Acta Cryst. (1995). D**51**, 190–198

X-ray Structures of the B-DNA Dodecamer d(CGCGTTAACGCG) with an Inverted Central Tetranucleotide and Its Netropsin Complex

BY K. BALENDIRAN

Department of Biochemistry, College of Agricultural Life Sciences, University of Wisconsin-Madison, Madison, WI 53706, *USA*

S. T. RAO AND C. Y. SEKHARUDU

Department of Biochemistry, College of Agricultural Life Sciences, University of Wisconsin-Madison, Madison, WI 53706, *USA, and Biological Macromolecular Structure, Department of Chemistry & Ohio State Biotechnology Center, The Ohio State University,* 1060 *Carmack Road, Columbus, OH* 43210, *USA*

G. ZON

Lynx Therapeutics Inc., 465 *Lincoln Center Drive, Forest City, CA* 94404, *USA*

AND M. SUNDARALINGAM*

Department of Biochemistry, College of Agricultural Life Sciences, University of Wisconsin-Madison, Madison, WI 53706, *USA, and Biological Macromolecular Structure, Department of Chemistry & Ohio State Biotechnology Center, The Ohio State University,* 1060 *Carmack Road, Columbus, OH* 43210, *USA*

(*Received* 29 *April* 1994; *accepted* 19 *September* 1994)

Abstract

The crystal structures of the B-DNA dodecamer d(CGCGTTAACGCG) duplex (T2A2), with the inverted tetranucleotide core from the duplex d(CGCGAATTCGCG) [A2T2, Dickerson & Drew (1981). *J. Mol. Biol.* **149**, 761–768], and its netropsin complex (T2A2–N) have been determined at 2.3 Å resolution. The crystals are orthorhombic, space group $P2_12_12_1$, unit-cell dimensions of $a = 25.7$, $b = 40.5$ and $c = 67.0$ Å for T2A2 and $a = 25.49$, $b = 40.87$, $c = 67.02$ Å for T2A2–N and are isomorphous with A2T2. The native T2A2 structure, with 70 water molecules had a final R value of 0.15 for 1522 reflections ($F > 2\sigma$), while for the netropsin complex, with 87 water molecules, the R value was 0.16 for 2420 reflections. In T2A2, a discontinuous string of zig-zagging water molecules hydrate the narrow A·T minor groove. In T2A2–N, netropsin binds in one orientation in the minor groove, covering the TTAA central region, by displacing the string of waters, forming the majority of hydrogen bonds with DNA atoms in one strand, and causing very little perturbation of the native structure. The helical twist angle in T2A2 is largest at the duplex center, corresponding to the cleavage site by the restriction enzymes *Hpa*I and *Hinc*II. The sequence inversion AATT→TTAA of the tetranucleotide at the center of the molecule results in a different path for the local helix axis in T2A2 and A2T2 but the overall bending is similar in both cases.

Introduction

Oligonucleotides containing A_nT_n tracts (with $n \geq 3$) behave anomalously in electrophoretic gels and migrate like curved (bent) molecules while those containing T_nA_n behave as straight rods (Hagerman, 1985, 1986; Koo, Wu & Crowthers, 1986). It was of interest to compare the bending in the known dodecamer d(CGCGAATTCGCG) (A2T2) (Dickerson & Drew, 1981) with that of the dodecamer d(CGCGTTAACGCG) (T2A2), which has a sequence inversion in the central tetranucleotide core, and both with $n = 2$. The hexanucleotide restriction fragment (GTTAAC) in the middle of the present dodecamer is recognized by the enzymes *Hpa*I and *Hinc*II, and is cleaved at the central TA step yielding blunt ends resulting in 5′-phosphoryl and 3′-hydroxyl ends (Kelly & Smith, 1970). However, in A2T2 the central hexanucleotide segment GAATTC is recognized by the enzyme *Eco*RI and is cleaved at the GA junctions with sticky ends (Goodman, Greene, Garfin & Boyer, 1977). In A2T2, the largest twist angle was observed at T20–C21, opposite the GA junction. It was of interest to see if a high twist is also characteristic of the cleavage site in T2A2. In addition, it was of interest to study the effects

* Author to whom correspondence should be addressed at Ohio State University.

of sequence inversion of helical parameters, groove widths and netropsin binding.

In this paper we report the reinvestigation of the structures of both the native T2A2 and its netropsin complex, T2A2–N, at 2.3 Å resolution because in our preliminary reports (Balendiran & Sundaralingam, 1991*a*,*b*), the intensity data sets for the two structures were inadvertently interchanged. The DNA structures in the native and the netropsin complex are very similar in the present analysis and also similar to those in our earlier analysis. However, the description of the minor grove interactions in both the native and the netropsin complex have been corrected; in particular netropsin is now found to bind in only one orientation. Since the first report of netropsin binding in the minor groove of 5Br-A2T2–N (Kopka, Yoon, Goodsell, Pjura & Dickerson, 1985), four other netropsin–DNA complexes belonging to the Drew–Dickerson family, have appeared: two complexes with an inner AATT sequence but different modifications on the major-groove side (5Br-A2T2–N, Kopka *et al.*, 1985; 6et-A2T2–N and A2T2–N, Sriram, van der Marel, Roelen, van Boom & Wang, 1992), one with an inner sequence of ATAT (ATAT–N, Coll *et al.*, 1989) and another with an extended inner sequence AAATTT (A3T3–N, Tabernero *et al.*, 1993). The structural features of all these complexes are compared.

Experimental procedures

The dodecamer d(CGCGTTAACGCT) was synthesized by solid-support phosphoramidite chemistry (Dorman, Noble, McBride & Caruthers, 1984) using an automated machine. The 5′-dimethoxytrityl derivative of the oligonucleotide was prepared on a 10 μmol scale, purified by reverse-phase high-pressure liquid chromatography (HPLC) (Zon & Thompson, 1986), followed by detritylation with acetic acid. The resultant 5′-hydroxyl product was isolated in the sodium form by precipitation with aqueous ethanol and NaCl with a 20–30% overall yield.

Native crystals of the dodecamer were grown from a solution buffered with Tris, containing 1 equivalent of oligonucleotide, 2.5 equivalents of $MgCl_2$, 2 equivalents of spermine tetrachloride and 40% 2-methyl-2,4-pentanediol (MPD) at pH 7.5, by the vapor-diffusion method against a 50% MPD reservoir at 282 K. Crystals of the netropsin complex were grown from a solution, at pH 7.5, containing 1 equivalent of oligonucleotide, 5 equivalents of $MgCl_2$, 1.8 equivalents of spermine tetrachloride, 1.2 equivalents of netropsin and 30% MPD, by the vapor-diffusion method against a 40% MPD reservoir at 282 K.

Intensity data collection

A crystal was mounted in a capillary tube and sealed with a plug of mother liquor, away from the crystal.

Table 1. *Crystal data and refinement statistics for d*(CGCGTTAACGCG) *dodecamer and its netropsin complex*

Crystal data*	Native	Netropsin complex
Crystal system	Orthorhombic	Orthorhombic
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters		
a (Å)	25.70	25.49
b (Å)	40.50	40.87
c (Å)	67.00	67.02
Volume/base-pair (Å^3)		1456
Intensity data collection		
X-ray source	Rigaku RU200 rotating anode	
Monochromator	Graphite	
Area detector	Nicolet X100	
Goniostat	Huber four circle	
Data-frame range (°)	0.20	
Exposure time per frame (s)	110	
Temperature (K)	258	
Resolution (Å)	2.3	
Crystal dimensions (mm)	0.15 × 0.25 × 0.5	0.30 × 0.50 × 0.90
Unique reflections	2214	2725
R_{sym} (%) based on I	6.2	5.1
Reflections with $F>2\sigma(F)$ (used in the refinement)	1522	2420
Refinement program	*X-PLOR*, 3.0	*X-PLOR*, 3.0
Final R value (%)	14.8	16.2
Model		
DNA atoms of residues	1–24	1–24
Solvents	70	87
Netropsin	—	Yes
Estimated positional error (Å) (Luzzati, 1952)	0.21	0.23
Parameter file	param11.dna	param11.dna
R.m.s. deviations in model from ideal geometry		
Bond lengths (Å)	0.015	0.015
Bond angles (°)	3.2	3.6
Torsion angles (°)	28.2	29.4
Improper angles (°)	2.1	2.2

* Similarity of these parameters with those of A2T2 suggest that they are isomorphous.

Intensity data on the native (T2A2) crystal (0.15 × 0.25 × 0.5 mm) were collected to 2.3 Å resolution at 258 K, on the area detector at Argonne National Laboratory using a Rigaku RU-200 rotating-anode generator equipped with a Cu anode and a graphite monochromator. The crystal dissolved before all the frames of data could be collected. The data frames were processed with *XENGEN* 1.3 (Howard *et al.*, 1987). Of the 3169 possible reflections, only 2214 reflections were recorded with an overall R_{sym} of 6%. The diffraction was strong to a resolution of 2.5 Å but rapidly became weak at higher resolution. 1522 reflections had $F > 2\sigma(F)$ and were used in the analysis. The unit-cell constants and the crystal settings used for date collection are summarized in Table 1.

The intensity data for the netropsin complex (T2A2–N) were collected by a similar procedure using the only crystal grown, which was considerably larger (0.33 × 0.5 × 0.9 mm). The diffraction from this crystal was stronger than from the native one. 2759 unique reflections were recorded (R_{sym} of 5%) and 2552

reflections had $F > 2\sigma(F)$ (81% of possible). Of these, 2420 reflections in the resolution range 6–2.3 Å were used in the analysis (Table 1). The R value between the native and drug-complex data sets was 0.25 for 1519 common reflections. The relatively small number of observed reflections from the native crystal is partly due to the incomplete coverage of the reciprocal lattice and also presumably due to the larger DNA mobility in the absence of netropsin binding.

Structure solution and refinement

Native (T2A2)

The coordinates of the A2T2 structure (Dickerson & Drew, 1981), Protein Data Bank (Bernstein *et al.*, 1977) entry 1BNA, was the starting model for refinement of the T2A2 structure. All observed reflections between 6 and 2.3 Å resolution were used in the refinement, carried out with the program *X-PLOR*, Version 3.0 (Brünger, 1992). The X-ray and energy terms were assigned equal weight and a mild restraint for C2′-*endo* sugar pucker was used by restraining the backbone dihedral angle δ to 160 (10)° with a barrier of 83.7 kJ mol^{-1} (20 kcal mol^{-1}). A rigid-body refinement dropped the R value to 0.32, confirming that the two structures are isomorphous. The four inner base pairs were now omitted one at a time, and the model was refined using the conjugate-gradient procedure. In the resulting $2F_o - F_c$ omit maps, the correct inner T2A2 sequence was evident for all the four inner base pairs. To minimize model bias, T2A2 was subjected to simulated annealing by first heating the system to 2273 K and slowly cooling it to 573 K in steps of 25 K. This dropped the R value to 0.27. Several rounds of model refitting were carried out, using omit $2F_o - F_c$ maps, in which one base pair was omitted at a time, and refitted on our Evans & Sutherland ESV-10 with the program *FRODO* (Jones, 1985). Several solvent molecules were identified in difference $F_o - F_c$ maps, using a procedure similar to that used in our previous studies (Jain, Zon & Sundaralingam, 1989). Discontinuous electron density corresponding to a string of water molecules was found in the narrow groove of the A·T tract (Fig. 1) and continued refinement, with the final model containing 70 solvent sites, converged to a final R value of 0.148. The electron density in the omit maps was strong for all the bases and phosphate groups, but somewhat weak for the sugars.

Netropsin complex (T2A2–N)

The refinement was started with the final coordinates of the native T2A2 model. All the solvents were removed and the model was annealed, in a manner similar to that used in the refinement of the native structure. The correct inner TTAA sequence was again recovered from the omit maps. The DNA model was refined and refitted. 48 well ordered solvent sites, away from the minor groove were identified and including them in the refinement dropped the R value to 0.25. A minimum-bias omit electron-density map (Read, 1986) showed a continuous electron density in the minor groove. The bulges in this electron density correspond to the carbonyl and *N*-methyl groups of the pyrrole ring, which are asymmetric about the molecular center and a model of the ordered netropsin in one orientation could be unambiguously built into the electron density (Fig. 1). Since the DNA sequence is self-complementary, netropsin could also bind in the

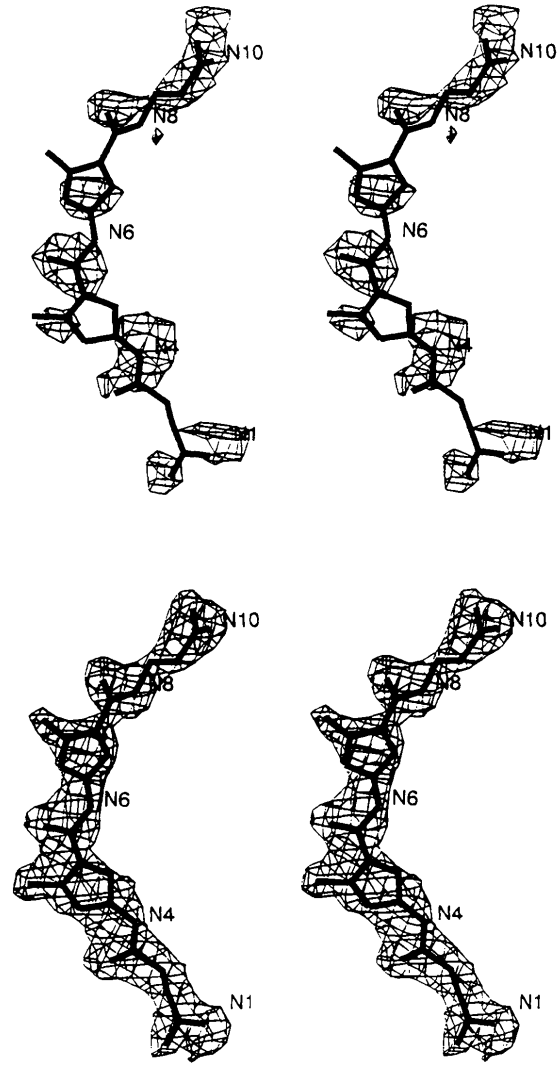


Fig. 1. Electron density in omit $2F_o - F_c$ maps using minimum bias coefficients (Read, 1986) in the native T2A2 (top) and the netropsin complex (bottom). View is into the the minor groove. The contours are drawn at 1.2σ. Waters in the native structure spanning the netropsin-binding region and netropsin atoms in the complex, respectively, were omitted from the phasing. The atomic model of netropsin after refinement is superposed. Notice the discrete electron density in the native for the water molcules bound in the mirror groove and the continuous electron density in the complex for netropsin with characteristic bulges corresponding to the C=O and N—CH_3 groups of the drug.

'opposite' orientation but this gave a poor fit to the electron density, and the model also refined to 1% higher *R* value. Thus, in the present T2A2–N crystal netropsin binds in one orientation to the duplex, in contast to the A2T2 crystal (Sriram *et al.*, 1992) where it binds in both orientations. Continued refinement, including netropsin and a total of 87 solvents in the model converged to an *R* value of 0.162 for 2420 reflections. In general, the omit electron density was stronger in the drug complex compared to the native, particularly for the sugars, reflecting the better quality of intensity data for the complex. The coordinates of both the native DNA and its netropsin complex have been deposited with the Protein Data Bank at Brookhaven.* The estimated coordinate errors (Luzzati, 1952) for the native and the drug complex are 0.21 and 0.23 Å, respectively.

Results and discussion

The structures of the T2A2 dodecamer (Fig. 2*a*) and its netropsin complex (Fig. 2*b*) are isomorphous with the A2T2 dodecamer and other members of this structural family. The main-chain backbone torsion angles α (P—O5′), β (O5′—C5′), γ (C5′—C4′), δ (C4′—C3′), ε (C3′—O3′) and ζ (O3′—P)† in T2A2 and the netropsin complex are in the common *gauche*$^-$ (−*sc*), *trans* (*ap*), *gauche*$^+$ (+*sc*), *trans* (*ap*), *trans* (*ap*), *gauche*$^-$ (+*sc*) ranges, respectively; all the glycosyl χ angles are in the *anti* range and correlate generally with the sugar pucker: lower χ values for the C(3′)-*endo*/C(4′)-*exo* sugars, and higher χ values for the C(2′)-*endo*/C(1′)-*exo* sugars (Sundaralingam, 1969). The first three torsion angles α, β, and γ on the 5′ side of the nucleotide residues show less variation than the corresponding three 3′-side torsion angles ζ, ε and δ. Except for the sugars of C23 of the native and C3, C23 and G24 of the drug complex, which are in the N-domain (C3′-*endo*), the remaining sugars are in the S-domain (C2′-*endo*) displaying a characteristically large latitude in the pseudorotation phase angle (110–190°) (Altona & Sundaralingam, 1972; Sudaralingam, 1982). The amplitudes of puckering of the sugars (τ_m) range from 25 to 50°, mean 37°; a wider spread in these values is seen in the second strand (residues 13–24) than in the first strand (residues 1–12).

The asymmetric unit is the duplex itself and the r.m.s. deviation between the two strands is 2.1 Å, for both T2A2 and the drug complex. The deviations progressively increase from the central step towards either end and also from the axis towards the sugar–phosphate perimeter of the duplex, resembling the trend found for the average atomic thermal parameters (Fig. 3).

In T2A2, 13 water molecules are found in the minor groove (Fig. 2*a*). This water spine (Drew & Dickerson, 1981) is somewhat broken in that some of the distances are rather long and may involve interactions with a second spell of hydration. The bases in only the central TTAA region form direct hydrogen bonds to the minor groove waters.

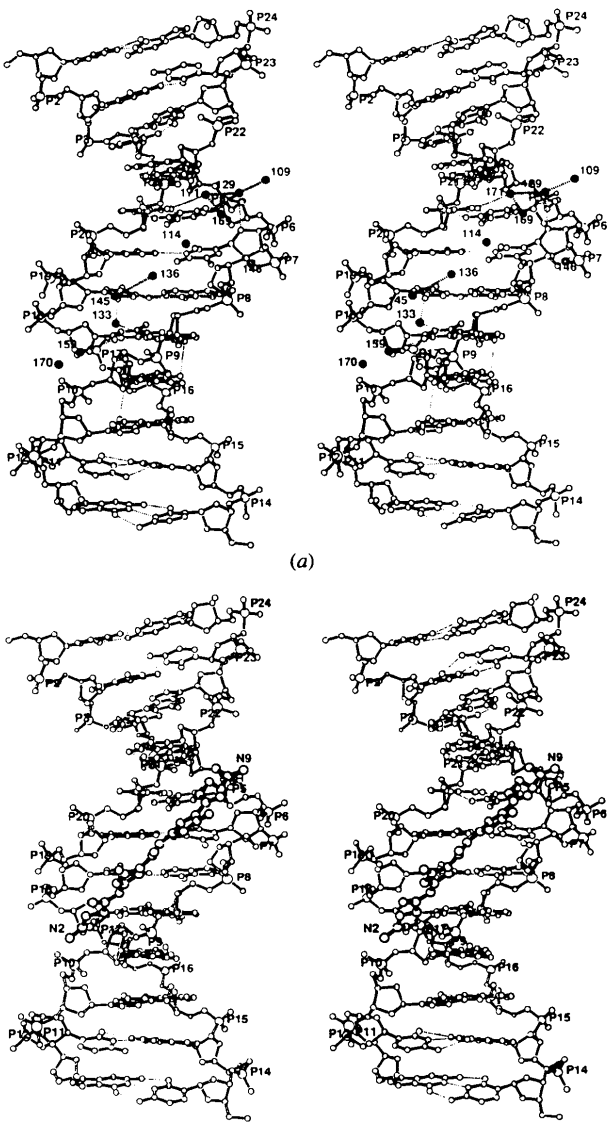


Fig. 2. Stereo diagram, looking into the minor groove of (*a*) T2A2 and (*b*) T2A2–N. (*a*) Minor groove waters (filled circles) are connected by dashed lines if the distances between them are less than 3.4 Å; (*b*) atoms N2 of the guanidinium group and N9 of the amidinium group of netropsin are labeled.

* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory. Free copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Reference: GR0385). A list of deposited data is given at the end of the issue.

† An earlier notation still in use for the backbone bonds: ω (P—O5′), φ (O5′—C5′), ψ (C5′—C4′), ψ' (C4′—C3′), φ' (C3′—O2′) and ω'(O3′—P) (Sundaralingam, 1969), respectively.

Guanidinium$\overset{17°}{—}$Peptide 1$\overset{9°}{—}$Pyrrole 1$\overset{24°}{—}$Peptide 2$\overset{17°}{—}$Pyrrole 2$\overset{19°}{—}$Peptide 3$\overset{15°}{—}$Amidinium

(I)

T2A2–Netropsin complex

The crescent-shaped netropsin molecule binds in the center of the narrow minor groove of the T2A2 duplex, displacing the minor groove water molecules. The drug binds in one orientation, with its amidinium group pointing towards the top C1·G24 base pair and its guanidinium group towards the bottom G12·C13 base pair. The dihedral angles between the seven successive planar groups are shown in (I) (for nomenclature see, for example, Kopka *et al.*, 1985).

The dihedral angle between the two pyrrole rings is 34° and that between the guanidinium and amidinium ends is 73°. Bifurcated (three-center) or direct hydrogen bonding of the three peptide N atoms, N4, N6 and N8, to the polar O2/N3 base atoms of the four inner base pairs, T5·A20, T6·A19, A7·T18 and A8·T17, help anchor the middle of the drug firmly into the DNA minor groove (Table 2). N1 of the NH_2 group of the charged guanidinium end, is hydrogen bonded to T17 and T18 of strand 2 (Fig. 4) while N10 of the NH_2 of the amidinium end, is hydrogen bonded to T5 of strand 1 and C21 of strand 2, which involves the G4·C21 base pair flanking the central tetramer. Thus, the drug netropsin hydrogen bonds to minor-groove base atoms of five residues, T17, T18, A19, A20 and C21 of strand 2 and only two bases, T5 and A8, in strand 1; the hydrogen-bonding interactions with the DNA are asymmetric. Two of the sugar O4′ atoms in the minor groove of TTAA participate in hydrogen bonds while the remaining sugars lining the minor groove are in van der Waals contact with netropsin. None of the sugar O4′ atoms stack on the pyrrole rings. Netropsin interacts with 2_1-screw-related duplexes on either side *via* water bridges. At the lower end, N1 of the guanidinium group is bridge to the 3′-OH group of G24 *via* a single water molecule while at the upper end, N10 of the amidinium group, is bridged *via* three water molecules to the 3′-OH group of G12. It is possible that this asymmetry reflects the asymmetry in the DNA duplex in the crystal and results in selecting a single orientation for the netropsin complex in the crystal.

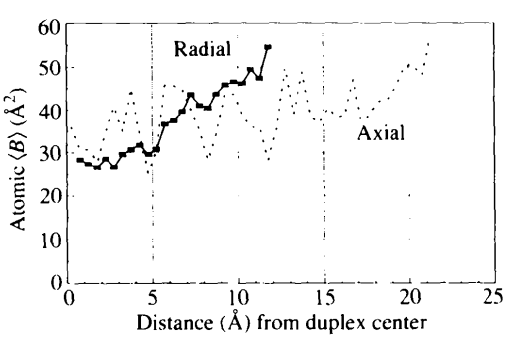


Fig. 3. Plot of the average atomic thermal parameter, ⟨*B*⟩, as a function of the radial and azimuthal distance from the molecular center for T2A2.

Table 2. *Selected distances between netropsin atoms and atoms on the floor of the DNA minor groove*

Netropsin atom		DNA atom		Strand	Distance (Å)	*X*—Net···DNA *X*	Angle (°)
Hydrogen bonds*							
N1	*T*	O2	(T17)	2	2.74	C1	132
		O4′	(T18)	2	3.18	C1	91
N4	*T*	O2	(T18)	2	3.04	C3	105
		N3	(A8)	1	3.04	C3	113
N6	*T*	N3	(A19)	2	3.14	C9	123
		O4′	(A20)	2	3.19	C9	119
N8	*D*	N3	(A20)	2	2.97	C15	130
N10	*T*	O2	(C21)	2	3.13	C18	133
		O2	(T5)	1	2.64	C18	109
N1		*W* 101			3.4	C1	103
N9		*W* 84			3.1	C18	90
		W 88			3.4	C18	151
N10		*W* 84			3.3	C18	83
O3		*W* 97			2.6	C15	157
Additional contacts							
N1		O4′	(T17)	2	3.7	C1	161
N3		O4′	(T18)	2	3.3	C1	88
O1		O4′	(C9)	1	3.5	C3	70
N4		O4′	(A19)	2	3.6	C3	91
N4		O4′	(C9)	1	3.6	C3	67
N8		O4′	(A21)	2	3.4	C15	106
N9		O4′	(T6)	1	3.7	C18	61
N10		O4′	(T6)	1	4.0	C18	49
N10		N2	(G4)	1	3.2	C18	166
Contacts with C2 atoms of adenines†							
C5		C2	(A7)	1	3.9	C4	112
C2		C2	(A8)	1	3.5	N3	145
N4		C2	(A8)	1	3.5	C3	103
N6		C2	(A19)	2	3.7	C9	138
C11		C2	(A19)	2	3.6	C10	100
N8		C2	(A20)	2	3.4	C15	145

* *T* = bifurcated (three-center) bond, *D* = direct bond.

† These contacts rationalize specificity of netropsin binding to AT (as opposed to GC) regions (Dervan, 1986).

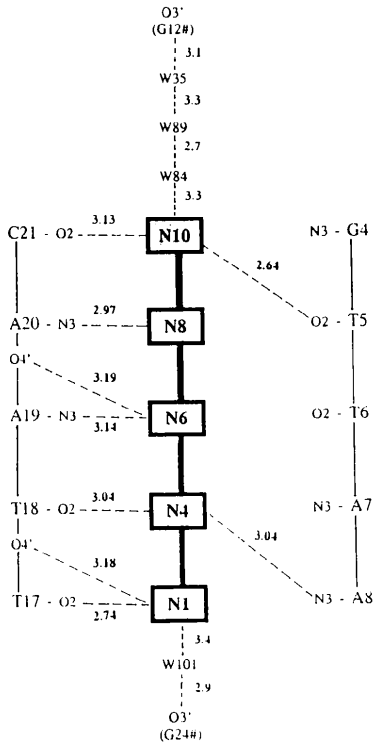
The native T2A2 structure is perturbed very little upon netropsin binding, r.m.s. 0.37 Å. Larger deviations are found for the anionic O atoms in the top half of the duplex, near the amidinium end of netropsin. The bases have much smaller deviations but the four inner base pairs show a concerted displacement. The T6·A19 and A8·T17 base pairs are displaced more into the minor groove than the other two, T5·A20 and A7·T18, and are pulled closer to the drug. The minor groove width shows no significant expansion or contraction.

Crystal structures of five other netropsin complexes in this dodecamer family are known: 5Br-A2T2–N, ATAT–N, 6et-A2T2–N, A2T2–N and A3T3–N. These isomorphous structures were also determined at medium resolution (2.1–2.5 Å) and have r.m.s. atomic deviations

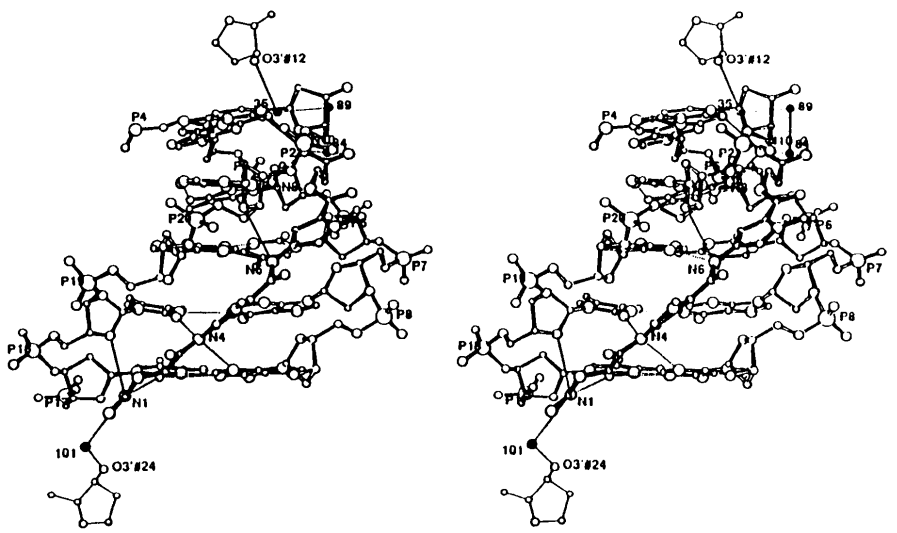
ranging up to 1.5 Å between any pair. Only in the present T2A2–N complex, the full complement of hydrogen bonds between netropsin and the duplex is seen. In the other complexes some of the sugar O4′ atoms are engaged in stacking interactions (Wang & Teng, 1990; Sriram *et al.*, 1992; Tabernaro *et al.*, 1993) presumably compensating for the lack of the full complement of hydrogen bonding. The O4′–pyrrole stacking seen in these cases, is similar to that found between sugar O4′ atoms and the bases in nucleotide X-ray structures (Bugg, Thomas, Rao & Sundaralingam, 1971). In these complexes, netropsin binding, based on the proposed hydrogen-bonding schemes, spans three to six base pairs. In 5Br-A2T2–N three base pairs are involved while in A3T3–N, six base pairs are involved and the binding locus is also shifted from the center of the duplex towards the lower G12·C13 base pair. Thus, the binding locus is somewhat variable (Table 3) *i.e.* the binding is sequence tolerant.

Netropsin binds in a single orientation in three of these complexes (T2A2–N this work, 5Br-A2T2–N and A3T3–N) and provides an opportunity to compare them and identify the common interactions (Table 3). The polarity of netropsin binding in T2A2–N is opposite to that in 5Br-A2T2–N, but is the same as in A3T3–N. In the remaining three cases, netropsin binds in a bidirectional fashion (disordered). In all six complexes, the guanidinium N1, the first amide N4 and the amidinium N10 are hydrogen bonded with DNA. However, it is noteworthy that when netropsin binds in one orientation, the third amide N8 is also hydrogen bonded to the DNA. In all these complexes, netropsin retains the register of NH groups on the guanidinium side better than on the amidinium side (Sriram *et al.*, 1992), *i.e.* the amidinium end is more variable. Anchoring the amide N8 in this variable end is apparently needed for the observed binding in a single orientation in the crystal. Further it appears that the polarity of netropsin (up or down) depends upon the presence (as in T2A2–N, A3T3–N) or absence (as in 5Br-A2T2–N) of hydrogen bonds involving the middle amide N6.

Besides the present work, the native structure of only A2T2 is known for a comparison of the native and its cognate drug complex. The distortions caused by netropsin binding are small in T2A2 (r.m.s. 0.38 Å) but are more substantial in A2T2 (r.m.s. 1.14 Å). However, both in T2A2 and A2T2, the hydration



(*a*)



(*b*)

Fig. 4. Netropsin interactions with T2A2 duplex. (*a*) A schematic diagram showing the hydrogen bonding with the bases and sugar O4′ atoms of DNA. (*b*) Stereoview of netropsin (filled bonds) with the duplex region spanning the binding site (open bonds). The single water bridging the guanidium end and three waters bridging the amidinium end (filled circles) with terminal deoxyriboses of two symmetry-related duplexes (bonds drawn as single lines) are also shown.

Table 3. *Comparison of netropsin binding in the dodecamer–netropsin complexes*

Identification*	Base sequences† ...4 5678 9...	Binding polarity‡	Hydrogen bonds to DNA atoms§ N1	N2	N3	N4	N6	N8	N9	N10	Reference
T2A2–N	CGCG **TTAA** CGCG	←	*WY*			*Y*	*y*	*Y*	*W*	*WY*	Present study
A3T3–N	CGCA **AATT** **TGCG**	←	*Y*			*Y*	*y*	*Y*	*Y*	*Y*	Tabernero *et al.* (1993)
5BrA2T2–N	CGCG **AATT** CGCG	→	*Y*			*Y*		*Y*		*Y*	Kopka *et al.* (1985)
A2T2–N	CGCG **AATT** CGCG	⟷	*WY*		*Y*	*Y*				*Y*	Sriram *et al.* (1992)
6etA2T2–N	CGCG **AATT** CGCG	⟷	*Y*			*Y*	*Y*			*Y*	Sriram *et al.* (1992)
ATAT–N	CGCG **ATAT** CGCG	⟷	*Y*			*Y*	*Y*		*Y*	*Y*	Coll *et al.* (1989)

* Abbreviations used: T2A2–N, netropsin complex of d(CGCGTTAACGCG); A2T2–N, Netropsin complex of d(CGCGAATTCGCG); ATAT–N, netropsin complex of d(CGCGATATCGCG); A3T3–N, netropsin complex of d(CGCAAATTTGCG); 5Br-A2T2–N, netropsin complex of d(CGCGTTAA5-BrCGCG): 6et-A2T2–N, netropsin complex of d(CGC$^{6\text{-et}}$GTTAACGCG).

† Netropsin bonding region is shown in bold.

‡ Polarity is the direction from the guandinium end to the amidinium end of netropsin. The top 1–24 base pair of the duplex is to the left and the bottom 12–13 base pair to the right. In the last three complexes, the binding is bidirectional (⟷).

§ *Y*,*y*, DNA atom; *W*, water; blank, no hydrogen bond.

pattern in the major groove is somewhat altered upon netropsin binding in the minor groove. Conversely, modification of a base in the major groove flanking the netropsin-binding locus (A2T2–N, 5Br–A2T2–N and 6et-A2T2–N) seems to affect the netropsin-binding mode (Table 3).

Comparison of T2A2 with A2T2

The atomic r.m.s. deviation between T2A2 and A2T2 is 0.73 Å. The helical parameters (EMBO Workshop, 1989) were calculated using all the available programs: *NEWHEL92* (Dickerson, 1989), *NUCPARM* (Bhattacharya & Bansal, 1989), *Dials & Windows* (Lavery & Skelnar, 1989) and *RNANEW8* (Babcock, Pednault & Olson, 1994). Even though the actual values are different, the trends in the derived parameters were similar and the values obtained from the fourth program are being used here. The helical twists for both T2A2 and its netropsin complex have similar trends: the average

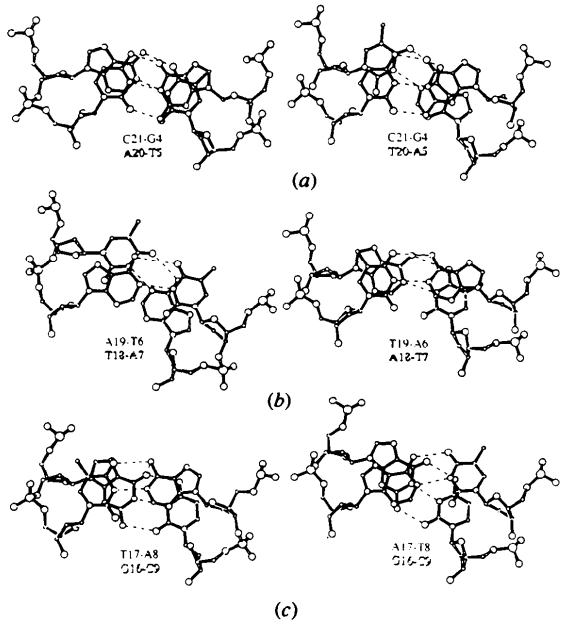


Fig. 5. Base-stacking interactions in T2A2 (left) and A2T2 (right) at steps where the step types (Py/Pu) have changed: (*a*) 4–5, (*b*) 6–7 and (*c*) 8–9. Dark bonds for the base pair on the top and open bonds for the base at the bottom. Notice the lack of stacking at the central Py–Py step 6–7 in T2A2, which is overwound (see text).

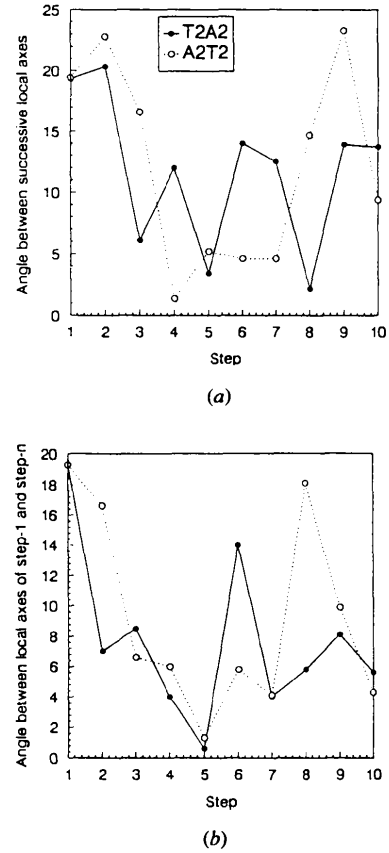


Fig. 6. Comparison of bending in T2A2 and A2T2. (*a*) Angles (°) between the local helical axes of the first step and at each succeeding step (i + 1) are plotted at the point i. (*b*) Angles (°) between local helix axes at each successive step are shown and the value for step $i \longrightarrow i + 1$ is plotted at i.

values are 35.5° (r.m.s. 6.0°) and 35.6° (r.m.s. 4.5°), and the same average rise per residue of 3.4 Å. The highest helical twist of 47.5° is at step 6 which corresponds to the site of the cleavage of T2A2 by the restriction enzymes *Hpa*I and *Hin*cII (Kelly & Smith, 1970). The high helical twist at step 6 is compensated by a low helical twist (25.2°) at the next step 7. On the other hand, in A2T2 a large helical twist of 40.4° is found at step 4, which also corresponds to the site of cleavage by *Eco*RI (Lomonossof, Butler & Klug, 1981). The sequence inversion between T2A2 and A2T2 results in changes at the three steps: 4, 6 and 8. Step 4 is Pu–Py in T2A2 and Pu–Pu in A2T2. In T2A2, the Pu–Py step is underwound (helical twist 27° *versus* 40°) leading to increased stacking (Fig. 5*a*). Upon the change from Pu–Py in A2T2 to a Py–Pu in T2A2, the central (sixth) step is overwound in T2A2 (47° *versus* 34°), leading to decreased stacking (Fig. 5*b*). However, upon a change from a Py–Py in A2T2 to Pu–Py in T2A2, the twist angle at step 8 does not change (39°) (Fig. 5*c*). Furthermore, there is an inversion of the roll angle at the central step (2.3° for T2A2 and −3.6° for A2T2) where the base sequence is inverted from TA to AT. The tilt angles at steps 4 and 8 are +3.8°; −2.6° in T2A2 and −1.0° and +3.2° in A2T2 and display an inverted profile between the two structures, as expected.

Bending in T2A2 and A2T2

The Crothers junction model (Levene & Crothers, 1983) predicts that A-tracts themselves are essentially straight and the bending occurs on either side of the A-tracts (Sundaralingam & Sekharudu, 1988). To investigate this, the bending in T2A2 and A2T2 has been examined in two ways. The local helix axis of each step was obtained and the angles between the local axes of successive steps (Fig. 6*a*), as well as the angles between the first and each of the remaining steps (Fig. 6*b*) were calculated. The T2A2 and A2T2 dodecamers exhibit a similar trend in their bend angles at steps 1–2, 2–3, 3–4, 5–6, 7–8, 9–10 and 10–11 but differ at 4–5, 6–7 and 8–9, precisely where the step types are also different. When the angles between the first and each of the remaining steps are examined, the effect of sequence inversion can be seen at two steps: in T2A2 the angle between steps 1 and 7 is high and between steps 1 and 9 is low, and an opposite trend is seen in A2T2. It is interesting that the angles between successive steps in the central region of the two duplexes (steps 3–8), where the sequence is inverted, are also anti-correlated, like roll and tilt. However, the first and last steps make similar angles in both T2A2 and A2T2, indicating the similarity in the overall curvature of the two duplex structures in the crystal. This is probably caused by the very similar tertiary interactions in the crystal, involving the two ends.

We wish to thank Dr Edward Westbrook and Ms Mary Westbrook for assistance in the use of the Siemens–Nicolet area detector at the Argonne National laboratory. We thank Dr Craig Bingman for intensity-data collection. This work was supported by a National Institutes of Health grant GM17378 from the United States Public Health Service.

References

ALTONA, C. & SUNDARALINGAM, M. (1972). *J. Am. Chem. Soc.* **94**, 8205–8212.

BABCOCK, M. S., PEDNAULT, E. P. D. & OLSON, W. (1994). *J. Mol. Biol.* **237**, 125–156.

BALENDIRAN, K. & SUNDARALINGAM, M. (1991*a*) *Int. J. Quantum Chem. Quantum Biol. Symp.* **18**, 199–203.

BALENDIRAN, K & SUNDARALINGAM, M. (1991*b*). *J. Biomol. Struct. Dynam.* **9**, 511–516.

BERNSTEIN, F. C., KOETZLE, T. F., WILLIAMS, G. J. B., MEYER, E. F. JR, BRICE, M. D., ROGERS, J. B., KENNARD, O., SHIMANOUCHI, T. & TASUMI, M. (1977). *J. Mol. Biol.* **112**, 535–542.

BHATTACHARYA, D. & BANSAL, M. (1989). *J. Biomol. Struct. Dynam.* **4**, 635–653.

BRÜNGER, A. T. (1992). *X-PLOR Manual*, Version 3.0, Yale Univ., New Haven, CT, USA.

BUGG, C. E., THOMAS, J. M., RAO, S. T. & SUNDARALINGAM, M. (1971). *Biopolymers*, **10**, 175–219.

COLL, M., AYMAMI, J., VAN DER MAREL, G. A., VAN BOOM, J. H., RICH, A. & WANG, A. H.-J. (1989). *Biochemistry*, **28**, 310–320.

DERVAN, P. B. (1986). *Science*, **232**, 464–471.

DICKERSON, R. E. (1989). *J. Biomol. Struct. Dynam.* **4**, 627–634.

DICKERSON, R. E. & DREW, H. R. (1981). *J. Mol. Biol*, **149**, 761–768.

DORMAN, M. A., NOBLE, S. A., MCBRIDE, L. J. & CARUTHERS, M. H. (1984). *Tetrahedron*, **40**, 95–102.

DREW, H. R. & DICKERSON, R. E. (1981). *J. Mol. Biol.* **151**, 535–556.

EMBO Workshop (1989). *EMBO J.* **8**, 1–4.

GOODMAN, H. M., GREENE, P. J., GARFIN, D. E. & BOYER, H. W. (1977). In *Nucleic Acid–Protein Recognition*, pp. 239–259. New York: Academic Press.

HAGERMAN, P. J. (1985). *Biochemistry*, **26**, 7033–7037.

HAGERMAN, P. J. (1986). *Nature (London)*, **321**, 449–450.

HOWARD, A. J., GILLILAND, G. L., FINZEL, B. C., POULOS, T. L., OHLENDORF, D. H. & SALEMME, F. L. (1987). *J. Appl. Cryst.* **20**, 383–387.

JAIN, S., ZON, G. & SUNDARALINGAM, M. (1989). *Biochemistry*, **28**, 2360–2364.

JONES, T. A. (1985). *Methods Enzymol.* **115**, 157–171.

KELLY, T. J. JR & SMITH, H. O. (1970). *J. Mol. Biol.* **51**, 393–409.

KOO, H. S., WU, H. M. & CROTHERS, D. F. (1986). *Nature (London)*, **320**, 501–506.

KOPKA, M. L., YOON, C., GOODSELL, D., PJURA, P. & DICKERSON, R. E. (1985). *Proc. Natl Acad. Sci. USA*, **82**, 1376–1380.

LAVERY, R. & SKLELNAR, H. (1989). *J. Biomol. Struct. Dynam.* **4**, 655–667.

LEVENE, S. D. & CROTHERS, D. M. (1983). *J. Biomol. Struct. Dynam.* **1**, 429–435.

LOMONOSSOFF, G. P., BUTLER, P. J. G. & KLUG, A. (1981). *J. Mol. Biol.* **149**, 745–760.

LUZZATI, V. (1952). *Acta Cryst.* **5**, 802–810.

READ, R. J. (1986). *Acta Cryst.* **A42**, 140–149.

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.

SUNDARALINGAM, M. (1979). *Biopolymers*, **7**, 821–860.

SUNDARALINGAM, M. (1982). In *Conformation Biology, the Festschrift Celebrating the Sixtieth Birthday of G. N. Ramachandran FRS*, edited by R. SRINIVASAN & R. H. SARMA, pp. 191–225. New York: Adenine Press.

SUNDARALINGAM, M. & SEKHARUDU, Y. C. (1988). *DNA Bending and Curvature*, edited by W. K. OLSON, M. H. SARMA, R. H. SARMA & M. SUNDARALINGAM, pp. 9–23. New York: Adenine Press.

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.

WANG, A. H.-J. & TENG, M.-K. (1990). In *Crystallographic and Modeling Methods in Molecular Design*, edited by C. E. BUGG & S. E. EALICK, pp. 123–150. New York: Springer-Verlag.

ZON, G. & THOMPSON, J. A. (1986). *Biochromatography*, **1**, 22–31.