Lienhard, G. E. (1973) *Science (Washington, D.C.) 180*, 149-154.
Liu, W. S., Smith, S. C., & Glover, G. I. (1979) *J. Med. Chem. 22*, 577-579.
Ondetti, M. A., & Cushman, D. W. (1980) *Annu. Rev. Biochem. 51*, 283-308.
Pearl, L. H., & Blundell, T. L. (1984) *FEBS Lett. 174*, 96-101.
Plattner, J. J., Greer, J., Fung, A. K. L., Stein, H., Kleinert, H. D., Sham, H. L., Smital, J. R., & Perun, T. J. (1986) *Biochem. Biophys. Res. Commun. 139*, 982-990.
Rich, D. H., & Sun, E. T. O. (1980) *J. Med. Chem. 23*, 27-33.
Rich, D. H., Bernatowicz, M. S., & Schmidt, P. G. (1982) *J. Am. Chem. Soc. 104*, 3535-3536.
Sawyer, T. K., Pals, D. T., Smith, C. W., Saneii, W. H., Epps, D. E., Duchamp, D. J., Hester, J. B., TenBrink, R. E., Staples, D. J., deVaux, A. E., Affholter, J. A., Skala, G. F., Kati, W. M., Lawson, J. A., Schuette, M. R., Kamdar, B. V., Emmert, D. E., Carlson, W. D., & Handsumacher, M. (1985) in *Proceedings of the 9th American Peptide Symposium* (Deber, C. M., Hruby, V. J., & Kopple, K. D., Eds.) pp 729-738, Pierce Chemical, Rockford, IL.
Sibanda, B. L. Blundell, T. L., Hobart, P. M., Fogliano, M., Bindra, J. S., Dominy, B. W., & Chirgwin, J. M. (1984) *FEBS Lett. 174*, 102-111.
Sibanda, B. L., Hemmings, A. M., & Blundell, T. L. (1985) in *Aspartic Proteinases and Their Inhibitors* (Kostka, V., Ed.) pp 339-349, de Gruyter, Berlin.
Skeggs, L. T., Lentz, K., Hochstrasser, H., & Kahn, J. R. (1964) *Can. Med. Assoc. J. 90*, 185-190.
Skeggs, L. T., Dover, F. E., Levine, M., Lentz, K. E., & Kahn, J. R. (1980) in the *The Renin-Angiotensin System* (Johnson, J. A., & Anderson, R. R., Eds.) p 1, Plenum, New York.
Szelke, M. (1985) in *Aspartic Proteinases and Their Inhibitors* (Kostka, V., Ed.) pp 421-441, de Gruyter, Berlin.
Szelke, M., Leckie, B., Hallett, A., Jones, D. M., Sueiras, J., Atrash, B., & Lever, A. F. (1982) *Nature (London) 299*, 555-557.
Thomas, K. A., Smith, G. M., Thomas, T. B., & Feldmann, R. J. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, (4843-4847.
Tree, M., Donovan, B., Gamble, J., Hallett, A., Hughes, M., Jones, D. M., Leckie, B., Lever, A. F., Morton, J. J., & Szelke, M. (1983) *J. Hypertens. 1*, 399-403.
Umezawa, H., Aoyagi, T., Morishima, H., Matzusaki, M., Hamada, H., & Takeuchi, T. (1970) *J. Antibiot. 23*, 259-262.
Wolfenden, R. (1972) *Acc. Chem. Res. 5*, 10-18.
Wood, J. M., Fuhrer, W., Buhlmayer, P., Riniker, B., & Hofbauer, K. G. (1985) in *Aspartic Proteinases and Their Inhibitors* (Kostka, V., Ed.) pp 463-466, de Gruyter, Berlin.
Workman, R. J., & Burkitt, D. W. (1979) *Arch. Biochem. Biophys. 194*, 157-164.

Accelerated Publications

Molecular Recognition between Oligopeptides and Nucleic Acids. Monocationic Imidazole Lexitropsins That Display Enhanced GC Sequence Dependent DNA Binding†

Koren Kissinger,‡ Krzysztof Krowicki,§ James C. Dabrowiak,*,‡ and J. William Lown*,§

Department of Chemistry, Syracuse University, Syracuse, New York 13244-1200, and Department of Chemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2G2

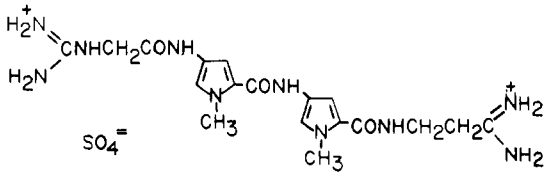
Received May 19, 1987; Revised Manuscript Received June 24, 1987

ABSTRACT: A series of monocationic lexitropsins, or information-reading oligopeptides, were synthesized to minimize and offset the AT bias for doubly cationic ligands bound in the minor groove of DNA. The compounds possess an *N*-formyl group in place of the guanidinium moiety normally present in netropsin. By systematic replacement of the *N*-methylpyrrole groups of the dipeptide with *N*-methylimidazole, a remarkably high degree of sequence specificity was obtained. One of the compounds having two *N*-methylimidazole residues was found to exhibit dramatically altered specificity when compared with netropsin and preferred to bind to the sequence $^{5'\text{-CCGT-}3'}_{3'\text{-GGCA-}5'}$. The structural elements underlying sequence recognition in terms of the model for the netropsin-DNA interaction are presented and discussed.

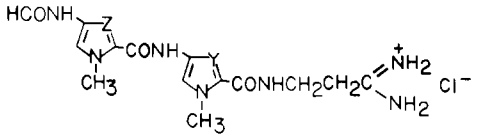
The combination of synthetic organic chemistry with DNA footprinting methodology is a powerful way to uncover the structural elements important in ligand-DNA sequence recognition. The antiviral agent netropsin (**1**) is particularly well suited for this approach since the compound has been extensively modified and the structure of its DNA complex at the atomic resolution level is known. A single crystal X-ray analysis of the compound bound to a dodecanucleotide showed that it interacts with four A·T base pairs by displacing the spine of hydration located in the minor groove of double-stranded DNA (Kopka et al., 1985). The drug lies in the center of the groove with its three amide hydrogen atoms engaged in bifurcated hydrogen bonds with O-2 of thymine and N-3 of

† This investigation was supported by grants (to J.W.L.) from the National Cancer Institute of Canada and the Biotechnology Strategic Grants Programme of the Natural Sciences and Engineering Research Council of Canada and by Grant GM 31895 (to J.C.D.) from the National Institutes of Health.
‡ Syracuse University.
§ University of Alberta.

0006-2960/87/0426-5590$01.50/0



I Netropsin



2 Z=CH, Y=CH
3 Z=N, Y=CH
4 Z=CH, Y=N
5 Z=N, Y=N

adenine. In exploring the factors controlling specificity, Dervan and co-workers have synthesized a number of poly(*N*-methylpyrrolyl)netropsin analogues (Dervan, 1986). When the length of the peptide was extended, it was found that the binding site size is one nucleotide longer than the number of amide groups present in the peptide framework.

In attempting to uncover the structural elements important in ligand–DNA sequence recognition, we have focused on a group of dipeptide netropsin analogues. Since the compounds require relatively modest length DNA binding sequences, it is possible to use short segments of DNA possessing many possible binding sequences and to employ conventional footprinting methods to access specificity.

In an earlier paper we described the sequence specificities of two *N*-methylimidazole analogues of netropsin, termed lexitropsins or information-reading oligopeptides (Kopka et al., 1985), toward a 139-base-pair segment of pBR322 DNA as probed by DNase I footprinting methodology (Lown et al., 1986a). Inclusion of one or two *N*-methylimidazole moieties into the dipeptide significantly relaxed its A·T binding specificity, suggesting that the compounds are engaging in hydrogen bonding to the 2-amino group of guanine located in the minor groove of DNA.

Recent ab initio calculations indicate that the deepest negative potential wells in DNA occur in the minor groove for AT runs and on the periphery of the major groove for GC runs (Laverly & Pullman, 1985). Since the prototype lexitropsins are dicationic, they exhibit a bias toward AT sequences which opposes the GC preference of the imidazole moiety (Lown et al., 1986a). Accordingly, we report the synthesis and DNA binding specificities of a group of second-generation lexitropsins that possess only one positive charge. These compounds have been modified by replacing the guanidinium group normally present on the amino terminus of netropsin with an *N*-formyl moiety.

One of the compounds studied shows a high affinity for the tetranucleotide sequence 5′-CCGT-3′. Since other sequences of type $(G\cdot C)_3(A\cdot T)$ were present in the restriction fragment studied, not only does the compound possess dramatically altered specificity when compared to netropsin but also, to a significant degree, the new agent is a reading DNA sequence.

Materials and Methods

Chemicals. The required compound **2** was already reported (Lown & Krowicki, 1985). The other compounds, **3**–**5**, were obtained by methods based essentially on our previous syntheses of **2** starting with the corresponding amino derivatives and using formylimidazole as the formylating agent. However, due to low solubilities of the amino derivatives containing imidazole moieties in methanol, Me_2SO[1] was used as the solvent. Melting points were determined on an Electrothermal melting point apparatus and are uncorrected. The IR spectra were recorded on a Nicolet 7199 FT spectrophotometer, and only the principal sharply defined peaks are reported. The ^1H NMR spectra were recorded on Bruker WH-200 and WH-400 cryospectrometers. FAB (fast atom bombardment) mass spectra were determined on Associated Electrical Industry (AEI) MS-9 and MS-50 double-focusing mass spectrometers. Kieselgel 60 (230–400 mesh) from E. Merck was used for flash chromatography, and precoated sheets of silica gel 60F-254 from E. Merck were used for TLC. The TLC system of methanol with some acetic acid was used.

[[*1-Methyl-4-*[[*1-methyl-4-(formylamido)imidazol-2-yl*]*carboxamido*]*pyrrol-2-yl*]*carboxamido*]*propionamidine Hydrochloride* (**3**). [[1-Methyl-4-[(1-methyl-4-aminoimidazol-2-yl)carboxamido]pyrrol-2-yl]carboxamido]propionamidine hydrochloride (147.2 mg, 0.4 mmol) was dissolved in Me_2SO (2 mL), and a solution of formylimidazole [from 162 mg (1 mmol) of carbonyldiimidazole and 38 μL (1 mmol) of formic acid] in Me_2SO was added. After half an hour at room temperature methanol was added to react with the excess formylimidazole, and the mixture was evaporated to dryness under reduced pressure. The residue was purified by flash column chromatography with MeOH as eluent. The resulting solid was dissolved in MeOH and precipitated with a mixture of acetonitrile and ethyl acetate to give 77 mg (48.5% yield) of pure product **3**: mp softens at 210 °C and melts at 230 °C; ^1H NMR ($[^2H_6]Me_2SO$) δ 2.65 (t, 2 H), 3.51 (q, 2 H), 3.82 (s, 3 H), 3.96 (s, 3 H), 7.07 (d, 1 H), 7.26 (d, 1 H), 7.48 (s, 1 H), 8.20 (s, 1 H), 8.30 (t, 1 H), 8.98 (br s, 4 H), 10.07 (s, 1 H), 10.43 (s, 1 H); IR (Nujol) 1284, 1377, 1441, 1465, 1549, 1569, 1669, 3300 cm^{-1}; FAB MS *m/z* 361 [$(M - Cl)^+$]. Anal. Calcd for $C_{15}H_{21}ClN_8O_3$: C, 45.4; H, 5.3; N, 28.2; Cl, 8.9. Found: C, 45.2; H, 5.5; N, 28.5; Cl, 9.1.

[[*1-Methyl-4-*[[*1-methyl-4-(formylamido)pyrrol-2-yl*]*carboxamido*]*imidazol-2-yl*]*carboxamido*]*propionamidine Hydrochloride* (**4**). [[1-Methyl-4-[(1-methyl-4-aminopyrrol-2-yl)carboxamido]imidazol-2-yl]carboxamido]propionamidine hydrochloride (147 mg, 0.4 mmol) was dissolved in Me_2SO (2 mL) and allowed to react with formylimidazole following the same procedure as for the compound **3**. After flash chromatography 80 mg (50.5% yield) of pure **4** was obtained: mp softens at 180 °C and melts at 210–215 °C; ^1H NMR ($[^2H_6]Me_2SO$) δ 2.66 (t, 2 H), 3.62 (q, 2 H), 3.86 (s, 3 H), 3.96 (s, 3 H), 7.00 (d, 1 H), 7.31 (d, 1 H), 7.55 (s, 1 H), 8.13 (s and t overlapped, 2 H), 9.06 (s, 4 H), 10.20 (s, 1 H), 10.36 (s, 1 H); IR (Nujol) 1314, 1366, 1377, 1447, 1461, 1530, 1594, 1642, 1678, 1687, 3102, 3135, 3330, 3400 cm^{-1}; FAB MS *m/z* 361 [$(M - Cl)^+$]. Anal. Calcd for $C_{15}H_{21}ClN_8O_3$: C, 45.4; H, 5.3; N, 28.2; Cl, 8.9. Found: C, 45.1; H, 5.5; N, 28.4; Cl, 9.0.

[[*1-Methyl-4-*[[*1-methyl-4-(formylamido)imidazol-2-yl*]*carboxamido*]*imidazol-2-yl*]*carboxamido*]*propionamidine Hydrochloride* (**5**). [[1-Methyl-4-[(1-methyl-4-aminoimidazol-2-yl)carboxamido]imidazol-2-yl]carboxamido]propionamidine hydrochloride (100 mg, 0.27 mmol) was suspended in dry Me_2SO (3 mL), and formylimidazole (from 162 mg of carbonyldiimidazole and 38 μL of formic acid in 1 mL of Me_2SO) was added. The mixture was stirred at room

[1] Abbreviations: NMR, nuclear magnetic resonace; MS, mass spectrometry; FAB, fast atom bombardment; TLC, thin-layer chromatography; IR, infrared; FT, Fourier transform; DMA, dimethylacetamide; Me_2SO, dimethyl sulfoxide; EDTA, ethylenediaminetetraacetic acid; Tris, tris(hydroxymethyl)aminomethane.

temperature until the solid dissolved. Methanol was added and solvents were evaporated to dryness in vacuo. The residue was purified by fractionated precipitation from a 2-propanol solution with ethyl acetate. The first fractions were impurities. A total of 87 mg (81% yield) of pure **5** was obtained: mp 161–164 °C; ^{1}H NMR ([$^2H_6$]Me_2SO) δ 2.63 (t, 2 H), 3.56 (q, 2 H), 3.98 and 4.00 (2 s, 6 H), 7.55 and 7.57 (2 s, 2 H), 8.25 (s, 1 H), 8.55 (br s, 1 H), 9.38 (vbr s, 6 H); IR (Nujol) 1376, 1465, 1540, 1566, 1675, 1747, 3260 cm^{-1}; FAB MS *m/z* 362 [(M – Cl)$^+$], 723 [(2M – Cl – HCl)$^+$]. Anal. Calcd for $C_{14}H_{20}ClN_9O_3$: C, 42.3; H, 5.1; N, 31.7; Cl, 8.9. Found: C, 42.0; H, 5.2; N, 31.9; Cl, 9.1.

Biochemicals. The plasmid pBR322 DNA was prepared as previously published (Maniatis et al., 1982). Calf thymus DNA obtained from Sigma was deproteinized by extraction with phenol and a 24:1 chloroform–isoamyl alcohol mixture (Maniatis et al., 1982). Bovine pancreatic DNase I (type DN-CS) was obtained from Sigma in a lyophilized form and dissolved in 50% glycerol–50 mM Tris-HCl, pH 7.0. The radiolabeled pBR322 DNA fragment was isolated, labeled at its 3′ end (position 33), and purified as earlier described (Lown et al., 1986b). The numbering index for the fragment is the standard numbering system for pBR322 (Maniatis et al., 1982).

DNase I Footprinting. Concentrations of the ligands were determined by weight. Solutions for footprinting experiments were prepared by mixing the following: 2 μL of calf thymus DNA, 770 μM base pairs in 50 mM Tris-HCl, pH 7.0, buffer; 2 μL of the fragment having a concentration of ~7 μM in buffer; 2 μL of buffer (control) or ligand; and 2 μL of the activated enzyme complex. The final DNA concentration in these studies was ~194 μM: 193 μM calf thymus DNA and ~1 μM fragment. The enzyme complex was prepared by the addition of 5 μL of a stock solution of the enzyme (2 units/μL) to 95 μL of buffer containing 50 mM Tris-HCl, 32 mM $MgCl_2$, and 8 mM $CaCl_2$, pH 7.5. Prior to the addition of the enzyme complex, the DNA–ligand mixtures were equilibrated for 30 min at 4 °C. For each compound the ratios of ligand to DNA base pairs studied, r_t, were 0, 0.004, 0.008, 0.015, 0.030, 0.06, 0.12, 0.25, 0.5, and 1. After 10 min at 37 °C, digestions were terminated by the addition of 10 μL of a solution containing 70% aqueous urea, 20 mM EDTA, and 0.025% each of bromophenol blue and xylene cyanol.

Electrophoresis, autoradiography, and microdensitometry were carried out as described previously (Dabrowiak et al., 1986; Ward et al., 1986). Owing to poor separation of the longer labeled oligomers in the gel, the resolution in the region <~110 was at the single nucleotide level, but in the region >~110, resolution was lower. Strong binding gave rise to inhibition at low values of ligand concentration while weak binding caused inhibition only at high values of ligand concentration. The densitometric scans selected for presentation were demonstrative of strong binding, r_t of 0.06, and weak binding, r_t of 1.0, to the fragment.

Results

Densitometric scans of selected autoradiographic data for compounds **2**–**5** are presented in Figures 1 and 2, and a summary of the observed inhibitions and enhancements on the sequence of the restriction fragment is given in Figure 3.

The strong inhibition sites for the bis(*N*-methylpyrrole) compound Py-Py (**2**) in AT-rich regions of the fragment are similar to those of netropsin (Lown et al., 1986a; Ward et al., 1987). Strong inhibitions (strong binding) were found at positions (all 5′ → 3′) 50–48 (ATA), 63–55 (CAATTTAAC), 100–80 (ATTGTTAGATTTCATACACGG), ~145 to ~135 (GCCTACAGCAT), and ~160 to ~153 (CATAACCA). For this compound weaker binding was observed at 54–51 (TGTG) and 68–64 (GTTAG). Enhancement in the DNase I cleavage rate occurred in GC-rich regions of the fragment at positions 78–69, ~134 to 101, and ~150 to ~145.

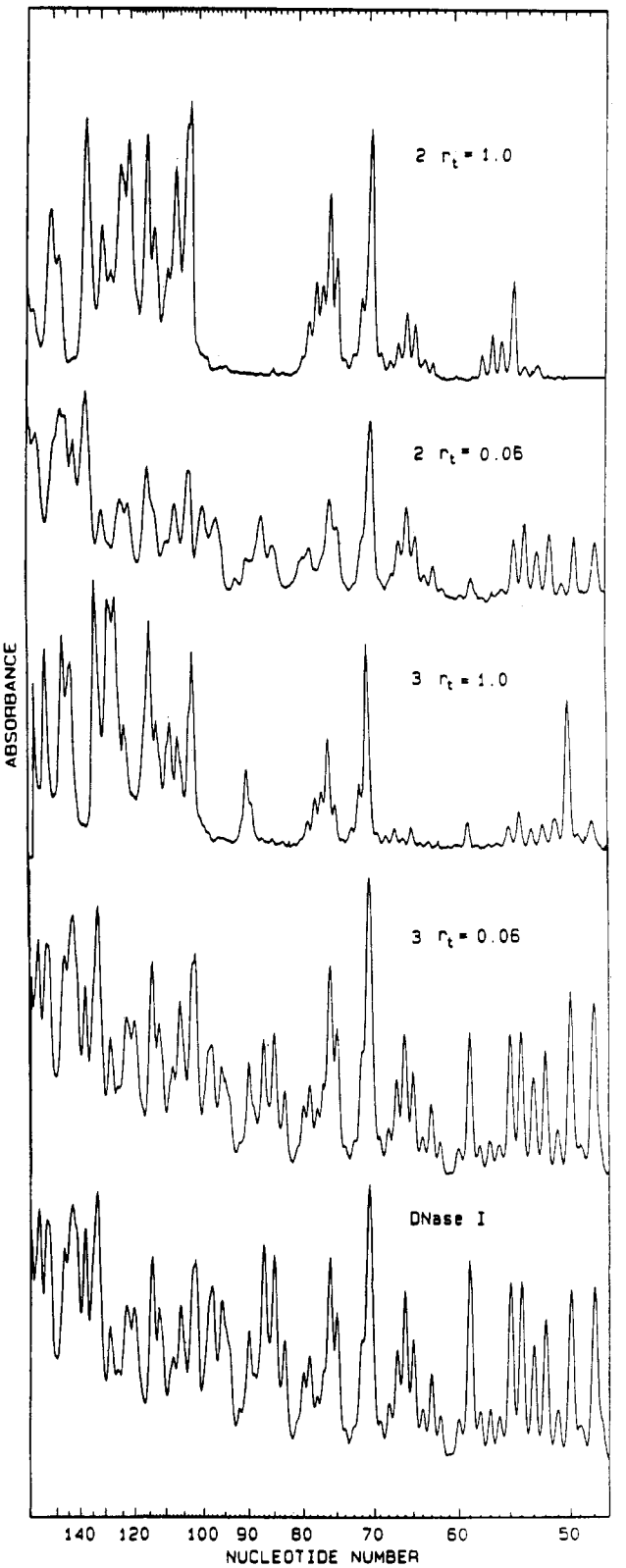


FIGURE 1: Densitometric scans of footprinting autoradiographic data for compound **2** (Py-Py) and compound **3** (Im-Py) and for DNase I alone.

The monoimidazole monopyrrole compound Im-Py (3) bound to a larger number sites on the 139-mer than did **2** and

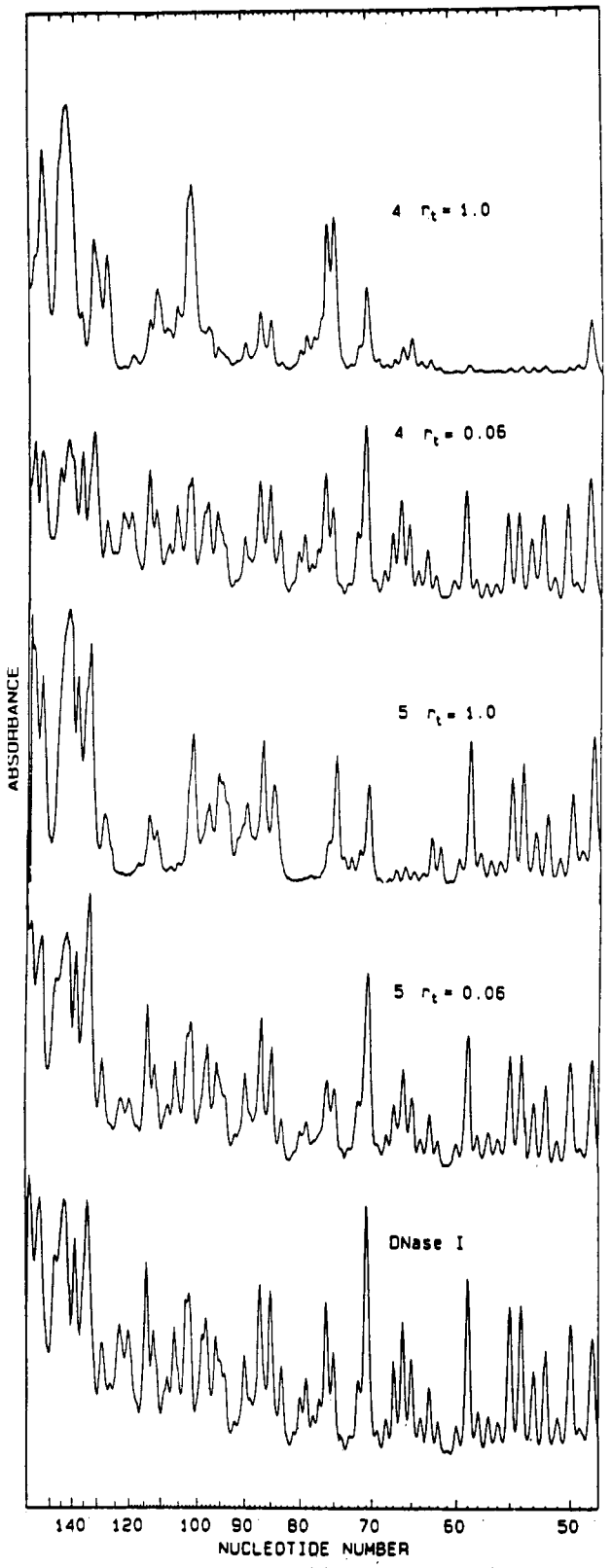


FIGURE 2: Densitometric scans of footprinting data for compound **4** (Py-Im), compound **5** (Im-Im), and DNase I.

netropsin. The compound exhibited weak binding in primarily GC-rich regions where **1** and **2** only exhibited enhancements (Figures 1 and 3). Noteworthy in connection with **3** is the observation that the sequence 5′-ATTT-3′ (92–89) is not a strong binding site as it is with **1** and **2**.

As is evident from Figures 2 and 3, changing the relative position of the imidazole and pyrrole residues of the dipeptide, i.e., Py-Im (**4**), has a dramatic effect on the binding sequence preference of the compound. Analogue **4** exhibits its strongest sites in regions of DNA having significant GC content, e.g., 62–50 (AATTTAACTGTGA) and ~127 to ~113 (GACGGTGCCGAGGAT).

Compound **5** was found to have only two strong binding sites on the fragment at positions 85–76 and ~127 to ~119. Except for the base at the 5′ end of the sequence, both binding sequences were identical, 5′-(C or G)ACGGTGCC-3′. This compound also exhibited weaker binding at positions 72–64 (CTGCGTTAG), ~119–104 (GAGGATGACGATGAG), and ~140 to ~128 (CAGCATCCAGGGT). The limited number of enhancements observed for **5** were found at positions 49–48, 75–73, and ~145 to ~141.

Preliminary analysis of the autoradiographic data (Brenowitz et al., 1986; Ward et al., 1987) shows that the binding constants for all the compounds are in the range 10^4–10^5 M^{-1}. While the apparent difference in the binding constants for the primary and secondary sites for compounds **2**–**4** is about 1 order of magnitude, the difference for the bis(imidazole) analogue **5** is ~10^2 M^{-1}.

Discussion

Replacement of the charged guanidinium group of the amino terminus of netropsin with a formyl group results in a reduction of total charge on the ligand from +2 to +1. As is evident from the netropsin–DNA crystal structure (Kopka et al., 1985), this substitution also affects a van der Waals contact between the ligand and DNA. The X-ray structural analysis of the drug bound to the sequence 5′-CGCGAATTCGCG-3′ established that the methylene group of the guanidinium residue enters into a van der Waals contact with the C-2 hydrogen of adenine in the minor groove of the duplex (Kopka et al., 1985). In addition to altered charge and van der Waals contacts, the lexitropsins possess one or two *N*-methylimidazole groups which allow the compounds to hydrogen bond to the 2-amino group of guanine located in the minor groove of DNA (Lown et al., 1986a).

In identifying the various ligand binding sequences on the 139-mer from the observed inhibition patterns, it was necessary to account for the cleavage geometry of DNase I on DNA. Previous quantitative footprinting studies with netropsin (Ward et al., 1987) and the crystal structure of DNase (Suck & Oefner, 1986) showed that a minor groove binding loop on the enzyme prevents it from cleaving within ~4 base pairs of the 3′ direction of the ligand on DNA. Since the drug itself is ~4 base pairs in length, the observed inhibition region for an isolated netropsin molecule on DNA is 8–9 nucleotides long with the drug positioned at the 5′ end of this region. Inhibition patterns having a minimum length of 8–9 base pairs were also found for the second-generation lexitropsins, **2**–**5** (Figures 1–3), indicating that for these compounds the enzyme is probably behaving in a manner similar to that observed with netropsin.

Compound **2** possessing two *N*-methylpyrrole groups exhibits binding that is very similar to the binding of netropsin. Like the antiviral agent, the lexitropsin strongly binds to sites of type $(A{\cdot}T)_4$ located at positions 50–46, 62–56, 92–89, and 159–156 of the fragment. However, sites of type $(A{\cdot}T)_3(G{\cdot}C)$ located in the region 100–80 are also among the initial loading events on the restriction fragment, indicating that the compound is more tolerant of a G·C base pair than is netropsin. This tolerance can be explained by considering the netropsin–DNA binding model and the structure of **2**. By reduction of the large methylene group of netropsin to a hydrogen atom in the *N*-formyl moiety of **2**, the steric interference caused by the presence of the 2-amino group of guanine no longer exists. The presence of the strong site at ~145 to ~135 for **2** dem-

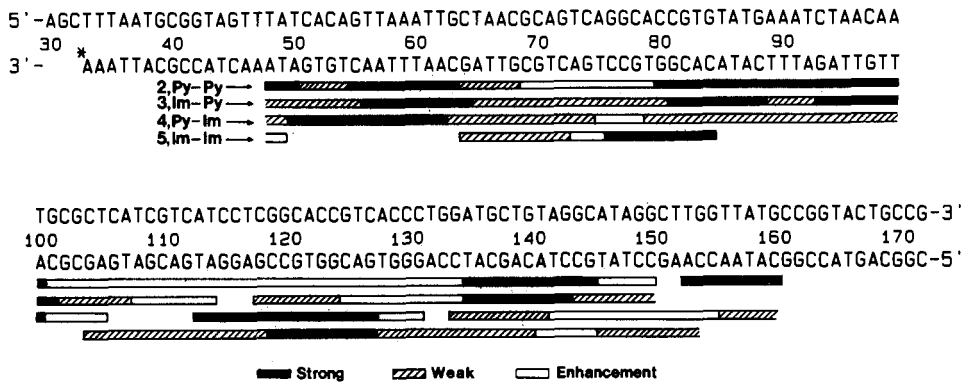


FIGURE 3: Sites of DNase I inhibition and enhancement shown on the sequence of the 139-mer used in the study.

onstrates that probably other factors are important in the lexitropsin–DNA contact. This region of the restriction fragment does not possess a site of type $(A{\cdot}T)_3$, which is required for the aforementioned analysis.

Substitutions of one of the *N*-methylpyrrole groups of **2** with *N*-methylimidazole further alters binding specificity. Compound **3** (Im-Py), which possesses an *N*-methylpyrrole residue attached to a propylamidinium moiety and thus is structurally similar to **1** and **2**, is most like the latter compound in its binding properties (Figures 1 and 3).

If the relative positions of the imidazole and pyrrole are reversed in the dipeptide, as in **4** (Py-Im), binding is altered in favor of GC-rich sites. The structural reason for this behavior is not yet clear, but it may be related to the inability of the pyrrole residue to contact DNA. When this residue is located between the charged propylamidinium moiety and the *N*-methylimidazole, electrostatic effects associated with the amidinium and the potential for hydrogen bonding between the imidazole and guanine may enhance the contact between the pyrrole and DNA. If, on the other hand, the pyrrole is positioned on the amino terminus of the ligand, the aforementioned effects may not be optimal for promoting pyrrole–DNA contact. Although analogy with netropsin suggests that the amide on the amino terminus of **4** can hydrogen bond to DNA, in the absence of the charged guanidinium moiety, this interaction alone may be insufficient to bring the pyrrole β-hydrogen into close contact with the adenine C-2 hydrogen of DNA. Subsequently, the ability to "read" DNA sequence is reduced.

When compared to **1**–**4** and all other known compounds of its size, the bis(imidazole) analogue **5** (Im-Im) exhibits a remarkably high degree of DNA binding specificity (Figures 2 and 3). Consideration of the regions of strong inhibition on the fragment and the cleavage geometry of DNase I strongly suggests that the compound is bound to the sequence 5′-CCGT-3′ (or 5′-ACGG-3′), which occurs twice on the restriction fragment. Although further study will be necessary to determine the details of the interaction, the structural features present in **5** and its binding sequence indicate that the imidazole residues of the lexitropsin are involved in hydrogen bonding to the guanines in the sequence. In addition, the propylamidinium group is most likely oriented toward the A·T base pair in the site (Figure 4).

Recent evidence in support of this orientation has been obtained from two-dimensional ^{1}H NMR analysis of the 1:1 complex of **5** with the duplex $d(CATGGCCATG)_2$. Unambiguous NOE results confirm unique binding to the overlined tetranucleotide sequence with the propylamidinium oriented over the T residue (Lee et al., 1987).

This orientation is preferred over the alternative, which would have the methylene group of the amidinium residue

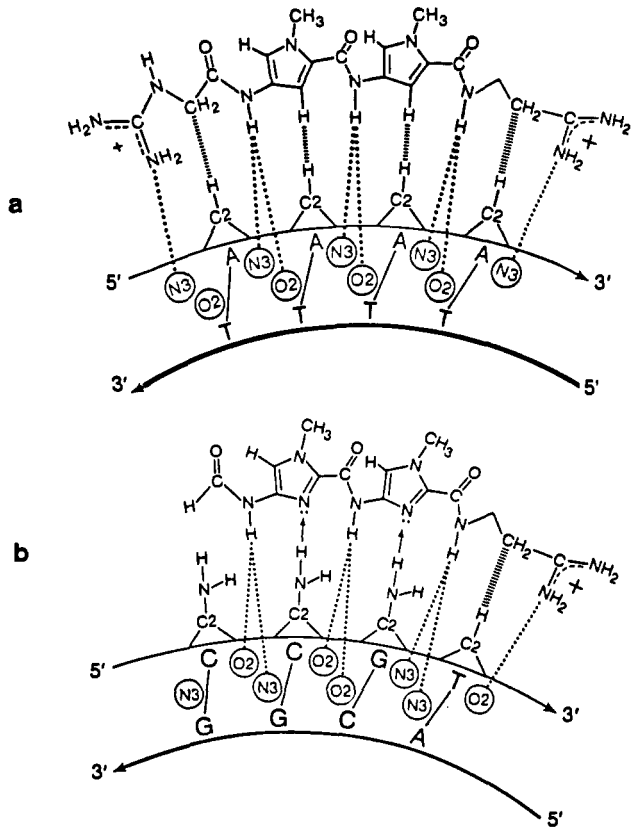


FIGURE 4: Depiction of molecular recognition components of (a) netropsin bound to 5′-AATT-3′ deduced from X-ray diffraction analysis and (b) **5** (Im-Im) bound to the sequence 5′-ACGG-3′: (···) hydrogen bonds; (≡) van der Waals contacts; (→) additional specific G-C(2)NH_2 hydrogen bonds directed from the minor groove.

sterically interacting with a hydrogen atom located on the 2-amino group of guanine. The preferred orientation requires that a G·C base pair be situated in the region of the *N*-formyl group of **5**. This raises the possibility that the carbonyl group on the amino terminus of the compound is directed toward, and hydrogen bonding with, the 2-amino group of guanine. More detailed work will be required to determine if this is actually the case.

The ability of the compound to discriminate between the preferred sequence and other possible binding sequences present, e.g. (all 5′ → 3′), TGCC, TGCG, AGCG, AGGG, and AGCC, appears related to structural differences between these sequences which are sensed by the lexitropsin. In this regard, the DNase I phosphodiester hydrolysis rate is significantly different for the ester linkages in the preferred site relative to those present in, for example, the weak site at positions 71–68 (5′-TGCG-3′) of the fragment (Figure 2). Since the enzyme is sensitive to the DNA twist angle and the

width of the minor groove (Drew, 1984; Lomonossoff et al., 1981; Drew & Travers, 1981; Suck & Oefner, 1986), the observed sequence-dependent affinities of **5** are probably due to significant differences in local DNA structure associated with sites of type $(G\cdot C)_3(A\cdot T)$.

Comparison of the results of this study with those involving the prototype series of lexitropsins (Lown et al., 1986a) suggests that the reduction of charge and/or alteration of one of the van der Waals contacts between the ligand and DNA has a significant effect on binding specificity. Studies currently in progress are designed to separate the contributions that each of these factors may have in controlling specificity.

Registry No. **2**, 23999-81-7; **3**, 109528-35-0; **4**, 109528-36-1; **5**, 109528-37-2; [[1-methyl-4-[(1-methyl-4-aminoimidazol-2-yl)-carboxamido]pyrrol-2-yl]carboxamido]propionamidine hydrochloride, 109528-38-3; formylimidazole, 3197-61-3; [[1-methyl-4-[(1-methyl-4-aminopyrrol-2-yl)carboxamido]imidazol-2-yl]carboxamido]propionamidine hydrochloride, 109528-39-4; [[1-methyl-4-[(1-methyl-4-aminoimidazol-2-yl)carboxamido]imidazol-2-yl]-carboxamido]propionamidine hydrochloride, 109528-40-7.

References

Brenowtiz, M., Senear, D. F., Shea, M. A., & Ackers, G. K. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 8462–8466.

Dabrowiak, J. C., Skorobogaty, A., Rich, N., Vary, C. P. H., & Vournakis, J. N. (1986) *Nucleic Acids Res. 14*, 489–499.

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

Drew, H. R. (1984) *J. Mol. Biol. 176*, 535–557.

Drew, H. R., & Travers, D. A. (1981) *Cell (Cambridge, Mass.) 14*, 287–296.

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

Laverly, R., & Pullman, B. (1985) *J. Biomol. Struct. Dyn. 2*, 1021–1027.

Lee, M., Pon, R. T., Hartley, J. A., & Lown, J. W. (1987) *Biochemistry* (submitted for publication).

Lomonossoff, G. P., Butler, P. J. G., & Klug, A. (1981) *J. Mol. Biol. 149*, 745–760.

Lown, J. W., & Krowicki, K. (1985) *J. Org. Chem. 50*, 3774–3779.

Lown, J. W., Krowicki, K., Bhat, U. G., Skorobogaty, A., Ward, B., & Dabrowiak, J. C. (1986a) *Biochemistry 25*, 7408–7416.

Lown, J. W., Sondhi, S. M., Ong, C.-W., Skorobogaty, A., Kishikawa, H., & Dabrowiak, J. C. (1986b) *Biochemistry 25*, 5111–5117.

Maniatis, T., Fritsch, E. T., & Sambrook, J. (1982) *Molecular Clonning—A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Suck, D., & Oefner, C. (1986) *Nature (London) 321*, 620–625.

Ward, B., Skorobogaty, A., & Dabrowiak, J. C. (1986) *Biochemistry 25*, 6875–6883.

Ward, B., Rehfuss, R., & Dabrowiak, J. C. (1987) *J. Biomol. Struct. Dyn. 4*, 685–695.

A Reexamination of the Reported B → Z DNA Transition in Nucleosomes Reconstituted with Poly(dG-m^{5}dC)·Poly(dG-m^{5}dC)†

J. Ausio, G. Zhou, and K. van Holde*

Department of Biochemistry and Biophysics, Oregon State University, Corvallis, Oregon 97331

Received June 19, 1987

ABSTRACT: Polynucleosomes with poly(dG-m^{5}dC)·poly(dG-m^{5}dC) have been reconstituted, and well-defined nucleosome core particles from these have been prepared. Upon addition of $MgCl_2$ to the levels used to induce the B to Z transition in this highly methylated DNA, significant changes in the circular dichroism spectrum are observed in solutions of these particles. However, such core particles also exhibit a noticeable instability when compared to chicken erythrocyte core particles under the same conditions. The change in circular dichroism can be entirely accounted for on the assumption that only free nucleotide, released by core particle dissociation, undergoes the B → Z transition. Therefore, no evidence has been found for "Z nucleosomes" in these solutions. In fact, the histone–DNA interaction in the nucleosome seems to partially inhibit the B to Z transition of the DNA. The analysis of our results is consistent with a model in which all of the DNA that remains bound to the histone octamer retains the B form.

Since of discovery of the left-handed conformation of DNA by analysis of poly(dG·dC) synthetic polymers in the presence of solutions of high ionic strength (Pohl & Jovin, 1972) and by X-ray crystallography (Wang et al., 1979), several attempts have been made to ascertain the biological relevance of this DNA form in vivo [see Rich et al. (1984) for a review]. At the structural level, the existence in eukaryote organisms of this left-handed DNA (also called Z DNA), as opposed to the canonical right-handed B form of DNA, would obviously be expected to introduce topological changes and constraints in those regions of chromatin associated with it. The first question thus arising is how the B to Z DNA transition (in those sequences with potential for such a transition) may affect the most elementary subunit of chromatin—the nucleosome.

Most of the studies undertaken to answer this question have taken advantage of the fact that poly(dG-m^{5}dC) can be induced to undergo the B to Z transition under very mild con-

† Supported by USPHS Grant GM22916. K.v.H. acknowledges the support of an American Cancer Society research professorship.

0006-2960/87/0426-5595$01.50/0