Hz, 1 H, lactam β-CH_2), 2.35-2.46 (m, 1 H, pro β-CH_2), 2.69 (dd, J = 14.1 and 7.8 Hz, 1 H, lactam β-CH_2), 2.77 (d, J = 4.8 Hz, 3 H, NCH_3), 3.46 (dd, J = 8.0 and 5.6 Hz, 2 H, pro δ-CH_2), 3.56-3.58 (m, 2 H, SCH_2), 4.80-4.85 (m, 1 H, thiazolidine α-CH), 5.18 (dd, J = 7.4 and 5.0 Hz, 1 H, SCHN), 7.49 (br s, 1 H, CONH); ^{13}C NMR (75.5 MHz, $CDCl_3$, pro = pyrrolidine) δ 23.61 (pro γ-C), 26.23 (pro β-C), 28.36 (Boc CH_3), 36.52 (lactam β-C), 39.38 and 39.54 (SCH_2 and NCH_3), 47.79 (pro δ-C), 57.42 (thiazolidine α-C), 62.69 (SCHN), 69.52 (pro α-C), 80.64 (Boc C—O), 153.72 (Boc C=O), 169.47 (CONH), 172.54 (lactam C=O); FAB MS m/z 356 $[MH]^+$. Anal. Calcd for $C_{16}H_{25}N_3O_4S$: C, 54.06; H, 7.09; N, 11.82; S, 9.02. Found: C, 53.84; H, 7.05; N, 11.64; S, 8.88.

[3′-(*S*)-[3′α,6′α(*R),7′aα]]-1-Acetyltetrahydro-5′-oxospiro[pyrrolidine-2,6′-(5′*H*)-pyrrolo[2,1-*b*]thiazolidine-3′-*N*-methylcarboxamide (6).** *N*-Methylamide **13** (40 mg, 0.11 mmol) was deprotected in 4 N HCl/dioxane (5 mL) at room temperature for 1 h. The solvent was removed in vacuo; and the residue was taken up in CH_2Cl_2, whereupon the solvent was again removed in vacuo. The residue was dried in vacuo overnight. It was then dissolved in Ac_2O (10 mL) and cooled in an ice bath, whereupon NEt_3 (1 equiv) and DMAP (cat.) were added. The solution was stirred at room temperature overnight. The solvent was evaporated in vacuo, and the residue was chromatographed on a 1.5 × 45 cm silica gel flash column using CH_2Cl_2/MeOH as the eluting solvent. The product (**6**) was isolated as an oil: $[\alpha]_D$ +139.2° (*c* 0.90, MeOH); ^1H NMR (300 MHz, $CDCl_3$, pro = pyrrolidine) δ 1.90-2.07 (m, 2 H, pro γ-CH_2), 2.09 (s, 3 H, $COCH_3$), 2.16-2.22 (m, 2 H, β-CH_2), 2.31-2.45 (m, 1 H, lactam β-CH_2), 2.72 (dd, J = 14.1 and 8.1 Hz, 1 H, lactam β-CH_2), 2.82 (d, J = 5.1 Hz, 3 H, NCH_3), 3.56-3.65 (m, 4 H, pro δ-CH_2 and SCH_2), 4.85 (dd, J = 8.5 and 5.0 Hz, 1 H, thiazolidine α-CH), 5.15 (dd, J = 8.0 and 4.4 Hz, 1 H, SCHN), 7.24 (br s, 1 H, CONH); ^{13}C NMR (75.5 MHz, $CDCl_3$, pro = pyrrolidine) δ 23.60 (pro γ-C), 24.13 ($COCH_3$), 26.50 (pro β-C), 36.31 (lactam β-C), 38.34 (SCH_2), 39.11 (NCH_3), 48.87 (pro δ-C), 57.70 (thiazolidine α-C), 62.67 (SCHN), 69.50 (pro α-C), 169.57 and 169.72 (CONH and $COCH_3$), 173.04 (lactam C=O); FAB MS m/z 298 $[MH]^+$. Anal. Calcd for $C_{13}H_{19}N_3O_3S$: C, 52.50; H, 6.44; N, 14.13; S, 10.78. Found: C, 52.29; H, 6.55; N, 13.98; S, 10.56.

NMR Studies. Nuclear Overhauser effects (NOE's) were detected using the Varian 300-MHz instrument. Difference NOE measurements were measured as negative enhancements upon irradiation of the proton of interest. The experiment was repeated in triplicate for each compound at 21 °C with varying d2 delay times. The decoupler power was set to 18 dB and the 90-deg pulse width was measured for each compound and set appropriately. The FIDs were multiplied by a heavy line broadening (lb = 5) apodization function before subtraction.

Temperature studies were performed on the GE 300-MHz instrument at a sample concentration of 25 mM. The temperature was set to 295 K, and a spectrum was taken at this temperature and at 5-deg intervals up to 325 K. The chemical shifts of the amide hydrogen versus temperature were plotted, and the temperature coefficients were measured as ppb/K from the slope of the line.

TFE titration studies were performed on the same instrument at 25 °C. A spectrum was first taken in DMSO; then TFE was added in portions so as to increase the amount of TFE by 10% with each addition. After each addition, another spectrum was recorded. This was repeated to a final concentration of 70% TFE in DMSO. The change in chemical shifts (Δδ) of the amide hydrogen was then measured.

Molecular Modeling Studies. Energy minimization studies were performed as described by Ferguson and Raber[16,17] using the Random Incremental Pulse Search (RIPS) method. Eight energy-minimized structures were generated, and the figures were displayed using a Macintosh Quadra 700 computer with Alchemy III software, TRIPOS Associates, Inc. 1992.

Acknowledgment. This research was supported in part by an NIH grant (NS20036) to RLJ and an NIH predoctoral traineeship (GM07994) to MJG. The authors gratefully acknowledge the assistance of Dr. David M. Ferguson in the molecular modeling studies.

Design of Peptides That Bind in the Minor Groove of DNA at 5′-(A,T)G(A,T)C(A,T)-3′ Sequences by a Dimeric Side-by-Side Motif

Warren S. Wade, Milan Mrksich, and Peter B. Dervan*

Contribution from the Arnold and Mabel Beckman Laboratories of Chemical Synthesis, California Institute of Technology, Pasadena, California 91125. *Received May 11, 1992*

Abstract: The designed peptides pyridine-2-carboxamide–netropsin (2-PyN) and 1-methylimidazole-2-carboxamide–netropsin (2-ImN) are crescent-shaped synthetic analogs of the natural products netropsin (N) and distamycin A (D). Footprinting experiments indicate that the peptides 2-PyN and 2-ImN bind specifically the 5 base pair sequence 5′-TGTCA-3′. Affinity cleaving data suggest that the complexes, 2-ImN·5′-TGTCA-3′ and 2-PyN·5′-TGTCA-3′, are composed of two equivalent orientations which disfavor a 1:1 model. The footprinting and affinity cleaving data are in accord with a 2:1 complex where a novel side-by-side antiparallel dimer binds in the minor groove of double-helical DNA.

Netropsin and Distamycin A. Netropsin (N) and distamycin A (D) are natural products that bind in the minor groove of double-helical DNA at sites of 4 or 5 successive A,T base pairs (Figure 1).[1-4] X-ray[5] and NMR[6] studies of netropsin–DNA and distamycin–DNA complexes reveal how sequence specificity is accomplished. The crescent-shaped di- and tripeptides are bound in the middle of the minor groove of an A,T-rich sequence. The amide hydrogens of the *N*-methylpyrrolecarboxamides form bifurcated hydrogen bonds with the N3 of adenine and the O2 of

(1) For reviews, see: (a) Dervan, P. B. *Science* **1986**, *232*, 464-471. (b) Zimmer, C.; Wähnert, U. *Prog. Biophys. Molec. Biol.* **1986**, *47*, 31-112.

(2) (a) Krylov, A. S.; Grokhovsky, S. L.; Zasedatelev, A. S.; Zhuze, A. L.; Gursky, G. V.; Gottikh, B. P. *Nucleic Acids Res.* **1979**, *6*, 289-304. (b) Zasedatelev, A. S.; Gursky, G. V.; Zimmer, Ch.; Thrum, H. *Mol. Biol. Rep.* **1974**, *1*, 337-342. (c) Zasedatelev, A. S.; Zhuze, A. L.; Zimmer, Ch.; Grokhovsky, S. L.; Tumanyan, V. G.; Gursky, G. V.; Gottikh, B. P. *Dokl. Akad. Nauk SSSR* **1976**, *231*, 1006-1009.

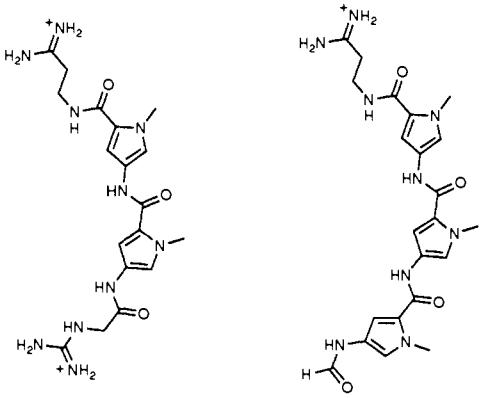
(3) (a) Van Dyke, M. W.; Hertzberg, R. P.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 5470-5474. (b) Van Dyke, M. W.; Dervan, P. B. *Cold Spring Harbor Symposium on Quantitative Biology* **1982**, *47*, 347-353. (c) Van Dyke, M. W.; Dervan, P. B. *Biochemistry* **1983**, *22*, 2373-2377. (d) Harshman, K. D.; Dervan, P. B. *Nucleic Acids Res.* **1985**, *13*, 4825-4835. (e) Fox, K. R.; Waring, M. J. *Nucleic Acids Res.* **1984**, *12*, 9271-9285. (f) Lane, M. J.; Dobrowiak, J. C.; Vournakis, J. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 3260-3264.

(4) (a) Schultz, P. G.; Taylor, J. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1982**, *104*, 6861-6863. (b) Taylor, J. S.; Schultz, P. G.; Dervan, P. B. *Tetrahedron* **1984**, *40*, 457-465. (c) Schultz, P. G.; Dervan, P. B. *J. Biomol. Struct. Dyn.* **1984**, *1*, 1133-1147.

(5) (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.

(6) (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.

0002-7863/92/1514-8783$03.00/0



Netropsin (N) Distamycin A (D)

Figure 1. Natural products netropsin (N) and distamycin A (D).

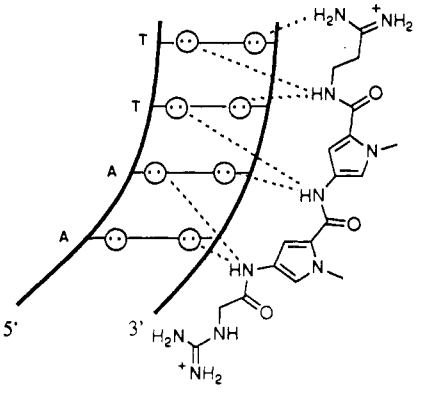
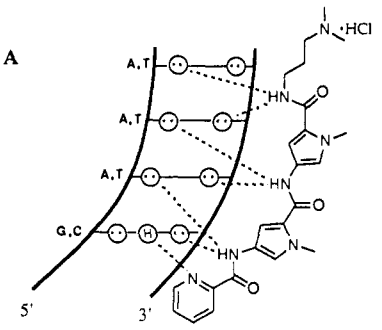


Figure 2. 1:1 Complex of netropsin with 5′-AATT-3′.[5a] Circles with dots represent lone pairs on N3 of purines and O2 of pyrimidines. Putative hydrogen bonds are illustrated by dashed lines.

thymine on the floor of the minor groove (Figure 2). The pyrrole rings fill the groove completely and form extensive van der Waals contacts with the walls of the minor groove. The aromatic hydrogens of the *N*-methylpyrrole rings are set too deeply in the minor groove to allow room for the guanine NH_2 of a G,C base pair, affording binding specificity for A,T sequences. Thermodynamic profiles reveal that for both netropsin and distamycin, the binding affinities at 25 °C are high and qualitatively similar ($\Delta G°$ = −12.7 kcal/mol) for complexation to poly[d(A-T)]·poly[d(A-T)] and poly(dA)·(dT).[7] However, for each peptide ligand, binding to the alternating copolymer duplex is enthalpy driven, whereas binding to the homopolymer duplex is entropy driven.[7] Two goals of the current research effort in this field are to increase the specificity of the natural products[8-10] and to alter the specificity to favor recognition of mixed sequences.[11-14]



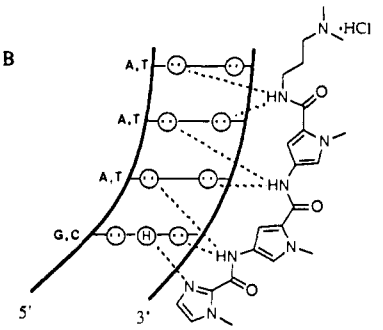


Figure 3. 1:1 Binding models for the complexes formed between (a) 2-PyN and (b) 2-ImN with 5′-(G,C)(A,T)$_3$-3′ sequences. Circles with dots represent lone pairs on N3 of purines and O2 of pyrimidines. Circles containing an H represent the N2 hydrogen of guanine. Putative hydrogen bonds are illustrated by dashed lines.

Oligopeptides for A,T Binding. Progress toward the former goal has been achieved with the design and synthesis of molecules that bind larger sequences of pure A,T-rich double helical DNA. Increasing the number of *N*-methylpyrrolecarboxamides increases both the binding site size and the sequence specificity for longer tracts of A,T-rich DNA.[8] Oligo(*N*-methylpyrrolecarboxamides) with incrementally increasing numbers of amino acid units from 2 to 9 contain 3 to 10 carboxamide NH's and are capable of binding from 4 to 9 A,T base pairs, respectively.[8-10] Oligopeptides have achieved recognition of sequences as long as 15 base pairs of A,T-rich DNA.[10d]

Peptides for Binding Sequences Containing G,C and A,T Base Pairs. Efforts to develop molecules with well-understood binding preferences which recognize sequences containing G,C base pairs have met with only limited success. Progress in this area is an important component in an overall strategy of coupling G,C units and A,T units into oligomers that uniquely identify long sequences of right-handed DNA in the minor groove. Early examples of hybrid molecules for mixed sequence recognition are the intercalator–groove binder bis(distamycin)phenoxazone (BEDP)[11] and the metalloregulated Ba^{2+}·bis(netropsin)-3,6,9,12,15-pentaoxoheptadecanediamide,[12] both of which bind the 10 base pair sequence 5′-TATAGGTTAA-3′. These hybrid molecules have similar functional architectures: a G,C binding unit flanked on both sides by an A,T-rich peptide binding unit. Although these can be considered encouraging benchmarks in the early development of the field of DNA recognition, they suffer from a lack of generalizable principles.

(7) (a) Markey, L. A.; Breslauer, K. J. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 4359–4363. (b) Breslauer, K. J.; Remeta, D. P.; Chou, W.-Y.; Ferrante, R.; Curry, J.; Zaunczkowski, D.; Snyder, J. G.; Marky, L. A. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 8922–8926.

(8) (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. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 2565–2569.

(9) (a) Gursky, G. V.; Zasedatelev, A. S.; Zhuze, A. L.; Khorlin, A. A.; Grokhovsky, S. L.; Streltsov, S. A.; Surovaya, A. N.; Nikitn, S. M.; Krylov, A. S.; Retchinsky, V. O.; Mikhaillov, M. V.; Beabealashvili, R. S.; Gottihk, B. P. *Cold Spring Harbor Symposium on Quantitative Biology* **1982**, *47*, 367. (b) Skamrov, A. V.; Rybalkin, I. N.; Bibilashvili, R. Sh.; Gottikh, B. P.; Grokhovskii, S. L.; Gurskii, G. V.; Zhuze, A. L.; Zasedatelev, A. S.; Nechipurenko, Yu. D.; Khorlin, A. A. *Mol. Biol.* **1985**, *19*, 153.

(10) (a) Schultz, P. G.; Dervan, P. B. *J. Am. Chem. Soc.* **1983**, *105*, 7748–7750. (b) Youngquist, R. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1985**, *107*, 5528–5529. (c) Griffin, J. H.; Dervan, P. B. *J. Am. Chem. Soc.* **1986**, *108*, 5008–5009. (d) Youngquist, R. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1987**, *109*, 7564–7566.

(11) Dervan, P. B.; Sluka, J. P. *New Synthetic Methodology and Functionally Interesting Compounds*; Elsevier: New York, 1986; pp 307–322.

(12) Griffin, J. H.; Dervan, P. B. *J. Am. Chem. Soc.* **1987**, *109*, 6840–6842.

(13) (a) Lown, J. W.; Krowicki, K.; Bhat, U. G.; 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.

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

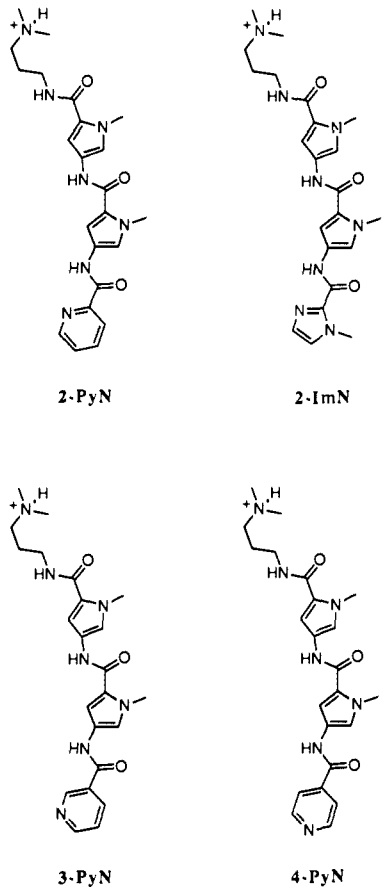


Figure 4. Synthetic analogs pyridine-2-carboxamide–netropsin (2-PyN), 1-methylimidazole-2-carboxamide–netropsin (2-ImN), pyridine-3-carboxamide–netropsin (3-PyN), and pyridine-4-carboxamide–netropsin (4-PyN).

An alternative approach is the systematic substitution of the tris (*N*-methylpyrrolecarboxamide) framework of distamycin to search for altered base pair specificity.[13,14] In the minor groove of DNA, the major difference between edges of A,T and G,C base pairs is the presence of the guanine 2-amino group protruding from the floor. By replacing the pyrrole CH with a heteroatom capable of forming a hydrogen bond to the guanine 2-amino group, for example, a steric clash might be turned into a stabilizing interaction. Molecules containing such a substitution display an increased tolerance for G,C base pairs in their binding sites, with an overall loss in specificity.[13]

Design Rationale: The 1:1 Complex. Our approach is illustrated in Figure 3A, wherein we expected that replacement of the terminal *N*-methylpyrrolecarboxamide of D with pyridine-2-carboxamide would afford an analog, pyridine-2-carboxamide–netropsin (2-PyN) (Figure 4), that would specifically bind to 1 G,C base pair followed by 3 A,T base pairs. In analogy with distamycin binding A,T-rich sites, it was anticipated that the amide hydrogens would form bifurcated hydrogen bonds with adenine N3 and thymine O2 atoms, while the pyridine nitrogen would participate in a key hydrogen bond with the NH_2 group of the G,C base pair (Figure 3). The 3-carboxamide and 4-carboxamide isomers (3-PyN and 4-PyN) serve as controls for the placement of the ring nitrogen (Figure 4). We found that 2-PyN bound 3 A,T-rich sites 5 base pairs in size, 5′-TTTTT-3′, 5′-AATAA-3′, 5′-CTTTT-3′, and an unanticipated mixed sequence 5′-TGTCA-3′.[14] Furthermore, affinity cleaving experiments revealed no orientation preference for binding to the 5′-TGTCA-3′ sequence. In contrast, pyridine-3-carboxamide–netropsin (3-PyN) and pyridine-4-carboxamide–netropsin (4-PyN) preferred A,T-rich sequences, suggesting that the placement of the ring nitrogen is important for the recognition of the sequence 5′-TGTCA-3′.[14] In an effort to explore the structural limitations of binding to the sequence 5′-TGTCA-3′, the pyridine ring was replaced with *N*-methylimidazole to afford 1-methylimidazole-2-carboxamide–netropsin (2-ImN) (Figures 3b and 4). Footprinting and affinity cleaving experiments indicate that this synthetic peptide binds the 5′-TGTCA-3′ site with higher sequence specificity than 2-PyN.

Results

Synthesis of 2-ImN, 2-PyN, 3-PyN, and 4-PyN. The syntheses of 2-ImN and the 2-, 3-, and 4-PyN analogs are outlined in Figure 5. *N*-Methyl-4-(*N*-methyl-4-nitropyrrole-2-carboxamide)-pyrrole-2-carboxylic acid (**9**) is available in five steps starting from *N*-methylpyrrole-2-carboxylic acid in 20% overall yield.[15,16]

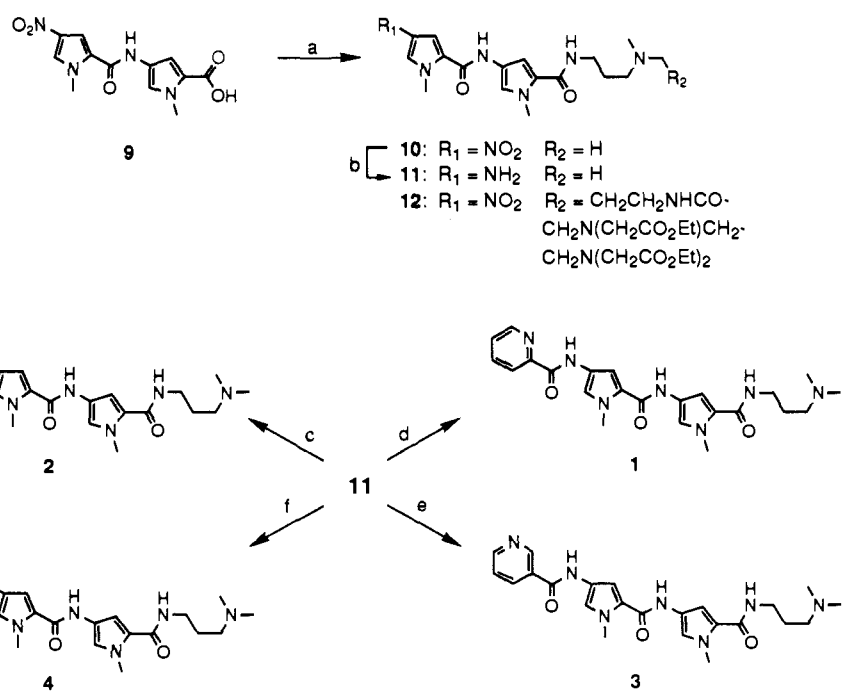


Figure 5. Synthetic scheme for 2-, 3-, and 4-PyN and 2-ImN. (a) (i) DCC, HOBt; (ii) 3-(dimethylamino)propylamine; (b) 100 psi H_2, 10% Pd/C; (c) *N*-methylimidazole-2-carboxylic acid, DCC, HOBt; (d) picolinic acid, DCC, HOBt; (e) nicotinic acid, DCC, HOBt; (f) isonicotinic acid, DCC, HOBt.

Activation of the acid (DCC, HOBt) and condensation with 3-(dimethylamino)propylamine affords nitrodipyrrole **10** in 98% yield. Subsequent reduction with hydrogen over palladium provides the corresponding aminopyrrole **11**, which is not stable as the free base. Although the hydrochloride salt can be stored for longer periods, we and others have found that coupling yields do not suffer when the amine is used without purification.[4b,16] *N*-Methylimidazole-2-carboxylic acid was synthesized from *N*-methylimidazole by treatment with 1 equiv of *n*-BuLi (THF, −78 °C), followed by quenching of the anion with CO_2(g). Coupling of this acid (DCC, HOBt) with aminopyrrole **11** afforded 2-ImN in 85% yield. The 2-, 3-, and 4-PyN analogs were synthesized in similar fashion. Picolinic acid was activated (DCC, HOBt) and coupled with **11** to afford 2-PyN in 85% yield. Likewise, nicotinic acid and isonicotinic acid were coupled with **11** to afford 3-PyN and 4-PyN, in 64% and 72% yields, respectively.

Footprinting. MPE·Fe(II) footprinting on the 517 base pair *Eco*RI/*Rsa*I restriction fragment from pBR322 (pH 7.9 in 40 mM Tris acetate at 37 °C) reveals that these 4 peptides protect distinct 5 base pair regions (Figures 6 and 7).[3] Distamycin, at 2 μM concentration, binds to the 5 base pair sites, in the order of decreasing affinity, 5′-TTTTT-3′ > 5′-AATAA-3′ > 5′-CTTTT-3′. 2-PyN, at 10 μM concentration, also binds these A,T-rich sites, in the order of decreasing affinity, 5′-TTTTT-3′ > 5′-AATAA-3′ > 5′-TTATT-3′ in addition to a 5′-TGTCA-3′ sequence. The 3-PyN and 4-PyN isomers do not bind the 5′-TGTCA-3′ site but have affinity for the distamycin sites, 5′-TTTTT-3′, 5′-TTAAT-3′,5′-AATAA-3′, and 5′-CTTTT-3′. In contrast, 2-ImN, at 10 μM concentration, binds preferentially the sequence 5′-TGTCA-3′ and has low affinity for distamycin sites of the type 5′-$(A,T)_5$-3′.

Synthesis of 2-ImNE, 2-PyNE, 3-PyNE, and 4-PyNE. Affinity cleavage experiments identify the sequence, groove location, and orientation preference with which a molecule binds to its target site.[4] Ethylenediaminetetraacetic acid (EDTA) was attached to the carboxy end of the four peptides 2-ImN, 2-PyN, 3-PyN, and 4-PyN to afford 2-ImNE, 2-PyNE, 3-PyNE, and 4-PyNE, respectively (Figure 8). The synthesis of 2-ImNE is outlined in Figure 9. In analogy to the synthesis of 2-ImN, dipyrrole carboxylic acid **9** is converted to the activated ester (DCC, HOBt) which is immediately treated with 3-(*tert*-butyloxy carbonylamino)-3′-amino-*N*-methyldipropylamine to afford the protected amine **13**. Nitropyrrole **13** was reduced with palladium under higher pressures of hydrogen (250–350 psi), and the corresponding amine was coupled with *N*-methylimidazole-2-carboxylic acid (DCC, HOBt) to afford **14** in 85% yield. Deprotection of the amine (TFA, CH_2Cl_2) and coupling with triethyl ethylenediaminetriacetate acetic acid (DCC, HOBt) provided the triethylester of 2-ImNE (**16**). Saponification (LiOH, H_2O, EtOH) yielded 2-ImNE. The 2-, 3-, and 4-PyNE analogs were synthesized by a similar route, wherein the triethylester of EDTA was coupled to the bis(*N*-methylpyrrole) earlier in the synthetic sequence. Dipyrrole **12** is available from *N*-methylpyrrole-2-carboxylic acid in eight steps in approximately 20% overall yield.[4b] Reduction of **12** (50 psi H_2, Pd/C) and coupling with an activated pyridine carboxylic acid followed by basic hydrolysis (LiOH, EtOH, H_2O) gave the affinity cleaving peptides in 20–30% yield.

Affinity Cleaving. Cleavage experiments on the 517 base pair restriction fragment were carried out in the presence of 2-ImNE·Fe(II) and 2-, 3-, and 4-PyNE·Fe(II) at 37 °C (pH 7.9 in 40 mM Tris acetate) (Figure 10). All of the peptides show cleavage patterns which are shifted to the 3′ side of the binding site, consistent with binding in the minor groove (Figure 11).[4b] Distamycin–EDTA·Fe(II) (ED), at 2 μM concentration, binds to the 5′-TTTTT-3′ and 5′-AATAA-3′ sites with modest orientation preference but to the 5′-CTTTT-3′ site with the EDTA·Fe(II) functionality strongly oriented to the G,C side of the sequence. 2-PyNE, at 50 μM concentration, also binds to the 5′-TTTTT-3′ and 5′-AATAA-3′ sites with little orientation preference but to the 5′-CTTTT-3′ site with a strong orientation preference in the opposite direction. Binding to the mixed se-

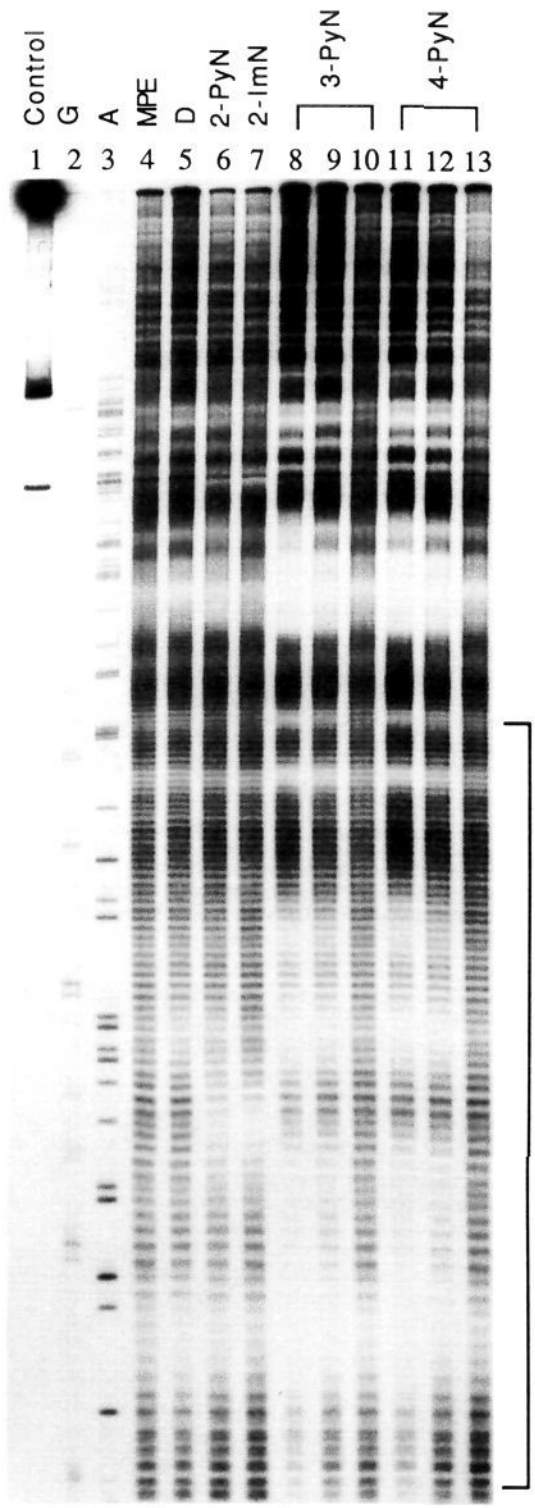


Figure 6. MPE·Fe(II) footprinting of D, 2-PyN, 2-ImN, 3-PyN, and 4-PyN. Autoradiogram of a 8% denaturing polyacrylamide gel. All reactions contain 4 mM DTT, 100 μM bp calf thymus DNA, and 12 kcpm 5′ labeled 517 base pair restriction fragment in 40 mM Tris acetate pH 7.9 buffer. Lane 1, intact DNA; lane 2, Maxam–Gilbert G reaction; lane 3, A reaction; lane 4, MPE·Fe(II) standard; lane 5, 4 μM D; lane 6, 20 μM 2-PyN; lane 7, 20 μM 2-ImN; lanes 8–10 contain 35 μM, 10 μM, and 1 μM 3-PyN, respectively; lanes 11–13 contain 35 μM, 10 μM, and 1 μM 4-PyN, respectively. Lanes 4–13 contain 4 μM MPE·Fe(II).

(15) Bialer, M.; Yagen, B.; Mechoulam, R. *Tetrahedron* **1978**, *34*, 2389–2391.
(16) Lown, J. W.; Krowicki, K. *J. Org. Chem.* **1985**, *50*, 3774–3779.

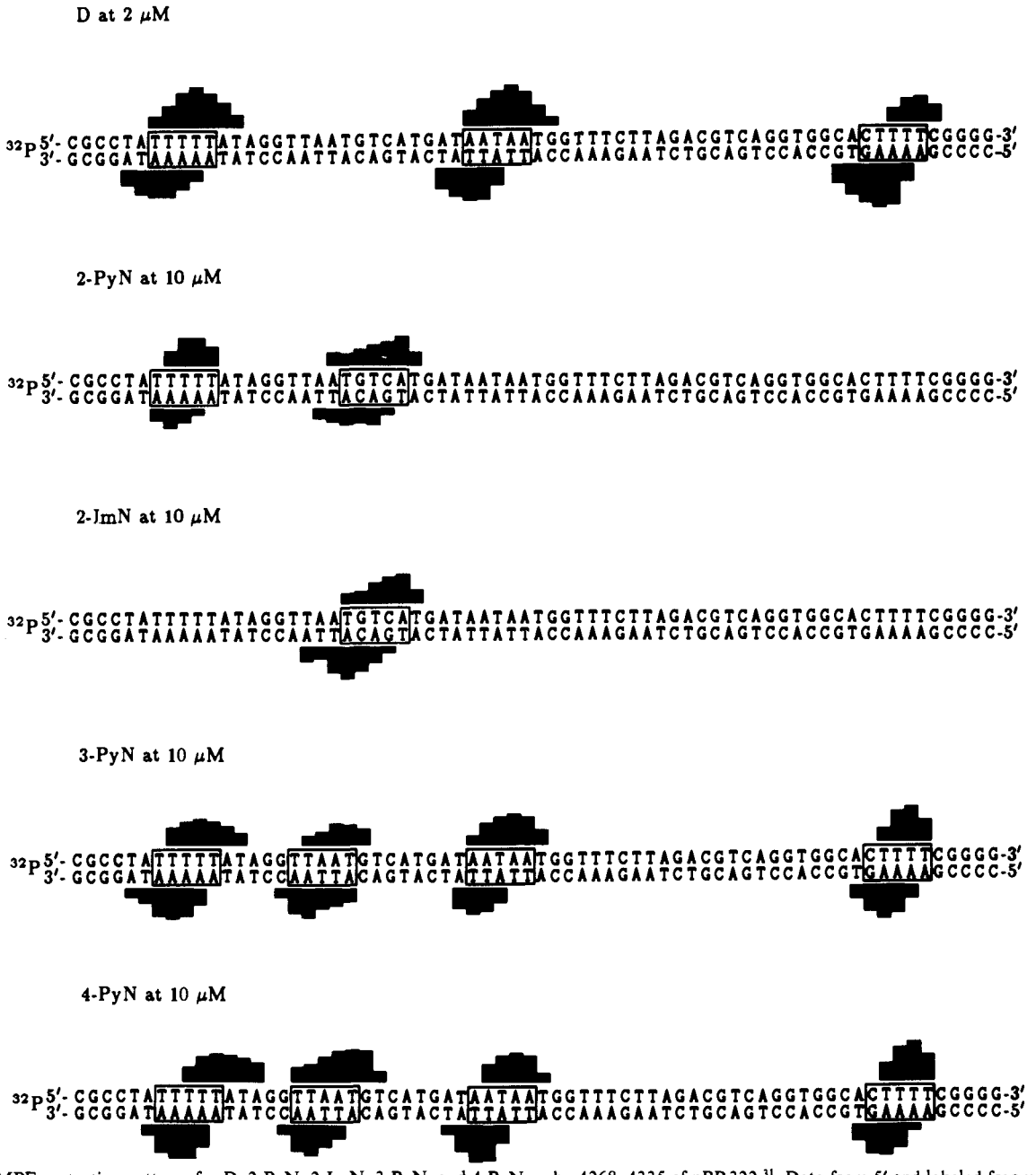


Figure 7. MPE protection patterns for D, 2-PyN, 2-ImN, 3-PyN, and 4-PyN on bp 4268–4335 of pBR322.[31] Data from 5′ end labeled fragment shown in Figure 5. Bar heights are proportional to the protection from cleavage at each band. Boxes represent equilibrium binding sites determined by the published model.[3]

quence, 5′-TGTCA-3′, results in equal cleavage intensities on each side of the binding site, suggesting two equivalent orientations. 2-ImNE, at 70 μM concentration, shows this same behavior at the 5′-TGTCA-3′ site, while binding to the 5′-CTTTT-3′ site occurs with the same orientation preference exhibited by 2-PyNE. In common with distamycin, 3-PyNE and 4-PyNE bind to the 5′-TTTTT-3′ and 5′-AATAA-3′ sites with similar orientation distributions but bind the 5′-CTTTT-3′ site with less orientation preference.

Sequence Specificity. In the resolvable region of the 517 base pair restriction fragment, only a small fraction of the 512 possible 5 base pair sites are present. To be confident that the observed strong binding sites are representative of all binding sites, the double-strand cleavage reactions were performed on DNA 4363 base pairs in length. Cleavage of *Sty*I linearized pBR322[17] by ED, 2-PyNE, and 2-ImNE reveals markedly different cleavage patterns for ED and 2-ImNE, while the 2-PyNE lane appears to be a combination of the ED and 2-ImNE specificities (Figure 12). To determine the specificity of 2-ImNE, the size in base pairs of each band is calculated by linear interpolation from the sizes of the standards (Figure 13). Sequence analysis of the plasmid reveals 27 resolvable 5′-WGWCW-3′ sites (W = A or T),[18] and remarkably, all are within 35 base pairs of observed 2-ImNE cleavage sites (Table I).[19] High-resolution analysis of eight of these sites confirms binding to a 5′-WGWCW-3′ sequence with two equivalent orientations.[20] With regard to binding to the 5′-WGWCW-3′ sites, A,T base pairs on either side of the sequence generally result in higher binding affinities.[20]

Discussion

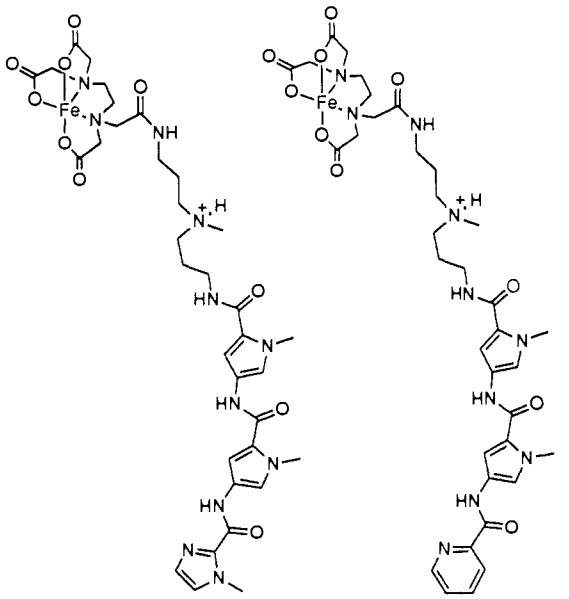
1:1 Model. Binding of 2-PyN and 2-ImN to the 5′-CTTTT-3′ site is consistent with the design rationale model (Figure 3). Both

(17) By actual count, pBR322 contains 99% of the 512 unique 5 bp sequences.

(18) Cornish-Bowden, A. *Nucleic Acids Res.* **1985**, *13*, 3021–3030.

(19) There are three other 5′-WGWCW-3′ sites, but these are within 700 base pairs of the labeled end and are not resolved on the gel. Calculation of the sizes of the bands has an error of ±40 base pairs.

(20) Wade, W. S. Ph.D. Thesis, California Institute of Technology, 1989.



2-ImNE 2-PyNE

Figure 8. Affinity cleaving analogs 1-methylimidazole-2-carboxamide–netropsin–EDTA·Fe(II) (2-ImNE) and pyridine-2-carboxamide–netropsin–EDTA·Fe(II) (2-PyNE).

compounds cleave this site with a strong preference for the orientation which places the heterocyclic nitrogen near the G,C base pair. The observations that the distamycin analog ED binds in the opposite orientation and that 3-PyNE and 4-PyNE do not have a strong orientation preference at this site support the 1:1 model. It is clear, however, that 5′-CTTTT-3′ is not the highest affinity binding site for 2-PyN and 2-ImN.

Both 3-PyN and 4-PyN bind strongly to A,T-rich sites with orientation preferences and relative site preferences similar to those of distamycin. This is presumably a consequence of the ring nitrogens being in the wrong positions to favorably hydrogen bond with guanine 2-amino groups. Binding of 2-PyN to these same sites likely arises from inadequate positioning of the pyridine nitrogen, with the pyridine ring rotated relative to the carboxamide. This conformation of 2-PyN, with the ring nitrogen facing away from the minor groove, more closely resembles the 3-PyN and 4-PyN analogs and can display analogous binding properties. Presumably the other conformer, with the pyridine nitrogen positioned along the concave edge of the molecule, is the conformation which recognizes the mixed sequence 5′-TGTCA-3′. The observation that 2-ImN does not recognize A,T-rich sites supports this model since positioning the imidazole nitrogen away from the groove places the bulky methyl group toward the floor of the minor groove, which is expected to disrupt binding.

Table I. Probable 2-ImNE Binding Sites on pBR322[a]

cleavage position[b]	site position[c]	sequence	relative strength
29	11[d]	ATGTTTGACAGCTTA	s
86	71[d]	AACGCAGTCAGGCAC	m
196	212	TCGCCAGTCACTATG	w
666	682[d]	AACCCAGTCAGCTCC	w
730	711[d]	GGGCATGACTATCGT	m
	732[d]	ACTTATGACTGTCTT	
	736[d]	ATGACTGTCTTCTTT	
1935	1969	AAGCCAGACATTAAC	m
1993	2021	CAGGCAGACATCTGT	m
2137	2106	ACCTCTGACACATGC	m
	2139	CAGCTTGTCTGTAAG	
	2163	GGAGCAGACAAGCCC	
2215	2224	GACCCAGTCACGTAG	s
2359	2374	CTCACTGACTCGCTG	w
2573	2585	GCTCAAGTCAGAGGT	m
2844	2861	CGGTAAGACACGACT	m
3269	3289	TGGTCTGACAGTTAC	s
3326	3335	CGATCTGTCTATTTC	m
	3362	TTGCCTGACTCCCCG	
3513	3534	CATCCAGTCTATTAA	m
3601	3624	CGTGGTGTCACGCTC	w
3749	3742	ATCGTTGTCAGAAGT	m
3823	3808	CTTACTGTCATGCCA	m
	3837	TTCTGTGACTGGTGA	
	3858	AACCAAGTCATTCTG	
4215	4199[d]	GTTATTGTCTCATGA	m
4332	4311[d]	TATCATGACATTAAC	m

[a] Underlined sequences represent putative or actual binding sites. [b] Calculated cleavage band positions averaged between labels, pBR322 numbering.[33] [c] Lowest numbered base pair of the putative binding site. [d] Cleavage site observed on a high-resolution sequencing gel.

2:1 Model. Binding of 2-PyN and 2-ImN to the 5′-TGTCA-3′ site was unanticipated. In early 1:1 models, the amide bonds were rotated, presenting carbonyl lone pairs which could participate in hydrogen bonds with guanine 2-amino groups to explain the presence of G,C base pairs in both the second and fourth positions of the binding site.[14] Another unusual feature of this complex was revealed in affinity cleaving experiments. The cleavage intensities on either side of the binding site were nearly equal, in contrast to distamycin analogs binding to A,T-rich sites, where it has been found that the orientation preferences are dependent on the flanking sequences.[4] Recognition by 2-ImN or 2-PyN of a DNA sequence containing 2 G,C base pairs separated by 1 A,T base pair with no orientation preference was difficult to rationalize with a 1:1 model.

NMR studies have revealed that distamycin, at millimolar concentrations, is capable of binding to the minor groove of an 5′-AAATT-3′ sequence as a side-by-side dimer.[21] Consideration

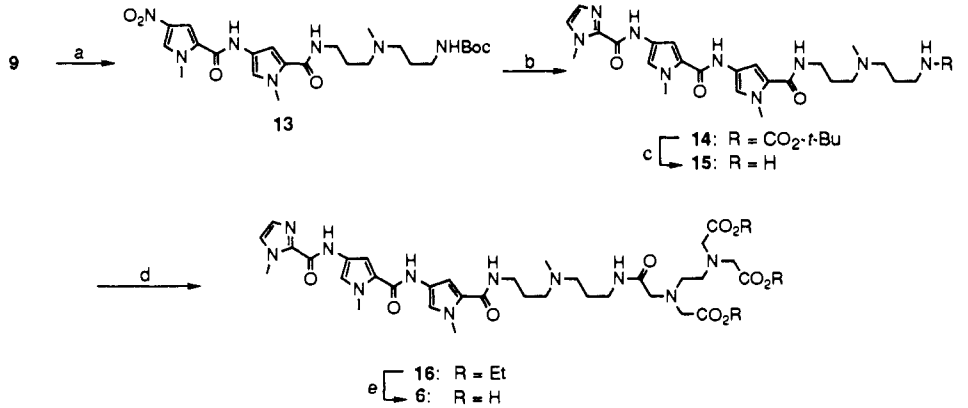


Figure 9. Synthetic scheme for 2-ImNE. (a) (i) DCC, HOBt; (ii) $H_2N(CH_2)_3NCH_3(CH_2)_3NHBoc$; (b) (i) 300 psi H_2, 10% Pd/C; (ii) *N*-methylimidazole-1-carboxylic acid, DCC, HOBt; (c) 20% TFA/CH_2Cl_2; (d) $(EtO_2CCH_2)_2N(CH_2)_2N(CH_2CO_2Et)(CH_2CO_2H)$, DCC, HOBt; (e) LiOH, EtOH, H_2O.

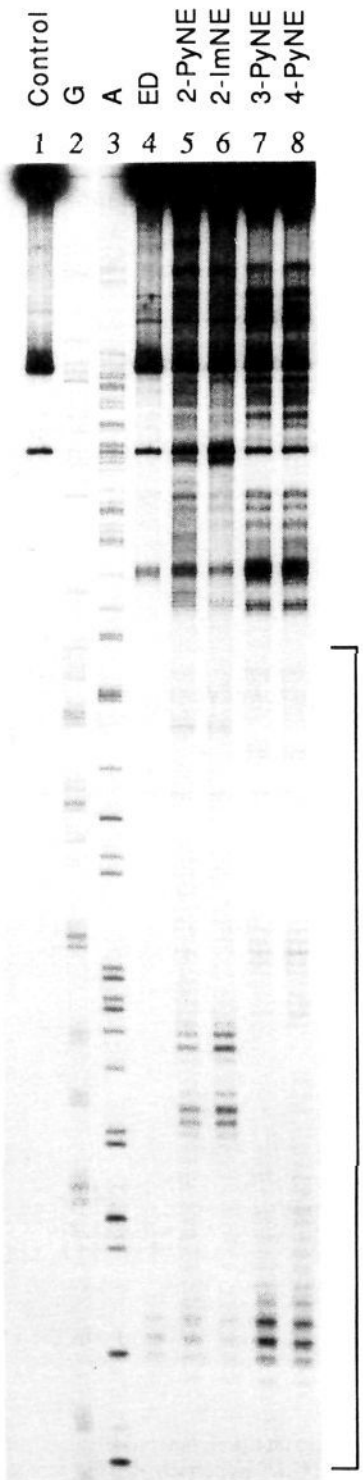


Figure 10. Specific cleavage by the ED, 2-PyNE, 2-ImNE, 3-PyNE, and 4-PyNE analogs. Autoradiogram of a 8% denaturing polyacrylamide gel. All reactions contain 4 mM DTT, 100 μM bp calf thymus DNA, and 12 kcpm 5′ labeled 517 base pair restriction fragment in 40 mM Tris acetate pH 7.9 buffer. Lane 1, intact DNA; lane 2, Maxam–Gilbert G reaction; lane 3, A reaction; lane 4, 2 μM ED; lane 5, 50 μM 2-PyNE; lane 6, 70 μM 2-ImNE; lane 7, 10 μM 3-PyNE; lane 8, 7 μM 4-PyNE.

of such a 2:1 binding model could explain the recognition of 2 G,C base pairs in the 5′-TGTCA-3′ site by 2-ImN or 2-PyN, with each guanine amino group hydrogen bonding to imidazole or pyridine ring nitrogens on different ligands (Figure 14). Furthermore, the $(2\text{-ImNE})_2$·5′-TGTCA-3′ and $(2\text{-PyNE})_2$·5′-TGTCA-3′ complexes position a cleaving function on each side of the binding site, consistent with the affinity cleaving results. In collaboration with the Wemmer group, we have recently characterized the complex between 2-ImN with d(GCATGACTCGG)·d(CCGAGTCATGC) by two-dimensional NMR spectroscopy.[22] 2-ImN binds as an antiparallel, side-by-side dimer with high cooperativity in the minor groove of DNA. Energy minimization of the complex with constraints from NOESY experiments affords a model wherein hydrogen bonds are likely formed between the imidazole nitrogens and the guanine 2-amino groups (Figure 14).[22]

Sequence Specificity for 5′-WGWCW-3′ Sites. 2-ImN and 2-PyN bind to 5′-WGWCW-3′ sites but not to the related sequences 5′-WGWGW-3′ or 5′-WCWGW-3′. This observation can be rationalized by analysis of the disposition of hydrogen binding sites in the minor groove of these sequences. The hydrogen-bonding character of an A·T base pair resembles that of a T·A base pair due to the approximate minor groove symmetry of the adenine N3 and thymine O2 atoms. In a G,C base pair, the guanine 2-amino group lies closer to the guanine-containing strand and as a result, the relative positions of the two guanine amino groups in the three sequences 5′-GTC-3′, 5′-GTG-3′, and 5′-CTG-3′ are very different. As a consequence of the helicity of B-DNA, *the amino groups are separated by a larger distance* in the 5′-GTC-3′ sequence relative to the 5′-CTG-3′ sequence. The 5′-GTG-3′ sequence is clearly different from the other two, since both amino groups derive from guanine residues on the same strand. This distinction, in combination with restrictions on ligand–ligand stacking interactions and other ligand–DNA contacts, favors binding to the 5′-WGTCW-3′ sequence over the similar 5′-WGTGW-3′ and 5′-WCTGW-3′ sequences.

Biological Implications of the 5′-WGWCW-3′ Sequence. A review of DNA-binding proteins revealed that a 5′-WGWCW-3′ sequence is present in the concensus recognition sites of the transcriptional activator GCN4[23] and the oncogenic proteins jun and fos.[24] Current models suggest that these leucine zipper proteins bind as Y-shaped dimers in the major groove of DNA, with no minor groove contacts by the proteins.[25,26] There exists the possibility that 2-ImN can bind as a dimer in the minor groove simultaneous with binding in the major groove by the dimeric proteins. Related work from this laboratory has shown that 2-ImN and GCN4(226–281)[26] can bind a 5′-CTGACTAAT-3′ sequence simultaneously.[27]

Implications for the Design of Minor Groove Binding Molecules. The most significant finding of this study is that 2-ImN shows an apparent preference for cooperative 2:1 binding to the 5′-TGTCA-3′ sequence over 1:1 binding. The width of the minor groove for B-DNA is sequence dependent and can vary from 3–4 Å for A,T-rich DNA to 5–6 Å for G,C-containing regions.[28] X-ray structures of 1:1 complexes formed between netropsin and distamycin with short oligonucleotides reveal that the natural products sit deeply in the minor groove and make extensive van der Waals contacts to both walls of the binding site, contributing favorably to the total free energy of binding.[5] The minor groove width for 1:1 complexes of A,T-rich DNA with distamycin and netropsin is approximately 3.4 Å[5] compared to 6.8 Å from models of 2:1 complexes.[21,22] Certain sequences with inherently wider

(21) (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.

(22) 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.

(23) (a) Hill, D. E.; Hope, I. A.; Macke, J. P.; Struhl, K. *Science* **1986**, *224*, 451–457. (b) Hope, I. A.; Struhl, K. *Cell* **1985**, *43*, 177–188.

(24) (a) Bohmann, D. T.; Bos, T. J.; Admon, A.; Nishimura, T.; Vogt, P. K.; Tjian, R. *Science* **1987**, *238*, 1386–1392. (b) Rauscher, F. J.; Sambucetti, L. C.; Curran, T.; Distel, R. J.; Spiegelman, B. M. *Cell* **1988**, *52*, 471–480.

(25) Vinson, C. R.; Sigler, P. B.; McKnight, S. L. *Science* **1989**, *246*, 911–916.

(26) Oakley, M. G.; Dervan, P. B. *Science* **1990**, *248*, 847–850.

(27) Oakley, M. G.; Mrksich, M.; Dervan, P. B. *Biochemistry*, in press.

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

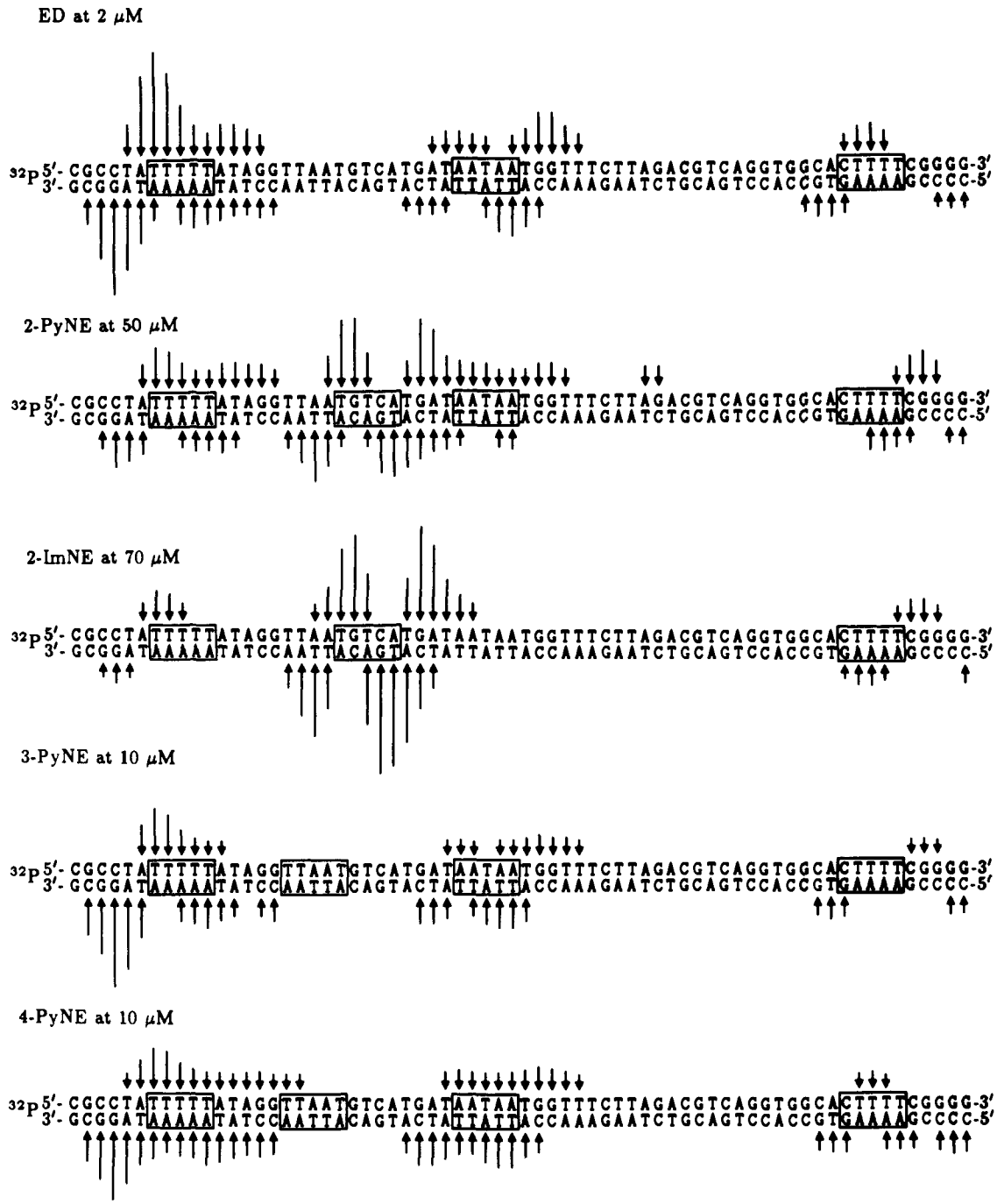


Figure 11. Cleavage of the 517 base pair fragment of pBR322, base pairs 4268–4335,[31] by the ED, 2-PyNE, 2-ImNE, 3-PyNE, and 4-PyNE analogs. Data from 5′ end labeled fragment shown in Figure 10. Arrows are proportional to the maximum densities of the cleavage bands. Boxes represent binding sites determined by the published method.[4]

minor groove widths may favor dimeric binding. We cannot distinguish at this time the importance of sequence dependent flexibility of the minor groove and sequence dependent groove width. We would anticipate that if the guanine NH_2 protruding from the floor of the minor groove is participating in a specific hydrogen bond with the imidazole N3 of 2-ImN, the peptide ligand is not set as deeply in the minor groove as distamycin. A shallow penetration into the minor groove for 2-ImN and 2-PyN in complex with 5′-TGTCA-3′ sequences would suggest that these 2:1 complexes would have overall lower free energy (and likely lower enthalpy) than that observed for distamycin in complex with poly[d(A-T)]·poly[d(A-T)].[7]

Finally, dimeric motifs seem unlikely to be limited to *N*-methylpyrrolecarboxamide analogs. For example, chromomycin, an oligosaccharide–chromophore, binds in the minor groove as a 2:1 complex.[29] Interestingly, the sugar of one ligand of chromomycin stacks on the chromophore unit of the other ligand. The increased width of the chromomycin dimer requires the DNA to undergo a major structural transition, resulting in a wider and more shallow minor groove.[29] It is reasonable to expect a larger class of structure types to form dimeric structures which can recognize double-helical DNA with high sequence specificity.

Experimental Section

^{1}H NMR spectra were recorded either at 400 MHz on a JEOL-GX 400 or at 300 MHz on a General Electric-QE 300 NMR spectrometer in DMSO-d_6. Chemical shifts are reported in parts per million relative to residual DMSO. IR spectra were recorded on a Perkin-Elmer FTIR spectrometer. High-resolution mass spectra (HRMS) were recorded using electron ionization (EI) or fast atom bombardment (FAB) tech-

(29) (a) Gao, X.; Patel, D. J. *Biochemistry* **1989**, *28*, 751–762. (b) Gao, X.; Mirau, P.; Patel, D. J. *J. Mol. Biol.* **1992**, *223*, 259–279.

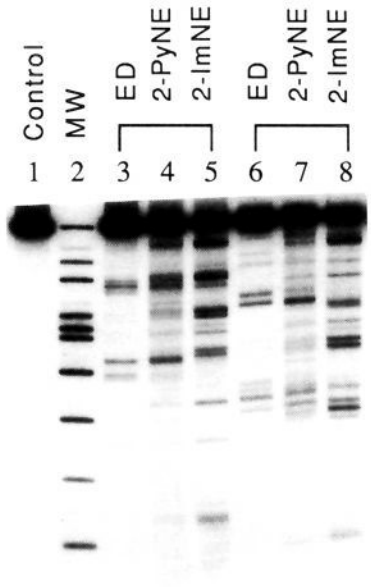


Figure 12. Site-specific double-stranded cleavage of pBR322 (4363 base pairs) by ED, 2-PyNE, and 2-ImNE. Autoradiogram of a 1% agarose gel. All reactions contain 1 mM sodium ascorbate, 100 μM bp calf thymus DNA, and 24 kcpm 3′ labeled pBR322 linearized with *Sty*I and labeled only at one terminus with Klenow fragment. Lanes 1–5 are labeled on the clockwise strand with [α-^{32}P]dATP. Lanes 6–8 are labeled on the counterclockwise strand with [α-^{32}P]TTP. Lane 1, intact DNA; lane 2, molecular weight standards; lanes 3 and 6, 2 μM ED; lanes 4 and 7, 50 μM 2-PyNE; lanes 5 and 8, 70 μM 2-ImNE.

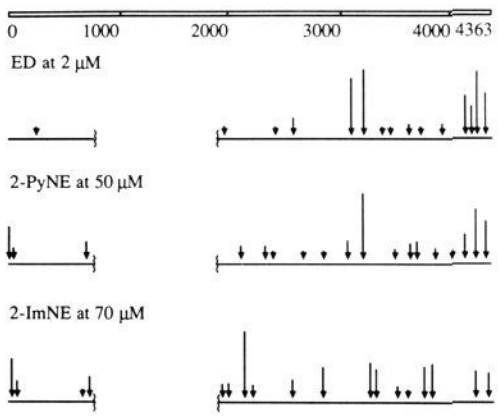


Figure 13. Cleavage sites on pBR322 by ED, 2-PyNE, and 2-ImNE. Arrows represent the extent of cleavage in lanes 3–8 of Figure 12. Band densities are averaged between labels and then normalized to the most intense band. Arrow positions represent approximate binding site locations accurate to within 40 base pairs. Cleavage bands within 700 base pairs of the labeled ends are not resolved on the gel.

niques at the Mass Spectrometry Laboratory at the University of California, Riverside. Reagent grade chemicals were used as received unless otherwise noted. Tetrahydrofuran (THF) was distilled under nitrogen from sodium/benzophenone ketyl. Dichloromethane and triethylamine were distilled under nitrogen from powdered calcium hydride. Dimethylformamide (DMF) was purchased as an anhydrous solvent from Aldrich. *N,N′*-Carbonyldiimidazole was sublimed at 100 °C and 1 Torr. Flash chromatography was carried out using EM Science Kieselgel 60 (230–400) mesh.[30] Thin-layer chromatography was performed on EM Reagents silica gel plates (0.2-mm thickness). All the pyrrole derivatives were visualized with short-wave ultraviolet light.

***N*-Methylimidazole-2-carboxylic Acid.** To a solution of *N*-methylimidazole (1.00 g, 12.2 mmol) in THF (20 mL), cooled to −70 °C, was added *n*-BuLi (5.2 mL, 13.0 mmol in hexanes), and the solution was allowed to stir for 90 min. Dry carbon dioxide was bubbled through the solution for 30 min, and the solution was allowed to stir while warming

(30) Still, W. C.; Kahn, M.; Mitra, A. *J. Org. Chem.* **1978**, *43*, 2923–2925.

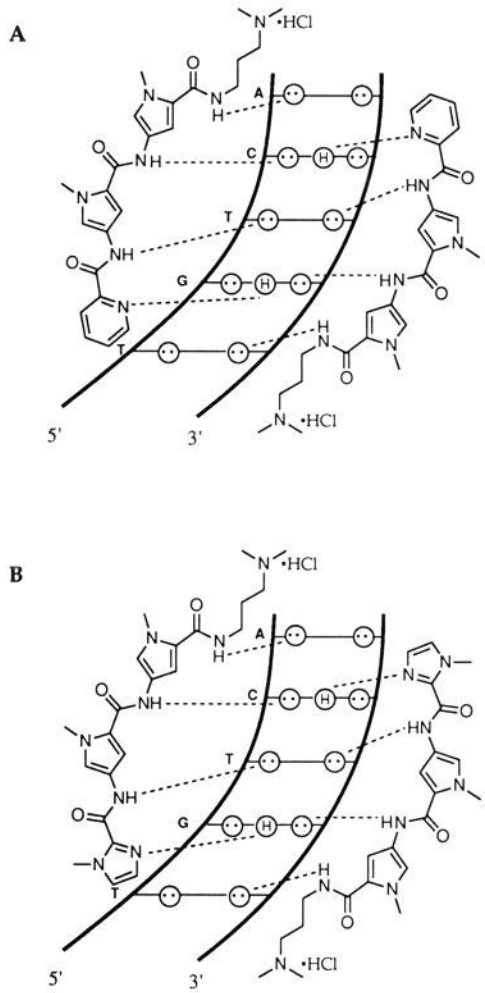


Figure 14. 2:1 Binding models for the complexes formed between (a) 2-PyN and (b) 2-ImN with the 5′-TGTCA-3′ sequence. Circles with dots represent lone pairs on N3 of purines and O2 of pyrimidines. Circles containing an H represent the N2 hydrogen of guanine. Putative hydrogen bonds are illustrated by dashed lines.

to room temperature. The reaction was quenched with HCl (1N, 3 mL), at which time a yellow oil separated. Lyophilization of this oil afforded the imidazolecarboxylic acid (1.3 g, 96%) as a white solid: ^{1}H NMR (400 MHz, DMSO-d_6) δ 3.99 (s, 3 H), 7.27 (s, 1 H), 7.50 (s, 1 H); IR (KBr) 3450 (w), 2600 (m), 1980 (w), 1652 (s), 1514 (m), 1462 (m), 1402 (s), 1354 (s), 1312 (s), 1276 (m), 1172 (w), 1014 (w), 906 (w), 800 (s), 778 (m), 632 (w) cm^{-1}; EIMS *m/e* (relative intensity) 126.0428 (17, M^+, 126.0430 calcd. for $C_5H_6N_2O_2$), 109 (6, M − OH), 108 (7, M − H_2O), 82 (100, M − CO_2).

3-[4-(*N*-Methyl-4-nitropyrrole-2-carboxamido)-*N*-methylpyrrole-2-carboxamido]dimethylpropylamine (10). To a solution of nitrodipyrrole **9** (2.5 g, 8.56 mmol), *N*-hydroxybenzotriazole hydrate (1.35 g, 10.0 mmol), and 3-(dimethylamino)propylamine (1.3 mL, 10.2 mmol) in DMF (70 mL) at 0 °C was added dicyclohexylcarbodiimide (1.90 g, 9.09 mmol) in CH_2Cl_2 (5 mL). The reaction mixture was allowed to warm to room temperature and was stirred for 20 h. The mixture was filtered through celite, and the solvent was removed under reduced pressure. The product was purified by flash column chromatography (1% concentrated aqueous ammonia in methanol) to give **10** (2.85 g, 88%): ^{1}H NMR (400 MHz, DMSO-d_6) δ 10.23 (s, 1 H), 8.13 (d, 1 H, J = 1.8 Hz), 8.07 (t, 1 H, J = 6.7 Hz), 7.53 (d, 1 H, J = 1.8 Hz), 7.17 (d, 1 H, J = 1.8 Hz), 6.77 (d, 1 H, J = 1.8 Hz), 3.90 (s, 3 H), 3.77 (s, 3 H), 3.2 (m, 2 H), 2.21 (t, 2 H, J = 7.4 Hz), 2.11 (s, 6 H), 1.59 (m, 2 H); FABMS *m/e* (relative intensity) 377.1925 (20, M + H, 377.1937 calcd. for $C_{17}H_{25}N_6O_4$).

Pyridine-2-carboxamide–netropsin (2-PyN) (1). To a solution of picolinic acid (0.075 g, 0.610 mmol) and *N*-hydroxybenzotriazole hydrate (0.090 g, 0.666 mmol) in DMF (2 mL) was added a solution of dicyclohexylcarbodiimide (0.130 g, 0.630 mmol) in CH_2Cl_2 (2 mL). The solution was allowed to stir for 40 min. Separately, a solution of nitrodipyrrole **10** (0.10 g, 0.266 mmol) and palladium on activated carbon (10%, 20 mg) in DMF (10 mL) was allowed to stir under a hydrogen

atmosphere (50 psi) for 5 h. The mixture was filtered through celite and added to the activated acid. The solution was allowed to stir for 4 h and was filtered through celite, and the solvent was removed under reduced pressure. The residue was partitioned between ethyl acetate (30 mL) and 10% aqueous $NaHCO_3$ (20 mL). The organic fraction was dryed ($MgSO_4$), and the solvent was removed under reduced pressure. The product was purified by flash column chromatography (gradient 0–1% concentrated aqueous ammonia in methanol) to afford 2-PyN (0.102 g, 85%): UV (H_2O) λ_{max} (ϵ) 223 (sh, 21 000), 238 (22 000), 296 (26 000) nm; ^{1}H NMR (400 MHz, DMSO-d_6) δ 10.73 (s, 1 H), 9.95 (s, 1 H), 8.71 (dd, 1 H, J = 6.5, 1.8 Hz), 8.17 (t, 1 H, J = 7.6 Hz), 8.11 (d, 1 H, J = 7.7 Hz), 8.05 (ddd, 1 H, J = 7.7, 7.8, 1.8 Hz), 7.64 (ddd, 1 H, J = 7.6, 4.8, 1.8 Hz), 7.37 (d, 1 H, J = 1.8 Hz), 7.24 (d, 1 H, J = 1.8 Hz), 7.18 (d, 1 H, J = 1.8 Hz), 6.93 (d, 1 H, J = 1.8 Hz), 3.85 (s, 3 H), 3.81 (s, 3 H), 3.23 (q, 2 H, J = 7.4 Hz), 3.02 (t, 2 H, J = 7.6 Hz), 2.74 (s, 6 H), 1.84 (quint, 2 H, J = 7.6 Hz); IR (KBr) 2910 (w), 2700 (w), 1640 (s), 1580 (s), 1530 (s), 1460 (m), 1432 (s), 1400 (m), 1260 (m), 1200 (w), 1120 (w) cm^{-1}; FABMS m/e (relative intensity) 452.2427 (60, M + H, 452.2410 calcd. for $C_{23}H_{30}N_7O_3$), 350 (28, M – $C_5H_{14}N_2$), 228 (46).

Pyridine-3-carboxamide–netropsin (3-PyN) (3). To a solution of nicotinic acid (0.057 g, 0.464 mmol) in DMF (3 mL) was added *N,N'*-carbonyldiimidazole (0.075 g, 0.463 mmol). The solution was allowed to stir for 1 h. Separately, a solution of nitrodipyrrole **10** (0.081 g, 0.215 mmol) and palladium on activated carbon (10%, 15 mg) in DMF (10 mL) was allowed to stir under a hydrogen atmosphere (50 psi) for 5 h. The mixture was filtered through celite and added to the activated acid. The resulting reaction mixture was allowed to stir for 18 h and was filtered through celite, and the solvent was removed under reduced pressure. The product was purified by flash column chromatography (gradient 0–1% concentrated aqueous ammonia in methanol) to afford 3-PyN (0.062 g, 64%): UV (H_2O) λ_{max} (ϵ) 213 (19 500), 246 (19 500), 300 (23 500) nm; ^{1}H NMR (400 MHz, DMSO-d_6) δ 10.64 (s, 1 H), 9.98 (s, 1 H), 9.77 (bs, 1 H), 9.14 (d, 1 H, J = 2.5 Hz), 8.77 (dd, 1 H, J = 6.2, 5.5 Hz), 8.38 (d, 1 H, J = 8.8 Hz), 8.17 (t, 1 H, J = 5.5 Hz), 7.64 (dd, 1 H, J = 8.8, 5.5 Hz), 7.35 (d, 1 H, J = 1.8 Hz), 7.19 (d, 1 H, J = 1.8 Hz), 7.10 (d, 1 H, J = 1.8 Hz), 6.93 (d, 1 H, J = 1.8 Hz), 3.87 (s, 3 H), 3.81 (s, 3 H), 3.24 (q, 2 H, J = 5.2 Hz), 3.05 (dt, 2 H, J = 8.3, 6.7 Hz), 2.76 (d, 6 H, J = 5.5 Hz), 1.84 (quint, 2 H, J = 7.8 Hz); IR (KBr) 3100 (w), 2940 (w), 2700 (w), 1670 (m), 1632 (s), 1575 (s), 1550 (s), 1530 (s), 1464 (m), 1434 (s), 1422 (s), 1260 (m), 1200 (w), 1120 (w) cm^{-1}; FABMS m/e (relative intensity) 452.2399 (2, M + H, 452.2410 calcd. for $C_{23}H_{30}N_7O_3$).

Pyridine-4-carboxamide–netropsin (4-PyN) (4). To a solution of isonicotinic acid (0.052 g, 0.423 mmol) in DMF (3 mL) was added *N,N'*-carbonyldiimidazole (0.068 g, 0.419 mmol). The solution was allowed to stir for 1 h. Separately, a solution of nitropyrrole **10** (0.080 g, 0.212 mmol) and palladium on activated carbon (10%, 15 mg) in DMF (10 mL) was stirred under a hydrogen atmosphere (50 psi) for 5 h. The mixture was filtered through celite and added to the activated acid. The solution was allowed to stir for 18 h and was filtered through celite, and the solvent was removed under reduced pressure. The crude residue was purified by flash column chromatography (gradient 0–1% concentrated aqueous ammonia in methanol) to afford 4-PyN (0.069 g, 72%): UV (H_2O) λ_{max} (ϵ) 209 (22 500), 241 (17 500), 293 (22 500) nm; ^{1}H NMR (400 MHz, DMSO-d_6) δ 10.82 (s, 1 H), 10.00 (s, 1 H), 8.84 (d, 2 H, J = 6.2 Hz), 8.18 (t, 1 H, J = 7.6 Hz), 8.01 (d, 2 H, J = 5.5 Hz), 7.38 (d, 1 H, J = 1.8 Hz), 7.20 (d, 1 H, J = 1.8 Hz), 7.13 (d, 1 H, J = 1.8 Hz), 6.93 (d, 1 H, J = 1.8 Hz), 3.87 (s, 3 H), 3.81 (s, 3 H), 3.23 (q, 2 H, J = 7.6 Hz), 3.02 (t, 2 H, J = 7.4 Hz), 2.74 (s, 6 H), 1.84 (quint, 2 H, J = 7.5 Hz); IR (KBr) 3100 (w), 2950 (w), 2700 (w), 1660 (m), 1640 (s), 1573 (s), 1542 (s), 1462 (m), 1434 (s), 1402 (s), 1288 (w), 1260 (m), 1202 (w) cm^{-1}; FABMS m/e (relative intensity) 452.2431 (6, M + H, 452.2410 calcd. for $C_{23}H_{30}N_7O_3$).

1-Methylimidazole-2-carboxamide–netropsin (2-ImN) (2). To a solution of *N*-methylimidazole-2-carboxylic acid (0.06 mg, 0.532 mmol) and *N*-hydroxybenzotriazole hydrate (0.075 g 0.55 mmol) in DMF (3 mL) was added a solution of dicyclohexylcarbodiimide (0.115 g, 0.55 mmol) in CH_2Cl_2 (2 mL). The mixture was allowed to stir for 12 h. Separately, a solution of nitropyrrole **10** (0.40 g, 1.06 mmol) and palladium on activated carbon (10%, 0.06 g) in DMF (5 mL) was allowed to stir under a hydrogen atmosphere (50 psi) for 6 h. The mixture was filtered through celite and added to the activated acid. The solution was allowed to stir for 12 h at which time methanol (4 mL) was added and solvent was removed under reduced pressure. The crude residue was taken up in ethyl acetate (25 mL) and washed once with 10% aqueous $NaHCO_3$. The organic layer was dried ($MgSO_4$), and solvent was removed under reduced pressure. The product was purified by flash column chromatography (gradient 0–1% concentrated aqueous ammonia in methanol) to afford 2-ImN (0.19 g, 79%) as a yellow oil: UV (H_2O) λ_{max} (ϵ) 255 (19 000), 302 (26 000) nm; ^{1}H NMR (400 MHz, DMSO-d_6) δ 10.92 (bs, 1 H), 10.05 (bs, 1 H), 9.89 (s, 1 H), 8.18 (t, 1 H, J = 6.6 Hz), 7.58 (s, 1 H), 7.35 (s, 1 H), 7.32 (d, 1 H, J = 1.8 Hz), 7.18 (d, 1 H, J = 1.8 Hz), 7.14 (d, 1 H, J = 1.8 Hz), 6.92 (d, 1 H, J = 1.8 Hz), 4.02 (s, 3 H), 3.85 (s, 3 H), 3.80 (s, 3 H), 3.24 (q, 2 H, J = 6.4 Hz), 3.04 (m, 2 H), 2.74 (d, 6 H, J = 5.8 Hz), 1.85 (quint, 2 H, J = 7.4 Hz); FABMS m/e (relative intensity) 455.2519 (0.8, M + H, 455.2519 calcd. for $C_{22}H_{31}N_8O_3$).

Pyridine-2-carboxamide–netropsin–EDTA (5). To a solution of picolinic acid (0.0095 g, 0.0772 mmol) in DMF (1 mL) was added *N,N'*-carbonyldiimidazole (0.0125 g, 0.0772 mmol), and the solution was allowed to stir for 1 h. Separately, a solution of nitropyrrole **12** (0.061 g, 0.0785 mmol) and palladium on activated carbon (10%, 0.030 g) was hydrogenated in a Parr Rocker (50 psi) for 18 h. The solution was filtered, and the solvent was removed under reduced pressure. A solution of the resulting amine in CH_2Cl_2 (2 mL) was added to the activated acid, and the mixture was allowed to stir for 18 h. The solvent was removed under reduced pressure, and the product was purified by flash column chromatography (0.25% concentrated aqueous ammonia in 1:1 methanol:ethanol) to afford 2-ImN-EDTA triester (14 mg, 62%). The triester was dissolved in ethanol (1 mL), 0.5 M LiOH (1 mL) was added, and the solution was allowed to stir for 1 h. Solvent was removed under reduced pressure, and the product was purified by flash column chromatography (1% concentrated aqueous ammonia in 1:1 methanol:ethanol) to afford 2-PyNE (3.9 mg, 62%): UV (H_2O) λ_{max} (ϵ) 245 (21 500), 305 (26 000) nm; ^{1}H NMR (400 MHz, DMSO-d_6 + TFA) δ 10.78 (s, 1 H), 9.95 (s, 1 H), 9.33 (bs, 1 H), 8.70 (dd, 1 H, J = 4.2, 1.7 Hz), 8.48 (t, 1 H, J = 6.8 Hz), 8.18 (m, 1 H), 8.15 (d, 1 H, J = 7.8 Hz), 8.06 (ddd, 1 H, J = 7.9, 7.7, 1.8 Hz), 7.64 (ddd, 1 H, J = 7.8, 4.9, 1.7 Hz), 7.37 (d, 1 H, J = 1.8 Hz), 7.24 (d, 1 H, J = 1.8 Hz), 7.17 (d, 1 H, J = 1.8 Hz), 6.96 (d, 1 H, J = 1.8 Hz), 4.01 (s, 2 H), 3.88 (s, 2 H), 3.85 (s, 3 H), 3.80 (s, 7 H), 3.30–3.00 (m, 12 H), 2.75 (d, 3 H, J = 5.4 Hz), 1.82 (m, 4 H); FABMS m/e (relative intensity) 769.3634 (2, M + H, 769.3633 calcd. for $C_{35}H_{49}N_{10}O_{10}$).

Pyridine-3-carboxamide–netropsin–EDTA (7). To a solution of nicotinic acid (0.0095 g, 0.0772 mmol) in DMF (1 mL) was added *N,N'*-carbonyldiimidazole (0.0125 g, 0.0772 mmol), and the solution was allowed to stir for 1 h. Separately, a solution of nitropyrrole **12** (0.061 g, 0.0785 mmol) and palladium on activated carbon (10%, 0.030 g) was hydrogenated in a Parr Rocker (50 psi) for 18 h, and filtered, and the solvent was removed under reduced pressure. A solution of the resulting amine in CH_2Cl_2 (2 mL) was added to the activated acid, and the mixture was allowed to stir for 18 h. The solvent was removed under reduced pressure, and the product was purified by flash column chromatography (0.25% concentrated aqueous ammonia in 1:1 methanol:ethanol) to afford 3-PyN-EDTA triester (13 mg, 59%). The triester was dissolved in ethanol (1 mL), 0.5 M LiOH (1 mL) was added, and the solution was allowed to stir for 1 h. Solvent was removed under reduced pressure, and the product was purified by flash column chromatography (1% concentrated aqueous ammonia in 1:1 methanol:ethanol) to afford 3-PyNE (3.8 mg, 66%): UV (H_2O) λ_{max} (ϵ) 238 (20 000), 296 (24 000) nm; ^{1}H NMR (400 MHz, DMSO-d_6 + TFA) δ 10.88 (s, 1 H), 9.99 (s, 1 H), 9.35 (bs, 1 H), 9.32 (s, 1 H), 9.01 (d, 1 H, J = 5.7 Hz), 8.88 (dd, 1 H, J = 7.7, 1.7 Hz), 8.47 (t, 1 H, J = 7.4 Hz), 8.18 (t, 1 H, J = 7.5 Hz), 8.08 (dd, 1 H, J = 7.7, 5.8 Hz), 7.36 (d, 1 H, J = 1.8 Hz), 7.17 (d, 1 H, J = 1.8 Hz), 7.12 (d, 1 H, J = 1.8 Hz), 6.96 (d, 1 H, J = 1.8 Hz), 4.01 (s, 3 H), 3.88 (s, 5 H), 3.81 (s, 4 H), 3.80 (s, 3 H), 3.30–3.00 (m, 12 H), 2.76 (d, 3 H, J = 5.2 Hz), 1.83 (m, 4 H); FABMS m/e (relative intensity) 769.3606 (10, M + H, 769.3633 calcd. for $C_{35}H_{49}N_{10}O_{10}$), 743 (20, M + H – CN).

Pyridine-4-carboxamide–netropsin–EDTA (8). To a solution of isonicotinic acid (0.0095 g, 0.0772 mmol) in DMF (1 mL) was added *N,N'*-carbonyldiimidazole (0.0125 g, 0.0772 mmol), and the solution was allowed to stir for 1 h. Separately, a solution of nitropyrrole **12** (0.061 g, 0.0785 mmol) and palladium on activated carbon (10%, 0.030 g) was hydrogenated in a Parr Rocker (50 psi) for 18 h and filtered, and the solvent was removed under reduced pressure. A solution of the resulting amine in CH_2Cl_2 (2 mL) was added to the activated acid, and the mixture was allowed to stir for 18 h. The solvent was removed under reduced pressure, and the product was purified by flash column chromatography (0.25% concentrated aqueous ammonia in 1:1 methanol:ethanol) to afford 4-PyN-EDTA triester (11 mg, 50%). The triester was dissolved in ethanol (1 mL), 0.5 M LiOH (1 mL) was added, and the solution was allowed to stir for 1 h. Solvent was removed under reduced pressure, and the product was purified by flash column chromatography (1% concentrated aqueous ammonia in 1:1 methanol:ethanol) to afford 4-PyNE (3.1 mg, 63%): UV (H_2O) λ_{max} (ϵ) 245 (18 000), 305 (22 500) nm; ^{1}H NMR (400 MHz, DMSO-d_6 + TFA) δ 11.09 (s, 1 H), 9.98 (s, 1 H), 9.33 (bs, 1 H), 9.08 (d, 1 H, J = 7.5 Hz), 8.49 (t, 1 H, J = 7.3 Hz), 8.46 (d, 1 H, J = 7.3 Hz), 8.18 (bs, 1 H), 7.38 (d, 1 H, J = 1.8 Hz), 7.15 (d, 1

H, J = 1.8 Hz), 6.97 (d, 1 H, J = 1.8 Hz), 4.04 (s, 2 H), 3.91 (s, 2 H), 3.88 (s, 3 H), 3.86 (s, 4 H), 3.79 (s, 3 H), 3.40–3.00 (m, 12 H), 2.75 (d, 3 H, J = 5.2 Hz), 1.82 (m, 4 H); FABMS m/e (relative intensity) 769.3638 (0.4, M + H, 769.3633 calcd. for $C_{35}H_{49}N_{10}O_{10}$).

3′-[4-(*N*-Methyl-4-nitropyrrole-2-carboxamido)-*N*-methylpyrrole-2-carboxamido]-3-[*tert*-butyloxycarbonylamino]methyldipropylamine (13). To a solution of nitrodipyrrole **9** (0.850 g, 2.71 mmol), *N*-hydroxybenzotriazole hydrate (0.371 g, 2.75 mmol), and 3-(*tert*-butyloxycarbonylamino)-3′-amino-*N*-methyldipropylamine (1.50 g, 6.12 mmol) in DMF (5 mL) at 0 °C was added dicyclohexylcarbodiimide (0.575 g, 2.75 mmol) in CH_2Cl_2 (5 mL). The reaction mixture was allowed to warm to room temperature and was stirred for 15 h. The mixture was filtered, and the solvent was removed under reduced pressure. The product was purified by flash column chromatography (25% methylene chloride in methanol) to give **10** (1.10 g, 78%): ^{1}H NMR (300 MHz, DMSO-d_6) δ 10.25 (s, 1 H), 10.12 (bs, 1 H), 8.25 (t, 1 H, J = 6.5 Hz), 8.11 (d, 1 H, J = 1.8 Hz), 7.58 (d, 1 H, J = 1.8 Hz), 7.19 (d, 1 H, J = 1.8 Hz), 6.90 (d, 1 H, J = 1.8 Hz), 3.94 (s, 3 H), 3.81 (s, 3 H), 3.38 (m, 2 H), 3.03–3.15 (b, 6 H), 2.70 (d, 3 H, J = 7.3 Hz), 1.89 (m, 2 H), 1.78 (m, 2 H), 1.40 (s, 9 H); IR 3317 (m), 2966 (m), 2358 (w), 1694 (s), 1652 (s), 1574 (m), 1538 (s), 1520 (s), 1310 (s), 1215 (m), 1167 (m), 1116 (w) cm^{-1}; FABMS m/e (relative intensity) 520.2862 (27, M + H, 520.2884 calcd. for $C_{17}H_{25}N_6O_4$), 420 (8, M + H − $C_5H_8O_2$).

1-Methylimidazole-2-carboxamide–netropsin–diaminoazaheptane-tBoc (14). To a solution of *N*-methylimidazole-2-carboxylic acid (0.465 g, 4.19 mmol) and *N*-hydroxybenzotriazole hydrate (0.930 g, 6.88 mmol) in DMF (10 mL), at 0 °C, was added a solution of 1,3-dicyclohexylcarbodiimide (1.00 g, 4.85 mmol) in CH_2Cl_2 (10 mL). The solution was allowed to warm to room temperature and was stirred an additional 10 h. Separately, a solution of nitropyrrole **13** (0.50 g, 0.962 mmol) and palladium on activated carbon (10%, 0.080 g) was allowed to stir under a hydrogen atmosphere (300 psi) in a Parr bomb apparatus for 5 h. The mixture was filtered through celite and added to the activated acid, and the solution was allowed to stir for 2 h. Methanol (3 mL) was added, and the solvent was removed under reduced pressure. The residue was slurried in 10% aqueous $NaHCO_3$ (50 mL) and washed with CH_2Cl_2 (2 × 50 mL). The combined organic layers were dried ($NaSO_4$), concentrated, and purified by flash column chromatography (methanol) to afford **14** (0.490 g, 85%) as a yellow solid: ^{1}H NMR (400 MHz, DMSO-d_6) δ 10.44 (s, 1 H), 9.89 (s, 1 H), 8.01 (t, 1 H, J = 7.3 Hz), 7.38 (s, 1 H), 7.27 (d, 1 H, J = 1.8 Hz), 7.17 (d, 1 H, J = 1.8 Hz), 7.14 (d, 1 H, J = 1.8 Hz), 6.82 (d, 1 H, J = 1.8 Hz), 6.77 (t, 1 H, J = 5.8 Hz), 3.98 (s, 3 H), 3.83 (s, 3 H), 3.78 (s, 3 H), 3.17 (q, 2 H, J = 6.8 Hz), 2.93 (q, 2 H, J = 6.9 Hz), 2.28 (t, 2 H, J = 7.2 Hz), 2.27 (t, 2 H, J = 7.3 Hz), 2.10 (s, 3 H), 1.59 (quint, 2 H, J = 7.2 Hz), 1.50 (quint, 2 H, J = 7.3 Hz), 1.35 (s, 9 H); IR 3322 (m), 2948 (m), 1643 (s), 1582 (m), 1538 (s), 1470 (s), 1428 (s), 1365 (m), 1254 (m), 1168 (m), 1122 (m) cm^{-1}; FABMS m/e (relative intensity) 598.3447 (41, M + H, 598.3464 calcd. for $C_{29}H_{44}N_9O_5$), 498 (19, M + H − $C_5H_8O_2$).

1-Methylimidazole-2-carboxamide–netropsin–diaminoazaheptane (15). To a solution of **14** (0.230 g, 0.385 mmol) in CH_2Cl_2 (8 mL) was added trifluoroacetic acid (2 mL), and the resulting mixture was allowed to stir 25 min. The reaction mixture was triturated with diethyl ether (100 mL), and the residue was dissolved in 1% concentrated aqueous ammonium hydroxide in methanol (30 mL), and the solvent was removed under reduced pressure. The product was purified by flash column chromatography (6% concentrated aqueous ammonium hydroxide in methanol) to afford the primary amine (185 mg, 97%): ^{1}H NMR (400 MHz, DMSO-d_6) δ 10.44 (s, 1 H), 9.92 (s, 1 H), 8.03 (t, 1 H, J = 7.2 Hz), 7.37 (d, 1 H, J = 1.8 Hz), 7.27 (d, 1 H, J = 1.8 Hz), 7.16 (d, 2 H, J = 1.8 Hz), 7.13 (s, 1 H), 7.02 (d, 1 H, J = 1.8 Hz), 3.97 (s, 3 H), 3.82 (s, 3 H), 3.77 (s, 3 H), 3.30 (q, 2 H, J = 7.2 Hz), 3.17 (q, 2 H, J = 6.8 Hz), 2.50 (m, 2 H), 2.28 (t, 2 H, J = 7.0 Hz), 2.10 (s, 3 H), 1.59 (quint, 2 H, J = 7.2 Hz), 1.47 (quint, 2 H, J = 7.1 Hz); IR 3315 (m), 2880 (w), 1642 (s), 1603 (m), 1462 (m), 1433 (m), 1259 (m) cm^{-1}; FABMS m/e (relative intensity) 498.2947 (22, M + H, 498.2941 calcd. for $C_{24}H_{36}N_9O_3$).

1-Methylimidazole-2-carboxamide–netropsin–EDTA, Triethyl Ester (16). A solution of EDTA triethyl ester (0.128 g, 0.34 mmol) and *N*,*N*′-carbonyldiimidazole (55 mg, 0.34 mmol) in CH_2Cl_2 (4 mL) was allowed to stir for 1 h. Primary amine **15** (0.066 g, 0.133 mmol) was added, and the resulting solution was allowed to stir for 12 h. Solvent was removed under reduced pressure, and the product was purified by flash column chromatography (0.25% concentrated aqueous ammonia in 1:1 methanol:ethanol) to afford the triester (0.017 g, 15%): ^{1}H NMR (400 MHz, DMSO-d_6) δ 10.43 (s, 1 H), 9.90 (s, 1 H), 8.01 (t, 1 H, J = 6.8 Hz), 7.95 (t, 1 H, J = 7.1 Hz), 7.39 (s, 1 H), 7.28 (d, 1 H, J = 1.8 Hz), 7.17 (d, 1 H, J = 1.8 Hz), 7.13 (d, 1 H, J = 7.3 Hz), 7.03 (d, 1 H, J = 1.8 Hz), 6.83 (d, 1 H, J = 1.8 Hz), 4.05 (q, 6 H, J = 6.9 Hz), 3.98 (s, 3 H), 3.88 (s, 3 H), 3.74 (s, 3 H), 3.49 (s, 4 H), 3.43 (s, 2 H), 3.18 (s, 2 H), 3.15 (m, 2 H), 3.11 (q, 2 H, J = 7.2 Hz), 2.69 (m, 4 H), 2.31 (t, 2 H, J = 7.1 Hz), 2.29 (t, 2 H, J = 7.1 Hz), 2.13 (s, 3 H), 1.6 (m, 4 H), 1.16 (t, 9 H, J = 7.3 Hz).

1-Methylimidazole-2-carboxamide–netropsin–EDTA (6). A solution of triethyl ester **16** (14 mg, 0.016 mmol) in ethanol (2 mL) and 0.5 M LiOH (1 mL) was allowed to stir for 16 h. The solvent was removed under reduced pressure, and the product was purified by flash chromatography (1% concentrated aqueous ammonia in 1:1 methanol:ethanol) to afford 2-ImNE (0.0034 g, 27%): UV (H_2O) λ_{max} (ϵ) 252 (19 000), 304 (26 500) nm; ^{1}H NMR (400 MHz, DMSO-d_6 + TFA) δ 10.96 (s, 1 H), 9.98 (s, 1 H), 9.36 (bs, 1 H), 8.47 (t, 1 H, J = 6.8 Hz), 8.19 (t, 1 H, J = 7.2 Hz), 7.70 (s, 1 H), 7.54 (s, 1 H), 7.32 (d, 1 H, J = 1.8 Hz), 7.16 (d, 1 H, J = 1.8 Hz), 7.12 (d, 1 H, J = 1.8 Hz), 6.95 (d, 1 H, J = 1.8 Hz), 4.02 (s, 5 H), 3.87 (s, 5 H), 3.81 (s, 4 H), 3.79 (s, 2 H), 3.78 (s, 3 H), 3.30–3.00 (m, 12 H), 2.75 (s, 3 H), 1.84 (m, 4 H); FABMS m/e (relative intensity) 772.3746 (14, M + H, 772.3742 calcd. for $C_{34}H_{50}N_{11}O_{10}$), 549 (39), 507 (17).

DNA Reagents and Materials. All water was distilled, filtered through an organic removal cartidge (Corning), and redistilled. For DNA manipulations, the water and all buffers used were autoclaved for 20 min at 160 °C. Acrylamide was purchased as a 30% solution from National Diagnostics. To each liter was added 15 g (1.5%) *N*,*N*′-methylenebisacrylamide. Calf thymus DNA was sonicated for 1 min and phenol extracted. The solution was exhaustively dialyzed against water and diluted to 1 mM bp. Sigma type XX tRNA was deproteinized by phenol extraction and diluted to 1 mg/mL. Plasmid pBR322 was grown in HB101 cells and purified by CsCl gradient.[31,32] Chemical sequencing reactions were performed according to published methods.[33,34] All other reagents and materials were used as received.

Sample Preparation. Milligram quantities of the compounds were weighed on a Sartorius microbalance and diluted to 100 mL with water. The average molar extinction coefficient from three measurements was used to determine the concentration of a stock solution, which was lyophilized in 600 μL double-lock eppendorf tubes and stored dry at −20 °C.

MPE·Fe(II) Footprinting.[3] A 20 μM MPE·Fe(II) solution was prepared by mixing 10 μL of a 1 mM MPE solution with 10 μM of a freshly prepared 1 mM ferrous ammonium sulfate solution, then diluting to 500 μL. Solutions were prepared containing 1 μL/tube 20X TA buffer (800 mM Tris, 100 mM sodium acetate, pH 7.9), 2 μL/tube 1 mM bp calf thymus DNA, labeled restriction fragment, and water to make 8 μL/tube total solution. 4 μL of a 5X solution of the compound was added, and the tubes were incubated in the dark for 30 min at 37 °C. To each tube was added 4 μL of the 20 μM MPE·Fe(II) solution followed by 4 μL of a freshly prepared 20 mM DTT solution. Final concentrations were 40 mM Tris acetate (pH 7.9), 5 mM sodium acetate, 100 μM bp DNA, 4 μM MPE·Fe(II), and 4 mM DTT, in 20 μL. The reactions were incubated at 37 °C for 10 min, lyophilized, and electrophoresed.

Affinity Cleavage.[4] Solutions were prepared containing 1 μL/tube 20X TA buffer, 2 μL/tube 1 mM bp calf thymus DNA, labeled restriction fragment, and water to make 12 μL/tube. The compounds to be examined were loaded with Fe^{2+} by adding 10 μL of a freshly prepared 2 mM ferrous ammonium sulfate solution to 10 μL of a 2 mM solution of the compound followed by appropriate dilution with water. 4 μL of the compound was added to each tube, and the solutions were incubated at 37 °C for 30 min in the dark. DTT (4 μL, 20 mM) was added, and the reactions were incubated in the dark at 37 °C for 30 min. Final concentrations were 40 mM pH 7.9 Tris acetate, 100 μM bp DNA, and 4 mM DTT, in a total volume of 20 μL. The reactions were lyophilized and electrophoresed.

Preparation of Linearized pBR322 Labeled at Only One Terminus. pBR322 plasmid was digested with *Sty*I and then labeled on the 3′ terminus of the clockwise strand with [α-^{32}P]dATP or on the 3′ terminus of the counterclockwise strand with [α-^{32}P]TTP using the Klenow fragment of DNA polymerase I.[32] The radiolabeled plasmid was purified on a 0.7% agarose gel containing ethidium bromide. The band containing linearized plasmid was excised, diluted with 500 μL TE buffer (20 mM Tris, 1 mM EDTA, pH 7.9), melted at 70 °C for 10 min, phenol extracted, and ethanol precipitated.

Molecular Weight Standards. Equal amounts of the [α-^{32}P]dATP and [α-^{32}P]TTP labeled plasmid were combined with 1 μg pBR322. Individual samples were digested with the restriction endonucleases *Eco*RI,

(31) (a) Sutcliffe, J. G. *Cold Spring Harbor Symposium on Quantitative Biology* **1978**, *43*, 77–90. (b) Peden, K. W. C. *Gene* **1982**, *22*, 277–280.
(32) Maniatis, T.; Fritsch, E. F.; Sambrook, J. *Molecular Cloning, A Laboratory Manual*; Cold Spring Harbor Laboratory: New York, 1982; pp 1–542.
(33) Iverson, B. L.; Dervan, P. B. *Nucleic Acids Res.* **1987**, *15*, 7823–7830.
(34) Maxam, A. M.; Gilbert, W. S. *Methods in Enzymol.* **1980**, *65*, 499–560.

*Xmn*I, *Bam*HI, and *Pst*I or *Pvu*II. The reactions were ethanol precipitated, counted, and diluted to the same concentration (cpm/μL), and equal volumes were combined with the same cpm of intact pBR322.

Cleavage of Linear pBR322. Solutions were prepared containing 0.75 μL/tube 20X TA buffer, 1.5 μL 1 mM bp calf thymus DNA, radiolabeled pBR322, and water to make 9 μL. 3 μL of a 5X solution of compound·Fe(II) was added, and the solutions were incubated at 37 °C for 30 min in the dark. 3 μL of a 5 mM sodium ascorbate solution was added, and the reactions were incubated at 37 °C for 2 h. Final concentrations were 40 mM pH 7.9 Tris acetate, 5 mM sodium acetate, 100 μM bp DNA, and 1 mM sodium ascorbate, in a total volume of 15 μL. The reactions were diluted with Ficoll loading buffer and electrophoresed on a 1% vertical agarose gel.

Densitometry. Footprinting and affinity cleaving autoradiograms were scanned on a LKB XL laser densitometer. The scans were output to an IBM printer, and footprints were determined by comparison to the MPE·Fe(II) standard lane. A horizontal line was drawn from the top of the unprotected band nearest the footprint, and the distance in absorbance units from this line to the maximum peak height was determined for each band and plotted as a histogram. Affinity cleavage patterns were measured in similar fashion, using peak heights of sites without specific cleavage as a baseline. To determine the sites of cleavage in the double strand cleavage assay, the cleavage lanes were scanned as above and output to the Hoeffer program GS370 through an analog to digital converter. Cleavage band molecular weights were determined by comparison to the molecular weight standard lane, and positions on pBR322 were calculated. The positions were averaged between labels and used, along with the average relative area of the peaks, to generate histograms.

Acknowledgment. We are grateful to the National Institutes of Health (GM-27681) and to Burroughs Wellcome for research support, to the National Science Foundation for a predoctoral fellowship to W.S.W., and to the National Institutes of Health for a Research Service Award to M.M.

What Causes Aqueous Acceleration of the Claisen Rearrangement?

Christopher J. Cramer*,† and Donald G. Truhlar*,‡

Contribution from the U.S. Army Chemical Research Development and Engineering Center, Aberdeen Proving Ground, Maryland 21010-5423, and Department of Chemistry and Supercomputer Institute, University of Minnesota, Minneapolis, Minnesota 55455-0431.
Received May 26, 1992

Abstract: We report the results of applying a new self-consistent-field solvation model to the Claisen rearrangement of allyl vinyl ether, all possible methoxy-substituted derivatives, two alkylated derivatives, and one carboxymethylated derivative in order to understand the effects of aqueous solvation on the reaction rates. We have employed the AM1-SM2 version of the model to calculate the changes in free energies of solvation in passing from the lowest-energy conformations of the starting materials to both chair and boat transition states. The hydrophobic effect is always accelerative but always small and not very structure sensitive. Other first-hydration-shell effects attributable to hydrophilic parts of the reagents are more sensitive to the substitution pattern. The polarization contributions to the activation energies are usually larger. A favorable polarization contribution is found to be associated with efficient sequestration of charges of opposite sign into separated regions of space. We conclude that aqueous acceleration of the Claisen rearrangement is caused by electric polarization and first-hydration-shell hydrophilic effects, with the relative magnitudes and even the signs of these effects being quite sensitive to substitution pattern.

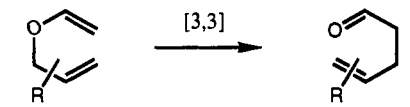
Introduction

Pericyclic reactions in aqueous solution can exhibit both rate acceleration and improved diastereoselection.[1] Increased rates have been in part attributed to hydrophobic interactions of nonpolar solutes with each other and with water. However, Blake and Jorgensen have suggested that improved solvation of a more polar transition state (TS) relative to starting material may also play a significant role in one particular class of pericyclic processes, Diels-Alder reactions.[2]

The Claisen rearrangement (Chart I) is a different example of a pericyclic reaction, a [3,3] sigmatropic shift, which also shows considerable acceleration on going from nonpolar to aqueous solvents.[3-6] By its very nature as an oxa-Cope rearrangement, even for the rearrangement of the prototype allyl vinyl ether, it is clear that both polar and hydrophobic (or hydrophilic) effects may readily contribute to the overall observed acceleration. Experimentally, it is difficult to discern the relative contributions to transition-state lowering from these effects.

We have recently described[7,8] a semiempirical, effective-Hamiltonian SCF approach[9] for modeling aqueous solvation. The

Chart I



method incorporates solvent-induced polarization[10] and structural relaxation[11] of the solute (whether equilibrium structure or

*Correspondence may be addressed to either author at University of Minnesota.
†U.S. Army Chemical Research Development and Engineering Center and Department of Chemistry, University of Minnesota.
‡Department of Chemistry and Supercomputer Institute, University of Minnesota.

(1) (a) Breslow, R. *Acc. Chem. Res.* **1991**, *24*, 159. (b) Grieco, P. A. *Aldrichim. Acta* **1991**, *24*, 59. (c) Blokzijl, W.; Blandamer, M. J.; Engberts, J. B. F. N. *J. Am. Chem. Soc.* **1991**, *113*, 4241. (d) Cativiela, C.; Garcia, J. I.; Mayoral, J. A.; Avenoza, A.; Peregrina, J. M.; Roy, M. A. *J. Phys. Org. Chem.* **1991**, *4*, 48.
(2) Blake, J. F.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1991**, *113*, 7430.
(3) White, W. N.; Wolfarth, E. F. *J. Org. Chem.* **1970**, *35*, 2196, 3585.
(4) Coates, R. M.; Rogers, B. D.; Hobbs, S. J.; Peck, D. R.; Curran, D. P. *J. Am. Chem. Soc.* **1987**, *109*, 1160.
(5) Gajewski, J. J.; Jurayj, J.; Kimbrough, D. R.; Gande, M. E.; Ganem, B.; Carpenter, B. K. *J. Am. Chem. Soc.* **1987**, *109*, 1170.
(6) Brandes, E.; Grieco, P. A.; Gajewski, J. J. *J. Org. Chem.* **1989**, *54*, 515.
(7) Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1991**, *113*, 8305, 9901(E).
(8) Cramer, C. J.; Truhlar, D. G. *Science* **1992**, *256*, 213.
(9) Tapia, O. In *Quantum Theory of Chemical Reactions*; Daudel, R., Pullman, A., Salem, L., Veillard, A., Eds.; Reidel: Dordrecht, 1980; Vol. 2, p 25.
(10) Aleman, C.; Maseras, F.; Lledós, A.; Duran, M.; Bertrán, J. *J. Phys. Chem.* **1989**, *2*, 611. Alagona, G.; Ghio, C.; Igual, J.; Tomasi, J. *J. Mol. Struct. (Theochem)* **1990**, *204*, 253.

0002-7863/92/1514-8794$03.00/0