Crystal structure of the DNA decamer d(CGCAATTGCG) complexed with the minor groove binding drug netropsin

ARTICLE *in* BIOCHEMISTRY · MAY 1997

Impact Factor: 3.02 · DOI: 10.1021/bi9628228 · Source: PubMed

3 AUTHORS, INCLUDING:

Christine Nunn
London Metropolitan University
110 PUBLICATIONS **2,383** CITATIONS
SEE PROFILE

Elspeth F Garman
University of Oxford
175 PUBLICATIONS **6,571** CITATIONS
SEE PROFILE

Crystal Structure of the DNA Decamer d(CGCAATTGCG) Complexed with the Minor Groove Binding Drug Netropsin[†]

Christine M. Nunn,*[,‡] Elspeth Garman,[§] and Stephen Neidle[‡]

The CRC Biomolecular Structure Unit, The Institute of Cancer Research, Sutton, Surrey SM2 5NG, U.K., and Laboratory of Molecular Biophysics, University of Oxford, Oxford OX1 3QU, U.K.

Received November 14, 1996; Revised Manuscript Received February 3, 1997[⊗]

ABSTRACT: The crystal structure of netropsin bound to the decamer d(CGCAATTGCG) has been determined at 2.4 Å resolution. This is the first example of a crystal structure of netropsin bound to decamer DNA. The central eight bases of each DNA single-strand base pair with a self-complementary strand to form an octamer B-DNA duplex. These duplexes lie end to end within the unit cell. The terminal 5′-C and G-3′ bases are unpaired and interact with the neighboring duplexes via interactions within both the major and minor groove to form base triplet interactions of the type C^+−G·C and G∗(G·C), respectively. The triplet interaction of the type C^+−G·C is known to exist within triplex DNA with the C^+ base oriented parallel with the Watson−Crick guanine base to which it hydrogen bonds. The netropsin molecule lies within the minor groove of the octamer duplex and assumes a class I type position, with bifurcated hydrogen-bonding interactions from the amide groups of the netropsin to the A·T base pairs of the minor groove. The netropsin molecule fits within a five base pair long minor groove site by bending of the flexible amidinium group at one end of the drug.

Netropsin is a naturally occurring antibiotic with experimental antitumor and antiviral activity. The wide range of biological effects exhibited by this drug are believed to result from its ability to bind to the minor groove of DNA. Netropsin is an AT-selective minor groove binding drug which shows DNA binding preference in the order AATT > TAAT = TTAA = ATAT > TATA (Abu-Daya et al., 1995). Crystallographic studies of netropsin bound to the minor groove of a number of B-DNA dodecamer sequences have been carried out. Those crystal structures studied to date are 1:1 complexes between the drug and duplexes comprising the two DNA strands d(CGCX_6GCG), where X_6 = $GA_2T_2{}^{Br}C$ (Kopka et al., 1985), GATATC (Coll et al., 1989), [e^6G]A_2T_2C and GA_2T_2C (Sriram et al., 1992), A_3T_3 (Tabernero et al., 1993), and GT_2A_2C (Balendiran et al., 1995). In all of these structures the accessible minor groove site for netropsin binding is six base pairs in length, and in all cases the four base pairs at the center of the duplex are A·T base pairs.

The crescent-shaped netropsin molecule binds within the AT-rich DNA minor groove, with the three amide groups facing the floor of the minor groove. The locations of the drug molecule within the minor groove of these structures have been divided into two categories (Goodsell et al., 1995). Class I structures have each of the three amide groups along the length of the netropsin positioned midway between the four AT-rich bases at the center of the dodecamer sequence, and these amide groups form bifurcated hydrogen bonds to the adenine and thymine bases on opposite strands of two adjacent base pairs. Class II structures, in contrast, have the netropsin position shifted by half a base pair such that the amide groups lie within the plane of the base pairs. Class I positioning of the drug molecule agrees with the results obtained from NMR studies (Patel, 1982) of netropsin−d(CGCGAATTCGCG) and explains the A·T sequence specificity of the drug, as each of the pyrrole rings rests against the adenine C2 atom of one base pair (Goodsell et al., 1995). The results obtained from the crystal structure analysis of class I netropsin−DNA structures have been important for the design and synthesis of "lexitropsins"—analogues of netropsin with altered DNA sequence specificity. By modification to the *N*-methylpyrrolecarboxamide framework and also by the linking of two or more molecules, analogues can be produced which are capable of recognizing longer binding sites and exhibit the potential for G·C base pair binding (Lown, 1995).

Two crystal structures of the native DNA decamer d(CGCAATTGCG) have been determined in this laboratory. These structures have different conformations by virtue of the different crystallizing conditions used. One structure is an octamer duplex with unpaired terminal 5′-C and G-3′ bases which interact with the terminal base pairs of a neighboring duplex within the unit cell (Spink et al., 1995). The other structure is a decamer duplex with no unpaired bases. The decamer duplexes within this structure stack end to end within the unit cell to form a crossed-helix packing arrangement (Wood, Nunn, Trent, and Neidle, to be published). A number of studies of ligands bound to the decamer d(CGCAATTGCG) have been undertaken using NMR spectroscopy. Lexitropsins form 1:1 complexes with the decamer (Singh et al., 1992; Kumar et al., 1991), as does Hoechst 33258 (Kumar et al., 1992), by noncovalent binding within the minor groove of the DNA duplex. In addition to

[†] Funding for this research was provided by Programme Grant SP1384 from the Cancer Research Campaign.

* Corresponding author. Tel: 44-181-6438901. Fax: 44-181-7707893. Email: c.nunn@icr.ac.uk.

[‡] Institute of Cancer Research.

[§] University of Oxford.

[⊗] Abstract published in *Advance ACS Abstracts,* April 1, 1997.

S0006-2960(96)02822-X CCC: $14.00

non-covalent binding, a covalent adduct forms upon binding of a cyclopropylpyrrolylindole−lexitropsin conjugate with the duplex d(CGCAATTGCG)$_2$ (Fregeau et al., 1995).

In this paper we present the first crystal structure of a complex between a drug molecule and the decamer d(CGCAATTGCG). This structure shows netropsin binding to a hitherto unobserved site within duplex DNA.

MATERIALS AND METHODS

Crystallization and Data Collection. The DNA decamer d(CGCAATTGCG) was purchased from the Oswel DNA Service (University of Southampton, U.K.). Crystals were grown by vapor diffusion at 291 K from sitting drops containing 3 μL of 3.3 mM DNA, 4 μL of 5 mM netropsin, 2 μL of 1 M $MgCl_2$ and 3 μL of 45% 2-methyl-2,4-pentanediol, equilibrated against a 1 mL reservoir of 45% 2-methyl-2,4-pentanediol. The DNA and MPD solutions were prepared using 30 mM sodium cacodylate buffer at pH 7.0. Crystals of the decamer complex grew over a period of 6−8 weeks.

Data collection was carried out at 289 K on a crystal of dimensions approximately 0.5 × 0.2 × 0.2 mm^3 using a Siemens multiwire area detector mounted on a Rigaku RU200H rotating anode X-ray generator operating at 40 mA, 40 kV, with a graphite monochromator. A crystal-to-detector distance of 10 cm and swing angle of 18° were used, giving a maximum resolution of 2.05 Å at the detector edge. Data were collected with χ fixed at 45°, while the crystal was rotated through ω. An oscillation angle of 0.2° and an exposure time of 180 s were used. Two data sets were collected for total ω rotations of 100° and 23.4°, respectively.

Data processing was carried out using the program XDS (Kabsch, 1993) and the CCP4 program suite. A total of 4063 reflections were collected. Of a possible 2280 unique reflections to 2.4 Å, 1864 reflections were observed representing an 82% complete data set. The percentage of unique observed reflections is 87% from 8 to 3.5 Å, 84% from 3.5 to 3 Å, 78% from 3 to 2.6 Å, and 71% from 2.6 to 2.4 Å. The final merged data set has cell parameters of $a = 24.84$ Å, $b = 39.80$ Å, $c = 54.08$ Å, $\alpha = \beta = \gamma = 90°$, and systematic absences indicating the space group $P2_12_12_1$. $R_{merge}(I)$ is 3.0% for all data to 2.4 Å.

Structure Solution and Refinement. The crystallographic asymmetric unit contains two decamer strands. Inspection of the diffracted intensities suggested a B-DNA structure with the base pair stacking lying perpendicular to the *c* axis of the unit cell. The *c* axis is 54.08 Å in length, which corresponds to a 16 base pair repeat along this axis within the unit cell.

Molecular replacement was carried out using the program AMoRe (Navaza, 1994). Duplexes of varying length and sequence were used as search models for molecular replacement, and an octamer duplex of sequence d(GCAATTGC)$_2$ provided an initial refinement model. These duplexes pack end to end within the unit cell, lying parallel with the crystallographic *c* axis. After rigid group fitting of the molecular replacement solution using data from 8 to 3 Å resolution, the crystallographic *R* value was 53% and the correlation coefficient 0.63. For 8−3 Å resolution data rigid group refinement of each nucleotide was carried out, followed by positional and simulated annealing refinement.

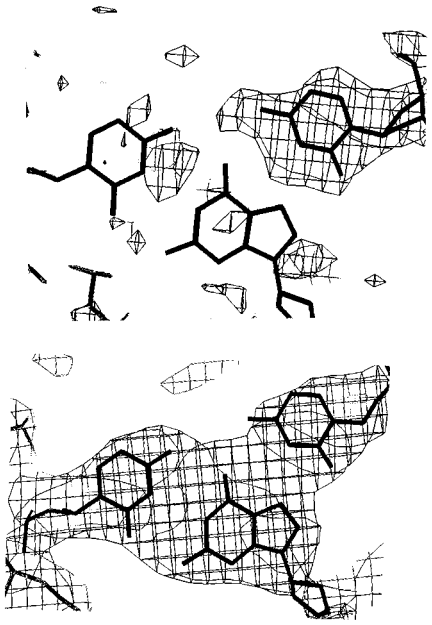


FIGURE 1: (top) $2\sigma\ F_o - F_c$ electron density map calculated prior to the addition of base C1 together with the final refined C1 position and (bottom) $1\sigma\ 2F_o - F_c$ electron density map calculated using the final model together with the final refined C1 position.

At this stage electron density maps were generated. The presence of extensive density within the DNA minor groove suggested that this model was worth pursuing.

Crystallographic refinement was carried out using the program X-PLOR (Brünger et al., 1987) with cross validation to direct the refinement protocol. Simulated annealing and positional refinement of the 16 bases and electron density map generation showed very poor $2F_o - F_c$ electron density and unrealistic geometries for one of the two terminal C·G base pairs. The molecular replacement had shown very poor discrimination along the *c* axis for the translation search, and the model obtained from molecular replacement was in fact out of register by one base pair from the correct DNA position. The register of the DNA was corrected.

At this stage of the refinement $R = 36.4\%$ ($R_{free} = 48.1\%$), and four DNA bases and a netropsin molecule were still to be located and included in the refinement. After accurate positioning of the octamer duplex, an $F_o - F_c$ electron density map was calculated. Electron density appeared to lie along the entire length of the DNA minor groove. As the octamer duplexes pack end to end within the unit cell, the terminal 5′-C and G-3′ bases of each single strand cannot be involved in Watson−Crick base-pairing interactions. Well-defined $F_o - F_c$ electron density extends from the 3′-end of the two DNA strands corresponding to the two terminal guanine bases. This density lies within the minor groove of a neighboring DNA duplex. The terminal G-3′ bases were fitted into the electron density maps, and the structure was refined to an *R* value of 24.6% ($R_{free} = 37.7\%$).

At the other end of each DNA strand the 5′-C bases are adjacent to the major groove of symmetry-related octamer duplexes. For C1 there was well-defined $F_o - F_c$ electron density lying in the plane of a G·C base pair of a symmetry-related duplex. This base pair is not the terminal base pair of the octamer duplex which lies closest to base G2, but the base pair one removed, C3·G18. Figure 1 shows the $F_o - F_c$ and $2F_o - F_c$ electron density maps for the C1 base

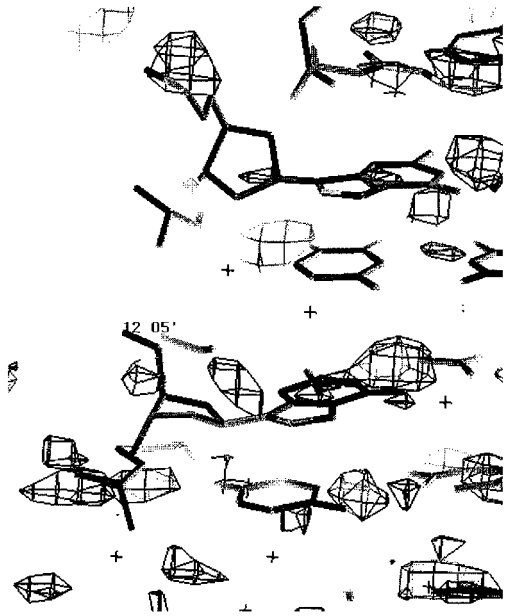


FIGURE 2: $2\sigma F_o - F_c$ electron density map calculated using the final model in the vicinity of atom 12O5′.

density together with the final refined position of base C1. At the other end of the molecule there is no evidence for any density for either base C11 or the associated phosphate group, and no C11 base was included in the refinement. Figure 2 shows the $F_o - F_c$ map generated for this region of the structure after the final refinement. Prior to the addition of the netropsin molecule into the DNA minor groove $R = 22.8\%$ ($R_{free} = 34.4\%$).

Netropsin Parametrization. To determine the best refinement parameters and topology for the netropsin molecule to use in X-PLOR refinement, the coordinates of the small molecule structure for netropsin (Berman et al., 1979) were obtained from the Cambridge Crystallographic Data Centre (Allen et al., 1979). The netropsin molecule was manipulated to assume the conformation of a minor groove binding drug before undergoing conformation minimization and an electrostatic potential calculation to determine atomic charges using the program MOPAC/ESP (Stewart, 1990). The minimized structure and the calculated charges were used as a starting model for X-PLOR refinement. For consistency with the parameter and topology files used for DNA refinement (Parkinson et al., 1996) only the polar hydrogen atoms were treated explicitly in the netropsin refinement.

Five base pair steps lie between the terminal two O3′ atoms of the minor groove binding guanosines G10 and G20, which lie close to coplanar with base pairs C3·G18 and G8·C13. Inspection of the $F_o - F_c$ electron density map prior to the addition of the energy-minimized netropsin molecule into the refinement showed almost continuous density in this region of the minor groove (Figure 3). The density was asymmetric in shape and similar to the conformation of the energy-minimized netropsin molecule with one end of the density clearly less planar and corresponding to the amidinium group. The netropsin was manually fitted into the $F_o - F_c$ density. Following positional and temperature factor refinement of the netropsin molecule within the DNA minor groove, the electron density was continuous over the entire drug molecule and $R = 19.3\%$ ($R_{free} = 32.2\%$).

After inclusion of the netropsin molecule into the refinement, solvent molecules were located and added to the

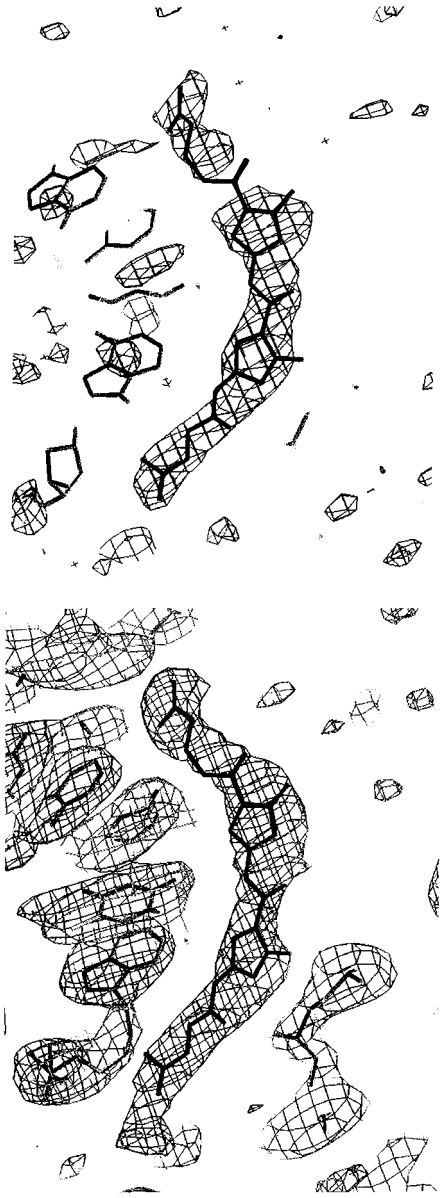


FIGURE 3: (top) $2\sigma F_o - F_c$ electron density map calculated prior to the addition of the netropsin molecule together with the final refined netropsin position and (bottom) 1σ $2F_o - F_c$ electron density map calculated at the end of the refinement showing the netropsin molecule.

model. At the conclusion of the refinement 85 water molecules had been added. The final R value is 20.1% for all data with $F > 2\sigma(F)$ in the resolution range 8−2.4 Å. The root mean square deviations from idealized values for bond lengths and angles within the DNA and netropsin molecules are 0.018 Å and 2.8°, respectively. Final atomic coordinates and experimental details have been deposited with the Nucleic Acid Database with code GDJ046. All graphics work was carried out using the programs O (Jones et al., 1991) and TOM/FRODO Alberta/Caltech v3.2.

RESULTS AND DISCUSSION

Overall Structure. The structure of the decamer d(CGCAATTGCG) comprises a Watson−Crick base-paired B-DNA octamer duplex with the two terminal G-3′ bases of each single strand lying within the minor groove of a symmetry-related duplex. These terminal G-3′ bases form hydrogen-bonding interactions across the minor groove−

Table 1: Helical Parameters for the Central Octamer Duplex for $d(CGCAATTGCG)_2$[a]

base pair	tip (deg)	inclin (deg)	buckle (deg)	propeller twist (deg)	base step	tilt (deg)	roll (deg)	helical twist (deg)	base pair rise (Å)
G2·C19	5	3	2	−5	G2·C19				
C3·G18	−2	5	0	−6	C3·G18	2	−7	34	3.2
A4·T17	2	5	−14	−10	A4·T17	−1	3	32	4.0
A5·T16	4	4	−4	−18	A5·T16	−1	2	37	3.2
T6·A15	−4	4	−2	−14	T6·A15	0	−8	36	2.9
T7·A14	−3	7	6	−8	T7·A14	3	0	37	3.1
G8·C13	1	6	0	−6	G8·C13	−1	5	33	3.7
C9·G12	−1	6	1	4	C9·G12	0	−3	37	3.2

[a] Calculations carried out using NEWHELIX with the terminal 5′-C and G-3′ bases removed.

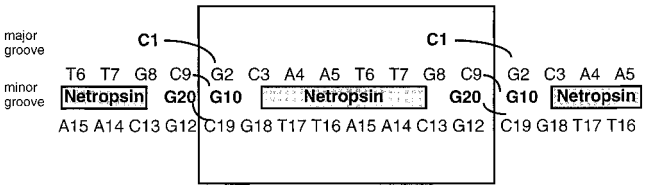


FIGURE 4: Schematic view of the overall structure together with the numbering scheme. The box contains the contents of the asymmetric unit.

minor groove interface with the terminal guanine base of a symmetry-related duplex. The terminal 5′-C base of one strand lies in the major groove of a symmetry-related duplex with hydrogen-bonding contacts to a guanine base. The other terminal 5′-C base was not located in the crystallographic refinement. This structure is very similar to the structure of the native sequence (Spink et al., 1995) which has a 2-fold symmetry axis through the center of the duplex and one DNA strand in the asymmetric unit. The position of the terminal 5′-C base was very poorly defined within the native structure and not located unambiguously.

A netropsin molecule binds to the DNA duplex along the minor groove, making hydrogen-bonding contact with the bases which line the groove. The bases are numbered from C1 to G10 for strand 1 and from G12 to G20 for strand 2. A schematic view of the overall structure together with the base numbering is shown in Figure 4.

DNA Structure. Within the unit cell, each of the octamer duplexes lies end to end with a symmetry-related octamer duplex to form pseudocontinuous DNA helices lying along the direction of the crystallographic *c* axis. At the intersection between duplexes there is a rise of 3.7 Å between the terminal C9·G12 and G2·C19 base pairs and a pseudohelical twist angle of 65°. At the junction of the two octamer duplexes, the 3′-end of each strand contacts the minor groove of a symmetry-related duplex, while the 5′-end contacts the major groove. These duplex junction interactions are illustrated in Figure 5.

The backbone torsion angles have standard values, and the end bases have values similar to those reported for the native structure (Spink et al., 1995). The intrastrand P···P separations along each strand vary between 6.0 and 7.1 Å with the smallest value occurring for 2P···3P. For the octamer duplex $d(GCAATTGC)_2$ the helical parameters are shown in Table 1. Within the central AT-rich region of the duplex there are large propeller twist values as typically seen in such regions of B-DNA.

Terminal Guanosines. The terminal guanosines lie within the minor groove of a neighboring duplex within the unit cell. Hydrogen bonding stabilizes this interaction via N2···N3 contacts between guanine bases across the minor

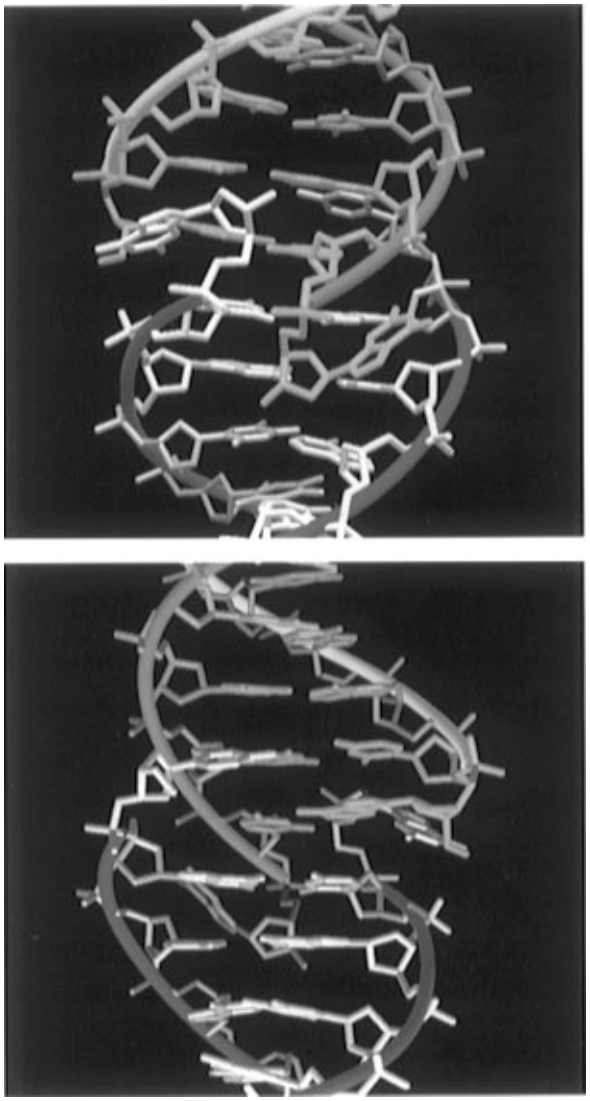


FIGURE 5: Interaction between duplexes within the unit cell viewed toward (top) the minor groove and (bottom) the major groove. The terminal guanosines lie in the minor groove of neighboring duplexes while the cytosine base at the 5′-end of one strand extends into the major groove of a neighboring duplex.

groove interface (Figure 6). G10 and G20 make hydrogen-bonding interactions with bases G2 and G12, respectively, via two interactions of the type N2···N3. In addition, atoms N2 of G10 and G20 both hydrogen bond to a neighboring O4′ deoxyribose sugar ring atom. The dihedral angle between bases G10 and G2 is 40° and between G20 and G12 is 36°. These guanine base interactions are similar to those observed in all the dodecamer structures of the type $d(CGX_8GC)$ where minor groove interlocking occurs for the terminal two C·G bases at each end of the duplex. For the

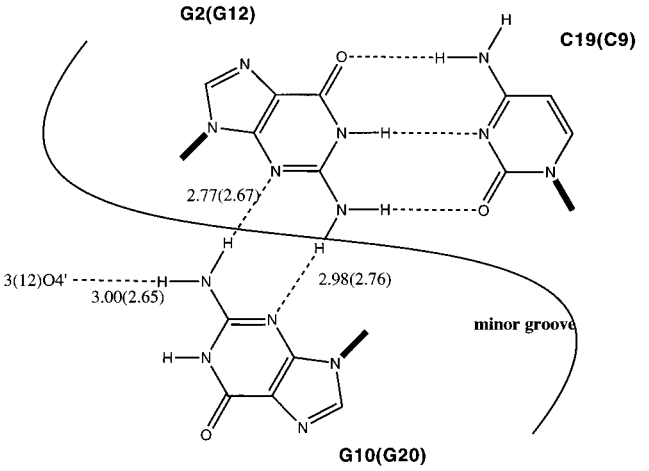


FIGURE 6: Schematic view of the guanine interactions across the DNA minor groove. Distances are in angstroms.

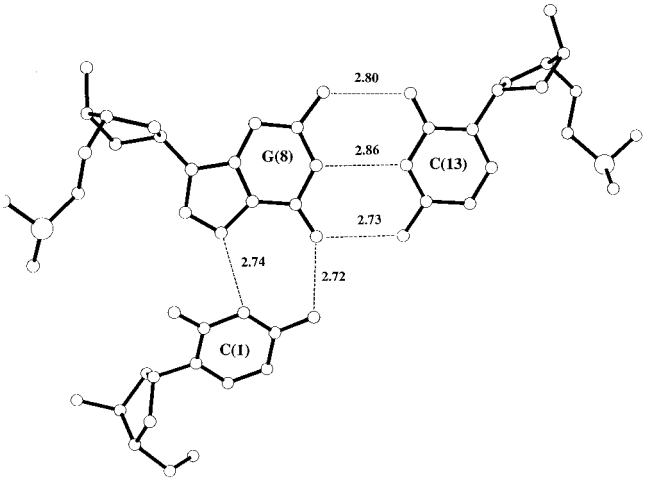


FIGURE 7: C^+−G·C triplet interaction, together with hydrogen-bonding distances in angstroms.

structure of netropsin bound to the dodecamer sequence d(CGCGAATT^{Br}CGCG) (Goodsell et al, 1995) there are similarly two N2···N3 hydrogen-bonding interactions between guanine bases across the minor groove, with dihedral angles between hydrogen-bonded guanine bases of 39° and 40°.

5′-Cytosine Forms a Triplex Triplet of the Type C^+−G·C. At the 5′ end of one DNA strand an unpaired cytosine base lies in the major groove of the neighboring duplex within the unit cell to form a base triplet of the type C^+−G·C. The position of this base is well defined by the omit difference density map and lies in the plane of the penultimate base pair C8·G13 of the octamer duplex. The hydrogen-bonding interactions between the three bases are shown in Figure 7. The dihedral angle between bases C1 and G2 is 39°, and the cytosine base C1 is postulated to be protonated. The resultant triplet is of the type known to occur in triplex DNA with parallel orientation of the third strand pyrimidine base with respect to the purine Watson-Crick strand and Hoogsteen base pairing.

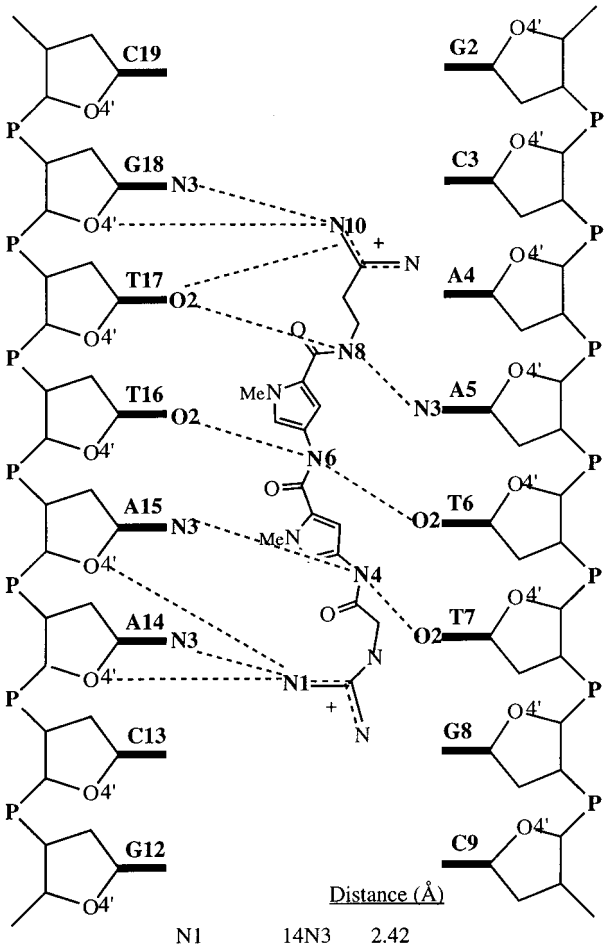
This crystal structure illustrates how crystal packing forces can be used to stabilize short pieces of triplex DNA. Recent crystal structure determinations have been carried out for the sequences d(GCGAATTCG) and d(GGCCAATTGG). Both of these structures contain base triplets of the type G−G·C (Van Meervelt et al., 1995; Vlieghe et al., 1996a,b) with a crystal packing arrangement within the unit cell similar to that seen within this structure. The overhanging 5′-G bases bind in the major groove of symmetry-related octamer duplexes via both Hoogsteen and reverse-Hoogsteen hydrogen-bonding interactions. In these structures the third strand guanine base is oriented parallel and antiparallel with respect to the guanine base of the duplex to which it hydrogen bonds.

Within the present crystal structure it is interesting to observe that the 5′-C base chooses not to hydrogen bond with the terminal base pair but with the penultimate base pair of the octamer duplex. Binding to the terminal base pair would require the cytosine base to hydrogen bond with the cytosine base of the terminal C·G base pair which is unfavorable, due to a lack of potential hydrogen-bonding interactions and the close proximity of two amine groups.

Binding of the Drug Netropsin within the Minor Groove. The netropsin molecule binds within the minor groove of the DNA duplex and assumes the isohelical nature of the groove. The dihedral angle between the two pyrrole rings is 21°. The netropsin molecule binds via hydrogen-bonding interactions to acceptor groups along the floor of the groove, in addition to making van der Waals contacts with the groove walls.

The netropsin−DNA interactions are shown in Figure 8. The three amide groups along the netropsin molecule are positioned approximately halfway between the four base pairs A4·T17∧A5·T16∧T6·A15∧T7·A14. These A·T base pairs display large propeller twist values (mean 12.5°), and each amide group makes bifurcated hydrogen-bonding interactions with O2 and N3 atoms of adenine and thymine bases from opposite strands on two adjacent base pairs. This pattern of bifurcated hydrogen bonding has previously been observed in other crystal structures with netropsin bound to the dodecamer DNA sequences of the type d(CGCGAATT^{Br}-CGCG) (Kopka et al., 1985; Goodsell et al., 1995) and d(CGCAAATTTGCG) (Tabernero et al., 1993) and are termed class I. In these dodecamer structures the amide groups have displaced the spine of hydration which is well defined in the native dodecamer structure d(CGCGAAT-TGCGC) (Kopka et al., 1983). No such hydration spine was located in the structure of the native decamer d(CGCAAT-TGCG) (Spink et al., 1995). A different type of structure has been reported (class II) for netropsin bound to the sequences d(CGCGAATTCGCG) (Sriram et al., 1992), d(CGCGATATCGCG) (Coll et al., 1989), and d(CGC^{6Et}-GAATTCGCG) (Sriram et al., 1992). In these structures the netropsin molecule assumes a different position within the minor groove shifted by one-half of a base pair.

As also observed in the dodecamer−netropsin complexes, the drug molecule extends beyond the central AT-rich region of the minor groove. In addition to the bifurcated hydrogen-bonding interactions, the terminal amidinium and guanidinium groups also hydrogen bond with the DNA minor groove. Interactions of this type were also observed in the structure of d(CGCGAATT^{Br}CGCG) (Kopka et al., 1985; Goodsell et al., 1995); however, the close contact seen between netropsin atom N10 and atom N3 of G(18) is unique to the present structure and results from the amidinium group sitting close to base G(18). In all of the dodecamer−netropsin structures studied by X-ray crystallography this G(18) position is occupied by a pyrimidine base, which precludes an interaction of the type observed here. The carbon atoms of the pyrrole rings which lie deepest within



N1	14N3	2.42
	14O4'	3.03
	15O4'	3.01
N4	7O2	2.76
	15N3	3.06
N6	6O2	3.20
	16O2	3.40
N8	5N3	3.08
	17O2	3.34
N10	17O2	3.12
	18N3	2.97
	18O4'	3.02

FIGURE 8: Schematic view of the netropsin−DNA interactions. **P** represents $CH_2PO_4^-$.

the minor groove are midway between the O4′ deoxyribose atoms of the thymine sugar rings T(6), T(17) and T(7), T(16), which face into the minor groove from the two DNA strands.

The accessible minor groove binding site for netropsin within this structure is different from any of the previously determined structures of 1:1 complexes between netropsin and DNA. In all of the other structures, netropsin binds within the minor groove of dodecamer DNA which exhibits an asymmetric minor groove geometry resulting from minor groove−minor groove interlocking of the native DNA structure in the crystal lattice. The dodecamer structures have a six base pair tract between the symmetry-related 12O3′ and 24O3′ atoms into which a netropsin molecule binds. In the case of the decamer d(CGCAATTGCG), the unit cell packing with the minor groove binding guanosines at each end of the duplex presents a symmetric and smaller DNA minor groove for netropsin binding than that observed for the dodecamer sequences. A five base pair tract lies between the terminal 10O3′ and 20O3′ atoms of the minor groove binding guanosines for netropsin binding with d(CGCAATTGCG). The terminal O3′ atoms for the decamer do not lie

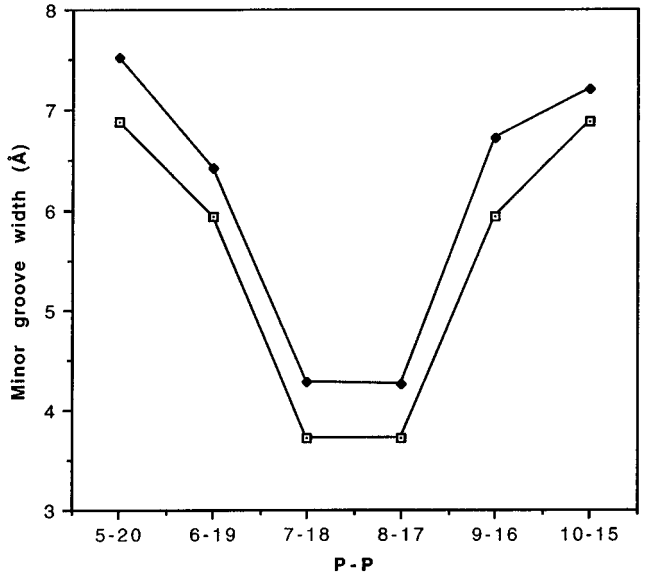


FIGURE 9: Minor groove width (intrastrand P−P distance minus 5.8 Å) for the native (⊡) (Spink et al., 1995) and netropsin-bound (◆) decamer sequence d(CGCAATTGCG) calculated with NEWHEL92.

symmetrically within the minor groove of the neighboring duplex. Atom 10O3′ lies between base pairs C3·G18 and A4·T17 while 20O3′ lies between base pairs G8·C13 and C9·G12.

The netropsin-accessible minor groove binding site within the native decamer sequence d(CGCAATTGCG) (Spink et al., 1995) contains a 2-fold axis of symmetry. Upon netropsin binding the minor groove becomes wider along its length while remaining essentially symmetric. Figure 9 shows the minor groove width for the native and netropsin-bound decamer sequences. The netropsin molecule fits snuggly within the DNA minor groove as shown in Figure 10. At one end of the groove the guanidinium group lies within the molecular surface of the DNA with the amidinium group at the other end of the netropsin bending toward base G(18). A similar type of bending was not seen in the structure of d(CGCGAATT^{Br}CGCG) due to the longer minor groove binding site of six base pairs, which allows the netropsin molecule ample space for minor groove binding.

For d(CGCAATTGCG) the netropsin is secured in position by the three amide groups of the netropsin hydrogen bonding to the adenine and thymine bases of the groove. The distance between amide atom N4 and the symmetry-related 20O3′ within the groove is 8.1 Å whereas the distance between amide atom N8 and 10O3′ is 6.4 Å. This asymmetry determines which way the netropsin lies within the groove. The longer guanidinium group lies toward the symmetry-related 20O3′ atom while the shorter and more flexible amidinium group at the other end of the netropsin bends toward base G(18) to enable fitting within the groove. In the case of the dodecamer structure with netropsin bound to d(CGCGAATT^{Br}CGCG) (Goodsell et al., 1995), the corresponding distances are larger with values of 9.9 Å between N4 and the symmetry-related 12O3′ atom and 9.7 Å between N8 and 24O3′.

Conclusions. This work presents the first crystal structure of netropsin bound within a DNA minor groove binding site

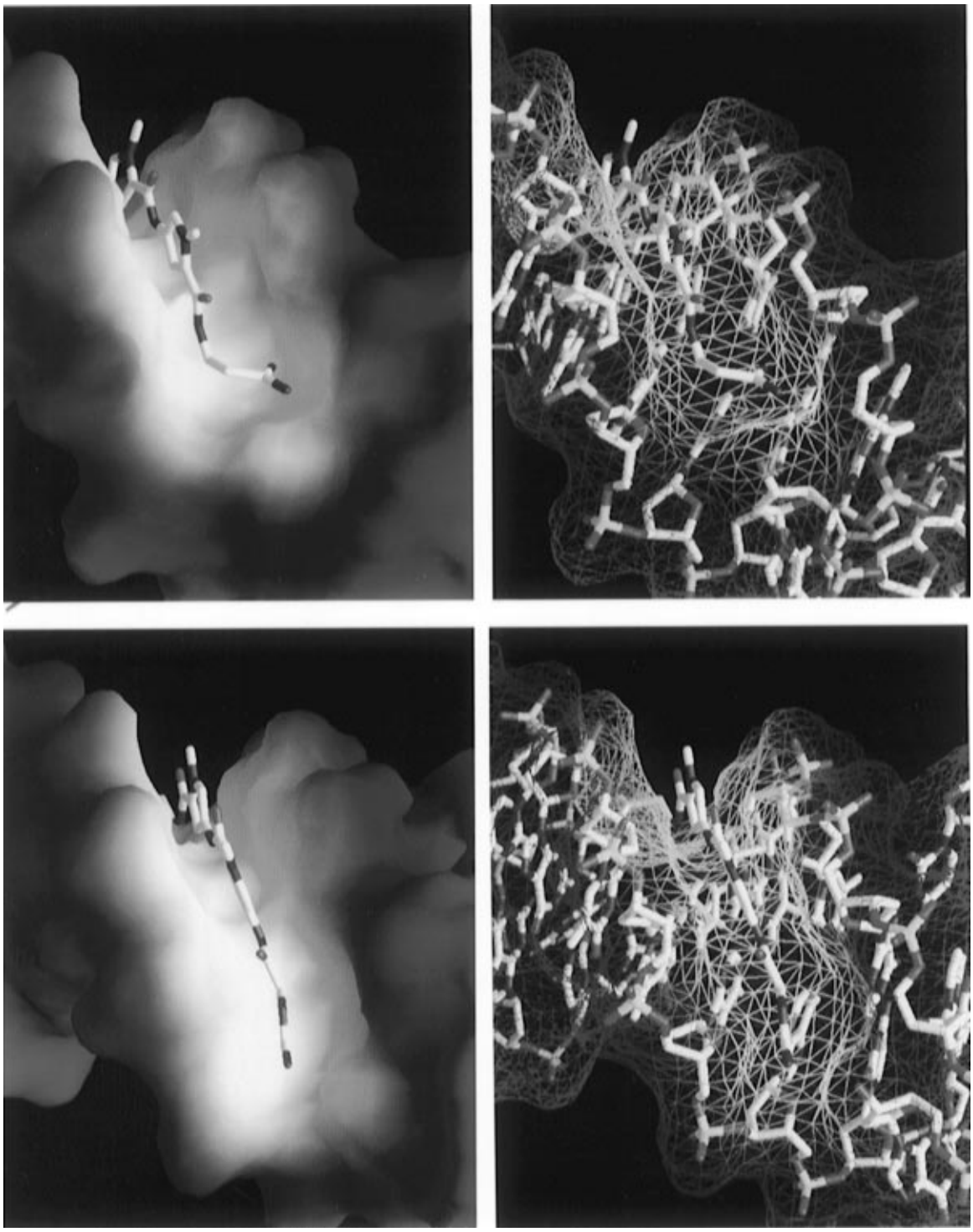


FIGURE 10: Netropsin molecule and DNA molecular surface, drawn using the program GRASP (Nicholls et al., 1991). The netropsin molecule fits snuggly within the minor groove surface at both the (top) amidinium and (bottom) guanidinium ends of the drug. Note the close position of the terminal 10O3′ and 20O3′ atoms and minor groove binding guanosines illustrated in the mesh views (right). The colors represent increasing distance from the netropsin molecule, from white to red.

of five base pairs in length and demonstrates how the netropsin molecule is accommodated by this shortened DNA minor groove site. The interactions of the netropsin amide groups anchor the drug in position while the ends of the drug fit within the groove by bending of the flexible amidinium group. The position of the netropsin molecule is unequivocally defined within this structure with the amide groups along the floor of the minor groove lying midway between A·T base pairs. This structure corresponds to netropsin binding of the class I type.

ACKNOWLEDGMENTS

We thank Dr. Charles A. Laughton, who assisted with the running of the program MOPAC, and Dr. Neil Spink, who was involved in preliminary work on this project.

REFERENCES

Abu-Daya, A., Brown, P. M., & Fox, K. R. (1995) *Nucleic Acids Res. 23*, 3385−3392.

Allen, F. H., Bellard, S., Brice, M. D., Cartwright, B. A., Doubleday, A., Higgs, H., Hemmelink, T., Hummelink-Peters, B. G.,

Kennard, O., Motherwell, W. D. S., Rodgers, J. R., & Watson, D. G. (1979) *Acta Crystallogr. B35*, 2331−2339.
Balendiran, K., Rao, S. T., Sekharudu, C. Y., Zon, G., & Sundaralingam, M. (1995) *Acta Crystallogr. D51*, 190−198.
Berman, H. M., Neidle, S., Zimmer, C., & Thrum, H. (1979) *Biochim. Biophys. Acta 561*, 124−131.
Berman, H. M., Olson, W. K., Beveridge, D. L., Westbrook, J., Gelbin, A., Demeny, T., Hsieh, S.-H., Srinivasan, A. R., & Schneider, B. (1992) *Biophys. J. 63*, 751−759.
Brünger, A. T., Kuriyan, J., & Karplus, M. (1987) *Science 235*, 458−460.
Coll, M., Aymami, J., van der Marel, G. A., van Boom, J. H., Rich, A., & Wang, A. H.-J. (1989) *Biochemistry 28*, 310−320.
Dickerson, R. E. (1992) *Methods Enzymol. 211*, 67−111.
Fregeau, N. L., Wang, Y., Pon, R. T., Wylie, W. A., & Lown, J. W. (1995) *J. Am. Chem. Soc. 117*, 8917−8925.
Goodsell, D. S., Kopka, M. L., & Dickerson, R. E. (1995) *Biochemistry 34*, 4983−4993.
Jones, T. A., Zou, J.-Y., Cowan, S. W., & Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
Kabsch, W. (1993) *J. Appl. Crystallogr. 26*, 795−800.
Kopka, M. L., Fratini, A. V., Drew, H. R., & Dickerson, R. E. (1983) *J. Mol. Biol. 163*, 129−146.
Kopka, M. L., Yoon, C., Goodsell, D., Pjura, P., & Dickerson, R. E. (1985) *J. Mol. Biol. 183*, 553−563.
Kumar, S., Bathini, Y., Joseph, T., Pon, R. T., & Lown, J. W. (1991) *J. Biomol. Struct. Dyn. 9*, 1−21.
Kumar, S., Joseph, T., Singh, M. P., Bathini, Y., & Lown, J. W. (1992) *J. Biomol. Struct. Dyn. 9*, 853−880.
Lown, J. W. (1995) *Drug Dev. Res. 34*, 145−183.
Nicholls, A., Sharp, K., & Honig, B. (1991) *Proteins 11*, 281−296.
Parkinson, G., Vojtechovsky, J., Clowney, L., Brünger, A. T., & Berman, H. M. (1996) *Acta Crystallogr. D52*, 57−64.
Patel, D. J. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 6424−6428.
Singh, M. P., Kumar, S., Joseph, T., Pon, R. T., & Lown, J. W. (1992) *Biochemistry 31*, 6453−6461.
Spink, N., Nunn, C. M., Vojtechovsky, J., Berman, H. M., & Neidle, S. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 10767−10771.
Sriram, M., van der Marel, G. A., Roelen, H. L. P. F., van Boom, J. H., & Wang, A. H.-J. (1992) *Biochemistry 31*, 11823−11834.
Stewart, J. P. P. (1990) *Mopac 6.0 (QCPE)*, Quantum Chemistry Program Exchange, Indiana University, Bloomington, IN 47405.
Tabernero, L., Verdaguer, N., Coll, M., Fita, I., van der Marel, G. A., van Boom, J. H., Rich, A., & Aymami, J. (1993) *Biochemistry 32*, 8403−8410.
Van Meervelt, L., Vlieghe, D., Dautant, A., Gallois, B., Precigoux, G., & Kennard, O. (1995) *Nature 374*, 742−744.
Vlieghe, D., Van Meervelt, L., Dautant, A., Gallois, B., Precigoux, G., & Kennard, O. (1996) *Acta Crystallogr. D52*, 766−775.
Vlieghe, D., Van Meervelt, L., Dautant, A., Gallois, B., Precigoux, G., & Kennard, O. (1996) *Science 273*, 1702−1705.

BI9628228