Structural Analysis of Covalent Peptide Dimers, Bis(pyridine-2-carboxamidonetropsin)$(CH_2)_{3-6}$, in Complex with 5′-TGACT-3′ Sites by Two-Dimensional NMR

Tammy J. Dwyer,† Bernhard H. Geierstanger,† Milan Mrksich,‡ Peter B. Dervan,*,‡ and David E. Wemmer*,†

Contribution from the 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 May 28, 1993®

Abstract: The peptide pyridine-2-carboxamidonetropsin (2-PyN) binds specifically in the minor groove of 5′-(A,T)G-(A,T)C(A,T)-3′ sequences as a side-by-side antiparallel dimer. Tethering two 2-PyN ligands through the nitrogens of the central pyrrole rings with propyl, butyl, pentyl and hexyl linkers affords covalent peptide dimers, bis(pyridine-2-carboxamide-netropsin)$(CH_2)_{3-6}$, which bind in the minor groove of DNA with increased binding affinities and improved sequence specificities. Two-dimensional NMR studies of the complexes formed upon binding of these covalent peptide dimers to an oligonucleotide containing a 5′-TGACT-3′ site reveal that the dimeric peptides bind as intramolecular dimers with nearly identical geometry and peptide-DNA contacts as in the $(2\text{-PyN})_2$·5′-TGACT-3′ complex.

Recent NMR spectroscopic studies led to the discovery and structural characterization of a novel binding motif for distamycin, in which two molecules bind in the minor groove of A,T-rich sites, effectively as a side-by-side antiparallel dimer.[1] Previous studies of 1:1 complexes by X-ray diffraction and NMR had shown that distamycin bound in unusually narrow minor groove regions, making close contacts both with the sides and bottom of the groove. The complementarity of shape and functional groups of the groove and the ligand was seen as the dominant determinant of sequence specificity and high binding affinity.[2,3] The 1:1 structures were useful in the design of oligopeptides for recognition of longer A,T tracts;[4] however, efforts to design peptides capable of sequence specific recognition of mixed A,T and G,C sequences using the 1:1 peptide–DNA models have been less successful.[5,6] NMR studies of a variety of different A,T-rich sequences indicated a coupling of the 2:1 binding mode to groove width, although in all complexes studied the basic features stabilizing the bound distamycin are similar to those present in 1:1 complexes: stacking of the ligand with DNA and with the other ligand, hydrogen bonds to DNA, and electrostatic interactions. Utilizing the 2:1 binding mode a second generation of peptides which can recognize mixed A,T and G,C sequences has now been designed.[7-10]

Peptides for 5′-TGTCA-3′ Recognition. Two analogs of distamycin have been prepared in which the terminal pyrrole ring has been replaced by an imidazole (2-ImN = 1-methylimidazole-2-carboxamidonetropsin) or pyridine (2-PyN = pyridine-2-carboxamidonetropsin) ring. Footprinting and affinity cleavage studies showed that these ligands bind to 5′-(A,T)G(A,T)C(A,T)-3′ sites, showing symmetric protection and cleavage patterns, which is expected in a 2:1 peptide–DNA complex.[7] Recently the $(2\text{-ImN})_2$·5′-TGACT-3′ complex was analyzed by two-dimensional NMR, and the antiparallel, side-by-side arrangement of the two ligands was directly verified.[7b] Examination of the model for this complex suggested that the two ligands could be covalently linked without disrupting any of the interactions with the DNA. It was expected that the binding affinities of such covalently linked peptides might be enhanced over binding by two monomeric ligands. Indeed, footprinting results demonstrate that the covalent peptide dimers bis(pyridine-2-carboxamidonetropsin)$(CH_2)_{3-6}$ ($(2\text{-PyN})_2$-C3, $(2\text{-PyN})_2$-C4, $(2\text{-PyN})_2$-C5, and $(2\text{-PyN})_2$-C6) bind the 5′-TGTCA-3′ site with higher affinity than does the monomeric peptide 2-PyN (Figure 1).[10] These results are consistent with intramolecular dimeric binding by the covalently linked peptides but do not exclude the possibility of intermolecular dimeric binding. We report here direct characterization by one- and two-dimensional NMR of the $(2\text{-PyN})_2$·5′-TGACT-3′ complex and the complexes formed upon binding of the four covalent peptide dimers $(2\text{-PyN})_2$-C3, $(2\text{-PyN})_2$-C4, $(2\text{-PyN})_2$-C5, and $(2\text{-PyN})_2$-C6 to an oligonucleotide containing a 5′-TGACT-3′ binding site.

† University of California.
‡ California Institute of Technology.
® Abstract published in *Advance ACS Abstracts*, October 1, 1993.

(1) (a) Pelton, J. G.; Wemmer, D. E. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 5723–5727. (b) Pelton, J. G.; Wemmer, D. E. *J. Am. Chem. Soc.* **1990**, *112*, 1393–1399.

(2) (a) Kopka, M. L.; Yoon, C.; Goodsell, D.; Pjura, P.; Dickerson, R. E. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 1376–1380. (b) Kopka, M. L.; Yoon, C.; Goodsell, D.; Pjura, P.; Dickerson, R. E. *J. Mol. Biol.* **1985**, *183*, 553–563. (c) Coll, M.; Frederick, C. A.; Wang, A. H.-J.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 8385–8389.

(3) (a) Patel, D. J.; Shapiro, L. *J. Biol. Chem.* **1986**, *261*, 1230–1240. (b) Klevitt, R. E.; Wemmer, D. E.; Reid, B. R. *Biochemistry* **1986**, *25*, 3296–3303. (c) Pelton, J. G.; Wemmer, D. E. *Biochemistry* **1988**, *27*, 8088–8096.

(4) (a) Schultz, P. G.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 6834–6837. (b) Youngquist, R. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1985**, *107*, 5528–5529. (c) Youngquist, R. S.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 2565–2569. (d) Youngquist, R. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1987**, *109*, 7564–7566.

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

(6) (a) Lown, J. W.; Krowicki, K.; Bhat, U. G.; Skorobogaty, A.; Ward, B.; Dabrowiak, J. C. *Biochemistry* **1986**, *25*, 7408–7416. (b) Kissinger, K.; Krowicki, K.; Dabrowiak, J. C.; Lown, J. W. *Biochemistry* **1987**, *26*, 5590–5595. (c) Lee, M.; Chang, D. K.; Hartley, J. A.; Pon, R. T.; Krowicki, K.; Lown, J. W. *Biochemistry* **1988**, *27*, 445–455.

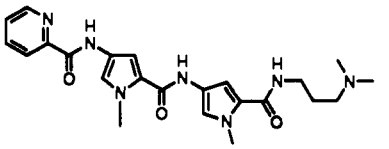
(7) (a) Wade, W. S.; Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 8783–8794. (b) Mrksich, M.; Wade, W. S.; Dwyer, T. J.; Geierstanger, B. H.; Wemmer, D. E.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 7586–7590. (c) Wade, W. S.; Mrksich, M.; Dervan, P. B. *Biochemistry*. In press.

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

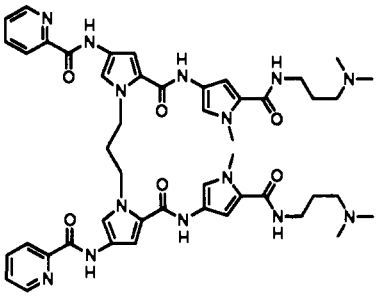
(9) (a) Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **1993**, *115*, 2572–2576. (b) Geierstanger, B. H.; Dwyer, T. J.; Bathini, Y.; Lown, J. W.; Wemmer, D. E. *J. Am. Chem. Soc.* **1993**, *115*, 4474–4482. (c) Geierstanger, B. H.; Jacobsen, J.-P.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E. Manuscript in preparation.

(10) Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.*, preceding paper in this issue.

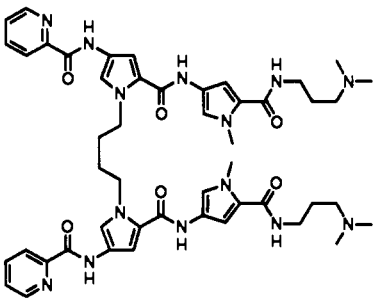
0002-7863/93/1515-9900$04.00/0



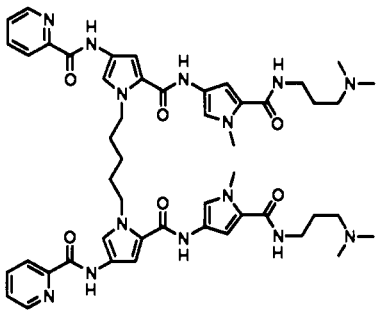
2-PyN



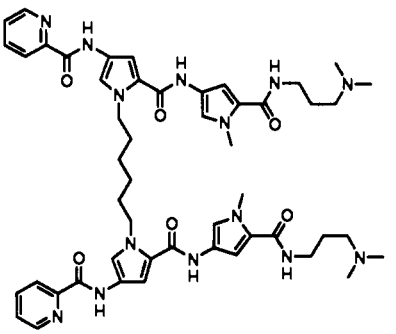
(2-PyN)$_2$-C3



(2-PyN)$_2$-C4



(2-PyN)$_2$-C5



(2-PyN)$_2$-C6

Figure 1. Synthetic peptide pyridine-2-carboxamidonetropsin (2-PyN) and covalent peptide dimers bis(pyridine-2-carboxamidonetropsin)$(CH_2)_{3-6}$ ((2-PyN)$_2$-C3, (2-PyN)$_2$-C4, (2-PyN)$_2$-C5, and (2-PyN)$_2$-C6).

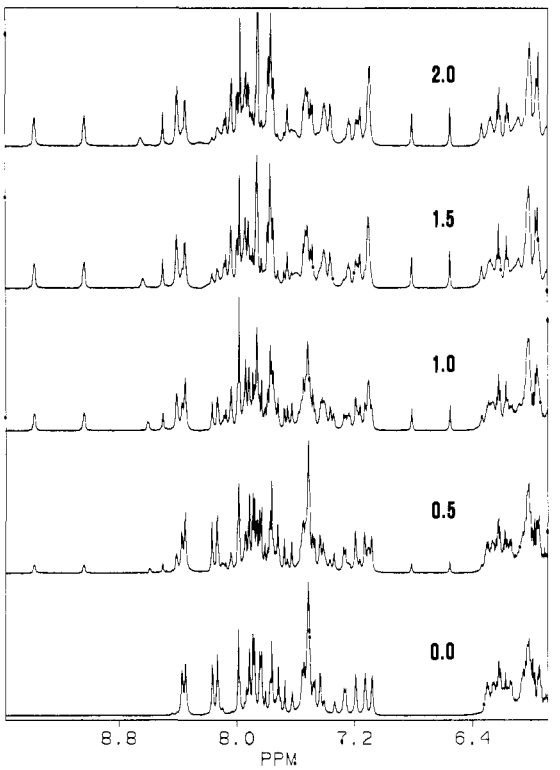


Figure 2. Aromatic region of NMR spectra acquired at several points in a titration of d(GCATGACTCGG):d(CCGAGTCATGC) with 2-PyN at 25 °C.

Experimental Section

Sample Preparation. DNA samples for the NMR titrations were prepared by dissolving each undecamer sample in 0.25 mL of 20 mM sodium phosphate buffer (pH 7.0) and lyophilizing to dryness. For experiments carried out in D_2O the resulting solid was lyophilized twice from 99.9% D_2O (Cambridge Isotope Laboratories) and subsequently redissolved in 0.5 mL of 99.96% D_2O (Cambridge Isotope Laboratories). For experiments in H_2O the solid was redissolved in 90% H_2O and 10% D_2O to a final volume of 0.5 mL.

Stock solutions for the NMR titrations were prepared by dissolving typically 1–2 mg of each ligand (2-PyN, (2-PyN)$_2$-C3, (2-PyN)$_2$-C4, (2-PyN)$_2$-C5, and (2-PyN)$_2$-C6) in 100–150 μL 99.96% D_2O containing 10 mM sodium phosphate buffer (pH 7.0). The concentration of each stock solution was determined independently by absorbance at 302 nm (an extinction coefficient of 3.5×10^4 M^{-1} cm^{-1} was assumed for the 2-PyN monomer and 7.0×10^4 M^{-1} cm^{-1} for the 2-PyN dimers). Extinction coefficients for d(GCATGACTCGG) and d(CCGAGTCATGC) were calculated to be 1.05×10^5 M^{-1} cm^{-1} and 1.02×10^5 M^{-1} cm^{-1}, respectively.[11] The concentrations of the double-stranded DNA samples were 2 mM, 1 mM, and 0.5 mM (as determined by absorbance at 260 nm at 80 °C) for the experiments involving 2-PyN, (2-PyN)$_2$-C6 and (2-PyN)$_2$-C3, (2-PyN)$_2$-C4, and (2-PyN)$_2$-C5, respectively.

1D NMR Titrations. NMR samples containing the DNA oligomer were titrated individually with 2-PyN or the covalently linked peptide dimers (in increments of 0.25 mol equivs) up to a ligand/DNA ratio of approximately 2:1 and 1:1, respectively. 1D spectra were acquired at 25 °C using 4096 complex points, 128–256 scans, and a spectral width of 6024 Hz. Suppression of the residual HDO resonance was achieved via application of a presaturation pulse during the 2.0-s recycle delay.

2D NOESY Spectra. NOESY spectra of the (2-PyN)$_2$·5′-TGACT-3′ complex and of the (2-PyN)$_2$-C6·5′-TGACT-3′ complex in D_2O were acquired on a Bruker AMX-600 spectrometer, while NOESY spectra of the (2-PyN)$_2$-C3·5′-TGACT-3′ complex in D_2O were acquired on a GN-500 spectrometer (General Electric Instruments). Spectra were acquired at ligand/DNA ratios of 2:1 for the monomer and 1:1 for the covalent peptide dimers (25 °C), and phase-sensitive detection was accomplished using the standard TPPI pulse sequence.[12] A mixing time of 200 ms was used for all D_2O NOESY spectra. For each t_1 value 64 scans were signal

(11) Warshaw, M.; Cantor, C. *Biopolymers* **1970**, *9*, 1079–1103.
(12) Drobny, G.; Pines, A.; Sinton, S.; Weitekamp, D. P.; Wemmer, D. E. *Faraday Symp. Chem. Soc.* **1979**, *13*, 49–55.

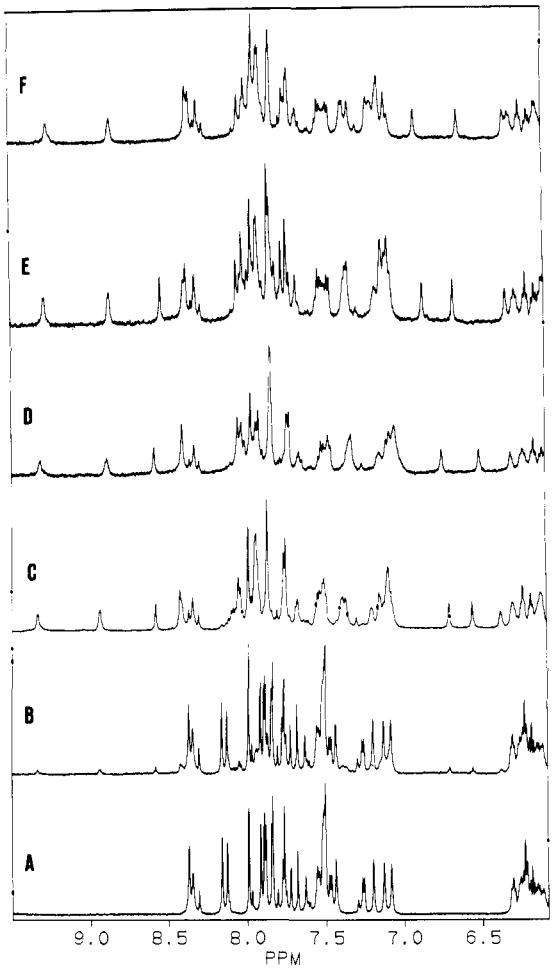


Figure 3. Aromatic region of NMR spectra acquired upon titration of d(GCATGACTCGG):d(CCGAGTCATGC) with the covalent peptide dimers: (a) free DNA, (b) 0.25:1 (2-PyN)$_2$-C6:DNA, (c) 1:1(2-PyN)$_2$-C6:DNA, (d) 1:1 (2-PyN)$_2$-C5:DNA, (e) 1:1 (2-PyN)$_2$-C4:DNA, and (f) 1:1 (2-PyN)$_2$-C3:DNA.

averaged, taking 1024 complex points with a recycle delay of 2 s. Again suppression of the residual HDO signal was achieved via presaturation during the recycle and mixing delays. Typically 512 t_1 experiments were recorded and zero-filled to 1 K. Spectral widths of 6024 Hz (600 MHz) and 5000 Hz (500 MHz) were used. NOESY spectra of the (2-PyN)$_2$·5′-TGACT-3′ complex and of the (2-PyN)$_2$-C6·5′-TGACT-3′ complex in water (25 °C) were acquired and processed as described previously.[7b]

Distance Restraints. Intermolecular distance restraints were generated from the volume integrals of the crosspeaks in the H_2O NOESY spectra of the (2-PyN)$_2$-C6·5′-TGACT-3′ complex as described previously.[7b] A total of 30 NOESY-derived restraints were evaluated (eight intramolecular ligand restraints and 19 intermolecular ligand–DNA restraints). Hydrogen bonds for standard Watson–Crick base pairing were included as NOE restraints. Listings of the restraints and the achieved distances are available as supplementary material.

Structure Refinement. Double-helical B-form DNA was constructed using the Biopolymer module of InsightII (Biosym) and was used as the starting structure for modeling. The (2-PyN)$_2$-C6 ligand was constructed from the coordinates of the 2-ImN ligands taken from the energy minimized NMR structure of the 2:1 complex with d(GCATGACTCGG):d(CCGAGTCATGC).[7b] The 2-ImN ligands were modified using the Builder module of InsightII to give 2-PyN ligands, and the hexyl linker was constructed and covalently attached to the ligands through the nitrogen atoms of the central pyrrole rings. The dimeric ligands were then docked in the minor groove of the binding site using the stereo feature of InsightII on the Silicon Graphics workstation. Energy minimizations incorporating the NOE distance restraints were performed using the Discover module of InsightII. A cutoff distance of 15 Å was used for nonbonded interactions, and the neighbor list was regenerated every 20 timesteps. Solvent effects were simulated through the use of a distance-dependent dielectric of the form $\epsilon = R$. The energy of the complex was minimized initially using 100 steps of a steepest descents algorithm and further using 15000 steps of conjugate gradient minimization with an NMR force constant of 25 (kcal/mol)/Å^2 (200 (kcal/mol)/Å^2 for Watson-Crick hydrogen bonds) to an rms derivative of <0.001 (kcal/mol)/Å^2.

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

Table I. Chemical Shift Assignments of the d(GCATGACTCGG):d(CCGAGTCATGC) Duplex[a]

	H6/H8			H1′		
	Free TGACT	(2-PyN)$_2$ TGACT	(2-PyN)$_2$-C6 TGACT	Free GACT	(2-PyN)$_2$ TGACT	(2-PyN)$_2$-C6 TGACT
			Strand 1			
G1	8.02	7.98	7.98	6.06	6.01	6.02
C2	7.57	7.52	7.51	5.77	5.85	5.84
A3	8.41	8.39	8.40	6.34	6.27	6.29
T4	7.12	7.36	7.36	5.75	5.24	5.35
G5	7.89	7.92	7.94	5.60	5.07	5.25
A6	8.20	8.03	8.04	6.26	5.61	5.52
C7	7.30	7.18	7.06	5.81	5.54	5.46
T8	7.48	7.16	7.15	6.11	5.31	5.35
C9	7.51	7.39	7.39	5.64	5.27	5.30
G10	7.92	7.86	7.86	5.73	5.64	5.64
G11	7.88	7.85	7.86	6.22	6.16	6.18
			Strand 2			
C12	7.81	7.76	7.75	6.07	6.01	6.00
C13	7.56	7.54	7.52	5.60	5.51	5.51
G14	7.95	7.93	7.93	5.58	5.64	5.67
A15	8.17	8.34	8.34	6.19	5.71	5.79
G16	7.54	7.99	8.03	5.88	6.03	6.11
T17	7.24	7.10	7.09	6.05	5.48	5.46
C18	7.59	7.10	7.10	5.69	5.44	5.40
A19	8.38	8.40	8.42	6.30	5.47	5.49
T20	7.17	7.10	7.11	5.80	5.39	5.40
G21	7.93	7.86	7.86	6.00	5.95	5.96
C22	7.56	7.49	7.49	6.27	6.22	6.22

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

Results

Titration of the 5′-TGACT-3′ Site with 2-PyN. Spectra obtained at several points in a titration of the 5′-TGACT-3′ duplex with 2-PyN reveal that upon addition of substoichiometric amounts of peptide, the spectra increase in complexity due to the appearance of a set of new resonances (Figure 2). The new signals that appear between 6.0 and 6.8 ppm are characteristic of the H3 pyrrole protons of the ligand. The peaks that appear downfield between 8.9 and 9.3 ppm belong to the pyridine H6 protons of the peptide. New resonances also appear in the aromatic region of the spectra (between 7.0 and 8.5 ppm) that correspond to DNA protons of the complex. The chemical shift assignments for protons in both free DNA and complexed DNA are presented in Table I.

The appearance of only a single set of pyrrole H3 and pyridine H6 protons, that continue to increase in intensity until 2 equiv of 2-PyN have been added, suggests that a unique peptide/DNA complex having a stoichiometry of 2:1 is formed. Furthermore, the behavior of 2-PyN in its titrations onto the 5′-TGACT-3′ site is completely analogous to the titration of 2-ImN on this site.[7b]

Titration of the 5′-TGACT-3′ Site with the Bis(pyridine-2-carboxamidonetropsin)(CH$_2$)$_{3-6}$ Peptides. Each of the covalently linked peptide dimers was titrated separately into samples of duplex DNA containing the 5′-TGACT-3′ binding site (Figure 3). The similarity of the spectra for all the peptides suggests that the length of the linker has little effect on complex geometry. As in the titration of 2-PyN onto the 5′-TGACT-3′ binding site, initial addition of a covalent peptide dimer leads to a doubling of resonances over much of the spectrum and the appearance of new signals. These new signals are identical in number (with the exception of the resonances associated with the linker) and similar in chemical shift, to those obtained on titration of the 5′-TGACT-3′ binding site with 2-PyN. In all cases, the complexes formed

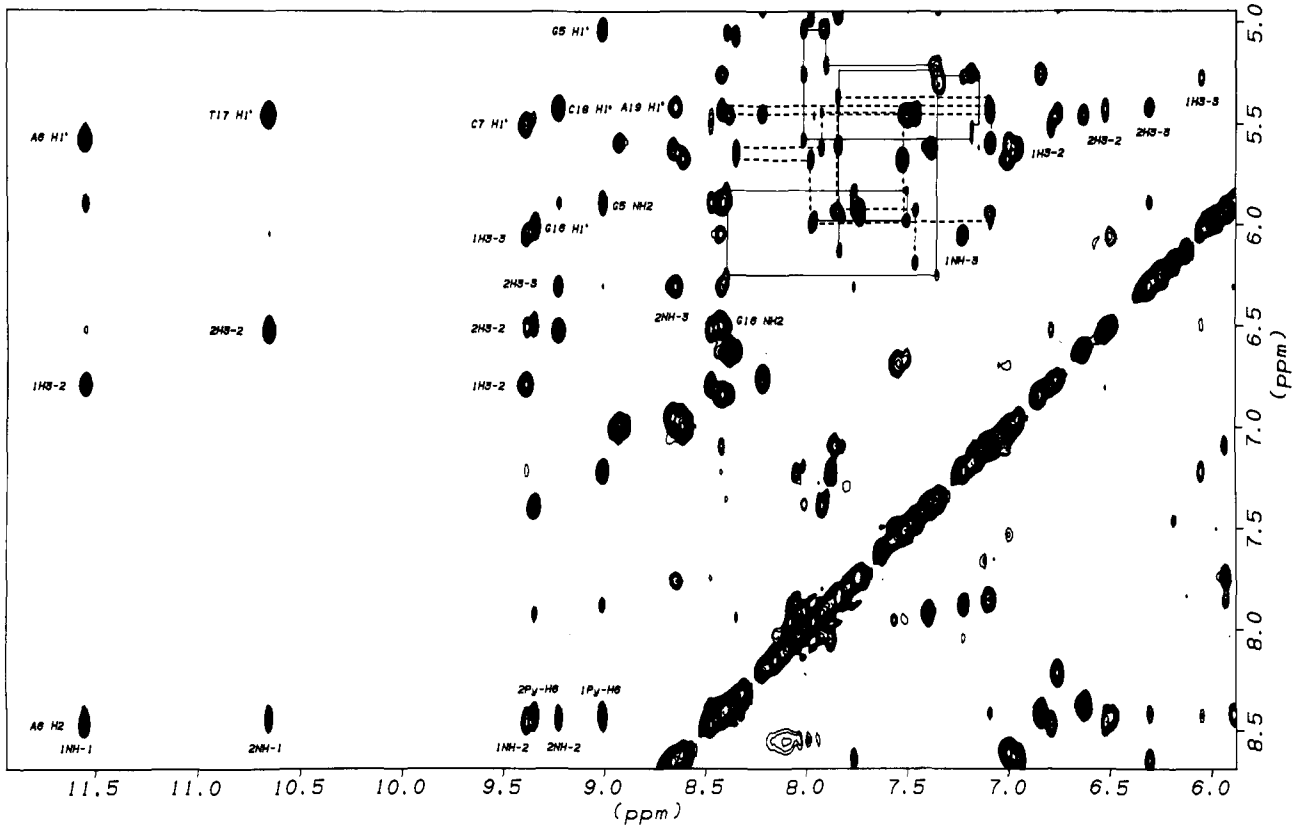


Figure 4. Expansion of the aromatic and amide region of a NOESY spectrum of the 2:1 2-PyN·d(GCATGACTCGG):d(CCGAGTCATGC) complex (in 90% H_2O/10% ^2H_2O; 15 °C; mixing time 200 ms). Sequential aromatic to C1′H connectivities for the 5′-TGACT-3′ strand are shown as solid lines; those for the 5′-AGTCA-3′ strand are shown as dashed lines. Crosspeaks are labeled according to their chemical shifts along ω_1 (vertical axis, label beside the peak) and along ω_2 (horizontal axis, label above or below the peak). Labeling conventions for peptides 1 and 2 are NH- for amide hydrogens, H3- for pyrrole hydrogens, and H6 for the pyridine hydrogen, numbered from the pyridine end to the ammonium end.

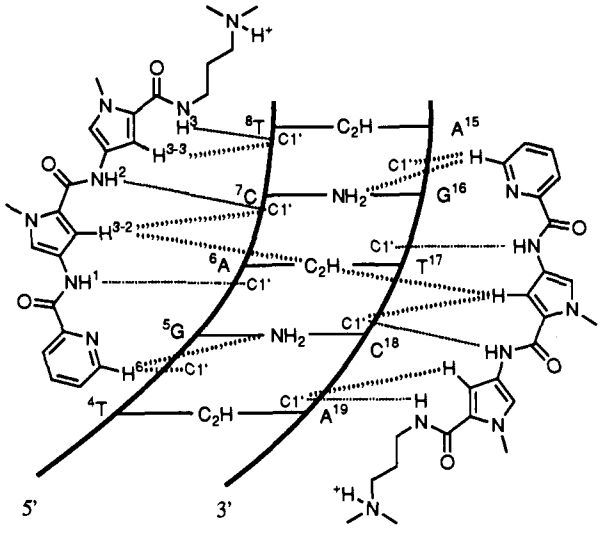


Figure 5. Schematic of selected intermolecular NOEs between 2-PyN peptide and d(GCATGACTCGG):d(CCGAGTCATGC) in the 2:1 2-PyN complex and the 1:1 (2-PyN)$_2$-C6 complex. Hashed lines represent contacts to peptide pyrrole C2H and pyridine C6H atoms and dotted lines represent contacts to peptide NH protons.

between the propyl-, butyl-, pentyl-, and hexyl-linked 2-PyN dimers and the 5′-TGACT-3′ duplex exhibit positive cooperativity since the singlet set of pyrrole H3 and pyridine H6 proton resonances appear upon the addition of as little as 0.25 mol equiv of dimer. It is evident that the DNA is completely complexed at a peptide dimer/DNA ratio of 1:1 (Figure 3). Since each dimeric 2-PyN ligand contains the equivalent of two monomeric 2-PyN ligands, it is apparent that both ends of the ligand are involved in binding. A comparison of the titration experiments suggests that the dimeric peptides bind in a mode nearly identical as in the (2-PyN)$_2$·5′-TGACT-3′ complex (Figures 2 and 3).

Signal Assignments. 2D NOESY spectra in D_2O were collected for the 2:1 complex between 2-PyN and the 5′-TGACT-3′ site, and for the 1:1 complex between (2-PyN)$_2$-C6 and the 5′-TGACT-3′ site. The DNA aromatic and C1′H protons in each complex were assigned from NOESY spectra via the sequential method.[13] The imino and adenine H2 protons in the complexes were assigned using 2D NOESY spectra obtained in H_2O. Peptide amide NH protons, pyrrole H3, and pyridine H6 protons were assigned by intramolecular connectivities among these protons in the H_2O NOESY spectra of the complexes. Ligand methylene protons of the propyl amine chain were assigned via intermolecular contacts to the pyridine H6 protons on the opposite ligand of the complexes. Table I presents the chemical shift assignments for the free oligonucleotide duplex, the 2:1 (2-PyN)$_2$·5′-TGACT-3′ complex, and the 1:1 (2-PyN)$_2$-C6·5′-TGACT-3′ complex.

Intermolecular Contacts in the (2-PyN)$_2$·5′-TGACT-3′ Complex. Figure 4 shows the NOESY spectrum of the 2:1 complex between 2-PyN and the 5′-TGACT-3′ site in H_2O. Similar to the 2:1 complex between 2-ImN and the 5′-TGACT-3′ site, the placement of the ligands on the DNA is facilitated by the observation of numerous intermolecular contacts.[7b] The orientations of the ligands are defined by the observation of strong NOE crosspeaks between the pyridine H6 protons and the C1′H and amino NH2 protons of G5 and G16 plus those between the H3-2 (ligand 1) and H3-3 (ligand 2) pyrrole protons and the H2 protons of A6 (Figure 5). 2D NOESY spectra obtained in H_2O provide further contacts between ligand amide protons and DNA protons along the minor groove of the central five base pair binding site, confirming the placement of the ligands.

Intermolecular Contacts in the (2-PyN)$_2$-C6·5′-TGACT-3′ Complex. The expanded region of a 2D NOESY spectrum of the

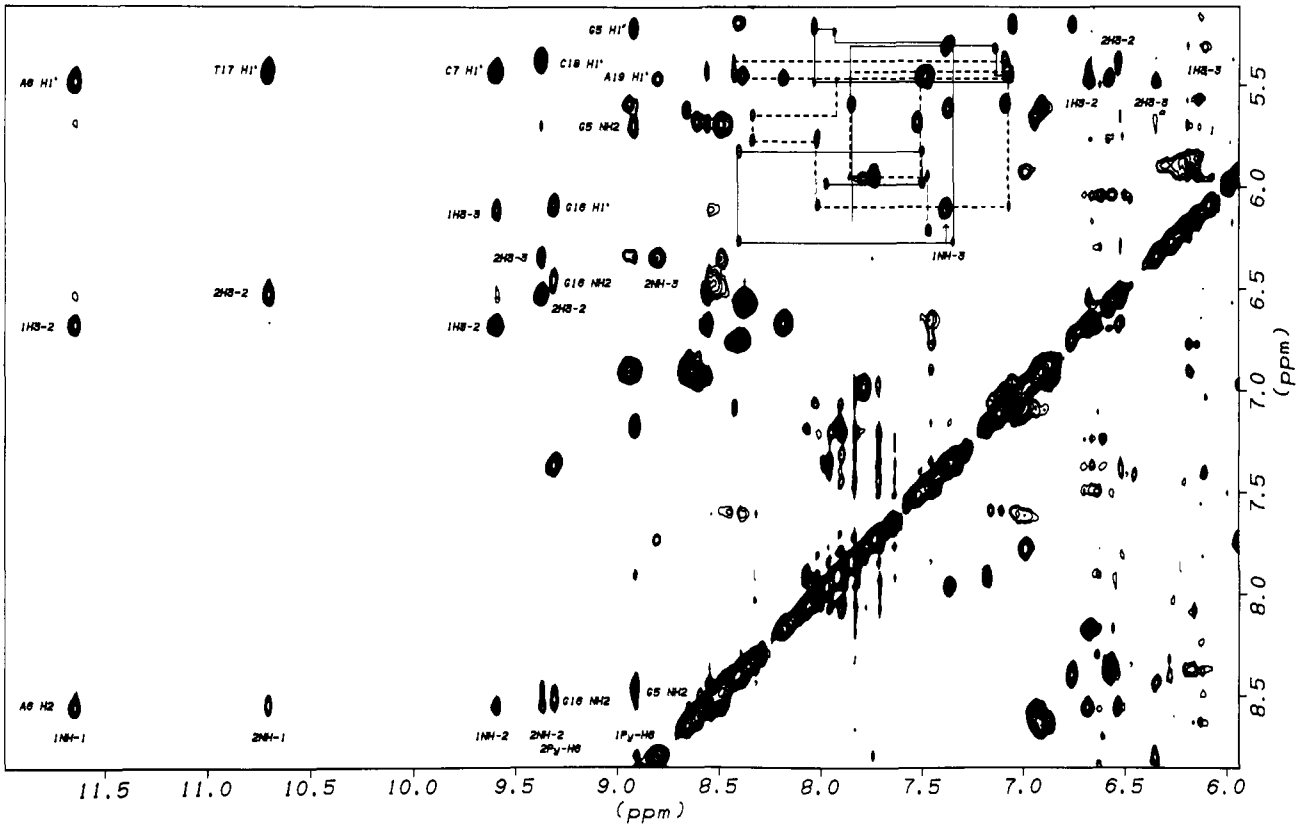


Figure 6. Expansion of the aromatic and amide region of a NOESY spectrum of the 1:1 (2-PyN)$_2$-C6·d(GCATGACTCGG):d(CCGAGTCATGC) complex. Spectral conditions and labeling conventions are the same as in Figure 4.

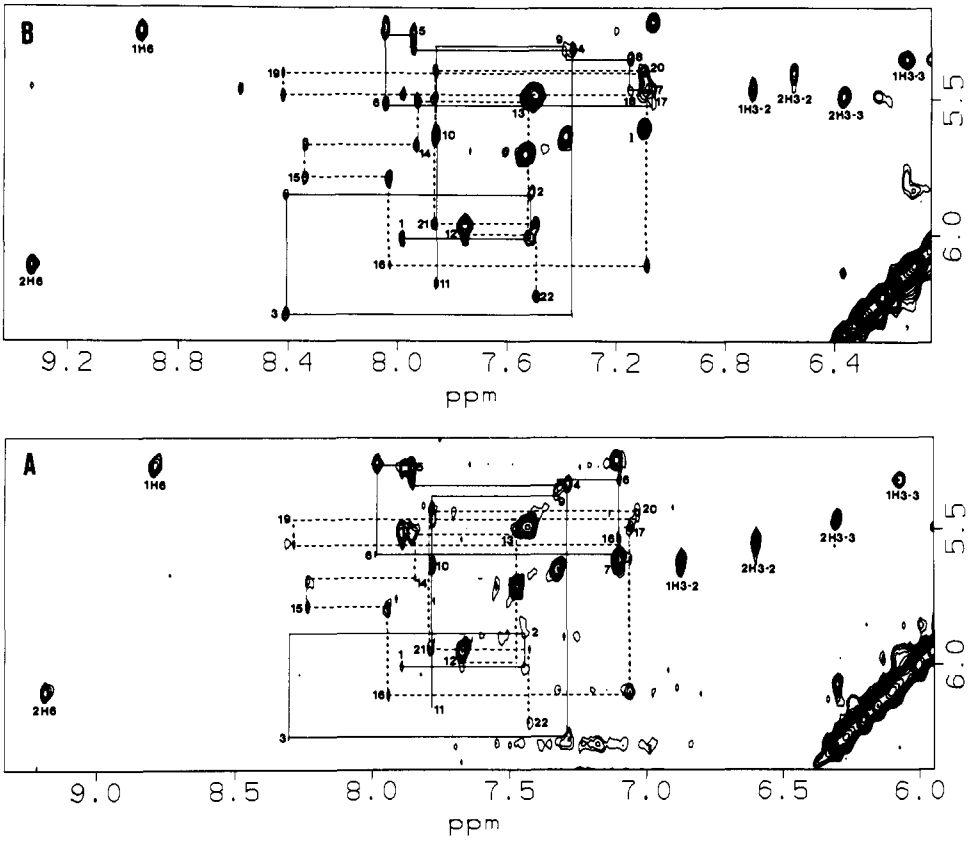


Figure 7. Expansions of the aromatic to C1'H regions of the NOESY spectra of (a) the 1:1 (2-PyN)$_2$-C3·d(GCATGACTCGG):d(CCGAGTCATGC) complex and (b) the 1:1 (2-PyN)$_2$-C6·d(GCATGACTCGG):d(CCGAGTCATGC) complex (both spectra in D_2O, 25 °C, mixing time 200 ms). The sequential connectivities of the 5'-TGACT-3' and 5'-AGTCA-3' strands are denoted by solid and dashed lines, respectively. Intraresidue aromatic to C1'H cross peaks are indicated with numbers. Peptide proton positions are labeled using the conventions noted in Figure 4.

(2-PyN)$_2$-C6·5'-TGACT-3' complex in H_2O (Figure 6) reveals that with no exceptions, nearly identical contacts are found when compared to the (2-PyN)$_2$·5'-TGACT-3' 2:1 complex. Therefore, the intermolecular contacts shown in Figure 5 are representative

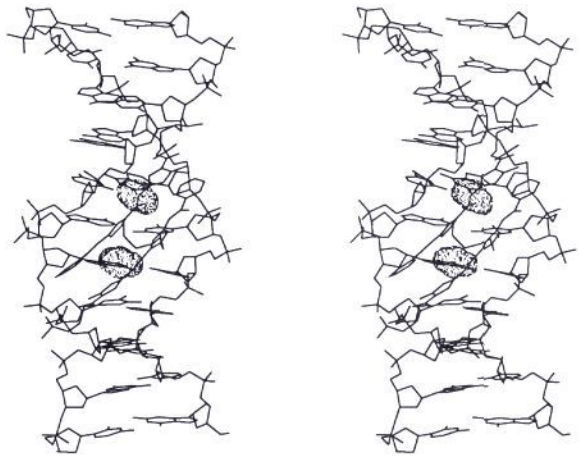


Figure 8. Stereodrawing of the 1:1 complex of $(2\text{-PyN})_2$-C6 with d(GCATGACTCGG):d(CCGAGTCATGC) obtained by energy minimization with semiquantitative distance restraints from NOESY spectra (see text). Hydrogens omitted for clarity. The amino groups of G5 and G16 specifically recognized by the pyridine nitrogens of the peptide are highlighted as van der Waals surfaces.

of both the 2:1 $(2\text{-PyN})_2$·5′-TGACT-3′ complex and the 1:1 $(2\text{-PyN})_2$-C6·5′-TGACT-3′ complex.

Geometry of the Covalent Peptide Dimer Complexes. The titration results implied that there are no significant differences in binding for the covalent 2-PyN dimers as a function of linker length. There are no large variations in chemical shifts of the proton resonances in the covalent dimer complexes. A 2D NOESY spectrum of $(2\text{-PyN})_2$-C3·5′-TGACT-3′ complex in D_2O was obtained for comparison with the corresponding spectrum of the $(2\text{-PyN})_2$-C6·5′-TGACT-3′ complex. A comparison of expansions of the aromatic to C1′H region of the NOESY spectra for both the $(2\text{-PyN})_2$-C3 and $(2\text{-PyN})_2$-C6 complexes shows that the intra- and intermolecular contacts observed for both complexes are nearly identical (Figure 7). Moreover, the similarity of the relative intensities of the NOE crosspeaks confirms that the geometry of the complexes is not significantly different.

Molecular Modeling. Due to the similarity of the contacts between the 2:1 complexes formed by 2-PyN and 2-ImN with the 5′-TGACT-3′ site, the $(2\text{-PyN})_2$·5′-TGACT-3′ complex was not modeled.[7b] However, semiquantitative modeling of the $(2\text{-PyN})_2$-C6·5′-TGACT-3′ complex was carried out using the InsightII/Discover model building and simulation package. All intermolecular contacts listed in Table II as well as eight intramolecular ligand contacts were used as restraints.

The covalently linked 2-PyN ligands stack side-by-side in the minor groove of the binding site in a manner identical to that found in all other peptide/DNA complexes.[1,7b,8,9b,c] Similar to the imidazole nitrogens of the $(2\text{-ImN})_2$·5′-TGACT-3′ complex, molecular modeling indicates that the pyridine nitrogens of the 2-PyN ligands in the $(2\text{-PyN})_2$-C6·5′-TGACT-3′ complex participate in hydrogen bonds with the 2-amino groups of G5 and G16 (Figure 8).[7b]

Discussion

Two-dimensional NMR experiments reveal that the monomeric ligand 2-PyN forms a single, unique complex with the DNA duplex d(GCATGACTCGG):d(CCGAGTCATGC). This complex is composed of two 2-PyN ligands bound simultaneously in the minor groove of the central five base pairs of the DNA. Titration experiments reveal that the 2:1 $(2\text{-PyN})_2$·5′-TGACT-3′ complex is formed with positive cooperativity. The observed intermolecular proton–proton contacts can be accounted for by a complex in which the ligands are oriented in a side-by-side

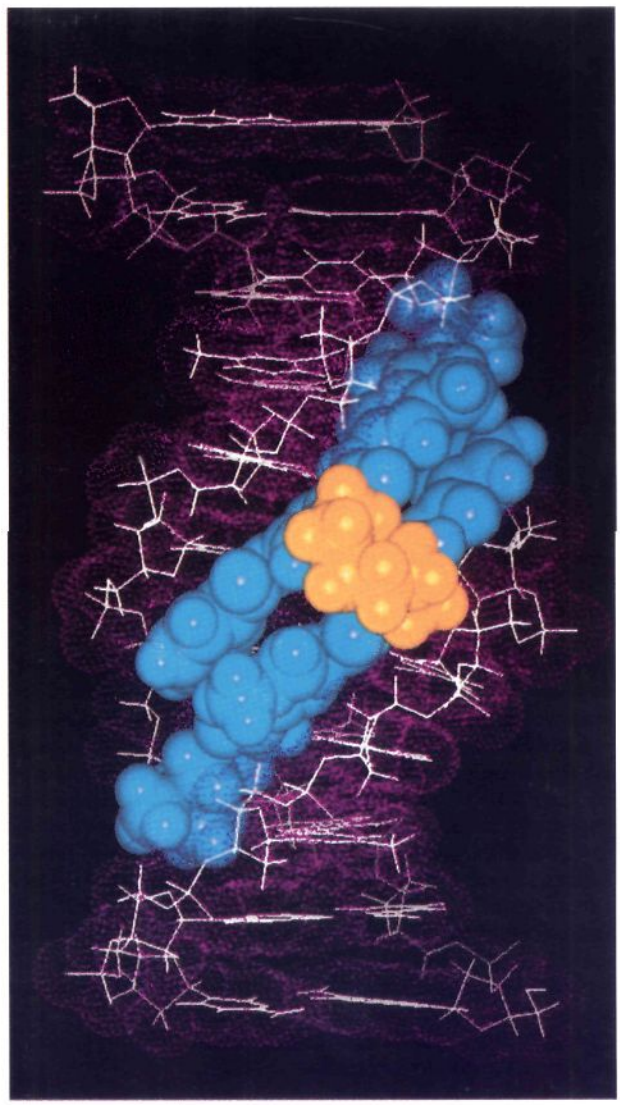


Figure 9. Van der Waals surface representation of the $(2\text{-PyN})_2$-C6·d(GCATGACTCGG):d(CCGAGTCATGC) complex shown in Figure 8. The two 2-PyN peptides are shown in blue and the hexyl linker is shown in yellow.

antiparallel arrangement and stacked such that the aromatic rings of one ligand overlap the amide bonds of the other ligand (Figures 8 and 9).

The four covalent peptide dimers bind to the 5′-TGACT-3′ site with nearly identical geometry and peptide–DNA contacts as in the $(2\text{-PyN})_2$·5′-TGACT-3′ complex. In all the complexes, one peptide spans the 5′-TGAC-3′ site, while the other peptide spans the 5′-AGTC-3′ site on the opposite strand. This geometry likely allows the pyridine nitrogen of each ligand to participate in a hydrogen bond with the 2-amino group of each guanine residue on the floor of the minor groove. This association via hydrogen bonds likely plays not only a role in the stabilization of the complex but also in the specific recognition of the 5′-TGACT-3′ site by the peptides. Further, the two cationic *N*,*N*-dimethylammonium groups of the ligands are directed toward opposite ends of the binding site, providing favorable electrostatic interactions with the minor groove.

A comparison of the binding affinities among the covalently linked 2-PyN dimers reveals that the length of the alkyl linker has little effect on overall complex stability.[10] The NMR results presented here suggest that the relative positions of the two side-by-side peptides and the specific peptide–DNA contacts are nearly identical for 2-PyN and the four linked peptides in complex with the 5′-TGACT-3′ site. While the alkyl tethers in these four dimers

Table II. Intermolecular Ligand–DNA Contacts Used in the Molecular Modeling of the (2-PyN)$_2$-C6-5′-TGACT-3′ Complex[a]

peptide 1	DNA	peptide 2
	A19 C1′H	NH-3, H3-3
H6	G5 C1′H	
H6	G5 NH_2	
	C18 C1′H	NH-2, H3-2
NH-1	A6 C1′H	
NH-1, NH-2, H3-2	A6 C2H	NH-1, H3-2
	T17 C1′H	NH-1
NH-2, H3-2	C7 C1′H	
	G16 C1′H	H6
	G16 NH_2	H6
NH-3, H3-3	T8 C1′H	

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

are able to accommodate the close contacts between the peptides and the minor groove, there remains the likely possibility that the linkers are not optimized. A second generation of linked peptides may display even higher binding affinities and improved sequence specificities.

Acknowledgment. We are grateful to the National Institutes of Health (GM-27681 to P.B.D. and GM-43219 to D.E.W.) and the National Foundation for Cancer Research for research support, a National Institutes of Health Research Service Award to M.M., and instrumentation grants from the U.S. Department of Energy, DEFG05-86ER75281, and the National Science Foundation, DMB 86-09305 and BBS87-20134.

Supplementary Material Available: Listings of the restraints and the achieved distances in the molecular modeling of the (2-PyN)$_2$-C6-5′-TGACT-3′ complex (1 page). Ordering information is given on any current masthead page. Coordinates of the molecular model are available upon request.