Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of the 1:1 netropsin–decamer d(CCIICICCII)$_2$ complex with a single bound netropsin

Ke Shi, Shome Nath Mitra† and Muttaiya Sundaralingam*

Biological Macromolecular Structure Center, The Ohio State University, Departments of Chemistry and Biochemistry, 100 West 18th Avenue, Columbus, OH 43210-1002, USA

† Department of Cell Biology, Cleveland Clinic Foundation, 9500 Euclid Avenue, Cleveland, OH 44195, USA.

Correspondence e-mail:
sundaral@chemistry.ohio-state.edu

Received 5 June 2001
Accepted 29 January 2002

PDB Reference: netropsin–decamer d(CCIICICCII)$_2$ complex, DD0024.

The crystal structure of the 1:1 complex of netropsin and the B-DNA decamer d(CCIICICCII)$_2$ has been elucidated and refined to an R factor of 19.6% and an $R_{\rm free}$ of 24.7% using 1790 reflections in the resolution range 8–2.4 Å. The complex crystallizes in space group $C2$, with unit-cell parameters $a = 62.40$, $b = 24.47$, $c = 36.31$ Å, $\beta = 110.09^\circ$ and one molecule of netropsin in the asymmetric unit; the rest of the minor groove is filled with six water molecules. The structure was solved by the molecular-replacement method using the DNA model d(CCCCCIIIII)$_2$ from the 2:1 netropsin complex by removing both bound netropsins (Chen *et al.*, 1998). Surprisingly, only one netropsin molecule is found to bind to the present decamer, covering residues 2–6 at the upper stream of the duplex. The positively charged guanidinium head is hydrogen bonded through N1H$_2$ to the O$_2$ of cytosine 2 and through N10H$_2$ to N$_3$ of inosine 6. The three amide N—H groups of the peptides face the minor groove and form three sets of bifurcated hydrogen bonds with the base atoms. The central part of the drug (C3–N8) is nearly conjugated. The preference of the cytosine carbonyl O2 atoms over the inosine N3 atoms in hydrogen bonding is seen. The drug-bound region has more uniform twists, roll angles, propeller twists and minor-groove widths compared with the water-bound region.

1. Introduction

Netropsin (Fig. 1*b*), a naturally occurring antibiotic isolated from *Streptomyces netropsis* (Finlay *et al.*, 1951; Thrum & Eine, 1959), has been demonstrated to possess antitumor and antiviral activities (Zimmer & Wahnert, 1986) arising from its interaction with DNA molecules. Netropsin has been shown to inhibit polymerase and helicase reactions *in vitro* by interacting with the DNA template (Tabernero *et al.*, 1993; Brosh *et al.*, 2000). It prefers AT-rich sites and does not bind to GC sites (Wartell *et al.*, 1974). Footprinting and affinity cleavage experiments on DNA fragments have shown conclusively that netropsin binding requires at least four consecutive A·T base pairs (Van Dyke *et al.*, 1982; Fox & Waring, 1984). The first crystal structure of a drug–DNA complex was that of a 1:1 complex of the Dickerson dodecamer (CGCGAATTBr^{5}CGCG)$_2$ with netropsin (Kopka *et al.*, 1985), which showed that the drug encompasses a stretch of four consecutive A·T base pairs and that the tail of netropsin additionally interacts with the O2 of cytosine 9 of a C·G base pair. Therefore, it interacts with five base pairs in the DNA. Several crystal structures of isomorphous complexes of netropsin–d(CGCGAATTCGCG)$_2$ (Coll *et al.*, 1989; Sriram *et al.*, 1992) and netropsin–d(CGCGTTAACGCG)$_2$ (Balendiran *et al.*, 1995) have been determined since then. The same DNA

© 2002 International Union of Crystallography
Printed in Denmark – all rights reserved

sequence with numerous other minor-groove-binding drugs, Hoechst 33258 (Pjura *et al.*, 1987; Teng *et al.*, 1988), berenil (Brown *et al.*, 1990), 4′-6-diamidine-2-phenylindole (DAPI; Larsen *et al.*, 1989), pentamidine (Edwards *et al.*, 1992) and propamidine (Nunn *et al.*, 1993) have also been studied. Quantification by footprinting methods has revealed a remarkable binding selectivity; for example, netropsin has a preference for homologous A(T) sequences (*e.g.* AAAA/AATT/TTAA) over alternating sequences (*e.g.* ATAT/TATA) (Bruzik *et al.*, 1987). Despite the large number of structures determined for netropsin bound to DNA fragments, there have been few studies comparing the interaction of netropsin with different combinations of I·C (A·T) pairs in the same oligomer. Oligomers with long binding sequences have been studied in order to shed light on this question (Abu-Daya *et al.*, 1997; Chen *et al.*, 1998; Hamdan *et al.*, 1998). Here, we have designed a decamer d(CCIICICCII) that can potentially bind to two netropsin molecules in tandem. The decamer can also be viewed as containing several different combinations of sequence motifs such as the alternating and non-alternating sequences (Fig. 1*a*).

Table 1
Crystal data and refinement statistics for the netropsin–d(CCIICICCII)$_2$ complex.

Crystal system	Monoclinic
Space group	*C*2
Unit-cell parameters	
a (Å)	62.40
b (Å)	24.47
c (Å)	36.31
β (°)	110.09
Volume per base pair (Å^3)	1293
Resolution (Å)	2.4
Crystal dimensions (mm)	0.2 × 0.2 × 0.3
No. of unique reflections	1795
$R_{\rm sym}$ (%) based on *I*	3.7
Data completeness for last shell (2.5–2.4 Å) (%)	62.7
$R_{\rm work}/R_{\rm free}$ (%)	19.6/24.5
Parameter file	Param_nd.dna
R.m.s. deviation from ideal geometry	
Bond lengths (Å)	0.004
Bond angles (°)	1.5
Torsion angles (°)	18.0
Improper angles (°)	1.3

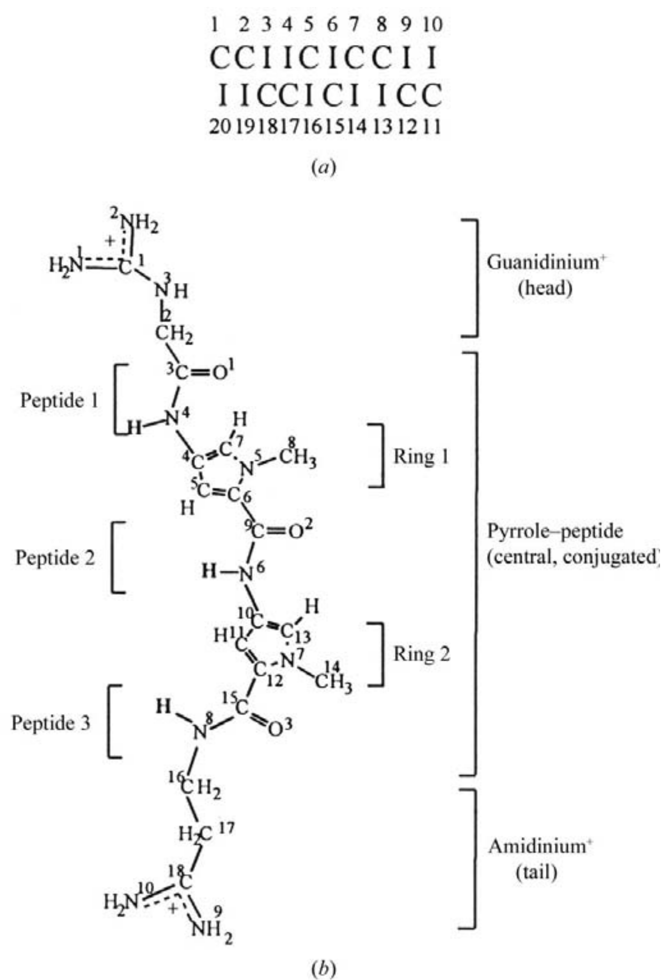

Figure 1
(*a*) The decamer DNA sequence and the numbering scheme. (*b*) Chemical structure and atom numbering of netropsin.

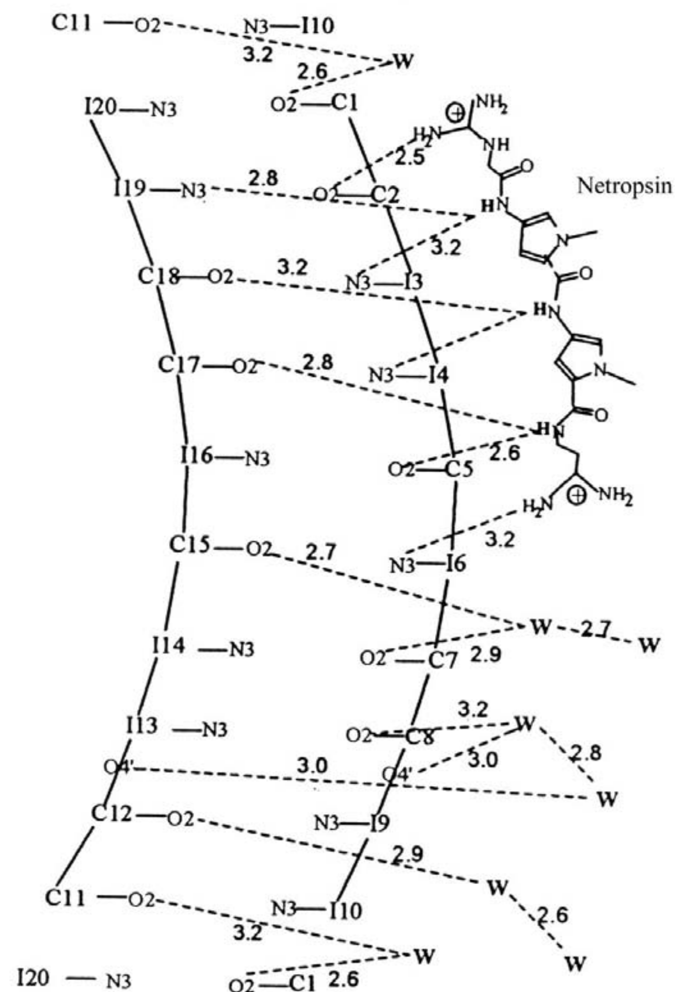

Figure 2
Hydrogen-bonding interactions between the drug, the water molecules and DNA. The three central amide groups of netropsin form bifurcated hydrogen bonds to N3/N3, O2/N3 and O2/O2, respectively. The terminal NH_2 of guanidinium and amidinium interact with O2 and N3, respectively.

2. Materials and methods

2.1. Synthesis, crystallization and data collection

The DNA decamer d(CCIICICCII) was synthesized by the phosphoramidite method on an Applied Biosystems 381 synthesizer and purified by ethanol precipitation in 2.5 *M* ammonium acetate at 248 K. The precipitate was dissolved in water and lyophilized. Crystals suitable for diffraction studies were obtained by the hanging-drop vapor-diffusion technique at 293 K. A drop containing 4 µl of the DNA in water (4 m*M* single-strand concentration), 4 µl 0.01 *M* cacodylate buffer pH 6.5, 2 µl 4 m*M* magnesium chloride, 4 µl 10 m*M* ammonium bicarbonate and 1 µl 2 m*M* netropsin on a cover slip was inverted over a well containing 500 µl 60% MPD on a Linbro tissue-culture plate. Crystals of dimensions 0.20 × 0.20 × 0.30 mm were obtained after three months. A crystal was mounted in a glass capillary sealed with mother liquor at either end. The unit-cell parameters are $a = 62.40$, $b = 24.47$, $c = 36.31$ Å, $\beta = 110.09^\circ$ and the space group is *C*2, with one molecule in the asymmetric unit. Three-dimensional intensity data were collected at 293 K on an R-AXIS IIc imaging-plate scanner at a 2θ setting of 0° with 50 kV and 100 mA graphite-monochromated Cu $K\alpha$ radiation. The crystal-to-detector distance was 100 mm. A frame width of 4° and an exposure time of 30 min were used for the φ scan. A total of 1795 unique reflections were collected to a resolution of 2.4 Å. The data completeness was 85.8%. The frames were processed using the software supplied by Molecular Structure Corporation.

2.2. Structure solution and refinement

The structure was solved by the molecular-replacement method using the program *AMoRe* (Navaza, 1994). The coordinates of the decamer alone (without the drug molecules) from the triclinic 2:1 netropsin–d(CCCCCIIIII)$_2$ complex was used as the search model (Chen *et al.*, 1998). The rotation search was carried out with an integration radius of 14 Å. The highest 11 peaks were significantly higher than the other peaks and were selected for translation searches. The *R* factor of the top peak was 48.2% using the rigid-body refinement of the FITTING subroutine of *AMoRe*; the correlation coefficient was 71.2%. This structure was refined by the rigid-body protocol of *X-PLOR* (Brünger, 1992) using 1790 reflections in the resolution range 8–2.4 Å. Mutating to the present [d(CCIICICCII)$_2$] sequence and positional and simulated-annealing refinement lowered $R_{\rm work}$ and $R_{\rm free}$ to 29.0 and 34.6%, respectively. The calculated $|F_o| - |F_c|$ electron-density map showed only one netropsin molecule, bound at the top of the sequence of the double helix in the minor groove. Positional and individual *B*-factor refinement of the netropsin complex lowered $R_{\rm work}$ and $R_{\rm free}$ to 24.4 and 28.8%, respectively. The final refinement, including 31 water molecules, lowered $R_{\rm work}$ and $R_{\rm free}$ to 19.6 and 24.7%, respectively. The model (asymmetric unit) contains the decamer duplex, one molecule of netropsin and 31 water molecules. The coordinates and structure factors have been deposited with the NDB (Berman *et al.*, 1992). Crystal data and refinement statistics are given in Table 1.

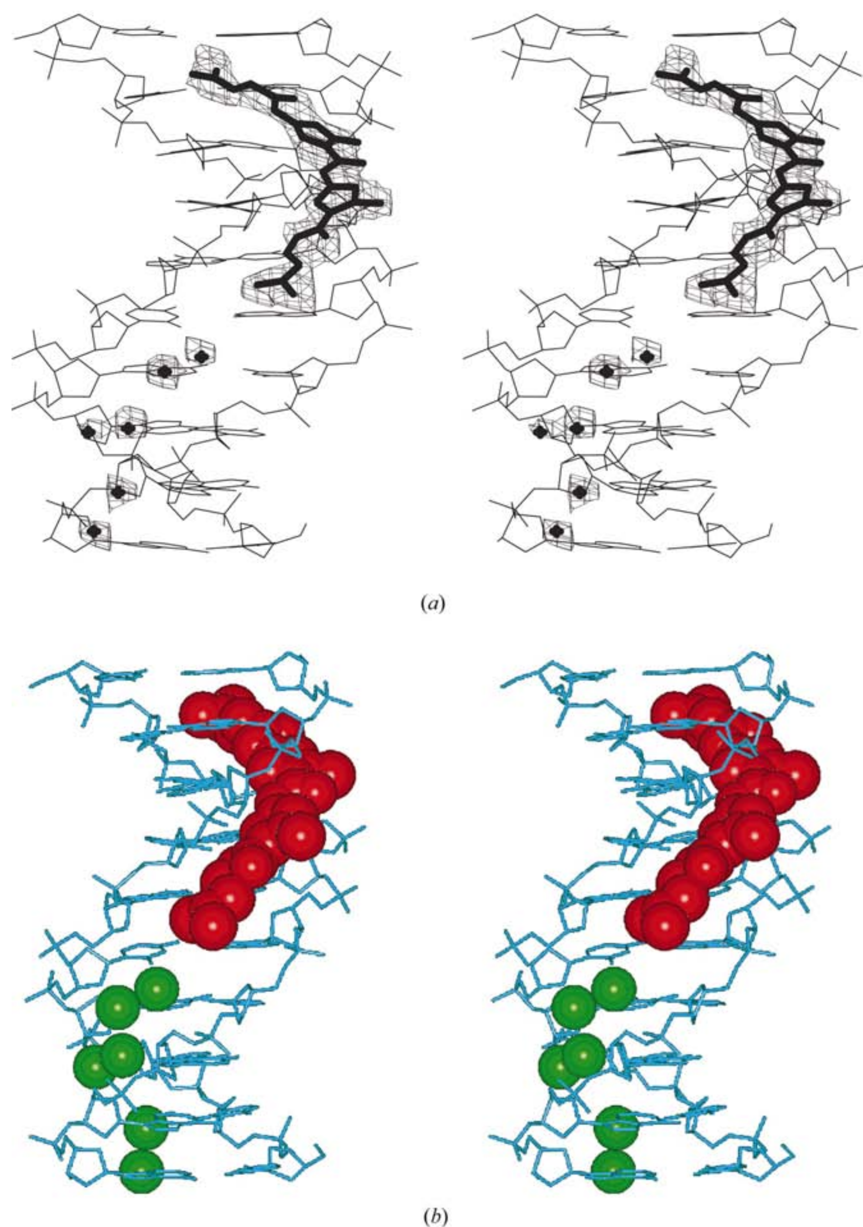

Figure 3
(*a*) Stereoview of the final $|F_o| - |F_c|$ density maps for the netropsin molecule and the minor-groove water molecules; starting contour at 3.0σ level. (*b*) Drawing of the netropsin–decamer complex. Cyan, DNA; red, netropsin; green, water.

3. Results and discussion

3.1. The novel netropsin–DNA complex

Although it appears that the entire minor groove of the present DNA decamer is accessible for two tandem netropsin molecules (Chen *et al.*, 1998), it binds only one netropsin molecule at the 5′ end of the DNA. The netropsin has slipped somewhat and the guanidinium head is positioned between cytosine 1 and cytosine 2. One of its amino groups forms a hydrogen bond with O2 of cytosine 2, while the amino group of the amidinium tail forms a hydrogen bond with O2 of cytosine 3 and N3 of inosine 6. Therefore, the rest of the decamer (CCII) is not long enough for the second netropsin molecule to bind. The crescent-shaped netropsin is isohelical with a B-DNA duplex and binds with its three amides facing the minor groove, forming three sets of bifurcated hydrogen bonds with the inosine N3 atoms and cytosine O2 atoms: N4—H with both N3 atoms of inosine 3 and inosine 19, N6—H with the N3 atom of inosine 4 and the O2 atom of cytosine 18 and N8—H with both O2 atoms of cytosine 5 and cytosine 17 (Fig. 2), similar to the bifurcated hydrogen bonds observed in earlier studies of 1:1 DNA–netropsin complexes (Kopka *et al.*, 1985). The drug fits tightly into the minor groove and was sandwiched between the sugar–phosphate chains. Large numbers of van der Waals interactions are formed between the methylpyrrole rings and the O4′ of the sugars as well as the phosphate backbone atoms. The rest of the minor groove is filled with water molecules that have roughly spherical electron densities, which distinguishes them from the netropsin electron density (Fig. 3*a*).

3.2. DNA helical parameters

A drawing of the crystal structure of the complex is shown in Fig. 3*b*. The DNA decamer adopts a typical B conformation with a broad major groove and a narrow minor groove. The parameters (Fratini *et al.*, 1982) of the DNA decamer show an average helical twist and rise of 36.6° and 3.2 Å, respectively. The number of residues per turn is 9.6. All the deoxyribose sugars have the C2′-*endo* pucker and the *anti* glycosyl

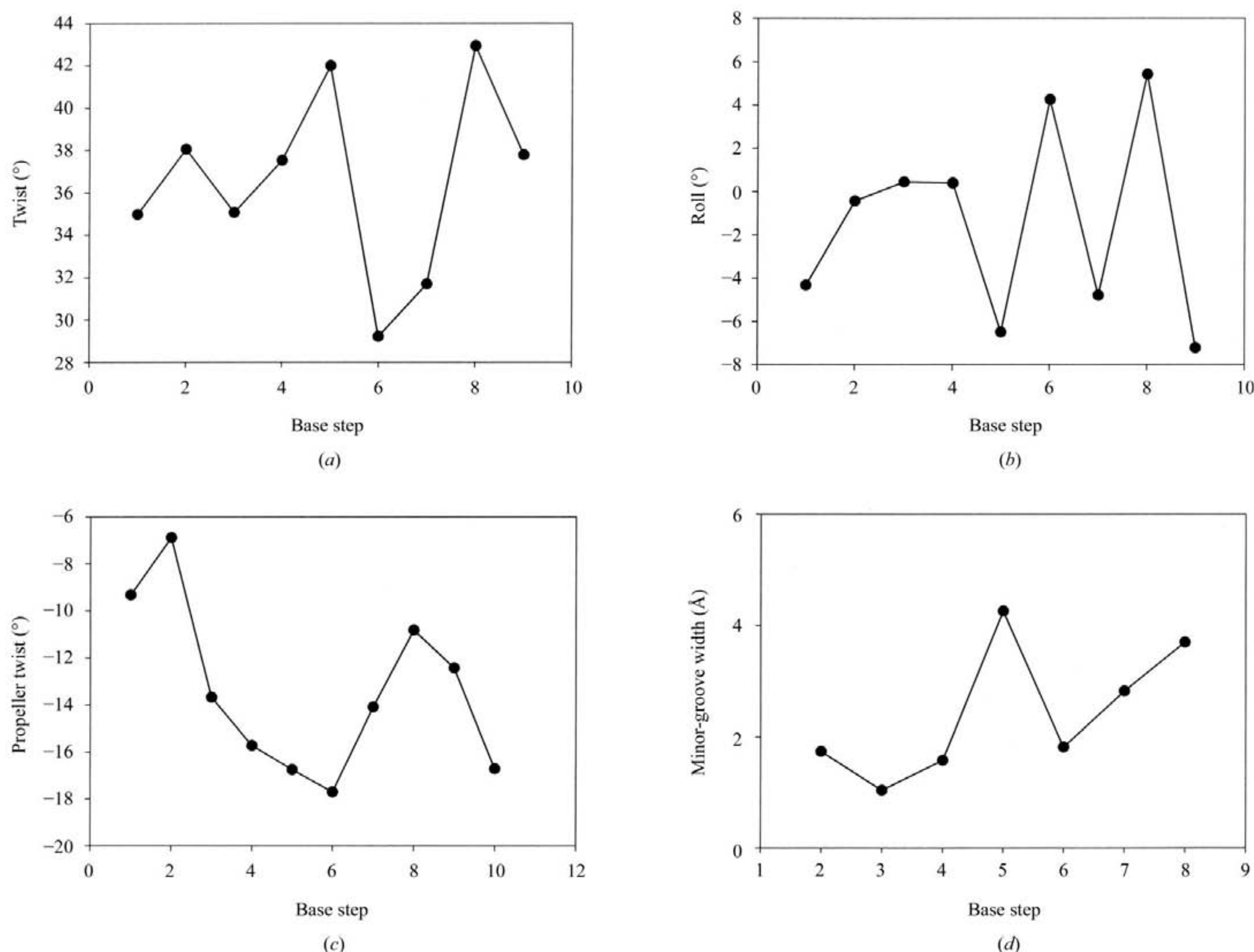

Figure 4
Base-pair twist (*a*), roll (*b*), propeller twist (*c*) and minor-groove width (*d*) at the different base steps. The values are more uniform in the netropsin-binding site and less uniform in the water region.

conformations. The present 1:1 netropsin–DNA complex enables us to compare conformational differences between the drug-bound and the water-bound portions of the decamer in the same crystal structure. Fig. 4 shows the conformational difference between the drug-bound and the water-bound parts. The roll angles in the drug-bound region, except the end base pairs, are close to 0°. The variations are more pronounced for the water-bound region. The twist angles also show similar trends. The average minor-groove width (measured by the interchain P—P distances across the narrow groove) are more uniform and slightly narrower in the drug-bound region than in the water-bound region. Comparison of different sequences studied by X-ray crystallography revealed that alternating AT regions have a wider minor groove and lower propeller twist values than the non-alternating sequences, which has been attributed to the weaker binding of the drug to the alternating sequences (Wang & Teng, 1990).

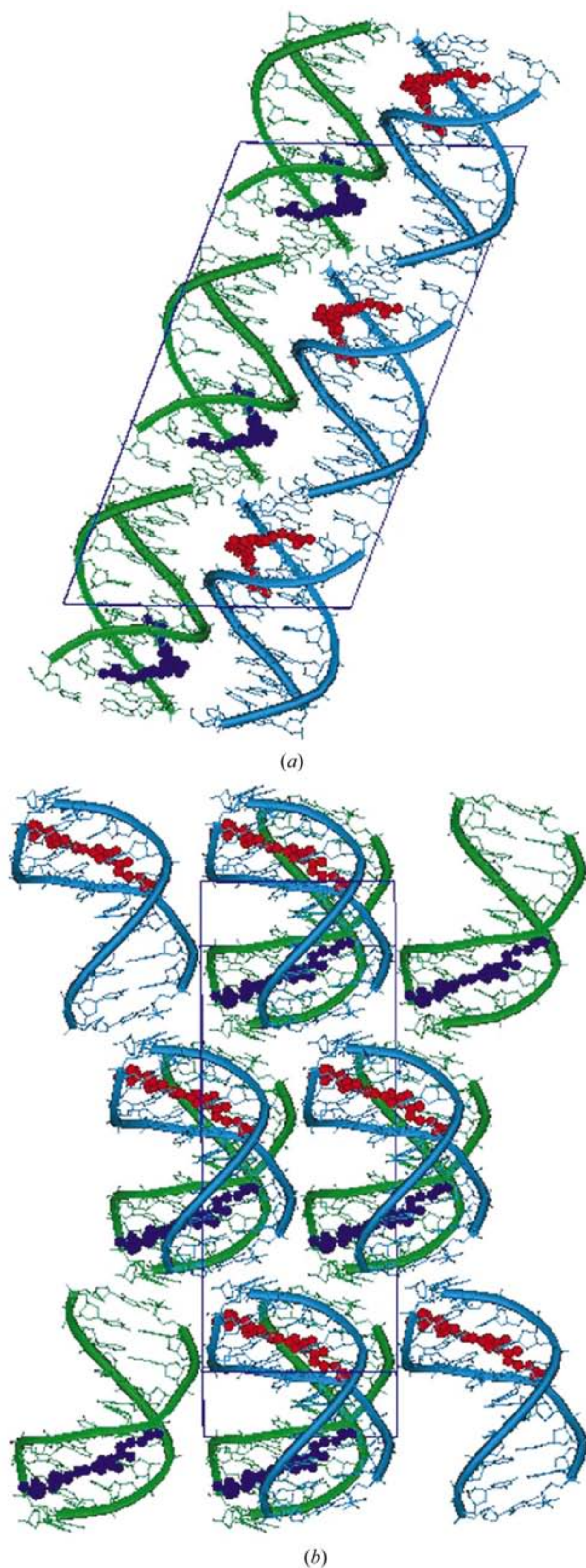

Figure 5
Crystal packing of the decamer–netropsin complexes in the unit cell.

3.3. Hydration

Most of the hydration occurs at the anionic phosphate and sugar backbone atoms and shows no regular pattern. The bottom part of the DNA decamer is filled by water molecules, while the top part is covered by the netropsin molecule. A total of 31 water molecules are in direct contact with the DNA atoms. In the drug-free region, all the carbonyl O2 atoms of cytosine bases form hydrogen-bonding interactions with minor-groove water molecules and there are no water molecules hydrogen bonded to inosine N3 atoms (Fig. 3). Interestingly, statistical analysis of previous crystal structures of 1:1 drug–DNA complexes has shown that O2 atoms of thymidines (pyrimidines) are better hydrogen-bonding sites than the N3 atoms of adenines (purines) (Wang & Teng, 1990) toward the drug peptide H—N groups. Similarly, in the present crystal structure, the O2 atoms of cytosines are shown to be better hydrogen-bonding acceptors than inosine N3 atoms. It has been shown that the I·C (C·I) base pairs in the B-DNA structure make the minor groove narrower, as do the A·T (T·A) pairs (Chen *et al.*, 1998), which promotes a single spine of water of hydration (Wang & Teng, 1990). The minor-groove water molecules lie midway between the base-pair planes and are hydrogen bonded to them. The water molecule W106 bridges the carbonyl O2 atoms of cytosine 7 and cytosine 15. The netropsin molecule does not bind at the junction (C·I) step; instead, a water molecule bridges O2 atoms of cytosine 1 and 11 (Fig. 3). In other words, all the water molecule bridges are at the I·C steps.

3.4. Crystal packing

The drug–decamer complexes are stacked in a pseudo-continuous fashion and the helical columns are inclined by 30° along the crystallographic *ac* plane (green and cyan are DNA, and red and blue are drugs in Fig. 5). The only direct interactions between the columns are intermolecular hydrogen bonds between O2P (inosine 6) and N4 (cytosine 7) of a symmetry-related molecule in the major groove and a CH···O hydrogen bond between O1P (cytosine 7) and C5′H (cytosine 7 of a symmetry-related duplex). These interactions would not pose a hinderance to the drug binding in the minor groove.

Volume per base pair is usually used to evaluate the tightness of packing of oligonucleotide crystals. The base-pair volume of the present structure is 1301 $Å^3$, which is within the range typically seen in B-DNA crystals (1100–1650 $Å^3$ per base pair).

We gratefully thank the NIH for grant GM-17378 and the Board of Regents of Ohio for an Ohio Eminent Scholar Chair and Endowment to MS. We also acknowledge the Hays Consortium Investment Fund by the Regions of Ohio for partial support for purchasing the R-AXIS IIc imaging-plate system.

References

Abu-Daya, A., Brown, P. M. & Fox, K. R. (1997). *Nucleic Acids Res.* **25**, 4962–4969.

Balendiran, K., Rao, S. T., Sekharudu, C. Y., Zon, G. & Sundaralingam, M. (1995). *Acta Cryst.* D**51**, 190–198.

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.

Brosh, R. M. Jr, Karow, J. K., White, E. J., Shaw, N. D., Hickson, I. D. & Bohr, V. A. (2000). *Nucleic Acids Res.* **28**, 2420–2430.

Brown, D. G., Sanderson, M. R., Skelly, J. V., Jenkins, T. C., Brown, T., Garman, E., Stuart, D. I. & Neidle, S. (1990). *EMBO J.* **9**, 1329–1334.

Brünger, A. T. (1992). *X-PLOR Version* 3.1. *A System for X-ray Crystallography and NMR.* Yale University, Connecticut, USA.

Bruzik, J. P., Auble, D. T. & deHaseth, P. L. (1987). *Biochemistry*, **26**, 950–956.

Chen, X., Mitra, S. N., Rao, S. T., Sekar, K. & Sundaralingam, M. (1998). *Nucleic Acids Res.* **26**, 5464–5471.

Coll, M., Aymami, J., Van Der Marel, G. A., Van Boom, J. H., Rich, A. & Wang, A. H.-J. (1989). *Biochemistry*, **28**, 310–320.

Edwards, K. J., Jenkins, T. C. & Neidle, S. (1992). *Biochemistry*, **31**, 7104–7109.

Finlay, A. C., Hochstein, F. A., Sobin, B. A. & Murphy, F. X. (1951). *J. Am. Chem. Soc.* **73**, 341–343.

Fox, K. R. & Waring, M. J. (1984). *Nucleic Acids. Res.* **12**, 9271–9285.

Fratini, A. V., Kopka, M. L., Drew, H. R. & Dickerson, R. E. (1982). *J. Biol. Chem.* **257**, 14686–14707.

Hamdan, I. I., Skellern, G. G. & Waigh, R. D. (1998). *Nucleic Acids Res.* **26**, 3053–3058.

Kopka, M. L., Yoon, C., Goodsell, D., Pjura, P. & Dickerson, R. E. (1985). *J. Mol. Biol.* **183**, 553–563.

Larsen, T., Goodsell, D. S., Cascio, D., Grzeskowiak, K. & Dickerson, R. E. (1989). *J. Biomol. Struct. Dyn.* **7**, 477–491.

Navaza, J. (1994). *Acta Cryst.* A**50**, 157–160.

Nunn, C. M., Jenkins, T. C. & Neidle, S. (1993). *Biochemistry*, **32**, 13838–13843.

Pjura, P. E., Grzeskowiak, K. & Dickerson, R. E. (1987). *J. Biol. Chem.* **197**, 257–271.

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.

Tabernero, L., Verdaguer, N., Coll, M., Fita, I., van der Marel, F. A., van Boom, J. G., Rich, A. & Aymami, J. (1993). *Biochemistry*, **32**, 8403–8410.

Teng, M.-K., Usman, N., Frederick, C. A. & Wang, A. H.-J. (1988). *Nucleic Acids Res.* **16**, 2671–2690.

Thrum, H. & Eine, N. (1959). *Naturwissenschaften*, **46**, 87.

Van Dyke, M. W., Hertzberg, R. P. & Dervan, P. B. (1982). *Proc. Natl Acad. Sci. USA*, **79**, 5470–5474.

Wang, A. H.-J. & Teng, M. (1990). *Crystallographic and Modeling Methods in Molecular Design*, edited by C. E. Bugg & S. E. Ealick, pp. 123–150. New York: Springer-Verlag.

Wartell, R. M., Larson, J. E. & Wells, R. E. (1974). *J. Biol. Chem.* **249**, 6719–6731.

Zimmer, C. & Wahnert, U. (1986). *Prog. Biophys. Mol. Biol.* **47**, 31–112.