The molecular origin of DNA–drug specificity in netropsin and distamycin

(antitumor drugs/DNA minor groove/DNA–drug adaptation/lexitropsins/DNA–drug binding)

MARY L. KOPKA, CHUN YOON, DAVID GOODSELL, PHILIP PIURA, AND RICHARD E. DICKERSON

Molecular Biology Institute and Institute of Geophysics and Planetary Physics, University of California, Los Angeles, Los Angeles, CA 90024

Communicated by Paul D. Boyer, October 26, 1984

ABSTRACT  X-ray analysis of the complex of netropsin with the B-DNA dodecamer of sequence C-G-C-G-A-A-T-T-B-C-G-C-G reveals that the antitumor antibiotic binds within the minor groove by displacing the water molecules of the spine of hydration. Netropsin amide NH furnish hydrogen bonds to bridge DNA adenine N-3 and thymine O-2 atoms occurring on adjacent base pairs and opposite helix strands, exactly as with the spine of hydration. The narrowness of the groove forces the netropsin molecule to sit symmetrically in the center, with its two pyrrole rings slightly non-coplanar so that each ring is parallel to the walls of its respective region of the groove. Drug binding neither unwinds nor elongates the double helix, but it does force open the minor groove by 0.5-2.0 Å, and it bends back the helix axis by $\theta^\circ$ across the region of attachment. The netropsin molecule has an intrinsic twist that favors insertion into the minor groove of B-DNA, and it is given a small additional twist upon binding. The base specificity that makes netropsin bind preferentially to runs of four or more A-T base pairs is provided not by hydrogen bonding but by close van der Waals contacts between adenine C2 hydrogens and CH groups on the pyrrole rings of the drug molecule. Substitution of one or more pyroles by imidazole could permit recognition of G-C base pairs as well, and it could lead to a class of synthetic "lexitropsins," capable of reading any desired short sequence of DNA base pairs.

Netropsin and its close relative distamycin (Fig. 1) are antiviral antitumor antibiotics that, although too toxic for clinical use, have received extensive study as the paradigms of base-specific yet non-intercalative DNA-binding drug molecules. First isolated from Streptomyces netropsis in 1951 (1, 2), netropsin exerts its biological activity by binding tightly to double-helical B-DNA, interfering with both replication and transcription (3, 4). It shows little or no affinity for single-stranded DNA or RNA or for double-stranded RNA or DNA–RNA hybrids (3–5), suggesting that it does not bind to the A helix. It also fails to bind to left-handed Z-DNA; in fact, binding of netropsin to DNA favors A-to-B and Z-to-B helix transitions (6, 7).

Chemical protection studies (3, 4, 8) and Overhauser NMR experiments (9) indicate that netropsin does not intercalate between base pairs, but it binds within the minor groove of the intact double helix, using hydrogen bonds between netropsin amide NH and exposed adenine N-3 and thymine O-2 on the floor of the minor groove. The drug molecule attaches to clusters of four or more A-T or I-C, but not to G-C, base pairs (3, 10, 11). Alternating A-T-A-T regions bind netropsin less well than continuous runs of A or T (12, 13). Binding involves both an electrostatic component from the two cationic ends and hydrogen bonds from the central three amide NH groups, although neither aspect is absolute-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.
der Waals contact with the helix, along the midlines of the major and minor grooves (15, 26).

Distances between netropsin atoms and adenine N-3 or thymine O-2 atoms of the DNA are listed in Table 1. The amidinium end of the drug molecule comes closer to the floor of the minor groove than does the guanidinium, perhaps as a result of the greater flexibility of its connecting CH₂-CH₂ chain. At each of the outer two netropsin amides, the hydrogen bond to thymine appears to be stronger than that to adenine. Distances from the central amide N to thymine O-2 are both very long for standard hydrogen bonds, but the reason for this can be seen schematically in Fig. 4. The geometry of the backbone chain (as determined from single-crystal structure analysis of netropsin alone; ref. 22) is such that steric hindrance between the DNA and pyrrole CH prevents the amide N from approaching close enough for hydrogen bonds of normal length. Although the N-to-N and N-to-O distances listed in Table 1 frequently are long by classical hydrogen bond criteria, the polar interactions certainly contribute to stabilizing the DNA–drug complex.

Influence of DNA Binding on Netropsin Structure. The netropsin molecule is not planar; it is twisted in a screw sense that matches that of the minor groove of B-DNA. Fig. 4 lists the observed torsional deviations from planarity about individual bonds as found in the netropsin molecule alone (below the bonds) and in the present DNA–drug complex (above the bonds). As mentioned earlier, normals to the pyrrole ring planes make an angle of 33° to one another in the DNA–drug complex, a consequence of fitting each ring symmetrically into the minor groove at its own site. In the free drug molecule, this angle is only 20°, which means that binding to the DNA actually induces a further twist in the netropsin molecule.

Each of these pyrrole plane angles is equal to the sum of the three torsion angle deviations in the bonds connecting the two pyrrole rings: +19° – 2° +16° = 33° for the DNA–netropsin complex and +26° – 4° – 2° = 20° for netropsin alone, indicating that the nonplanarity of pyrrole rings arises only from rotation about these three bonds. The twists about the two pyrrole-to-CO bonds are observed whether DNA is present or not and probably are to be ascribed to steric clash between pyrrole CH and adjacent amide NH groups, schematized via van der Waals radii about hydrogen atoms in Fig. 4. The extra twist produced upon binding to DNA occurs in the central NH-to-pyrrole bond, which rotates from –2° to +16°. Hence, the netropsin molecule is both predisposed to bind to the minor groove of B-DNA and is subjected to an additional Koshland-like induced fit when binding occurs.

Influence of Netropsin Binding on the DNA Structure. The binding of netropsin to DNA produces no striking systematic changes in helix rotation or in base stacking—it neither unwinds nor elongates the double helix. Mean and individual values of these and other helical parameters for the B-DNA dodecamer (16, 17) are scarcely to be distinguished from those for DNA without netropsin. In only two respects does the binding of netropsin appear to alter the DNA conformation: the minor groove is widened by 0.5–2.0 Å by the entry
of netropsin (table 3 of ref. 16), and the helix axis is bent backward at the site of attachment by 8° per netropsin molecule. Both deformations illustrate the tightness of fit of the drug molecule into the minor groove.

**Energetics of Netropsin Binding.** Breslauer and co-workers (27, 28) have measured the thermodynamic parameters for binding of netropsin to two sequences of continuous fibers: poly(dA)-poly(dT) and poly(dA-dT)·poly(dA-dT), with the results listed in Table 2. Since the reaction involves displacement of water molecules from both the free DNA and the free netropsin molecule, the overall reaction is best written as shown at the bottom of Table 2. The free energy of binding is comparable for both the alternating (A-T)·(A-T), and the homogeneous A·T. Binding of netropsin to the homogeneous sequence appears to be strongly entropy driven, with

![Diagram](image_url)

**FIG. 3.** Diagrammatic representation of netropsin binding to DNA. DNA is shown only by a minor groove ladder, with adenine N-3 and thymine O-2 atoms indicated. The numbering of C, N, and O atoms of netropsin from the guanidinium end corresponds to that used in the single-crystal x-ray analysis of netropsin alone (22). Dotted lines indicate N-to-N or N-to-O distances short enough to be standard hydrogen bonds, whereas dotted lines indicate distances of 3.2 Å or more.

10 kcal/mol (1 cal = 4.184 J) coming from an increase in disorder upon binding, and only 2 kcal/mol being contributed by the energy of binding. In contrast, binding to the alternating copolymer is enthalpy driven with a relatively small entropy contribution (Table 2). The net free energy of binding is similar in either case.

This large difference in thermodynamic behavior undoubtedly reflects a difference in conformation or in hydration between the homo- and heteropolymers, as does the observation (12, 13, 29, 30) that netropsin binds less well to alternating poly(dA-dT) than to poly(dA)·poly(dT). The nature of this difference is unknown, although Klug et al. (31) have proposed that poly(dA-dT) has a noncanonical alternating B helix structure. The thermodynamic data in Table 2 can be explained by assuming that, when netropsin binds to the homopolymer, it must displace a more extensive hydration shell or spine than with the alternating heteropolymer. This would release more entropy in the form of displaced water molecules, but it also would result in a less negative overall enthalpy change because part of the energy liberated on binding DNA to drug would be required to break the bonds between the DNA or the drug and the solvent that was displaced upon binding.

The structure of our DNA–netropsin complex offers no obvious explanation for the difference between binding thermodynamics of homo- or heteropolymers, nor is it clear which of the polymers is the better limiting case for our A-A-T-T binding site. The single crystal x-ray analysis of netropsin alone (22) suggests that number y in the equation in Table 2 might be = 5, and the structure analysis of the B DNA do-

![Diagram](image_url)

**FIG. 4.** Flattened drawing of central region of netropsin, using bond lengths and angles as determined from the x-ray analysis of netropsin alone (22). Guanidinium is off the diagram to the left, and amidinium is to the right. The floor of the minor groove of DNA is schematized by curved dashed line at bottom. The van der Waals radii drawn around the hydrogens illustrate how steric hindrance between DNA and pyrrole CH prevents the NH of the central amide from approaching close enough for hydrogen bonds of normal length. Numbers alongside bonds are torsional deviations from planarity about these bonds: small numbers below bonds are for the netropsin molecule alone, and larger numbers above bonds are for the present netropsin–DNA complex. Positive torsion angles indicate clockwise rotation of the most distant end of a bond when viewing along the bond. Steric clash between pyrrole CH and adjacent amide NH probably is responsible for the torsional nonplanarity about the two pyrrole-to-CO bonds observed even in the absence of DNA. Twist in central NH-to-pyrrole bond is induced by binding to DNA.

Table 1. Distances from netropsin in backbone atoms to DNA base edge atoms on bottom of minor groove

<table>
<thead>
<tr>
<th>Netropsin atom</th>
<th>DNA atom</th>
<th>Distance, Å</th>
</tr>
</thead>
<tbody>
<tr>
<td>N-1</td>
<td>N-3 of A-5</td>
<td>3.66</td>
</tr>
<tr>
<td>N-4</td>
<td>O-2 of T-20</td>
<td>2.68</td>
</tr>
<tr>
<td>N-4</td>
<td>N-3 of A-6</td>
<td>3.28</td>
</tr>
<tr>
<td>N-6</td>
<td>O-2 of T-19</td>
<td>3.86</td>
</tr>
<tr>
<td>N-6</td>
<td>O-2 of T-7</td>
<td>3.54</td>
</tr>
<tr>
<td>N-8</td>
<td>N-3 of A-18</td>
<td>3.45</td>
</tr>
<tr>
<td>N-8</td>
<td>O-2 of T-8</td>
<td>2.56</td>
</tr>
<tr>
<td>N-10</td>
<td>N-3 of A-17</td>
<td>2.65</td>
</tr>
<tr>
<td>C-2</td>
<td>C-2 of A-5</td>
<td>3.90</td>
</tr>
<tr>
<td>C-5</td>
<td>C-2 of A-6</td>
<td>4.06</td>
</tr>
<tr>
<td>C-11</td>
<td>C-2 of A-18</td>
<td>3.87</td>
</tr>
<tr>
<td>C-16</td>
<td>C-2 of A-17</td>
<td>3.94</td>
</tr>
<tr>
<td>C-2</td>
<td>&quot;N-2&quot; of A-5</td>
<td>3.05</td>
</tr>
<tr>
<td>C-5</td>
<td>&quot;N-2&quot; of A-6</td>
<td>3.11</td>
</tr>
<tr>
<td>C-11</td>
<td>&quot;N-2&quot; of A-18</td>
<td>2.87</td>
</tr>
<tr>
<td>C-16</td>
<td>&quot;N-2&quot; of A-17</td>
<td>2.92</td>
</tr>
</tbody>
</table>

Table 2. Thermodynamics of netropsin (Nt) binding to B-DNA*

<table>
<thead>
<tr>
<th></th>
<th>ΔG°, kcal/mol</th>
<th>ΔH°, kcal/mol</th>
<th>ΔS°, kcal/°mol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poly(dA)=poly(dT)</td>
<td>-12.2</td>
<td>-2.2</td>
<td>-10.0</td>
</tr>
<tr>
<td>Poly(dA-dT)</td>
<td>-12.7</td>
<td>-11.2</td>
<td>-1.5</td>
</tr>
</tbody>
</table>

Reaction: DNA+H₂O + NtyH₂O → DNA-Nt + (x + y)H₂O

*Data are from ref. 28.
decamer indicates that \( x \) might be close to 12. Hence, 17 water molecules would be set loose per molecule of netropsin bound to DNA, contributing to the observed entropy of binding. A detailed structural explanation will have to remain open until structure results are available for complexes of netropsin with other types of A-T sequences.

**Specificity of Netropsin Binding.** When the subject of base specificity in DNA-binding molecules is raised, it frequently is suggested that a large part of this specificity could arise from formation of particular hydrogen bonds between atoms on DNA bases and the recognition molecules: drugs, repres-
sors, or other control proteins. But because the principal difference between A-T and G-C base pairs in the minor groove is the presence of the N-2 amine groups on guanine, it was suggested many years ago that the lack of binding of netropsin to G-C might arise from steric hindrance by this NH\(_2\) group (3).

The present x-ray analysis shows this to be the case. Care-
ful examination of Fig. 2b reveals that, although hydrogen bonds may aid in positioning the netropsin molecule correctly, the actual sensing or reading of A vs. G is done by close contacts between adenine ring C-2 atoms and two categories of netropsin atoms: either pyrrole ring CH, or the CH\(_2\) flan-
king the outer amides on the wings of the drug molecule. These are denoted in Fig. 3 as carbon atoms C-2, C-5, C-11, and C-16. Distances between these netropsin and adenine atoms are listed in Table 1, and all are seen to represent close van der Waals contacts. If one were to add an NH\(_2\) group with proper geometry to each adenine C-2 atom, then the separations between these hypothetical N and the same netropsin atoms would be as shown at the bottom of Table 1. These are impermissibly short for nonbonded contacts and, hence, netropsin does not readily bind to regions of DNA containing G-C base pairs. The hydrogen bonds from netropsin amides supply a certain amount of binding energy and provide a correct reading frame by positioning the netropsin properly along the DNA molecule. But the actual reading of A vs. G bases is performed by van der Waals non-
bonded contacts. This division of labor between H bonds and nonbonded contacts is schematized in Fig. 5a.

The closely related distamycin molecule shows a greater tolerance for isolated G-C base pairs. Distamycin results if the guanidinium tail of netropsin, from atom C-2 upward in Fig. 3, is removed and replaced by another pyrrole ring and by a terminal NH-CHO. It therefore has one more amide NH than netropsin and loses one of netropsin’s two cationic ends. As expected from the Dervan binding model (23) and our structure results, the added amide NH makes the optimal distamycin binding site 5 base pairs rather than 4 as with netropsin. If the natural distamycin molecule is designated as Dst-3 (for its three pyrrole rings), synthetic analogues with as few as two or as many as seven pyrroles can be prepared. Comparative studies have been made of binding prefer-
cences of netropsin and Dst-2 through Dst-5 (23, 32–34). Although netropsin and Dst-2 apparently demand 4-base-pair sites containing only A-T base pairs, the longer distamycin analogues become increasingly tolerant of occasional isolated G-C base pairs, particularly near the ends of a binding site. This may arise from two factors. The loss of one cationic

**Tailoring Base Specificity: Lexitropsins.** Although the hypo-
thetical amine-N-to-netropsin distances at the bottom of Table 1 are too short to be nonbonded contacts, they all would be quite acceptable as hydrogen bond lengths if netropsin provided an atom with which NH\(_2\) could form a hydrogen bond. Substitution of methylimidazole for methylpyrrole in the netropsin molecule would replace the ring CH by N, providing both space for an NH\(_2\) on the purine and a hydrogen bond acceptor. The drug molecule then would favor a G-C base pair at that position, as indicated in Fig. 5b.

It should be possible to prepare a series of netropsin ana-
logues of various lengths using known synthetic techniques (35–37), and by selecting between pyrrole or imidazole at specific loci, to create artificial "lexitropsins" capable of reading and binding to DNA double helices of any desired short sequence of A-T and G-C base pairs. One might envision attaching to them EDTA·Fe(H), as Dervan has done with distamycin and its analogues of varying length (24), thereby creating miniature restriction enzyme analogues tail-
ored to particular base sequences. Alternatively, the ability to synthesize drug molecules that would target specific short DNA sequences could prove to be of pharmacological or medical value in directing agents toward intruder or neoplastic cells and away from normal host cells.

Netropsin and distamycin apparently exert their pharma-

---

*W. Lown (University of Alberta) has independently arrived at the idea of altering netropsin base specificity by substituting imida-
zole for pyrrole, on the basis of a study of space-filling models. He has synthesized a series of lexitropsin molecules for binding to DNA and collaborative examination by NMR and x-ray methods.
colocological effects by binding tightly to double-helical DNA, interfering both with replication and transcription. In doing so, they displace the spine of hydration that normally helps to stabilize the B form of DNA, substituting for a string of hydrogen-bonded water molecules the covalently bonded backbone chain of the drug molecule. The ejection of hydrating waters makes an entropic contribution to DNA–drug binding; and electrostatic, H-bonding, and hydrophobic interactions probably all contribute to the enthalpy of binding. From an informational standpoint, the role of hydrogen bonds is not information-bearing element of the DNA backbone, rather of base sequence, but rather of proper orientation of the reading frame. The information flow from DNA to drug may be involve a different channel, rather of the major groove. It may be that this picture of two independent channels of information readout has real biological significance, permitting a degree of independence in the control of DNA expression by protein repressors and by smaller drug molecules.

This work was carried out with the aid of National Institutes of Health Grant GM-31299. One of us (D.G.) is the holder of a Predoctoral Traineeship under Public Health Service National Research Service Award GM-07104.