3, and NH-4 of 2-ImN form hydrogen bonds with T17 O2, T18 O2, and A19 N3. Most strikingly, the imidazole nitrogen of 2-ImN responsible for the specificity of the complex forms a hydrogen bond with one amino proton of G16. The N–N distance is 2.94 Å (H–N 1.92 Å), well within the limits of an idealized hydrogen bond. Reproducibly, the imidazole nitrogen of 2-ImN and the amino group of G16 are positioned near the idealized hydrogen bond geometry (Figure 7).

Characterization of the Complexes Formed by 2-ImN and Dst with d(CGCAAACTGGC)·d(GCCAGTTTGCG)

In parallel with footprinting and affinity cleavage studies (Mrksich & Dervan, 1993a), the formation of a 1:1:1 heterocomplex to the designed AAACT·AGTTT site was investigated by NMR. Dst forms a single 2:1 Dst·DNA complex with the AAACT·AGTTT site (at 25 °C) (Figure 8a) as apparent from the analysis of 2D NOESY spectra in both D_2O and H_2O. The sharpness of the lines indicates slow exchange between complex and free DNA although there is evidence for Dst ligands exchanging positions (Figure 6d) at a rate of less than 0.5 s^{-1} at 25 °C. As 2-ImN is added to the AAACT·AGTTT sample already saturated with two equivalents of Dst, one Dst ligand is stoichiometrically replaced from the 2:1 Dst·DNA complex, and 1:1:1 2-ImN·Dst·DNA heterocomplex is formed (Figure 8b–d). At higher ratios, lines corresponding to free ligands appear. At a molar 2-ImN/Dst/DNA ratio of approximately 1.5:2:1 (Figure 8d) the analysis of a 2D NOESY spectrum in D_2O indicates that the 1:1:1 2-ImN·Dst·DNA heterocomplex is the major complex in solution. The analysis was obscured by the presence of 2:1 Dst·DNA complex and exchange processes with free Dst ligand. Therefore, excess ligand was removed by flow dialysis (24 h against deionized water using a membrane with a molecular cutoff of 1 kDa). After dialysis the 1:1:1 2-ImN·Dst·DNA heterocomplex is the predominant complex in solution (Figure 8e), with only small amounts of the 2:1 Dst·DNA complex present. A 2D NOESY spectrum in D_2O reveals that one Dst ligand and one 2-ImN ligand are combined in the antiparallel side-by-side motif similar to the 1:1:1 2-ImN·Dst·DNA heterocomplex with d(GCCTAACAAGG)·d(CCTTGTTAGGC). The Dst ligand lies along the 5′-AACT-3′ strand of the AAACT·AGTTT site with the charged tail group pointing to the 3′-end. The 2-ImN ligand

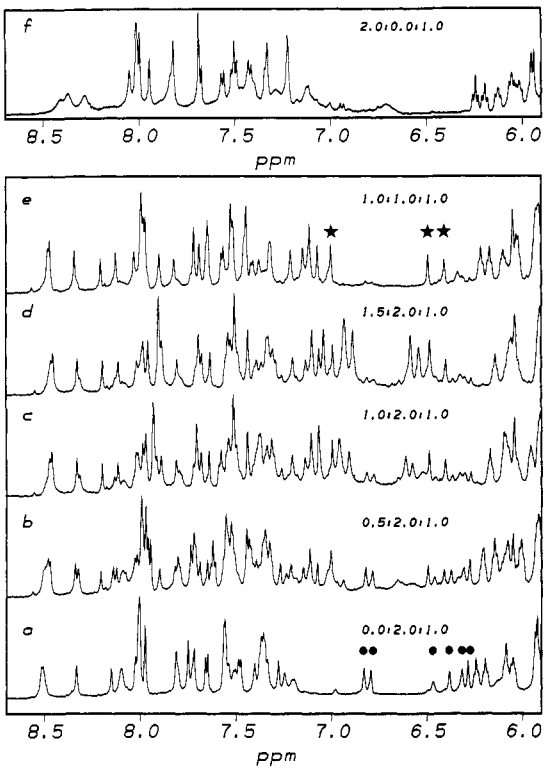


FIGURE 8: Aromatic region of ^{1}H NMR spectra acquired at several points in a titration of d(GCGAAACTGGC)·d(GCCAGTTTGCG) with Dst and 2-ImN at 25 °C (in D_2O). The molar 2-ImN·Dst·DNA ratios are indicated for each spectrum. Ligand resonances of the 2:1 Dst·DNA complex (a) are indicated by dots. Ligand resonances of the 1:1:1 2-ImN·Dst·DNA complex are indicated by stars. Spectrum e was obtained after removal of excess ligand by dialysis. Spectrum f was acquired at 500 MHz on a GN 500 spectrometer.

spans the 5′-GTTT-3′ strand in the opposite direction relative to Dst (Figure 6a). No exchange between ligand positions is observed at 25 °C. Two equivalents of 2-ImN titrated onto the AAACT·AGTTT site results in a spectrum (Figure 8f) with the resonances of the ligands and of the binding site significantly broader than in the case of TAACA·TGTTA (Figure 2). Ligand exchange similar to the 2:1 2-ImN·DNA complex with TAACA·TGTTA is likely to occur but a detailed characterization was not attempted.

DISCUSSION

Intermolecular ligand–DNA and ligand–ligand NOE contacts unambiguously confirm the formation of the 1:1:1 2-ImN·Dst·DNA heterocomplex at a five base pair binding site in the minor groove of DNA. The Dst molecule spans 5′-AACA-3′ and the 2-ImN ligand lies along 5′-GTTA-3′ of the TAACA·TGTTA binding site, while the respective sites are 5′-AACT-3′ and 5′-GTTT-3′ in the AAACT·AGTTT site. The positively charged tail group of each ligand points to the 3′-end of its DNA strand. The ligand arrangement is determined by a specific hydrogen bond between the imidazole nitrogen of 2-ImN and the guanine amino group in the binding site. Evidence for this hydrogen bond is provided indirectly by NOE cross peaks between the guanine amino protons and neighboring ligand protons in the H_2O NOESY spectrum of the 1:1:1 2-ImN·Dst·DNA heterocomplex with the TAACA·TGTTA duplex. The appearance of these cross peaks indicates that the rotation of the amino group about the N–C bond is slowed significantly due to interactions with the imidazole nitrogen (Geierstanger et al., 1993; Boelens et al., 1985; Mrksich et al., 1992). Molecular modeling of this heterocomplex suggests that the 2-ImN imidazole nitrogen is positioned almost ideally with respect to geometry and distance for the formation of a hydrogen bond to the guanine amino group (Figure 7). Because of the strong similarities of the NOE contacts in the 1:1:1 2-ImN·Dst·DNA heterocomplex with AAACT·AGTTT modeling was felt to be redundant and was not attempted.

Competition titrations show that the 2:1 ligand·DNA homocomplexes are converted to 1:1:1 2-ImN·Dst·DNA heterocomplexes in the presence of an excess of both ligands. Furthermore, both of the 2:1 homocomplexes with the TAACA·TGTTA site show clear evidence of two different types of ligand exchange processes: As summarized in Figure 6c the two 2-ImN ligands of the 2:1 2-ImN·DNA complex with TAACA·TGTTA exchange positions with a rate of about 70 s^{-1} at 25 °C. The lines of the 2-ImN complex with AAACT·AGTTT were significantly broader than with TAACA·TGTTA suggesting a 2-ImN homocomplex of lower affinity. Interestingly, only one 2:1 Dst homocomplex is observed with AAACT·AGTTT (Figure 6d) while the Dst ligands slide between different positions in the TAACAA·TTGTTA site (Figure 6b), a situation similar to Dst bound to AAATT·AATTT vs AAATTT·AAATTT (Pelton & Wemmer, 1989, 1990). This can be attributed to the fact that the TAACAA·TTGTTA site resembles a six base pair site for Dst, while in the five base pair AAACT·AGTTT site sliding is restricted. In contrast to both 2:1 ligand·DNA homocomplexes, no evidence for ligand exchange was found for the 1:1:1 2-ImN·Dst·DNA heterocomplexes. On the basis of the competition titrations and the analysis of the exchange dynamics, the 1:1:1 2-ImN·Dst·DNA heterocomplex is believed to be the most stable complex on the TAACA·TGTTA as well as on the AAACT·AGTTT site.

Titrations of d(GCCTAACAAGG)·d(CCTTGTTAGGC) and d(CGCAAACTGGC)·d(GCCAGTTTGCG) with Dst and/or 2-ImN yield only complexes with two ligands bound cooperatively in the antiparallel side-by-side motif. No 1:1 ligand·DNA complexes are observed during the titrations. This is in agreement with the previously characterized homodimeric complex of 2-ImN with TGACT·AGTCT (Mrksich et al., 1992) and homodimeric and heterodimeric complexes of the pyrrole–imidazole–pyrrole analog 2-ImD (3) and Dst with various GC-containing sequences (Dwyer et al., 1992; Geierstanger et al., 1993; Wemmer et al., 1994). The later studies also suggest that the guanine amino group is not a major obstacle for Dst binding in the 2:1 ligand·DNA motif. Moreover, the relative positions and geometries of the dimeric peptides bound in the minor groove are very similar in all of these complexes. The peptides stack such that the amides of one ligand overlap the aromatic rings of the other ligand. Furthermore, only a single orientation of peptide has been observed, with the amino terminus of the peptide lying toward the 5′-end of the DNA strand with which it interacts. The 2:1 ligand·DNA binding motif appears to be a general feature of the complexes formed by distamycin and its analogs with the minor groove of mixed GC/AT DNA sequences, probably arising from the intrinsically wider grooves seen in GC-containing sequences (Yoon et al., 1988).

The most important feature of the binding site for the 1:1:1 2-ImN·Dst·DNA heterocomplexes is the guanine in the sequence 5′-(A,T)G(A,T)$_3$-3′. Recognition of this guanine by 2-ImN determines the ligand arrangement. Flanking A and T's may be interchangeable while maintaining the hydrogen-bonding interactions between the peptides and DNA. Although the 1:1:1 2-ImN·Dst·DNA heterocomplexes with

TAACA·TGTTA and AAACT·AGTTT are very similar in structure, the binding affinties of the peptides for these two sites are likely different. A complete characterization of the sequence specificity of these peptides for 5′-(A,T)G(A,T)$_3$-3′ binding sites requires independent determination of the binding affinities of Dst and 2-ImN for each of the 16 sequences.

When compared to a heterocomplex of 2-ImN and the pyrrole–pyrrole–pyrrole system P3 (**4**), the 1:1:1 2-ImN·Dst·DNA heterocomplex exhibits a significantly higher stability (unpublished results). P3 differs from Dst by having the formyl group replaced by an acetyl moiety and the amidinium group by a dimethylammonium group. Differences in shape, charge distribution, or hydrogen bonding could be responsible for the difference in binding affinity. Molecular modeling of the 1:1:1 2-ImN·Dst heterocomplex suggests hydrogen-bonding interactions of the Dst amidinium group with DNA, supported also by recent ^{15}N NMR data (Rhee et al., 1993). This suggests that the nature of the cationic ligand tail group may be important for specificity and overall stability of these complexes.

The most striking feature of the 1:1:1 2-ImN·Dst·DNA heterocomplexes is the recognition of the guanine amino group by the imidazole nitrogen of 2-ImN. This is analogous to the recognition of the G–C base pair in the AAGTT·AACTT site by Dst and 2-ImD (Geierstanger et al., 1993) and recognition of two G–C base pairs in the TGACT·AGTCA site by two side-by-side 2-ImN peptides (Wade & Dervan, 1987; Mrksich et al., 1992). In these three examples, hydrogen bonding of one imidazole nitrogen per guanine amino group is sufficient for sequence-specific recognition of G–C base pairs in the minor groove. We believe that the higher affinity of the heterocomplex arises directly from this hydrogen bond, an idea supported by recent free energy perturbation calculations (Singh et al., 1993). The potential for recognition of specific DNA sequences by the antiparallel side-by-side motif of distamycin-like ligands is still only partially explored. The design of sequence-specific peptides requires consideration of sequence-dependent groove width, as a major determinant of binding motif, as well as specific hydrogen bonds to the floor of the double helix by the pyrrolecarboxamides and imidazolecarboxamides. The exploration of covalent dimers of these ligands for enhanced specificity and selectivity has already been started (Mrksich & Dervan, 1993b; Dwyer et al., 1993).

SUPPLEMENTARY MATERIAL AVAILABLE

Listings of the intermolcular ligand–DNA and ligand–ligand restraints and the achieved distances (1 page). Ordering information is given on any current masthead page.

REFERENCES

Boelens, R., Scheek, R. M., Dijkstra, K., & Kaptein, R. (1985) *J. Magn. Reson. 62*, 378–386.

Coll, M., Fredrick, C. A., Wang, A. H.-J., & Rich, A. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8385–8389.

Dervan, P. B. (1986) *Science 232*, 464–471.

Dwyer, T. J., Geierstanger, B. H., Bathini, Y., Lown, J. W., & Wemmer, D. E. (1992) *J. Am. Chem. Soc. 114*, 5911–5919.

Dwyer, T. J., Geierstanger, B. H., Mrksich, M., Dervan, P. B., & Wemmer, D. E. (1993) *J. Am. Chem. Soc. 115*, 9900–9906.

Fagan, P., & Wemmer, D. E. (1990) *J. Am. Chem. Soc. 114*, 1080–1081.

Geierstanger, B. H., Dwyer, T. J., Bathini, Y., Lown, J. W., & Wemmer, D. E. (1993) *J. Am. Chem. Soc. 115*, 4474–4482.

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

Harris, R. K. (1986) *Nuclear Magnetic Resonance Spectroscopy*, Longman Scientific & Technical, Harlow, England.

Klevit, R. E., Wemmer, D. E., & Reid, B. R. (1986) *Biochemistry 25*, 3296–3303.

Kopka, M. L., Yoon, C., Goodsell, D., Pjura, P., & Dickerson, R. E. (1985a) *Proc. Natl. Acad. Sci. U.S.A. 82*, 1376–1380.

Kopka, M. L., Yoon, C., Goodsell, D., Pjura, P., & Dickerson, R. E. (1985b) *J. Mol. Biol. 183*, 553–563.

Mrksich, M., & Dervan, P. B. (1993a) *J. Am. Chem. Soc. 115*, 2572–2576.

Mrksich, M., & Dervan, P. B. (1993b) *J. Am. Chem. Soc. 115*, 9892–9899.

Mrksich, M., Wade, W. S., Dwyer, T. J., Geierstanger, B. H., Wemmer, D. E., & Dervan, P. B. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 7586–7590.

Pelton, J. G., & Wemmer, D. E. (1988) *Biochemistry 27*, 8088–8096.

Pelton, J. G., & Wemmer, D. E. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 5723–5727.

Pelton, J. G., & Wemmer, D. E. (1990) *J. Am. Chem. Soc. 112*, 1393–1399.

Rhee, Y., Wang, C., Gaffrey, B. L., & Jones, R. A. (1993) *J. Am. Chem. Soc. 115*, 8742–8746.

Schultz, P. G., & Dervan, P. B. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 6834–6837.

Schultz, P. G., & Dervan, P. B. (1984) *J. Biomol. Struct. Dyn. 1*, 1133–1147.

Schultz, P. G., Taylor, J. S., & Dervan, P. B. (1982) *J. Am. Chem. Soc. 104*, 6861–6863.

Singh, S. B., Ajay, Wemmer, D. E., & Kollman, P. A. (1994) *Proc. Natl. Acad. Sci. U.S.A.* (submitted for publication).

Sklenàr, V., Brooks, B. R., Zon, G., & Bax, A. (1987) *FEBS Lett. 216*, 249–252.

Taylor, J. S., Schultz, P. G., & Dervan, P. B. (1984) *Tetrahedron 40*, 457–465.

Wade, W. S., & Dervan, P. B. (1987) *J. Am. Chem. Soc. 109*, 1574–1575.

Wade, W. S., Mrksich, M., & Dervan, P. B. (1992) *J. Am. Chem. Soc. 114*, 8783–8794.

Wade, W. S., Mrksich, M., & Dervan, P. B. (1993) *Biochemistry 32*, 11385–11389.

Warshaw, M., & Cantor, C. (1970) *Biopolymers 9*, 1079–1103.

Wemmer, D. E., Geierstanger, B. H., Fagan, P. A., Dwyer, T. J., Jacobsen, J. P., Pelton, J. G., Ball, G. E., Leheny, A. R., Chang, W.-H., Bathini, Y., Lown, J. W., Rentzeperis, D., Marky, L., Singh, S., & Kollman, P. (1994) in *Proceedings of the Eighth Conversation on Biomolecular Stereodynamics*, Adenine Press, New York.

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

Yoon, C., Privé, G. G., Goodsell, D. S., & Dickerson, R. E. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 6332–6336.

Youngquist, R. S., & Dervan, P. B. (1985a) *Proc. Natl. Acad. Sci. U.S.A. 82*, 2565–2569.

Youngquist, R. S., & Dervan, P. B. (1985b) *J. Am. Chem. Soc. 107*, 5528–5529.

Zimmer, C., & Wähnert, U. (1986) *Prog. Biophys. Mol. Biol. 47*, 31–112.