VOLUME 114, NUMBER 15
JULY 15, 1992

JOURNAL
OF THE
AMERICAN
CHEMICAL
SOCIETY

Design and Binding of a Distamycin A Analog to d(CGCAAGTTGGC)·d(GCCAACTTGCG): Synthesis, NMR Studies, and Implications for the Design of Sequence-Specific Minor Groove Binding Oligopeptides

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

Contribution from the Department of Chemistry and Graduate Group in Biophysics, University of California, Berkeley, California 94720, and Department of Chemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2G2. *Received January 15, 1992*

Abstract: An oligopeptide has been synthesized in which an imidazole ring is substituted for the central pyrrole ring of distamycin A. Two-dimensional NMR spectroscopy was used to characterize the complex formed between 3-[1-methyl-4-[1-methyl-4-[1-methyl-4-(formylamino)pyrrole-2-carboxamido]imidazole-2-carboxamido]pyrrole-2-carboxamido]propionamidine hydrochloride (2-ImD) and d(CGCAAGTTGGC)·d(GCCAACTTGCG). Titration of the AAGTT duplex with 2-ImD yielded a single complex with a ligand:DNA stoichiometry of 2:1. The nuclear Overhauser effect (NOESY) experiment in D_2O was used to assign the aromatic and C1'H DNA protons and to identify intermolecular ligand–DNA contacts between nonlabile protons. The NOESY experiment in H_2O was used to assign the imino and amino DNA protons and the amide protons of the ligands and to identify ligand–DNA contacts involving these labile protons. These data indicate that two ligand molecules bind simultaneously to the minor groove of the central 5'-AAGTT-3' sequence in a head-to-tail orientation. Molecular modeling, using 35 ligand–DNA distance constraints derived from a semiquantitative analysis of the NOESY data, shows that the imidazole N3 of one of the ligands forms a hydrogen bond with the C2 amino group of the guanine in the binding site. Additionally, the titration of d(CGCAAATTGGC)·d(GCCAATTTGCG) with 2-ImD was performed. No specific complex was detected by NMR spectroscopy between 2-ImD and the AAATT duplex. This result emphasizes the importance of the imidazole N3 atom of 2-ImD in the recognition of the AAGTT binding site.

Introduction

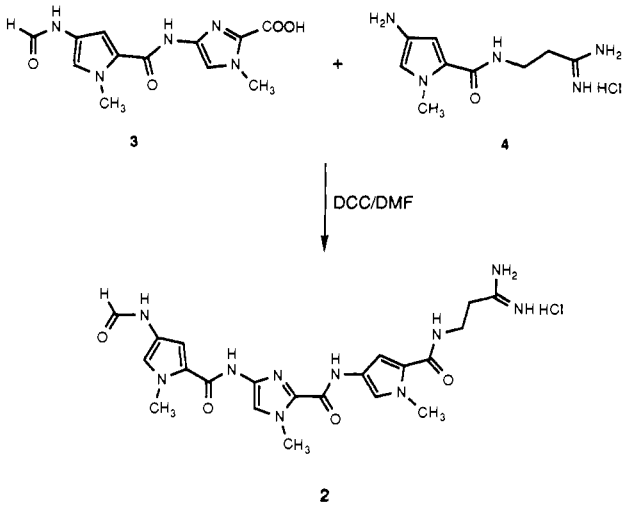
Many biological processes, such as the regulation of gene expression, involve specific interactions of proteins with DNA. Sequence-specific recognition of DNA by these proteins is achieved primarily via hydrogen bonds and van der Waals contacts between protein side chains and the major groove of DNA.[1] It is argued that the accessibility of the functional groups in the major groove enables the information encoded in the DNA sequence to be "read" by the protein.[1] More recent findings,[2] however, provide evidence that contacts with the minor groove are also important for sequence-specific recognition of DNA by a variety of DNA-binding proteins.

* Address correspondence to this author at the Department of Chemistry, University of California, Berkeley, CA 94720.
[†] Department of Chemistry, University of California.
[‡] Graduate Group in Biophysics, University of California.
[§] Department of Chemistry, University of Alberta.

(1) Steitz, T. A. *Q. Rev. Biophys.* **1990**, *23*, 205. Pavletich, N. P.; Pabo, C. O. *Science* **1991**, *252*, 809. Jordan, S. R.; Pabo, C. O. *Science* **1988**, *242*, 893. Aggarwal, A. K.; Rodgers, D. W.; Drottar, M.; Ptashne, M.; Harrison, S. C. *Science* **1988**, *242*, 899. Vinson, C. R.; Sigler, P. B.; McKnight, S. L. *Science* **1989**, *246*, 911.
(2) Kissinger, C. R.; Liu, B.; Martin-Blanc, E.; Kornberg, T. B.; Pabo, C. O. *Cell* **1990**, *63*, 579. Sluka, J. P.; Horvath, S. J.; Glasgow, A. C.; Simon, M. I.; Dervan, P. B. *Biochemistry* **1990**, *29*, 6551. Otting, G.; Qian, Y. Q.; Billeter, M.; Müller, M.; Affolter, M.; Gehring, W. J.; Wüthrich, K. *EMBO J.* **1990**, *9*, 3085. Percival-Smith, A.; Müller, M.; Affolter, M.; Gehring, W. J. *EMBO J.* **1990**, *9*, 3967.

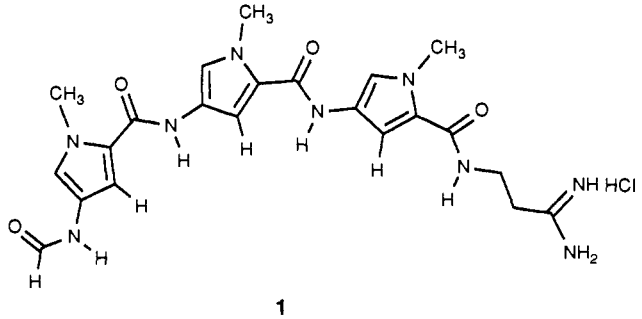
0002-7863/92/1514-5911$03.00/0

Scheme I. Initial Convergent Route to 2-ImD (**2**)



3 4 2

In contrast to most proteins, peptide-linked polypyrroles such as the natural antibiotic agents netropsin and distamycin A (**1**) bind exclusively to the minor groove of DNA. These ligands



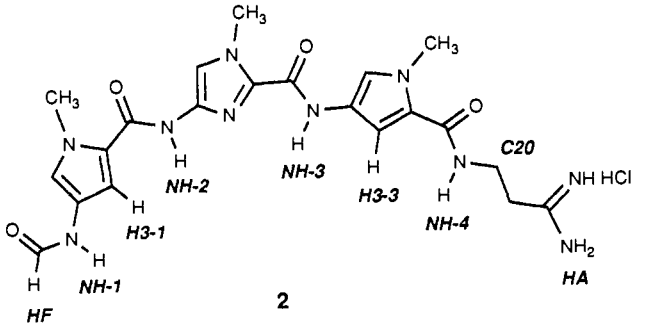
1

recognize stretches containing at least four AT base pairs and replace a network of well-ordered water molecules with a similar pattern of hydrogen bonds to the DNA. A combination of hydrogen bonding, van der Waals contacts with the walls of the minor groove, and electrostatic interactions between the cationic ligands and the DNA contributes to the high affinity for ligand binding.[3] There are three properties of netropsin/distamycin A binding sites that may lead to the preference for AT-rich sequences: (1) the width of the minor groove of AT-rich regions of B-DNA is narrower than standard B-DNA;[4] (2) the site of most negative electrostatic potential of AT-rich regions of B-DNA is near N3 of adenine at the bottom of the minor groove;[5] and (3) the presence of the C2 amino group at the bottom of the minor groove in GC base pairs introduces a steric hindrance to ligand binding to GC-containing sequences.[6]

Dickerson[6] and Lown[7] independently proposed that, by chemical modification of peptide analogs, the exclusion of binding to GC base pairs may be overcome. For example, by replacing a pyrrole ring of distamycin A with an imidazole ring, one might transform the steric interference with the C2 amino group of a guanine into an energetically favorable hydrogen bond. Initial attempts to design information-reading "lexitropsin" molecules have yielded ligands with an increased tolerance for GC base pairs, although not GC specificity.[7]

Recent NMR studies have indicated that the minor groove of AT-rich regions of DNA is flexible enough to accommodate two distamycin A molecules side-by-side.[8] These results suggest that the minor groove width may affect the binding affinity of distamycin A. Presumably the introduction of a GC base pair into an AT-rich sequence of DNA should widen the minor groove. Therefore, a lexitropsin capable of forming a 2:1 complex would optimize the van der Waals contacts with the wider groove and enhance binding.

On the basis of these considerations, the distamycin A analog 2-ImD (**2**) was designed to specifically recognize the DNA sequence AAGTT. We describe here NMR studies that demon-



2

strate that this compound forms a specific 2:1 complex with the duplex d(CGCAAGTTGGC)·d(GCCAACTTGCG). While 2-ImD binds to the predicted site with high affinity and cooperativity, no specific binding to the sequence AAATT was observed. This suggests that, indeed, the imidazole nitrogen acts as a hydrogen-bond acceptor for the C2 amino group of the guanine base and that this interaction determines binding specificity.

Synthesis

Lexitropsin **2** (2-ImD) is an analog of distamycin A (**1**) in which the central *N*-methylpyrrolecarboxamide was replaced by an *N*-methylimidazolecarboxamide. Initially it was planned to synthesize the lexitropsin **2** following a highly convergent approach in which *N*-formyl and amidinium groups were introduced at early stages as described in Scheme I. Preparation of the acid **3** and the amine **4** followed by condensation should give the targeted oligopeptide **2**.

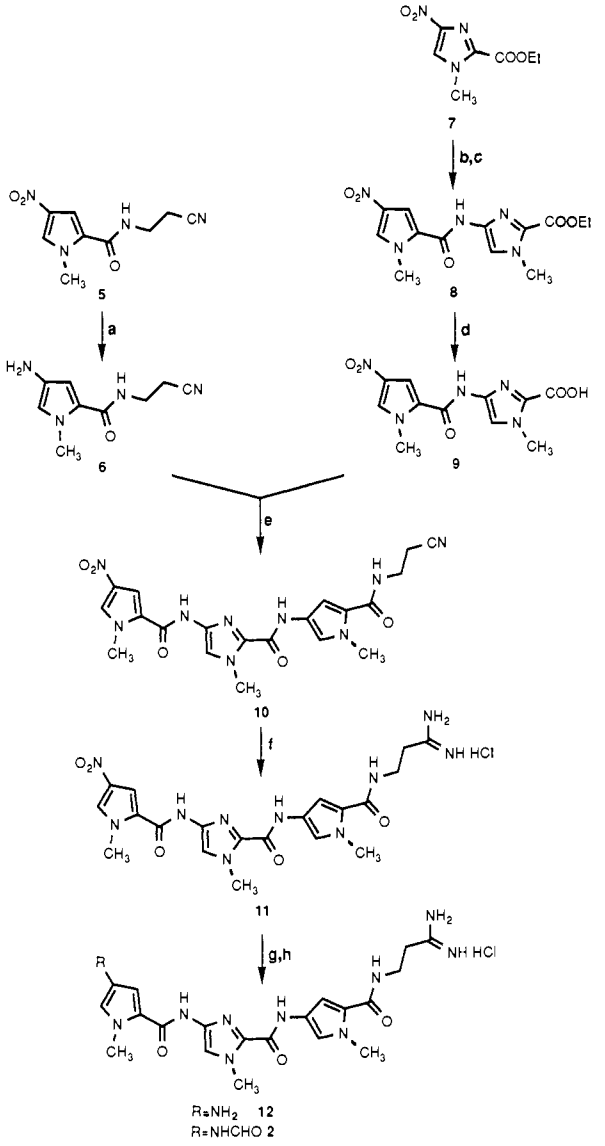
Attempted condensation of acid **3** and amine **4** with dicyclohexylcarbodiimide (DCC) at 0 °C and room temperature did not give the required product. It appears that the carboxylic acid is very unreactive under the above conditions and the amine is very reactive toward DCC, and as a result an amine–DCC addition product is formed. Condensation in the presence of a catalytic amount of 4-(dimethylamino)pyridine (DMAP) afforded the required product in low yield, and the product was contaminated with unreacted amine **4** and a DCC–amine addition product.

In view of the above problems, the synthesis of the desired lexitropsin by an alternative route as described in Scheme II was attempted. Condensation of amine **6** and acid **9** followed by modification of a nitrile group to an amidine and a nitro group to an *N*-formyl group should give the desired lexitropsin **2**.

The amine **6** was prepared following the reported procedure[10,11] as described in Scheme II. The second key synthon, **9**, was prepared starting from ethyl 1-methyl-4-nitroimidazole-2-carboxylate (**7**) as described in Scheme II. Required starting

(3) Zimmer, C.; Wähnert, U. *Prog. Biophys. Mol. Biol.* **1986**, *47*, 31.
(4) Fratini, A. V.; Kopka, M. L.; Drew, H. R.; Dickerson, R. E. *J. Biol. Chem.* **1982**, *257*, 14686. Nelson, H. C. M.; Finch, J. T.; Luisi, B. F.; Klug, A. *Nature (London)* **1987**, *330*, 221.
(5) Pullman, B. *J. Biomol. Struct. Dyn.* **1983**, *1*, 773.
(6) Kopka, M. L.; Yoon, C.; Goodsell, D.; Pjura, P.; Dickerson, R. E. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 1376.
(7) 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. Mol. Recognit.* **1989**, *2*, 84.
(8) (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.
(9) Lown, J. W.; Krowicki, K. *J. Org. Chem.* **1985**, *50*, 3774.
(10) Rao, K. E.; Bathini, Y.; Lown, J. W. *J. Org. Chem.* **1990**, *55*, 728.
(11) Krowicki, K.; Lown, J. W. *J. Org. Chem.* **1987**, *52*, 3493.

Scheme II. Successful Convergent Route to 2-ImD (**2**)[a]



[a] (a) H_2/Pd–C, 10%; (b) H_2/Pd–C, 10%; (c) 1-methyl-4-nitropyrrole-2-carbonyl chloride/triethylamine; (d) LiOH; (e) EDCI/DMF; (f) dry HCl/EtOH, dry ammonia/EtOH; (g) H_2/Pd–C, 10%; (h) acetic formic anhydride.

material **7** was prepared following the reported procedure.[9] Nitroimidazole was reduced to aminoimidazole by catalytic hydrogenation in the presence of 10% Pd–C, and this amine was acylated with the acid chloride of 1-methyl-4-nitropyrrole-2-carboxylic acid to give the amide **8** in overall 87% yield. The required acid, **9**, was obtained from lithium hydroxide hydrolysis of ester **8**. Condensation of acid **9** and amine **6** in dimethylformamide with 1-ethyl-3-[3-(dimethylamino)propyl]carbodiimide hydrochloride (EDC) afforded the tripeptide **10** in 35% yield (Scheme II). This tripeptide was also prepared by an alternative route via acylation of amine **6** with the acid chloride of **9** in the presence of triethylamine in 48% yield. The nitrile group in **10** was converted to amidinium hydrochloride **11** under Pinner reaction conditions.[10] Catalytic hydrogenation of the nitropyrrole tripeptide **11** in the presence of 10% palladium on charcoal at room temperature and at atmospheric pressure did not give the desired amine **12**. Reduction with large amounts of catalyst, acid catalysis (drop of acetic acid or hydrochloric acid), and longer reaction times (24 h) led only to decomposed material. It appears that the rate of reduction of the nitro group to amino is slow and the amine group also slowly decomposes under the reaction conditions. When the reduction was carried out at a pressure of 3 atm and room temperature, the nitro compound **11** was smoothly converted to the amine **12**. This amine was formylated with formic acetic anhydride to give the lexitropsin **2**. The overall yield on going from the nitro to the *N*-formyl derivative is 58%.

Experimental Section

Synthesis of 2-ImD (2). Melting points were determined on an Electrothermal melting point apparatus and are uncorrected. The IR spectra were recorded on a Nicolet 7199 FT spectrophotometer, and only the principal bands are reported. The ^{1}H-NMR spectra were recorded on Bruker WH-200 and WH-400 spectrometers. Fast atom bombardment mass spectra (FABMS) using glycerol as the matrix were determined on Associate Electrical Industries (AEI) MS-9 and MS-50 focusing high-resolution mass spectrometers. Kieselgel 60 (230–400 mesh) of E. Merck and Florisil (60–100 mesh) were used for chromatography, and precoated silica gel 60F-254 sheets of E. Merck were used for TLC, with the solvent system indicated in the procedure. TLC plates were visualized using UV light or 2.5% phosphomolybdic acid in methanol with heating.

All compounds obtained commercially were used without further purification unless otherwise stated. Ethanol and methanol were freshly distilled from magnesium turnings; THF was distilled from Na/benzophenone under an atmosphere of argon; ether was dried over Na; methylene chloride was distilled from P_2O_5 and stored over molecular sieves, 3 Å; Et_3N was treated with KOH, then distilled from barium oxide, and stored over molecular sieves, 3 Å; DMF was distilled from barium oxide and stored over molecular sieves, 3 Å.

Ethyl 1-Methyl-4-(1-methyl-4-nitropyrrole-2-carboxamido)imidazole-2-carboxylate (8). A solution of **7** (995 mg, 5 mmol) in a 1:1 mixture of ethyl acetate and methanol (60 mL) was hydrogenated over 10% Pd–C (120 mg) at room temperature and atmospheric pressure. The catalyst was removed by filtration, and the filtrate was concentrated to give ethyl 4-amino-1-methylimidazole-2-carboxylate. This amine was used immediately in the next step.

To the cooled solution of the above amine in CH_2Cl_2 (50 mL) and Et_3N (2 mL) was added dropwise 1-methyl-4-nitropyrrole-2-carbonyl chloride (940 mg, 5 mmol) in dichloromethane (30 mL). The reaction mixture was slowly warmed to room temperature and stirred for 6 h, then diluted with $CHCl_3$, and washed successively with water and brine. Solvent was removed under reduced pressure, and the resulting solid was purified on silica gel (flash chromatography, hexane/EtOAc, 1:1) to give **8**, 1.4 g (87% yield): mp 236–238 °C; IR (KBr) ν_{max} 3420, 1715, 1665, 1560, 1530, 1500, 1310 cm^{-1}; ^{1}H-NMR ($CDCl_3$) δ 1.4 (t, 3 H, OCH_2CH_3), 4.02 (br s, 6 H, 2 × NCH_3), 4.38 (q, 2 H, OCH_2CH_3), 7.1 (d, 1 H, Ar H), 7.54 (s, 1 H, Ar H), 7.6 (d, 1 H, Ar H), 8.66 (s, 1 H, NH, exch); MS, *m/z* (rel intensity) calcd for $C_{13}H_{15}N_5O_5$ 321.1073, found 321.1071 (45), 220 (25), 153 (100).

1-Methyl-4-(1-methyl-4-nitropyrrole-2-carboxamido)imidazole-2-carboxylic Acid (9). A cooled suspension (0–5 °C) of **8** (642 mg, 2 mmol) in ethanol (20 mL) was treated with 20 mL of 1 M lithium hydroxide solution. Progress of the hydrolysis was followed by TLC. After completion of the hydrolysis (1.5–2 h), the reaction mixture was diluted with water and extracted with ether. The aqueous layer was cooled and acidified with 2 M HCl. The precipitated solid was collected and washed with water and ether to give acid **9**, 520 mg (89% yield): mp 250 °C dec; IR (KBr) ν_{max} 3420, 3135, 1680, 1584, 1507, 1370, 1350 cm^{-1}; ^{1}H-NMR (DMSO-d_6) δ 3.91 (s, 3 H, NCH_3), 3.95 (s, 3 H, NCH_3), 7.63 (s, 1 H, Ar H), 7.80 (d, 1 H, Ar H), 8.2 (d, 1 H, Ar H), 11.1 (s, 1 H, NH, exch); MS, *m/z* (rel intensity) 249.0862 (M^+ – C02, 58), 221 (16), 153 (100), 107 (50).

3-[1-Methyl-4-[1-methyl-4-(1-methyl-4-nitropyrrole-2-carboxamido)-imidazole-2-carboxamido]pyrrole-2-carboxamido]propionitrile (10). A solution of **5** (111 mg, 0.5 mmol) in methanol (15 mL) was hydrogenated over 10% Pd–C (15 mg) at room temperature and atmospheric pressure. The catalyst was removed by filtration, and the filtrate was concentrated to give amine **6** as an oily residue, which was used immediately in the following step.

To a cooled (–10 °C) mixture of acid **9** (147 mg, 0.5 mmol) and the above amine **6** in dry DMF (20 mL) was added EDCI (153 mg, 0.8 mmol) in DMF (6 mL) over 30 min. The reaction mixture was allowed to attain room temperature and was stirred for 12 h. The solvent was removed under reduced pressure, and water (20 mL) was added to the resulting residue. This was extracted with a 1:1 mixture of ethyl acetate and THF (3 × 20 mL). The combined organic extract was successively washed with water and brine, dried, and evaporated. The resulting solid was purified on silica gel (flash chromatography, EtOAc/hexane, 5:1) to give **10**, 82 mg (35% yield): mp 263–265 °C; IR (KBr) ν_{max} 3400, 3380, 3100, 2240, 1660, 1550 cm^{-1}; ^{1}H-NMR (DMSO-d_6) δ 2.74 (t, 2 H, CH_2), 3.4 (q, 2 H, CH_2), 3.82 (s, 3 H, NCH_3), 3.96 (2 s merged, 6 H, 2 × NCH_3), 7.0 (d, 1 H, Ar H), 7.26 (d, 1 H, Ar H), 7.58 (s, 1 H, Ar H), 7.76 (d, 1 H, Ar H), 8.22 (d, 1 H, Ar H), 8.38 (t, 1 H, NH, exch), 10.05 (s, 1 H, NH, exch), 10.83 (s, 1 H, NH, exch); MS, *m/z* (rel

intensity) calcd for $C_{20}H_{29}N_9O_5$ 467.1665, found 467.1665 (M^+, 77), 370 (17), 344 (22), 191 (21), 153 (100).

This oligopeptide, **10**, was also prepared by an alternative route from the amine **6** and the acid chloride of **9** in the presence of triethylamine in 48% yield following a procedure similar to that described for **8**. The acid chloride was obtained from the reaction of **9** with thionyl chloride in benzene for 3 h under refluxing conditions.

3-[1-Methyl-4-[1-methyl-4-(1-methyl-4-nitropyrrole-2-carboxamido)-]imidazole-2-carboxamido]pyrrole-2-carboxamido]propionamidine Hydrochloride (11). A suspension of **10** (94 mg, 0.2 mmol) in absolute ethanol (30 mL) was cooled to −15 °C and saturated with dry HCl gas. The reaction mixture was stirred at 0 °C for 1 h and at room temperature for an additional 1 h. The solvent was removed in vacuo, and the resulting residue was washed with dry ether to give an intermediate imidate ester. This solid was suspended in absolute ethanol (20 mL), and then 15 mL of ammonia was condensed into the reaction vessel. After the reaction mixture was stirred for 8 h at room temperature, the solvent was removed in vacuo, and the resulting solid was recrystallized from methanol/ethyl acetate to give **11** as a light yellow crystalline solid, 85 mg (82% yield): mp 193–195 °C; IR (KBr) ν_{max} 3380, 3120, 1650, 1545, 1404, 1312 cm^{-1}; ^{1}H-NMR (DMSO-d_6) δ 2.64 (t, 2 H, CH_2), 3.5 (q, 2 H, CH_2), 3.82 (s, 3 H, NCH_3), 3.98 (2 s merged 6 H, 2 × NCH_3), 7.04 (d, 1 H, Ar H), 7.26 (d, 1 H, Ar H), 7.58 (s, 1 H, Ar H), 7.76 (d, 1 H, Ar H), 8.21 (d, 1 H, Ar H), 8.3 (t, 1 H, NH, exch), 8.65 (br s, 2 H, NH_2, exch), 9.02 (br s, 2 H, NH_2, exch), 9.9 (s, 1 H, NH, exch), 10.85 (s, 1 H, NH, exch); FABMS, m/z (rel intensity) 485 (M^+ − Cl, 12). Anal. Calcd for $C_{20}H_{25}ClN_{10}O_5$: C, 46.11; H, 4.80; N, 26.89. Found: C, 46.34; H, 4.68; N, 27.28.

3-[1-Methyl-4-[1-methyl-4-[1-methyl-4-(formylamino)pyrrole-2-carboxamido]imidazole-2-carboxamido]pyrrole-2-carboxamido]propionamidine Hydrochloride (2). A solution of **11** (52 mg, 1 mmol) in methanol (25 mL) was hydrogenated over 10% Pd–C at room temperature and 3 atm of pressure using a Paar hydrogenator for 4 h. The catalyst was removed by filtration. The filtrate was concentrated, and the residue was coevaporated with hexane (2 × 15 mL) to give amine **12**. Owing to the instability of this amine, it was used immediately in the following step. The above amine **12** in 98% formic acid (2 mL) was treated with 2 mL of acetic formic anhydride. The reaction mixture was stirred for 12 h at room temperature. The solvent was evaporated to dryness in vacuo, and the resulting residue was crystallized from methanol and ethyl acetate to give **2**, 30 mg (58% yield): mp 210 °C dec; IR (KBr) ν_{max} 3420, 1655, 1540, 1470, 1440, 1400 cm^{-1}; ^{1}H-NMR (DMSO-d_6) δ 2.63 (t, 2 H, CH_2), 3.5 (q, 2 H, CH_2), 3.81 (s, 3 H, NCH_3), 3.84 (s, 3 H, NCH_3), 3.96 (s, 3 H, NCH_3), 6.99 (d, 1 H, Ar H), 7.05 (d, 1 H, Ar H), 7.24 (d, 1 H, Ar H), 7.28 (d, 1 H, Ar H), 7.54 (s, 1 H, Ar H), 8.11 (d, 1 H, CHO), 8.3 (t, 1 H, NH), 8.64 (br s, 2 H, amidinium NH_2), 9.02 (br s, 2 H, amidinium NH_2), 9.93 (s, 1 H, NH), 10.13 (s, 1 H, NH), 10.31 (s, 1 H, NH). D_2O exchange resulted in the disappearance of peaks at 8.3, 8.64, 9.02, 9.93, 10.13, and 10.31. Peaks at 3.5 and 8.11 became a triplet and a singlet, respectively. FABMS: 483 (M^+ − Cl, 15). Anal. Calcd for $C_{21}H_{27}ClN_{10}O_4$: C, 48.60; H, 5.21; N, 27.00. Found: C, 48.78; H, 5.36; N, 26.74.

Synthesis of Oligonucleotides. The oligomers d(CGCAAGTTGGC), d(GCCAACTTGCG), d(CGCAAATTGGC), and d(GCCAATTTGCG) were synthesized and purified as described previously.[8b]

Sample Preparation. NMR samples were prepared by dissolving the undecamer oligonucleotide 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 99.99% 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.

A stock solution of 2-ImD·HCl was prepared by dissolving 1.93 mg of the ligand in 100 μL of 99.99% D_2O containing 10 mM sodium phosphate buffer (pH 7.0). The concentration of the stock solution was determined to be 19.3 mM by UV absorbance at 304 nm ($\epsilon = 3.6 \times 10^4$ M^{-1} cm^{-1}). Extinction coefficients for d(CGCAAGTTGGC) and d(GCCAACTTGCG) were calculated[12] to be 1.03×10^5 and 9.93×10^4 M^{-1} cm^{-1}, respectively. The concentration of the double-stranded DNA sample was determined to be 1 mM by UV absorbance at 260 nm at 80 °C.

1D NMR Titration. 1D NMR titration spectra of 2-ImD/AAATT were acquired at 500 MHz on a General Electric GN-500 spectrometer, while those for 2-ImD/AAGTT were acquired at 600 MHz on a Bruker AMX-600 spectrometer. 2-ImD·HCl was titrated into the NMR sample containing the DNA in 0.25 molar equiv/addition. 1D spectra were acquired at 25 °C using 4096 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.

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[13] pulse sequence. 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. Typically 512 t_1 experiments were recorded and zero-filled to 1K. For each t_1 value 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.

NOESY spectra in water were acquired at 25 °C, the last 90° pulse being replaced by a 1:1 jump and return sequence[14] to suppress the solvent 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. Typically, 547–567 t_1 experiments were recorded and zero-filled to 2K. 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 software (Hare Research) on a Vax 11/785 computer or with FELIX software (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; 512–567 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 Constraints. Intermolecular distance constraints 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 is more likely at longer mixing times, we calculated distance constraints only for crosspeaks that were present in the 100-ms NOESY. We are confident that the model for the complex that results from this type of analysis accurately depicts the binding site. Volume integrals were measured for each mixing time using FELIX software. The crosspeak volumes were scaled according to $V_{corr} = V_{obsd}/[\sin(\Delta\omega\tau)]$, where V_{corr} is the corrected crosspeak volume, V_{obsd} is the measured crosspeak volume, τ 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 F2. 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, 32 intermolecular ligand–DNA constraints and three intermolecular ligand–ligand constraints were used.

Structure Refinement. A starting B-form oligonucleotide model was constructed using the Biopolymer module of Insight II (Biosym). For the ligands, the coordinates for the two distamycin A molecules in the energy-minimized NMR structure of the 2:1 complex with d(CGCAAATTGGC)·d(GCCAATTTGCG)[8] were read into Insight II. Using the Builder module of Insight II, the distamycin A molecules were transformed into 2-ImD molecules. The Silicon Graphics workstation facilitated a simple docking procedure to merge the ligands and the DNA. Energy minimizations were performed using the Discover module of Insight II. NOE distance constraints were incorporated into the structure prior to each simulation. All hydrogen bonds in the standard Watson–Crick base-pairing scheme were included as constraints using a force constant of 300 (kcal/mol)/Å^2. A cutoff distance of 15 Å was used for nonbonded interactions, and the neighbor list was regenerated every 20 time steps. A distance-dependent dielectric of the form $\epsilon = R$ was used to account for solvent effects. The energy of the complex was minimized initially using 100 steps of a steepest descents algorithm and further using 15 000 steps of conjugate gradient minimization with an NMR force constant of 50 (kcal/mol)/Å^2 to a root mean square derivative of <0.01 (kcal/mol)/Å^2.

Results

Titration of the AAGTT Site with 2-ImD. The results of the titration of 2-ImD into an NMR sample of the duplex of d(CGCAAGTTGGC)·d(GCCAACTTGCG) 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

(12) Warshaw, M.; Cantor, C. *Biopolymers* **1970**, *9*, 1079.

(13) Drobny, G.; Pines, A.; Sinton, S.; Weitekamp, D. P.; Wemmer, D. E. *Faraday Div. Chem. Soc. Symp.* **1979**, *13*, 49.

(14) Skelenar, V.; Brooks, B. R.; Zon, G.; Bax, A. *FEBS Lett.* **1987**, *216*, 249.

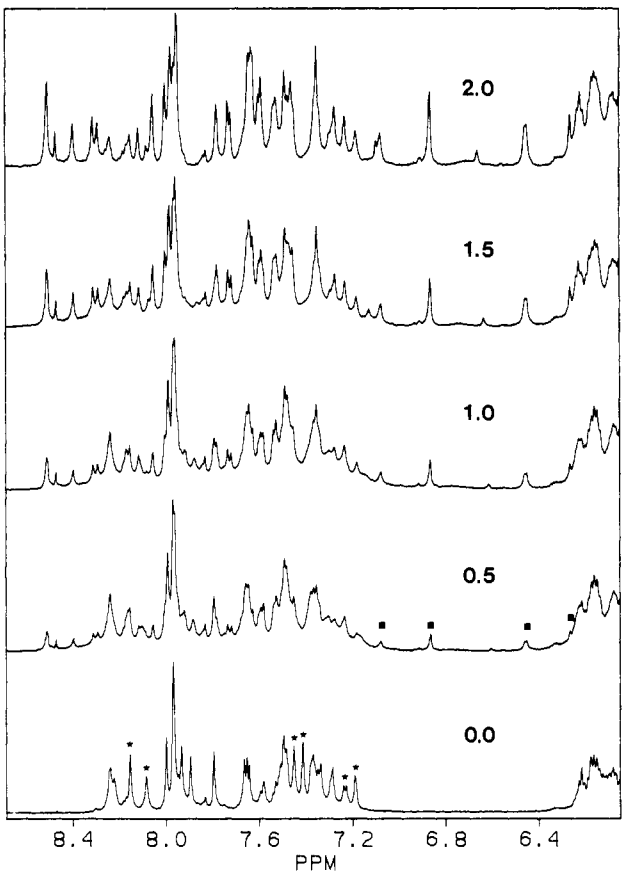


Figure 1. Titration of d(CGCAAGTTGGC)·d(GCCAACTTGCG) with 2-ImD at 25 °C. The ligand:DNA ratio is indicated for each spectrum. Stars denote the aromatic proton resonances of uncomplexed DNA for the central three base pairs of the ligand binding site. Squares indicate proton resonances of the bound ligands.

assigned in the NOESY spectrum of the free duplex by standard sequential methods.[15] Upon the addition of substoichiometric amounts of 2-ImD, the NMR spectrum increases in complexity as evidenced by a doubling of the number of resonances. At a ligand:DNA ratio of 0.5:1 new resonances appear (indicated by squares), whose intensities are one-third the intensity of the free DNA signals. At a ligand:DNA stoichiometry of 1:1 the intensity ratio of complexed DNA to free DNA is 1:1, indicating that one-half of the DNA exists as complex. Subsequent additions of 2-ImD to the sample increase intensities of the signals from complexed DNA as well as the signals from the bound ligand without appearance of new peaks. At a ligand:DNA ratio of 2:1 there are only one set of DNA resonances and one set of bound ligand resonances, which implies that a single ligand–DNA complex has been formed. Further, a comparison of the peak intensities of complexed to uncomplexed DNA as a function of concentration of added ligand shows that this single complex has a stoichiometry of 2:1. The sharpness of the ligand and DNA signals throughout the titration indicates that the ligands are in slow exchange on the NMR time scale.

Signal Assignments. 2D NOESY spectra of the complex in D_2O were acquired at 2 molar equiv of added ligand. 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[16] facilitated the assignment of the adenine H2 protons in the complex. Ligand amide NH protons and pyrrole H3 protons were assigned by intramolecular connectivities among these protons in the NOESY spectrum of the complex in H_2O. Ligand methylene protons (C20) were assigned via intermolecular contacts to the formyl proton (HF) on the opposite ligand of the 2:1 complex, as well as to A4 H2 (ligand 1) and A15 H2 (ligand 2). The chemical shift assignments for the ribose ring C1′ protons and aromatic protons of the DNA in the free duplex and in the 2:1 complex are shown in Table I. The

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

	H6/H8			H1′		
	free duplex	2:1 complex	Δδ	free duplex	2:1 complex	Δδ
strand 1						
C1	7.65	7.64	−0.01	5.80	5.79	−0.01
G2	7.96	8.00	+0.04	5.90	5.90	0.00
C3	7.35	7.36	+0.01	5.49	5.82	+0.33
A4	8.25	8.51	+0.26	5.93	5.51	−0.42
A5	8.08	8.31	+0.23	6.02	5.80	−0.22
G6	7.45	7.64	+0.19	5.82	5.16	−0.66
T7	7.18	7.08	−0.10	6.02	5.64	−0.38
T8	7.28	7.23	−0.05	5.85	5.21	−0.64
G9	7.89	7.72	−0.17	5.65	5.63	−0.02
G10	7.79	7.73	−0.06	6.01	5.92	−0.09
C11	7.48	7.46	−0.02	6.21	6.22	+0.01
strand 2						
G12	7.99	7.97	−0.02	6.02	6.01	−0.01
C13	7.50	7.53	+0.03	6.08	6.08	0.00
C14	7.48	7.48	0.00	5.32	5.80	+0.48
A15	8.24	8.51	+0.27	5.95	5.60	−0.35
A16	8.15	8.39	+0.24	6.11	6.10	−0.01
C17	7.23	7.29	+0.06	5.71	5.48	−0.23
T18	7.41	7.18	−0.23	6.03	5.81	−0.22
T19	7.33	7.27	−0.06	5.85	5.24	−0.61
G20	7.93	7.78	−0.15	5.85	5.81	−0.04
C21	7.37	7.34	−0.03	5.81	5.72	−0.09
G22	7.96	7.95	−0.01	6.17	6.17	0.00

[a] Chemical shifts are given in ppm relative to the residual HDO signal at 4.80 ppm (25 °C).

Table II. Chemical Shift Assignments of the 2-ImD Ligands in the 2:1 Complex with d(CGCAAGTTGGC)·d(GCCAACTTGCG)

proton	ligand 1	ligand 2
HF	8.03	8.09
NH-1	10.23	10.41
H3-1	6.05	6.27
NH-2	9.65	8.94
NH-3	10.54	10.43
H3-3	6.46	6.44
NH-4	7.90	7.88
C(20)H[a]	1.23	1.23
HA[a,b]	9.37	9.32
	8.23	8.17

[a] Not stereospecifically assigned. [b] HA = amidinium NH_2 protons.

Table III. Ligand–DNA and Ligand–Ligand Intermolecular Contacts[a]

ligand 1	DNA	ligand 2
	Ligand–DNA	
	A4 C1′H	HA
	T19 C1′H	NH-4, H3-3
HF, NH-1	A5 C1′H	
H3-1	A5 C2H	H3-3
	T18 C1′H	NH-3
	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
H3-3	A16 C2H	H3-1
NH-4, H3-3	T8 C1′H	
HA	A15 C1′H	
	Ligand–Ligand	
C(20)H		HF
HF		C(20)H

[a] Identified in the H_2O NOESY acquired at 200-ms mixing time.

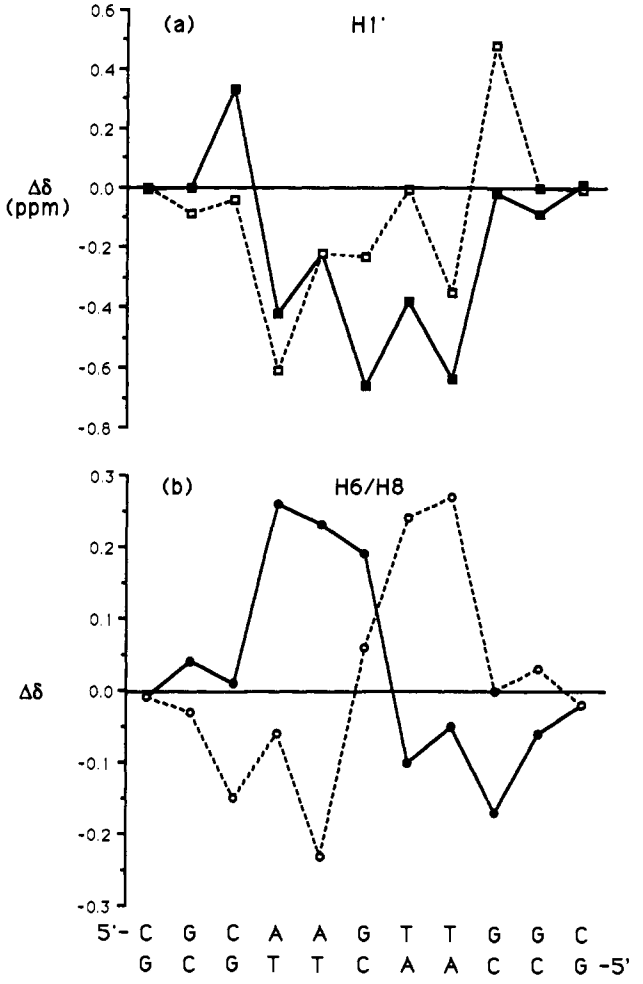


Figure 2. Graphical presentation of the ligand-induced chemical shift changes for binding of 2-ImD to the AAGTT duplex. The 5'-C1 to C11 strand is represented by solid lines and the 5'-G12 to G22 strand by dashed lines.

ligand-induced chemical shift changes for these protons are shown graphically in Figure 2. The chemical shifts for the bound ligands are given in Table II.

Intermolecular Contacts. The NOESY spectrum of the 2:1 complex in H_2O is shown in Figure 3. 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 ligand 1 and the H2 protons of A5 and A16, respectively, and between the H3-1 and H3-3 protons of ligand 2 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 C20 methylene protons on the opposite ligand confirm the antiparallel arrangement. These intermolecular contacts are shown diagrammatically in Figure 4. 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.

Molecular Modeling. Semiquantitative modeling of the complex was carried out using the Insight II/Discover model building and simulation package. All intermolecular contacts listed in Table III plus intramolecular ligand contacts were used as constraints. The structure obtained from the energy-minimization procedure described above is presented in Figure 5. As in similar models of distamycin A complexes with AT-rich DNA oligomers,[8] the charged end of each 2-ImD molecule lies deep in the minor groove, providing favorable contributions to the total energy of the system. For ligand 1, hydrogen bonds are formed between NH-1 and A5 N3, NH-2 and G6 N3, NH-3 and T7 O2, and NH-4 and T8 O2 and between N1 of the imidazole and the NH_2 group of G6. Modeling indicates only one hydrogen bond between ligand 2 and DNA (NH-3 to T18 O2).

The stacking of the ligand molecules with respect to one another, as shown by molecular modeling, is similar to the staggered arrangement found in the 2:1 complexes formed between distamycin A and d(CGCAAATTGGC) and d(CGCAAATTTGCG)$_2$.[8]

Titration of the AAATT Site with 2-ImD. Figure 6 shows the results of adding up to 2 molar equiv of 2-ImD to an NMR sample containing the duplex d(CGCAAATTGGC)·d(GCCAATTTGCG). The spectrum of uncomplexed DNA was acquired at 25 °C. Upon the addition of 0.25 molar equiv of ligand, all of the DNA signals broaden significantly. Raising the temperature to 40 °C results in sharper resonances; hence, the remainder of the titration was carried out at this temperature. This behavior indicates that the ligands are near fast exchange at 40 °C and lowering the temperature would broaden the lines as the ligands enter intermediate exchange. It is unlikely that the ligands would approach slow exchange at any experimentally obtainable temperature. All subsequent additions of ligand up to a stoichiometry of 2:1 continue to broaden the DNA signals. The ligand signals that appear in the region 6.5–7.2 ppm are extremely broad throughout the titration. Complete characterization of the complex in this exchange regime is extremely difficult.

At low ratios of 2-ImD it is possible that intermolecular exchange of the drug is responsible for the broadening. However, the persistence of broadening of a 2:1 molar ratio, a saturating level for distamycin A with AAATT or 2-ImD with AAGTT, indicates that binding is either weaker than distamycin A on this site or nonspecific or both. Further quantitative measurements of binding affinity are needed to settle this point.

Discussion

In our laboratory we have recently found[17] highly cooperative binding of distamycin A to the sequence d(CGCIIICCGGC)·d(GCCIICCCGCG). This result implies that the minor groove of a DNA oligomer containing a sequence of IC base pairs is probably wider than an equivalent sequence of AT base pairs. Since a very narrow groove is required for distamycin A to be in van der Waals contact with both sides of the groove, it is reasonable that a wider groove would favor the 2:1 complex. If, as expected, the groove width in AAGTT is significantly larger than in AAATT, then it should be possible to design a ligand that can be expected to bind in the 2:1 mode rather than the conventional 1:1 mode.

Titration of the duplex d(CGCAAGTTGGC)·d(GGCAACTTGCG) with 2-ImD yields only one set of new resonances in the 1D NMR spectrum. These resonances are associated with the 2:1 mode as was revealed by the analysis of 2D NMR spectra. This shows that even at low ligand:DNA stoichiometries (e.g., 0.25:1) the 2:1 form is the dominant binding mode, indicating that the two ligands bind highly cooperatively. This behavior is in contrast to complexes between distamycin A and the AAATT and AAATTT duplexes[8] in which the 2:1 binding mode is not significantly populated until 0.5–0.75 equiv of drug has been added. This is consistent with our expectations based on groove width arguments.

The ligand binding sites of the 2-ImD complex are clearly defined by the intermolecular ligand–DNA NOEs given in Table III. For ligand 1, the pyrrole protons H3-1 and H3-3 give strong NOEs to A5 C2H and G6 C1'H, and A16 C2H and T8 C1'H, respectively. Therefore, the pyrrole–imidazole–pyrrole ring system of ligand 1 spans the A5-G6-T7 sequence with the formyl group oriented on the 5' side of the strand. Similarly, for ligand 2, H3-1 and H3-3 pyrrole protons are spatially close to A16 C2H and C17 C1'H, and A5 C2H and T19 C1'H, respectively. These data indicate that the pyrrole–imidazole–pyrrole ring system of ligand 2 spans the same segment (A5-G6-T7) as ligand 1 with the formyl group oriented to the 3' side of this strand. NOEs between the methylene protons on C20 of each ligand and the formyl proton

(17) Fagan, P. A.; Wemmer, D. E. *J. Am. Chem. Soc.* **1992**, *114*, 1080.

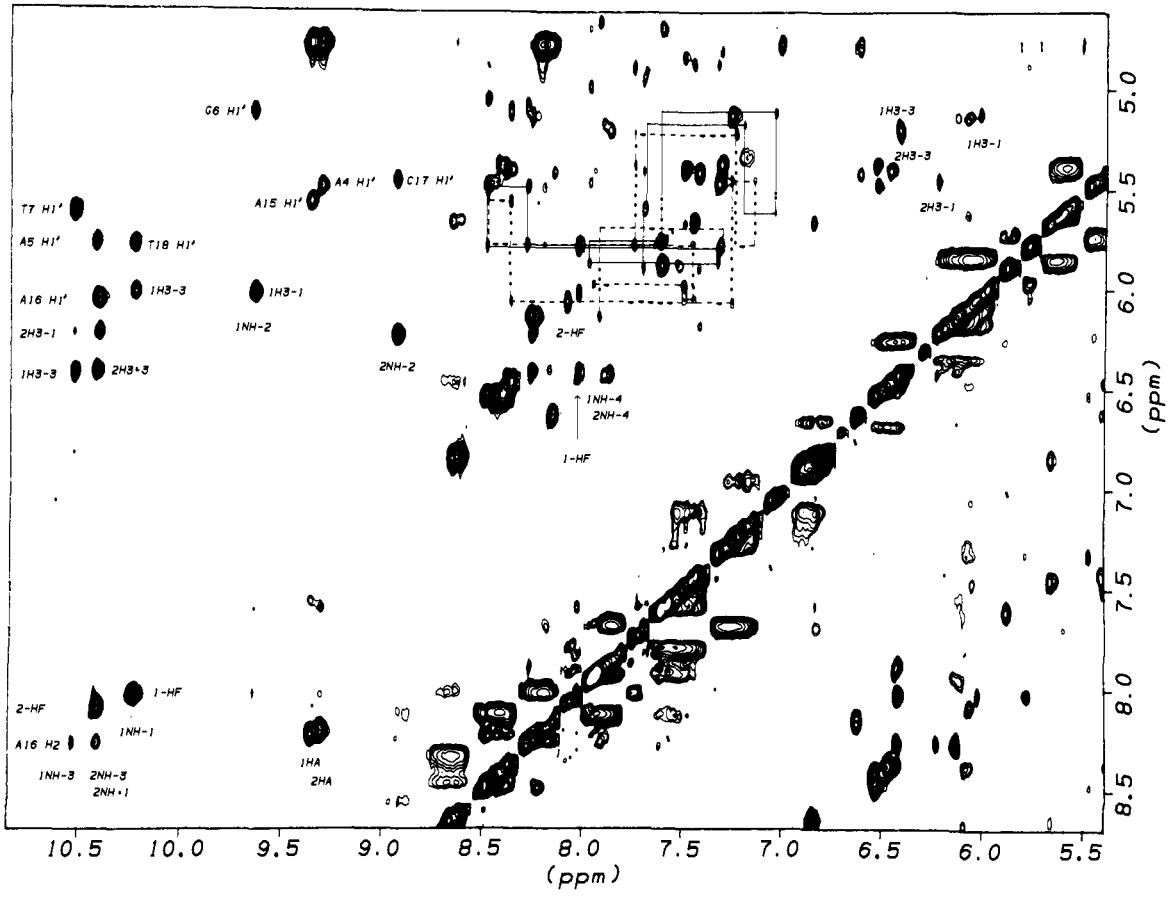


Figure 3. Expansion of the aromatic and amide region of a NOESY spectrum of the 2:1 complex of 2-ImD with d(CGCAAGTTGGC)·d(GCCAACTTGCG) (in 90% H_2O/10% D_2O, 25 °C, τ_m = 200 ms). Sequential aromatic to C1'H connectivities for the AAGTT strand are denoted by solid lines while for the AACTT strand they are shown as dashed lines. Crosspeaks are labeled according to structure **2** for ligand 1 and ligand 2. Labels above or below a crosspeak denote the chemical shift along the ω_2-axis (horizontal) while labels to the left or right of a peak indicate the chemical shift along the ω_1-axis (vertical).

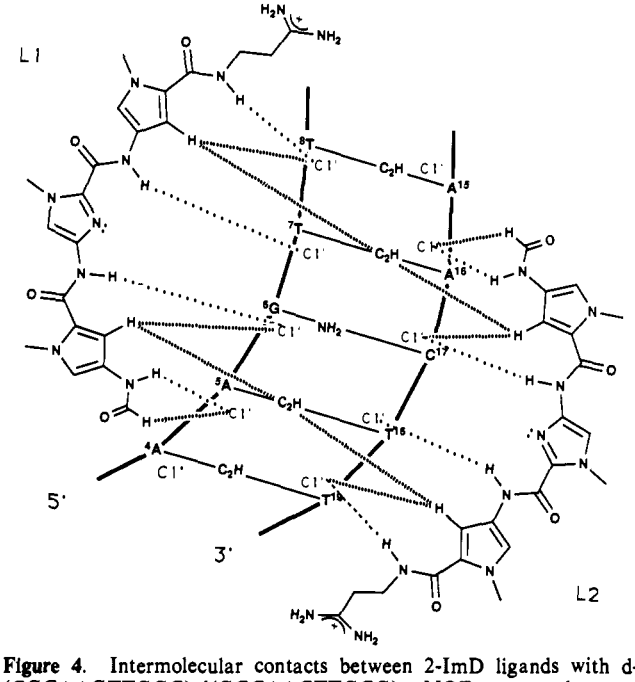


Figure 4. Intermolecular contacts between 2-ImD ligands with d(CGCAAGTTGGC)·d(GCCAACTTGCG). NOE contacts between amide protons of the ligands and DNA protons are indicated as dashed lines while contacts between other ligand protons and DNA are shown as stippled lines. Ligand–ligand contacts are not included.

on the opposite ligand confirm the relative orientations of the ligands. These data indicate that the two-drug binding site is in the minor groove of the central portion of this DNA oligomer.

The stability of the ligand–DNA complex is determined by several factors. The amide NH protons contribute favorably to the binding energy through hydrogen-bond formation to thymine O2 and adenine N3 atoms. Pyrrole H3 protons make close van der Waals contacts with C1'H and A C2H protons of the DNA. The positively charged amidinium group facilitates the attraction of the ligand to the polyanionic DNA; however, the extent to which electrostatic effects contribute to sequence specificity is uncertain. Stacking of the conjugated ring systems of the ligands in the 2:1 complex provides additional binding energy. In the 2:1 complex between 2-ImD and the AAGTT duplex, the number and nature of contacts between the ligands and the DNA are preserved when compared to those found in 2:1 complexes between distamycin A and AAATT and AAATTT duplexes.[8] Similar to these 2:1 complexes, there is no obvious distortion of the DNA helix or conformation as a result of ligand binding.

Minor groove binders such as distamycin A and netropsin bind with a high preference to AT-rich regions of B-DNA. These drugs show a significantly reduced affinity for binding sites that contain GC base pairs, due to the presence of the C2 amino group of guanine that protrudes from the floor of the minor groove preventing close contacts between the drug and the DNA. The most striking observation in the binding of 2-ImD to the AAGTT duplex is its recognition of a mixed GC- and AT-containing sequence.

Intermolecular ligand–DNA NOEs combined with molecular modeling confirm that, as was intended, the imidazole nitrogens of the two 2-ImD ligands are in close proximity to the C2 amino group of G6. However, only the imidazole N3 of ligand 1 is positioned optimally for hydrogen bonding. Comparison of 2-ImD binding to AAGTT and AAATT duplexes provides additional insight into the importance of this interaction.

Titration of the duplex d(CGCAAATTGGC)·d(GCCAATTTGCG) with 2-ImD results in increasingly broad

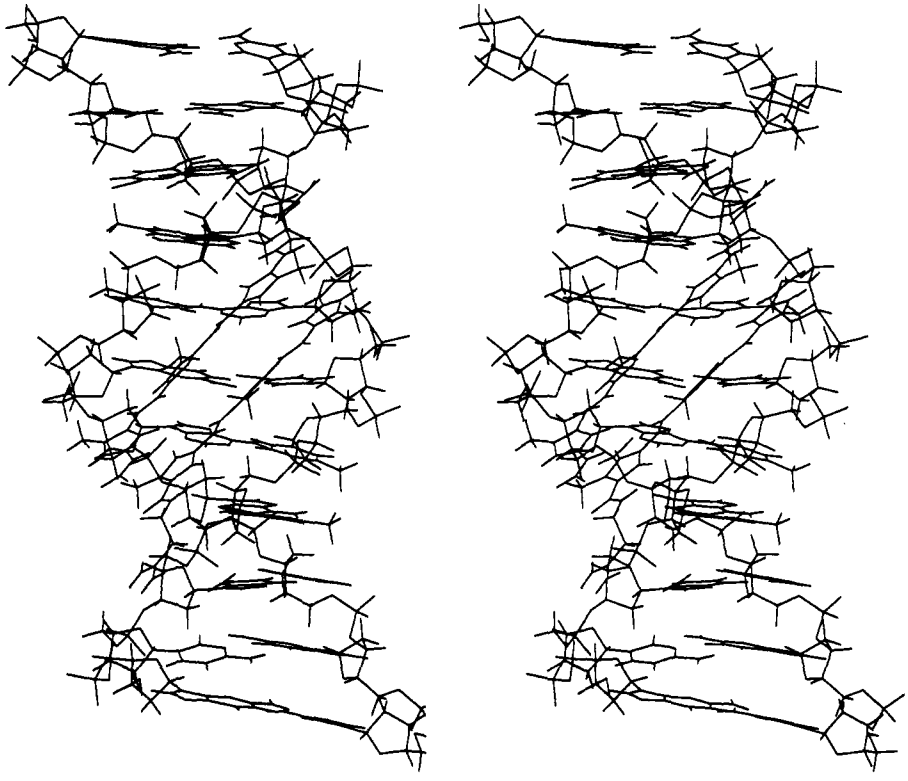


Figure 5. Stereo diagram of the molecular model of the 2:1 complex of 2-ImD with d(CGCAAGTTGGC)·d(GCCAACTTGCG) obtained by energy minimization with semiquantitative distance constraints from NOESY spectra (see text).

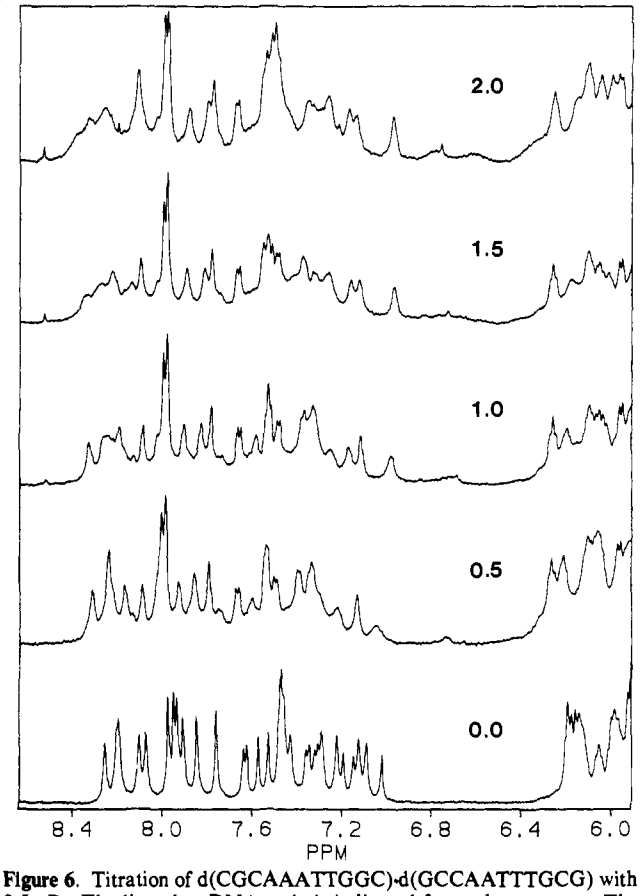


Figure 6. Titration of d(CGCAAATTGGC)·d(GCCAATTTGCG) with 2-ImD. The ligand to DNA ratio is indicated for each spectrum. The spectrum for uncomplexed DNA was acquired at 25 °C, while the titration was completed at 40 °C to sharpen the resonance lines.

lines in the 1D spectra. Rapid exchange of ligands, leading to broadening, indicates that 2-ImD does not bind specifically to the AAATT site. The absence of downfield adenine H8 protons, which are characteristic in a 2:1 complex, indicates that no 2:1 complex is formed between 2-ImD and AAATT. Distamycin A and 2-ImD are identical except for the replacement of a C–H of the central pyrrole ring with a nitrogen atom. Binding of 2-ImD to AAATT would place the electronegative imidazole nitrogen in the vicinity of N3 of adenine, close to the most negative electrostatic potential on the DNA. This suggests that the resulting repulsive electrostatic interaction between the DNA and the ligand(s), as well as the requirement for desolvation of the imidazole, disfavors binding of 2-ImD to AAATT.

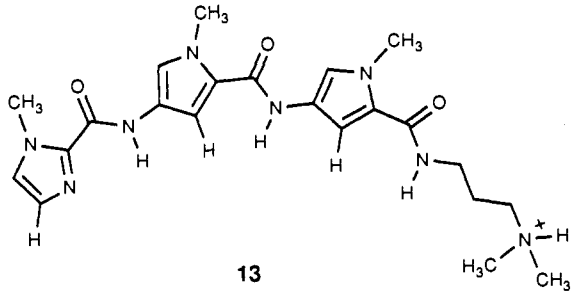
As pointed out above, distamycin A binds to the AAATT site with 2:1 ligand:DNA stoichiometry when greater than 0.5 equiv of ligand is added to the oligomer sample. Replacing an AT base pair in the center of the ligand binding site with a GC base pair will affect the local DNA structure, and a wider minor groove for the AAGTT oligomer is expected. The result that 2-ImD binds with high cooperativity in the 2:1 mode to AAGTT suggests that a variation in groove width alone cannot explain the difference in binding specificity of 2-ImD for the two DNA sites.

Due to steric interference, the presence of a C2 amino group of guanine in the minor groove lowers the binding affinity of distamycin A to sites containing GC base pairs. Recent results in our laboratory suggest that the presence of a GC base pair in an otherwise AT-rich site does not preclude formation of a 2:1 distamycin A complex.[18] In contrast, in the 2-ImD complex with AAGTT the nitrogen of the imidazole ring supplies a lone pair of electrons as a hydrogen-bond acceptor for the guanine C2 amino protons. Formation of a 2:1 complex between 2-ImD and the d(CGCAAGTTGGC)·d(GCCAACTTGCG) duplex necessarily places the imidazole ring of each ligand, and more specifically each imidazole N3 atom, in the vicinity of G6. This provides two potential hydrogen-bond acceptors for the G6 amino protons. Modeling of the complex suggests that only the imidazole N3 of ligand 1 is positioned for hydrogen-bond formation with the C2 amino protons of guanine G6. Nevertheless, the comparison of 2-ImD binding to AAATT and AAGTT indicates that, indeed, the imidazole nitrogens proximal to the NH_2 of guanine are responsible for the recognition of the GC-containing sequence.

(18) Geierstanger, B. H. Unpublished observations.

This is in contrast to previous studies of less specific imidazole-containing oligopeptides where the position of the GC base pairs within the ligand binding sites did not correlate with the prediction of hydrogen-bond formation between an imidazole N3 and guanine NH_2.

While this work was in progress, the distamycin A analog 2-ImN, (**13**), designed by Wade and Dervan,[19] was shown to bind specifically to TGACT·AGTCA.[20] This molecule binds with high



13

cooperativity to the minor groove of DNA (2:1 mode). The imidazole nitrogens of the two ligands each recognize a different guanine amino group. The results thus far contribute to our present understanding of sequence-specific recognition of DNA by minor groove binding ligands: (1) the 2:1 binding mode allows the ligands to optimize tight van der Waals contacts and hydrogen bonding with the minor groove surface independent of sequence-dependent variations in groove width; (2) the positive charge on the ligand is important for the initial attraction to DNA, but two positive charges at opposite ends of the ligand prevent the formation of the 2:1 complex; (3) the spacing of hydrogen donor and acceptor groups on the ligands must match those on the DNA allowing for stabilizing interactions similar to the "spine of hydration" in the free duplex.

Both 2-ImD and 2-ImN specifically recognize mixed GC- and AT-containing DNA sequences. The basic idea in the design of lexitropsins has therefore proven successful. We propose that binding to GC-containing sequences can be enhanced when a single hydrogen-bond acceptor per G amino group is strategically positioned for complexation. We emphasize the importance of the 2:1 binding mode for optimization of van der Waals contacts and of hydrogen bonding between ligands and DNA. On the basis of these findings, lexitropsin molecules can now be designed to specifically recognize many DNA sequences of four to five base pairs in length.

(19) Wade, W. S. Ph.D. Thesis, California Institute of Technology, 1989.
(20) Mrksich, M.; Wade, W. S.; Dwyer, T. J.; Geierstanger, B. H.; Wemmer, D. E.; Dervan, P. B. Manuscript in preparation.

Acknowledgment. This work was supported by NIH Grant GM-43129 (D.E.W.), 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.

Novel Dipyridophenazine Complexes of Ruthenium(II): Exploring Luminescent Reporters of DNA

Richard M. Hartshorn† and Jacqueline K. Barton*

Contribution from the Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125. *Received January 30, 1992*

Abstract: A series of ruthenium(II) complexes have been prepared which contain two phenanthroline ligands and a third bidentate ligand which is one of a set of derivatives of the parent dipyrido[3,2-*a*:2′,3′*c*]phenazine (DPPZ) ligand. The spectroscopic properties of these complexes in the presence and absence of DNA have also been characterized. The derivatives have been prepared by condensation of different diaminobenzenes or diaminopyridines with the synthetic intermediate bis(1,10-phenanthroline)(1,10-phenanthroline-5,6-dione)ruthenium(II). $[Ru(phen)_2DPPZ]^{2+}$, like $[Ru(bpy)_2DPPZ]^{2+}$, acts as a molecular "light switch" for the presence of DNA, displaying no detectable photoluminescence in aqueous solution but luminescing brightly on binding to DNA. None of the DPPZ derivatives prepared show comparable "light switch" enhancements, since some luminescence may be detected in aqueous solution in the absence of DNA. For some complexes, however, luminescence enhancements of a factor of 20–300 are observed on binding to DNA. For these and the parent DPPZ complexes, the large enhancements observed are attributed to a sensitivity of the ruthenium–DPPZ luminescent charge-transfer excited state to quenching by water; although these complexes show little or no luminescence in water, appreciable luminescence is found in acetonitrile. Other derivatives show little solvent sensitivity in luminescence, and these, like $Ru(phen)_3^{2+}$, display moderate enhancements (20–70%) on binding to DNA. $[Ru(phen)_2DPPZ]^{2+}$ and its derivatives all show at least biexponential decays in emission. Two binding modes have been proposed to account for these emission characteristics: a perpendicular mode where the DPPZ ligand intercalates from the major groove such that the metal–phenazine axis lies along the DNA dyad axis, and another, side-on mode where the metal–phenazine axis lies along the long axis of the base pairs.

There has been increasing attention given to the design of novel transition metal complexes which recognize and react with nucleic acids so as to develop new diagnostic and therapeutic agents.[1-7] Our laboratory has focused, in part, on the preparation of ruthenium complexes which bind to DNA by intercalation.[4-7] Ruthenium complexes provide very sensitive reporters of DNA in aqueous solution and may become particularly useful in de-

* To whom correspondence should be addressed.
† Present address: School of Chemistry, University of Melbourne, Parkville, Victoria 3052, Australia.

(1) Pyle, A. M.; Barton, J. K. *Prog. Inorg. Chem.* **1990**, *38*, 413.
(2) (a) Moser, H. E.; Dervan, P. B. *Science* **1987**, *238*, 645. (b) Sigman, D. S. *Acc. Chem. Res.* **1986**, *19*, 180. (c) Hecht, S. M. *Acc. Chem. Res.* **1986**, *19*, 383. (d) Pyle, A. M.; Long, E. C.; Barton, J. K. *J. Am. Chem. Soc.* **1989**, *111*, 4520. (e) Pyle, A. M.; Morii, T.; Barton, J. K. *J. Am. Chem. Soc.* **1990**, *112*, 9432.

0002-7863/92/1514-5919$03.00/0