
Design of a G \cdot C-Specific DNA Minor Groove-Binding Peptide

Author(s): Bernhard H. Geierstanger, Milan Mrksich, Peter B. Dervan and David E. Wemmer

Source: *Science*, Oct. 28, 1994, New Series, Vol. 266, No. 5185 (Oct. 28, 1994), pp. 646-650

Published by: American Association for the Advancement of Science

Stable URL: <http://www.jstor.com/stable/2885448>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



American Association for the Advancement of Science is collaborating with JSTOR to digitize, preserve and extend access to *Science*

JSTOR

baboon with a high virus load in its PBMCs, widespread infection of lymphoid tissues, and a dramatic loss of CD4⁺ lymphocytes, clinical signs and symptoms including extensive fibromatosis were observed. Moreover, a second HIV-2-infected animal appears to be following a similar clinical course. These features closely resemble those in human and simian AIDS. Although other investigators have shown that baboons are susceptible to infection with HIV-2 strains (23), our results show viral persistence and pathogenesis in this animal. HIV-2 infection of baboons thus offers a promising, reproducible, and affordable animal model for studies of HIV persistence and pathogenesis. This model is particularly important when protection against infection, as well as disease development, must be assessed.

REFERENCES AND NOTES

1. H. J. Alter *et al.*, *Science* **226**, 549 (1984); P. N. Fultz *et al.*, *J. Virol.* **58**, 116 (1986); W. J. W. Morrow *et al.*, *AIDS Res. Hum. Retrovir.* **5**, 233 (1989).
2. M. B. Agy *et al.*, *Science* **257**, 103 (1992).
3. Current prices for chimpanzees are \$30,000 to \$65,000 per animal; rhesus monkeys and pig-tailed macaques are \$3,000 to \$4,000 each, and the latter are in limited supply. In contrast, baboons presently cost \$1100 to \$1800 and are in plentiful supply. Moreover, they are easily imported into the United States.
4. L. R. Frumkin *et al.*, *Virology* **195**, 422 (1993).
5. E. Cortes *et al.*, *N. Engl. J. Med.* **320**, 953 (1989); T. R. O'Brien, J. R. George, S. D. Holmberg, *J. Am. Med. Assoc.* **267**, 2775 (1992); P. G. Babu, N. K. Saraswathi, F. Devapriya, T. J. John, *Indian J. Med. Res.* **97**, 49 (1993); M. Grez *et al.*, *J. Virol.* **68**, 2161 (1994).
6. V. M. Hirsch, R. A. Olmsted, M. Murphey-Corb, R. H. Purcell, P. R. Johnson, *Nature* **339**, 389 (1989); M. Peeters, P. Piot, G. van der Groen, *AIDS* **5**, S29 (1991).
7. R. C. Desrosiers, *Annu. Rev. Immunol.* **8**, 557 (1990); M. B. Gardner and P. A. Luciw, *Fed. Am. Soc. Exp. Biol. Proc.* **3**, 2593 (1989).
8. J. Li, C. I. Lord, W. Haseltine, N. L. Letvin, J. Sordroski, *J. Acquired Immune Defic. Syndr.* **5**, 639 (1992); R. Shibata and A. Adachi, *AIDS Res. Hum. Retrovir.* **8**, 403 (1992).
9. P. Putkonen *et al.*, *J. Acquired Immune Defic. Syndr.* **2**, 366 (1989); C. Stahl-Hennig *et al.*, *AIDS* **4**, 611 (1990); J. Livartowski *et al.*, *Cancer Detect. Prev.* **16**, 341 (1992).
10. J. McClure *et al.*, abstract, Tenth Annual Symposium on Nonhuman Primate Models for AIDS, San Juan, Puerto Rico, 1992.
11. B. A. Castro, M. Nepomuceno, N. W. Lerche, J. W. Eichberg, J. A. Levy, *Virology* **184**, 219 (1991).
12. The UC2 strain was originally recovered from a dually infected individual after its passage in the GEM T lymphoblastoid cell line [L. Evans, *Lancet* **ii**, 1389 (1988)]. The UC3 strain more closely resembles a primary isolate in that it has been passaged only in human PBMCs. UC3 is described elsewhere (17).
13. UC12 and UC14 are primary HIV-2 isolates recovered from patients from the Gambia (S. W. Barnett, H. Whittle, J. A. Levy, unpublished results).
14. P. Racz *et al.*, *Prog. Allergy* **37**, 81 (1986); L. V. Chalifoux *et al.*, *Am. J. Pathol.* **128**, 104 (1987); D. J. Ringle *et al.*, *ibid.* **134**, 373 (1989); M. S. Wyand *et al.*, *ibid.*, p. 385.
15. D. Blackbourn, S. W. Barnett, J. A. Levy, unpublished observations.
16. G. Pantaleo *et al.*, *Nature* **362**, 355 (1993); J. Embretson *et al.*, *ibid.*, p. 359.
17. Y. J. Rosenberg *et al.*, *AIDS Res. Human Retrovir.* **9**, 639 (1993).
18. K. Stromberg *et al.*, *Science* **224**, 289 (1984); P. A. Marx *et al.*, *J. Virol.* **56**, 571 (1985).
19. A. S. Teirstein and M. J. Rosen, *Clin. Chest Med.* **9**, 467 (1988); S. A. Schroeder *et al.*, *Chest* **101**, 1065 (1992).
20. S. M. Schnittman *et al.*, *Science* **245**, 305 (1989); K. Hsia and S. A. Spector, *J. Infect. Dis.* **164**, 470 (1991).
21. V. M. Hirsch, P. M. Zack, A. P. Vogel, P. R. Johnson, *J. Infect. Dis.* **163**, 976 (1991); Y. K. Donaldson *et al.*, *Lancet* **343**, 382 (1994).
22. DNA sequence analysis of UC2 and UC14 in the C2-V3 region of their Env coding regions revealed only 82% amino acid identity.
23. N. L. Letvin *et al.*, *J. Infect. Dis.* **156**, 406 (1987); I. Nicol *et al.*, *Intervirology* **30**, 258 (1989).
24. A. D. Hoffman, B. Banapour, J. A. Levy, *Virology* **147**, 326 (1985).
25. Baboons were inoculated intravenously with 5000 TCID₅₀ of UC2 or 10,000 TCID₅₀ of UC14 on the given dates. HIV-2 strains were prepared and titrated in human PBMCs obtained from human seronegative donors [S. Barnett, *J. Virol.* **67**, 1006 (1993)]. The PBMCs from all animals were prescreened for in vitro susceptibility to virus infection (17). All inoculations and animal manipulations were performed according to NIH guidelines at the Southwest Foundation for Biomedical Research (San Antonio, TX). Every 2 weeks for 12 weeks, then at 4-week intervals, animals were sedated with ketamine hydrochloride (10 mg per kilogram of body weight) and examined for hepatomegaly or splenomegaly, lymphadenopathy, fever, weight loss, and cutaneous signs. At these times, venipuncture was performed and blood specimens collected. Virus isolations were performed by cocultivation of the PBMCs of infected animals with phytohemagglutinin (PHA)-stimulated PBMCs from seronegative human donors as previously described (17). Culture supernatants were monitored for virus at 3- to 4-day intervals with either the HIV-1 p24 enzyme-linked immunosorbent assay (ELISA) (Coulter, Hialeah, FL) or the reverse transcriptase (RT) assay (24). In some cases, CD4⁺ baboon PBMCs were purified with CD4 antibody immunomagnetic beads (Dynal, Lake Success, NY) before cocultivation with human PBMCs [C. Mackewicz, *J. Clin. Invest.* **87**, 1462 (1991)]. Fresh plasma specimens were assayed for infectious virus within 3 hours after collection by inoculation onto human PBMCs [L. Pan, *J. Clin. Microbiol.* **31**, 283 (1993)]. T cell subsets were measured by flow cytometry with the leu3a (CD4) and leu2a (CD8) monoclonal antibodies (Becton Dickinson, San Jose, CA).
26. B. G. Herndler *et al.*, *AIDS* **8**, 575 (1994).
27. A. Albin, C. D. Mitchell, E. W. Thompson, *J. Cell. Biochem.* **36**, 369 (1988).
28. The authors thank Y.-D. Sun, S. Lunceford, S. Rouse, D. Johnson, and N. Abbey for their excellent technical assistance. We also thank D. Kilhorn for administrative assistance and C. Beglinger for her help in preparing the manuscript. The HIV-2_{ST} gp130 antibody was provided by SmithKline Beecham, King of Prussia, PA. Supported by NIH grant U01-AI26471 and the UCSF AIDS Clinical Research Center, funded by the California State University AIDS Research Program. S.W.B. was a Scholar of the American Foundation for AIDS Research.

8 June 1994; accepted 30 August 1994

Design of a G·C-Specific DNA Minor Groove-Binding Peptide

Bernhard H. Geierstanger, Milan Mrksich, Peter B. Dervan,*
David E. Wemmer*

A four-ring tripeptide containing alternating imidazole and pyrrole carboxamides specifically binds six-base pair 5'-(A,T)GCGC(A,T)-3' sites in the minor groove of DNA. The designed peptide has a specificity completely reversed from that of the tripyrrole distamycin, which binds A,T sequences. Structural studies with nuclear magnetic resonance revealed that two peptides bound side-by-side and in an antiparallel orientation in the minor groove. Each of the four imidazoles in the 2:1 ligand-DNA complex recognized a specific guanine amino group in the GCGC core through a hydrogen bond. Targeting a designated four-base pair G·C tract by this synthetic ligand supports the generality of the 2:1 peptide-DNA motif for sequence-specific minor groove recognition of DNA.

During the past decade, efforts to alter the sequence specificity of naturally occurring small molecules that bind to the minor groove of DNA have met with limited success (1, 2). The natural products netropsin and distamycin are *N*-methylpyrrole-containing di- and tripeptides that bind in the minor groove at sites of at least four successive A·T base pairs (1, 3-6). Inspired by the x-ray structure of the 1:1 complex of netropsin with A,T-rich

DNA (4), efforts were initiated to synthesize ligands capable of recognizing G,C-containing sequences through specific hydrogen bonding to the guanine amino group in the minor groove (7, 8). Initial attempts based on the 1:1 complex led to molecules with increased tolerance for G·C base pairs in the binding site, but not to high specificity (7). The limitations of this approach became apparent when it was realized that two distamycin molecules could be combined side-by-side in the minor groove of DNA (9-11). The 2:1 binding mode had been observed for peptide analogs containing a single imidazole or pyridine ring, which bind sequences containing both A·T and G·C base pairs, emphasizing the importance of the sequence-dependent

B. H. Geierstanger and D. E. Wemmer, Graduate Group in Biophysics and Department of Chemistry, University of California, Berkeley, CA 94720, USA.
M. Mrksich and P. B. Dervan, Arnold and Mabel Beckman Laboratories of Chemical Synthesis, California Institute of Technology, Pasadena, CA 91125, USA.

*To whom correspondence should be addressed.

width of the minor groove (12–20). We used this motif to design a molecule that binds to the minor groove of 5'-(A,T)-GCGC(A,T)-3' sequences with high specificity and affinity, completely reversing the natural binding specificity of distamycin for A,T sequences.

An imidazole-pyrrole-pyrrole dipeptide, 1-methylimidazole-2-carboxamide-*netropsin* (2-ImN), was shown to specifically bind the mixed sequence 5'-(A,T)G(A,T)C(A,T)-3' as a side-by-side antiparallel dimer (12–14) (Fig. 1A). Each ligand of the homodimer interacts with one of the DNA strands in the minor groove. The imidazole nitrogen of each ligand hydrogen bonds specifically with one guanine amino group. The side-by-side combination of one imidazole ring on one ligand and a pyrrole ring on the second ligand recognizes G·C, whereas a pyrrole-imidazole pair targets a C·G base pair (12–14). A pyrrole-pyrrole combination is partially degenerate and binds to either A·T or T·A base pairs (12–14). The generality of the 2:1 motif has been demonstrated by targeting of other designated sequences of mixed G,C and A,T composition with heterodimers and covalently linked homo- and heterodimers (15–20). 2-ImN can be combined with distamycin in a heterodimer to target a 5'-(A,T)G(A,T)₃-3' site (Fig. 1B) (17, 18). A pyrrole-imidazole-pyrrole peptide and distamycin bind the sequence 5'-(A,T)₂G(A,T)₂-3' (Fig. 1C) (16). However, a designed ligand targeted to a pure G,C sequence has never been documented.

From previous work, general features of the 2:1 peptide motif for recognition of mixed A,T and G,C sequences emerge (12–20). The side-by-side ligands are arranged in an antiparallel head-to-tail orientation, with a single charged end pointing toward the 3' end of the contacted DNA strand. Each imidazole ring contacts G on the adjacent DNA strand by forming a specific hydrogen bond with the guanine amino group, and a pyrrole ring from the adjacent ligand optimizes the fit. Ligands with *n* rings pair to recognize an *n* + 2 base-pair site. Therefore a four-ring ligand, ImPImP, consisting of alternating imidazole-pyrrole-imidazole-pyrrole carboxamides, should form a homodimeric complex recognizing the 5'-GCGC-3' core of six base-pair sites of sequence composition 5'-(A,T)GCGC(A,T)-3' (Fig. 1D). We report here the direct structural characterization of the synthetic ligand ImPImP (Fig. 2) in complex with a pure G,C DNA sequence.

The ImPImP ligand was synthesized in eight steps from readily available starting materials (12) and purified by reversed phase chromatography (21). Footprinting experiments show that ImPImP specifically

binds the six base-pair sites 5'-TGCGCA-3' and 5'-AGCGCT-3'. Affinity cleavage studies of ImPImP-EDTA-Fe indicate that the ligand binds these two sites in two equal orientations, which is consistent with an antiparallel side-by-side dimer in the minor groove (21). The ligand fails to bind the 5'-TGGCGT-3' and 5'-TCGGCA-3' sites, which demonstrates that G·C and C·G base pairs can be discriminated in the minor groove by side-by-side ligands. However, ImPImP binds the single hydrogen bond mismatch site 5'-AACGCA-3' site with a third of its affinity for 5'-(A,T)GCGC(A,T)-3' (21).

The complex of ImPImP with d(CG-TAGCGCTACG)₂ was characterized by ¹H nuclear magnetic resonance (NMR) (22, 23). This oligomer duplex is symmetric, leading to one set of resonances for the two strands. Binding of a single ligand would break the symmetry and double the number of DNA resonance lines (5). If the resulting complex is symmetric (as anticipated for a 2:1 complex), the number of DNA lines remains the same. When substoichiometric amounts of ImPImP are added to d(CGTAGCGCTACG)₂, a new set of DNA lines appears and the intensities of the free DNA lines decrease (Fig. 3A). At an ImPImP:DNA ratio of 2:1, the number

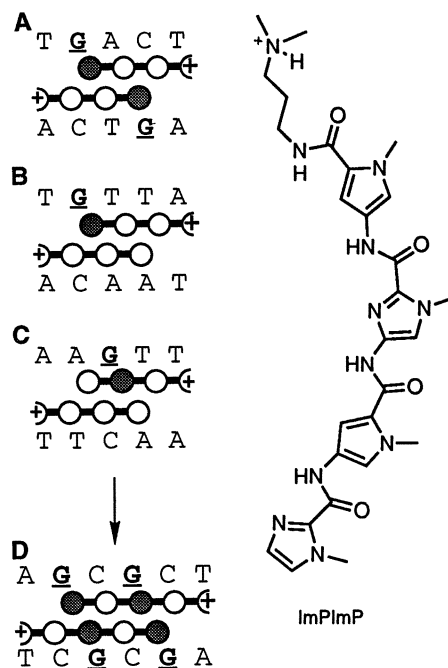


Fig. 1. (left) Specific 2:1 complexes with mixed A,T and G,C sequences. (A) 2-ImN homodimer (12–14). (B) 2-ImN-Dst heterodimer (17, 18). (C) 2-ImD-Dst heterodimer (16). (D) the designed four-ring ImPImP homodimer. The shaded and light circles represent imidazole and pyrrole rings, respectively. The top strands of the sites read 5' to 3', with the specifically targeted guanines highlighted. **Fig. 2.** (right) Structure of four-ring crescent-shaped tripeptide ImPImP.

of lines clearly indicates a symmetric complex. The sharpness of the lines indicates that the exchange between free DNA and the complex is very slow on the NMR time scale. Even at a molar ratio of 0.2:1 ImPImP:DNA, no lines from a 1:1 complex are observed, which indicates high cooperativity for the formation of the 2:1 complex. This behavior has been seen with all complexes of distamycin and its analogs at mixed G,C and A,T sites (10, 15, 16, 18, 19). The 2:1 stoichiometry and the symmetric binding mode indicate an antiparallel side-by-side arrangement that is confirmed by the 2D nuclear Overhauser effect spectroscopy (NOESY) data presented below.

In order to create a detailed molecular model of the ImPImP complex with d(CG-TAGCGCTACG)₂, 2D NOESY spectra in D₂O and H₂O were recorded at a molar 2:1 ImPImP:DNA ratio (Fig. 3B) (22, 23). The intermolecular ligand-DNA and intermolecular ligand-ligand nuclear Overhauser effect (NOE) crosspeaks observed are summarized in Table 1 and Fig. 3C. NOE crosspeaks between amide and pyrrole protons of the ligand and DNA protons place the ImPImP ligand in the minor groove spanning nucleotides 5'-G5-C6-G7-C8-T9-3'. The first imidazole ring is located near guanine G5, whereas the second imidazole contacts G7. The presence of NOE crosspeaks involving the guanine amino protons indicates the formation of hydrogen bonds, as previously discussed (16, 18). The methylene protons and the *N,N*-dimethyl group show NOEs to the C2H protons of adenines A4 and A10, which places the tail of the ligand deep in the groove. The antiparallel ligand arrangement is confirmed by NOEs between the H4-1 and H5-1 protons of one ImPImP molecule and the protons of the methylene tail of the second, and by crosspeaks between H3-2 and NH-4 as well as between pyrrole *N*-methyl NCH₃-4 protons and the H5-1 proton (Fig. 3C and Table 1).

NMR data from NOESY spectra in H₂O and D₂O at 100 ms mixing time were used to derive semiquantitative distance restraints (23, 24). Two sets of 39 intermolecular ligand-DNA, 3 ligand-ligand, and 6 intra-ligand restraints for each of the symmetry-related ligands were used to model the 2:1 ImPImP:DNA complex. The results are presented in Fig. 4. The ligands sit deep in the minor groove, with the positively charged tail pointing toward the 3' end of the adjacent DNA strand. The two ligands are stacked in an antiparallel orientation, as in previous 2:1 DNA complexes of three-ring ligands (13, 15, 16, 18). Molecular modeling indicates hydrogen bonds between each ligand amide and acceptor groups on the bases of the adjacent DNA

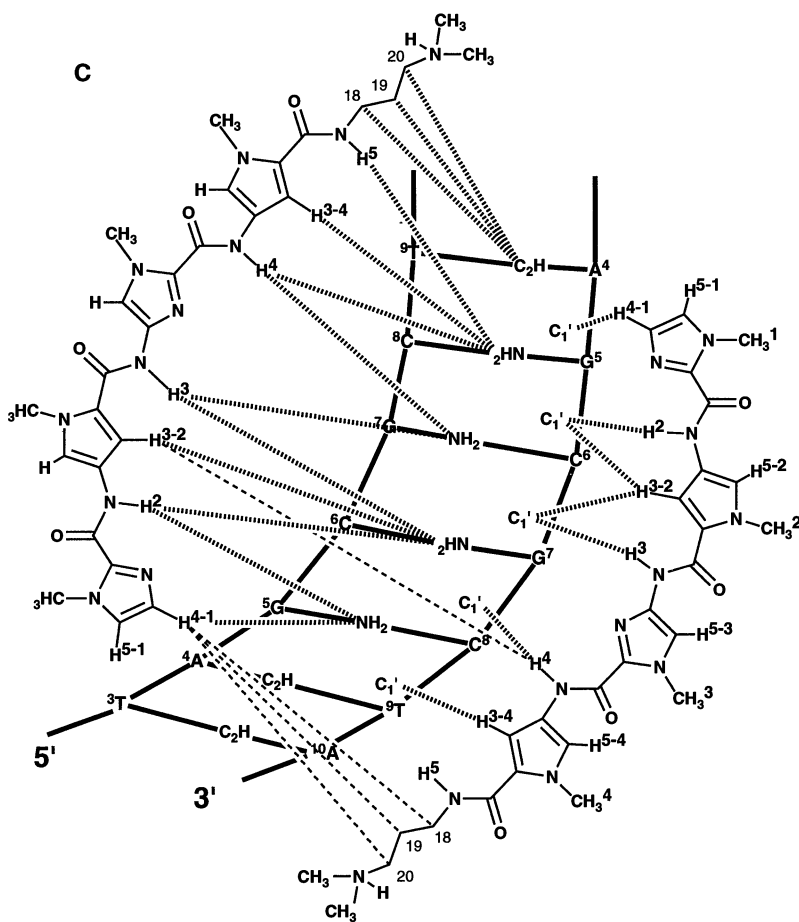
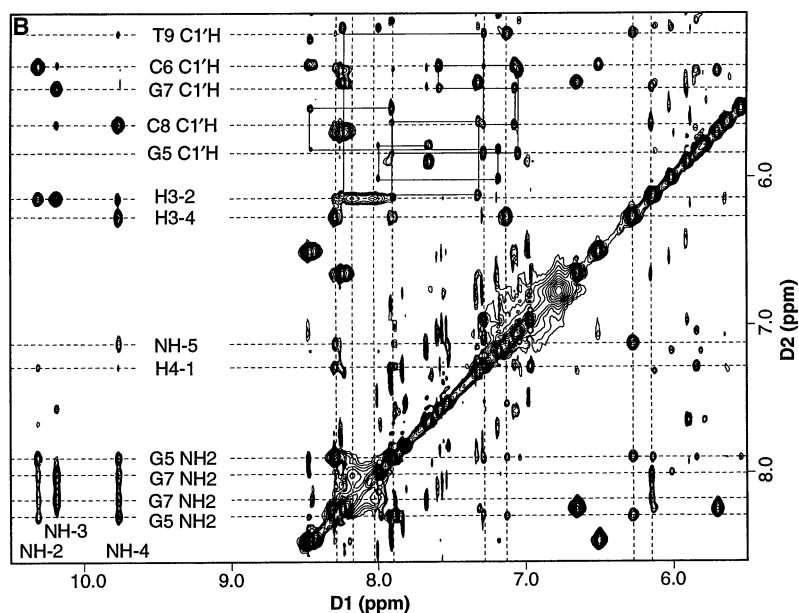
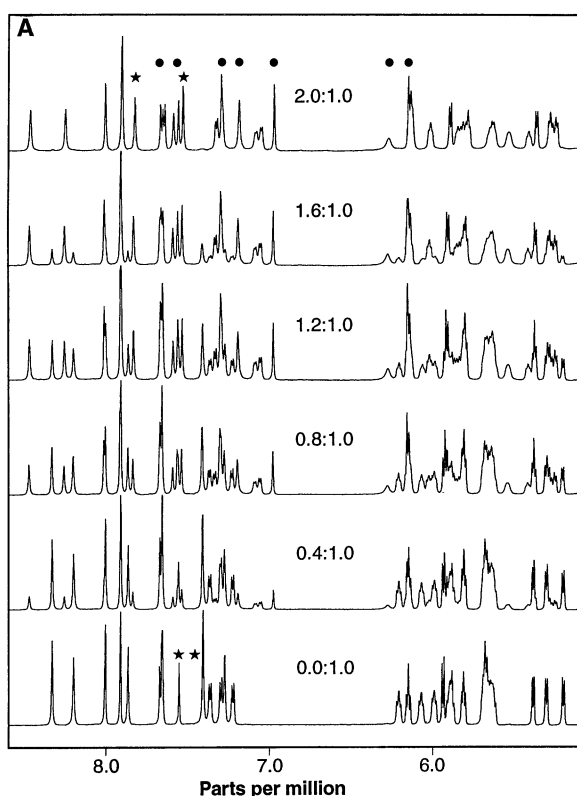


Fig. 3. NMR data (22, 23) [protons of ImPImP are labeled according to C]. **(A)** Titration of d(CGCTAGCGCTACG)₂ with ImPImP (25°C, 600 MHz). The aromatic and C1'H regions of spectra in D₂O are shown. The approximate ligand:DNA ratio is indicated. Stars mark the lines of adenine C2 protons in the free duplex and in the complex. Dots indicate ImPImP ligand resonances in the complex. The H3-4 resonance line (second dot from the right) is broadened by motions of the ligand tail group. **(B)** Expansion of the aromatic and amide region of a NOESY spectrum of the 2:1 complex of ImPImP with d(CGCTAGCGCTACG)₂ (in 90% H₂O and 10% D₂O, 25°C, 600 MHz, $\tau_{\text{mix}} = 200$ ms). Sequential aromatic-to-C1'H connectivities are shown as solid lines. **(C)** Summary of selected intermolecular ligand-DNA and ligand-ligand NOEs, confirming the ligand arrangement in the 2:1 motif. For clarity, symmetrically related NOEs are not shown.

strand (25). Each of the imidazole nitrogen atoms forms a hydrogen bond with the guanine amino proton that does not participate in Watson-Crick base pairing (26). The center part of the complex, with the 5'-GCGC-3' sequence, is essentially symmetrical, although the symmetry-related ligands were free to move independently (with an

identical set of restraints applied) during the energy minimization. However, the positively charged tail groups adopt distinctly different conformations, which reflects the uncertainties of the modeling and of the NOESY data as well as the dynamic behavior of the tail suggested by temperature-dependent spectra (27).

According to the NMR data and the molecular model, the curvature of the ImPImP ligand in the 2:1 ImPImP-DNA complex closely follows the curvature of the DNA minor groove surface. The present molecular model of the 2:1 ImPImP-DNA complex does not indicate any major distortions of the DNA, which suggests that

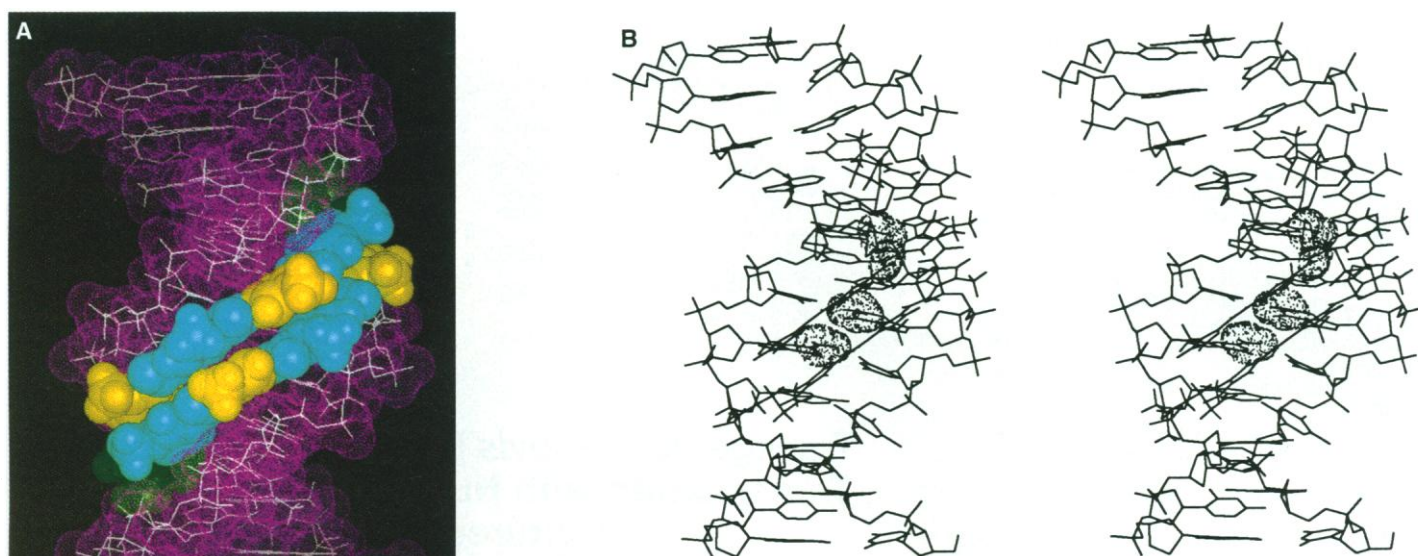


Fig. 4. Molecular model of the 2:1 ImPImP-DNA complex with d(CGTAGCGCTACG)₂. **(A)** Molecular model obtained by energy minimization with the use of semiquantitative distance restraints derived from NOESY data (24). ImPImP ligands are shown as a space-filling model in blue, with imidazole rings and the positively charged end group highlighted in yellow and green, respectively. The DNA is shown as van der Waals surface in cyan, with the model overlaid in white. **(B)** Stereodiagram of the same model. For clarity, hydrogen atoms are omitted from the DNA but not from the ligands. The guanine amino groups recognized by the imidazole nitrogens of the ImPImP peptides are highlighted as van der Waals surface.

der Waals surface in cyan, with the model overlaid in white. **(B)** Stereodiagram of the same model. For clarity, hydrogen atoms are omitted from the DNA but not from the ligands. The guanine amino groups recognized by the imidazole nitrogens of the ImPImP peptides are highlighted as van der Waals surface.

Table 1. Selected ligand-DNA and ligand-ligand intermolecular NOE contacts.

ImPImP	Ligand-DNA: DNA	Ligand-ligand: ImPImP 2
H4-1	G5 C1'H, G5 NH ₂ , C6 C4'H	C19, C18, C20 Hs, NH(CH ₃) ₂ *
H5-1	C6 C5'Hs, C6 C4'H	NCH ₃ -4
NH-2	G5 NH ₂ , C6 C1'H, G7 NH ₂	NH-4
H3-2	G7 C1'H, G7 NH ₂ , C6 C1'H	NH-4
H5-2	G7 C5'Hs, G7 C4'H	
NH-3	G7 NH ₂ , G7 C1'H	
H5-3	C8 C5'Hs, C8 C4'H	
NH-4	G7 NH ₂ , G5 NH ₂ , C8 C1'H	
H3-4	G5 NH ₂ , T9 C1'H	
NH-5	G5 NH ₂ , T9 C1'H	
C18, C19 Hs	A4 C2H	
C20 Hs	A4 C2H, A10 C2H	
NH(CH ₃) ₂ *	A10 C2H, A4 C1'H, A4 C2H	

*Contact to methyl protons.

this motif might be expanded to longer sequences using oligopeptides of this type (28).

The most striking feature of the ImPImP complex is its sequence specificity. Undoubtedly, a combination of electrostatic, van der Waals, and hydrogen bonding interactions between DNA and ligand contribute to the free energy of binding. However, the specificity of binding is determined by hydrogen bonds between the imidazole nitrogens of the ligands and guanine amino groups in the minor groove of DNA. Molecular modeling shows that each imidazole ring of the ImPImP dimer is positioned so that it only interacts with one unique amino group of the guanine on the proximal DNA strand. Guanine amino resonances are usually broadened by exchange through rotation about the N-C bond (29). NOE cross-

peaks are observed to the guanine amino protons of G5 and G7, resonating at four different chemical shifts, providing indirect evidence for the specific hydrogen bonds to the ligand imidazole nitrogens (16, 18). The crosspeaks to the amino protons of G7 are distinctly broader than equivalent crosspeaks to the G5 amino protons. This may reflect a slightly more favorable geometry for the formation of a hydrogen bond at the first ring than at the third ring position. In the molecular model of the 2:1 ImPImP-DNA complex, the lengths of the four imidazole to guanine amino group hydrogen bonds are similar (26). The relative strengths of the hydrogen bonds are probably determined by the details of the interaction and cannot be directly determined. Imidazole nitrogens in the different rings may not contribute equally to the binding energy.

The four-ring ligand ImPImP forms a 2:1 complex with the six base-pair sites 5'-(A,T)GCGC(A,T)-3' with positive cooperativity. A decade after the determination of the seminal 1:1 x-ray structure of the natural product netropsin with an A,T tract (4), a design strategy based on a 2:1 structure has been identified for synthetic, sequence-specific, minor groove-binding ligands targeted to a pure G,C sequence. Until more is understood about the sequence-dependent variation in minor groove width and structure in double-helical DNA, it remains to be seen whether it is now possible to target any G-C base pair within a designated sequence with the 2:1 peptide-DNA motif.

REFERENCES AND NOTES

1. P. B. Dervan, *Science* **232**, 464 (1986); C. Zimmer and U. Wähnert, *Prog. Biophys. Mol. Biol.* **47**, 31 (1986).
2. J. A. Hartley, J. W. Lown, W. B. Mattes, K. W. Kohn, *Acta Oncol.* **27**, 503 (1988).
3. J. S. Taylor, P. G. Schultz, P. B. Dervan, *Tetrahedron* **40**, 457 (1984); P. G. Schultz and P. B. Dervan, *J. Biomol. Struct. Dyn.* **1**, 1133 (1984).
4. M. L. Kopka, C. Yoon, D. Goodsell, P. Pjura, R. E. Dickerson, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 1376 (1985); *J. Mol. Biol.* **183**, 553 (1985).
5. R. E. Klevit, D. E. Wemmer, B. R. Reid, *Biochemistry* **25**, 3296 (1986); J. G. Pelton and D. E. Wemmer, *ibid.* **27**, 8088 (1988).
6. M. Coll, C. A. Fredrick, A. H.-J. Wang, A. Rich, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 8385 (1987).
7. J. W. Lown *et al.*, *Biochemistry* **25**, 7408 (1986); K. Kissinger, K. Krowicki, J. C. Dabrowiak, J. W. Lown, *ibid.* **26**, 5590 (1987); M. Lee *et al.*, *ibid.* **27**, 445 (1988); M. Lee, D. M. Coulter, R. T. Pon, K. Krowicki, J. W. Lown, *J. Biomol. Struct. Dyn.* **5**, 1059 (1988); M. Lee *et al.*, *J. Mol. Recognit.* **2**, 84 (1989); M. Lee, A. L. Rhodes, M. D. Wyatt, S. Forrow, J. Hartley, *Biochemistry* **32**, 4237 (1993); G. Burckhardt *et al.*,

- Biochim. Biophys. Acta* **1009**, 11 (1989).
8. W. S. Wade and P. B. Dervan, *J. Am. Chem. Soc.* **109**, 1574 (1987).
 9. J. G. Pelton and D. E. Wemmer, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 5723 (1989); *J. Am. Chem. Soc.* **112**, 1393 (1990).
 10. D. E. Wemmer *et al.*, in *Structural Biology: The State of the Arts, Proceedings of the Eighth Conversation on Biomolecular Stereodynamics*, R. H. Sarma and M. H. Sarma, Eds. (Adenine Press, Guilderland, NY, 1994), vol. 2, p. 301.
 11. X. Chen, B. Ramakrishnan, S. T. Rao, M. Sundaralingam, *Nat. Struct. Biol.* **1**, 169 (1994).
 12. W. S. Wade, M. Mrksich, P. B. Dervan, *J. Am. Chem. Soc.* **114**, 8783 (1992).
 13. M. Mrksich *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 7586 (1992).
 14. W. S. Wade, M. Mrksich, P. B. Dervan, *Biochemistry* **32**, 11385 (1993).
 15. T. J. Dwyer, B. H. Geierstanger, Y. Bathini, J. W. Lown, D. E. Wemmer, *J. Am. Chem. Soc.* **114**, 5911 (1992).
 16. B. H. Geierstanger, T. J. Dwyer, Y. Bathini, J. W. Lown, D. E. Wemmer, *ibid.* **115**, 4474 (1993).
 17. M. Mrksich and P. B. Dervan, *ibid.*, p. 2572.
 18. B. H. Geierstanger, J. P. Jacobsen, M. Mrksich, P. B. Dervan, D. E. Wemmer, *Biochemistry* **33**, 3055 (1994).
 19. M. Mrksich and P. B. Dervan, *J. Am. Chem. Soc.* **115**, 9892 (1993); T. J. Dwyer, B. H. Geierstanger, M. Mrksich, P. B. Dervan, D. E. Wemmer, *ibid.*, p. 9900.
 20. M. Mrksich and P. B. Dervan, *ibid.* **116**, 3663 (1994).
 21. ———, unpublished results.
 22. DNA oligomers were synthesized and purified as reported previously (9). NMR samples contained 10 mM sodium phosphate buffer (pH 7) in 0.5 ml 99.96% D₂O or 90% H₂O and 10% D₂O. ImPImP-HCl was dissolved in 99.96% D₂O (with 10 mM sodium phosphate buffer), yielding a stock solution of approximately 26 mM as determined by ultraviolet (UV) absorbance ($\epsilon_{312} \approx 4.5 \times 10^4 \text{ M}^{-1} \text{ cm}^{-1}$). Stock solutions were stored at -70°C . The concentration of d(CGTAGCGCTACG)₂ was 1.4 mM duplex as determined by UV absorbance at 80°C ($\epsilon_{260} = 1.13 \times 10^5 \text{ M}^{-1} \text{ cm}^{-1}$).
 23. NMR experiments were done on a Bruker AMX-600 (Bruker Instruments, Billerica, MA) at 600 MHz or on a GE Omega-500 (General Electric, Fremont, CA) at 500 MHz. ImPImP was titrated into the NMR sample containing duplex DNA in approximately 0.2 mole equivalents per addition. 1D spectra in D₂O (128 scans), 2D NOESY spectra (100 ms and 200 ms mixing time; 550 to 743 t₁ experiments with 32 or 48 scans) in D₂O and H₂O, and double-quantum filtered correlation spectroscopy spectra in D₂O (775 or 810 t₁ experiments with 16 or 32 scans) were acquired as described previously (13, 15). The data were processed with FELIX 2.30 (Biosym, San Diego, CA) on Silicon Graphics workstations. DNA and ligand resonances were assigned by standard sequential methods and as previously described (13, 15).
 24. Intermolecular distance restraints for the modeling of the 2:1 ImPImP:DNA complex with d(CGTAGCGCTACG)₂ were generated from crosspeak volumes in the 100 ms H₂O NOESY spectrum as described previously (13, 15). The crosspeak volumes were classified semiquantitatively into three categories: strong (1.8 to 2.5 Å), medium (2.5 to 3.7 Å), or weak (3.7 to 4.2 Å), relative to the cytosine C5H-C6H crosspeak volumes. In all, two identical sets of 39 intermolecular ligand-DNA restraints, 3 intermolecular ligand-ligand restraints, and 6 intramolecular restraints for each of the symmetry-related ligands were used. Hydrogen bonds for standard Watson-Crick base pairing were included as NOE restraints as described previously (18). The model of the d(CGTAGCGCTACG)₂ duplex was constructed with the use of the Biopolymer module of InsightII (Biosym) from standard B-form DNA. The ImPImP ligand model was built based on coordinates of the 2-ImN-Dst heterocomplex model (18). Two ImPImP molecules were roughly arranged in the head-to-tail orientation and manually docked into the minor groove. Restrained energy minimizations were done with Discover (Biosym) (with AMBER forcefield) as de-

- scribed previously (18). The model fulfills all restraints to within 0.1 Å.
25. The ligand amide NH-2 hydrogen bonds (as assigned by InsightII) with C6 cytosine C2 oxygen, NH-3 with G7 guanine N3 nitrogen, NH-4 with C8 cytosine C2 oxygen, and NH-5 with T9 thymine C2 oxygen.
 26. Hydrogen bonds were assigned by the InsightII program. Nitrogen-nitrogen distances varied between 2.8 and 3.1 Å, with nitrogen-hydrogen distances between 1.8 and 2.2 Å.
 27. B. H. Geierstanger, M. Mrksich, P. B. Dervan, D. E. Wemmer, unpublished results.
 28. R. S. Youngquist and P. B. Dervan, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 2565 (1985).
 29. R. Boelens, R. M. Scheek, K. Dijkstra, R. Kaptein,

- J. Magn. Reson.* **62**, 378 (1985); V. Sklenar, B. R. Brooks, G. Zon, A. Bax, *FEBS Lett.* **216**, 249 (1987).
30. We are grateful to NIH (grants GM-27681 to P.B.D. and GM-43129 to D.E.W.) and the National Foundation for Cancer Research for research support, to the Ralph M. Parsons Foundation for a graduate fellowship to M.M., and to the U.S. Department of Energy (DE FG05-86ER75281) and NSF (DMB 86-09305 and BBS 87-20134) for instrumentation grants. The authors especially thank T. J. Dwyer for her contributions to the project. B.H.G. and D.E.W. also thank J. P. Jacobsen, H. P. Spielmann, and P. A. Fagan for helpful discussions.

6 July 1994; accepted 2 September 1994

Transgenic *X. laevis* Embryos from Eggs Transplanted with Nuclei of Transfected Cultured Cells

Kristen L. Kroll* and John C. Gerhart

Transgenic *Xenopus laevis* embryos were produced by transplantation of transfected cultured cell nuclei into unfertilized eggs. A *Xenopus* cell line, X-C, was stably transfected with plasmids containing a hygromycin-resistance gene and genes for either β -galactosidase with a heat shock promoter or chloramphenicol acetyltransferase (CAT) with a muscle-specific actin promoter. Nuclei transplanted from these cells into unfertilized eggs directed development of embryos containing stably integrated copies of the plasmids in each cell. Transgenic embryos showed somite-specific expression of CAT and uniform expression of β -galactosidase. Transgenic embryos produced by nuclear transplantation should be useful for testing the function of cloned genes in amphibian development.

Xenopus laevis (the African clawed toad) is widely used for the study of vertebrate development because the embryos are easily manipulated and accessible at all stages. Gene products that mediate inductive and morphogenetic events in the early *Xenopus* embryo have been identified. However, it has only been possible to study the function of products of these cloned genes by analyzing transiently expressed RNA or DNA injected into fertilized eggs or early cleavage stage embryos. Injected RNAs are translated immediately and are often degraded before the inductive interactions and morphogenesis of the gastrula stage have begun. Thus, although RNA injection can be effectively used to study maternally expressed genes, the endogenous products of which are present at early stages in the embryo, it is unfavorable for the study of zygotic gene products expressed after the mid-blastula transition, which occurs when the embryo has about 4000 cells. DNA injection has been used for analyzing zygotic gene expression, but the usefulness of this approach has been limited by the extremely mosaic expres-

sion of such DNAs in the embryos after injection.

Transgenesis has been attempted in *Xenopus* by injecting plasmid DNA into fertilized eggs, raising embryos to adults, breeding the offspring, and screening them for transmission of injected sequences (1). However, the long generation time of *Xenopus* (8 months minimum) makes this approach cumbersome, and in the test case, the adults produced were mosaic for the introduced sequences, integrated the DNA into the germ line only at low frequency, and produced offspring that failed to express the introduced reporter gene.

We have developed an alternate approach for transgenesis in *Xenopus*. Cultured cells are transfected with reporter sequences, stable integrants are selected, and these cells are used as donors for nuclei that are then transplanted into unfertilized eggs. Nuclei from various *Xenopus* cells are able to support embryonic development after such transplantations, although the extent of development depends on the proliferative and differentiated state of the donor cells (2).

The *Xenopus* cell line X-C was generated from stage 34–38 whole tadpoles (3). Like other established lines, the line is slightly aneuploid (4). We used lipofection to stably

Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720, USA.

*To whom correspondence should be addressed.