Articles

Molecular Structure of the A-Tract DNA Dodecamer d(CGCAAATTTGCG) Complexed with the Minor Groove Binding Drug Netropsin†

Lydia Tabernero,‡ Nuria Verdaguer,‡ Miquel Coll,‡ Ignasi Fita,‡ Gijs A. van der Marel,§ Jacques H. van Boom,§ Alexander Rich,⊥ and Joan Aymamí*,‡

Grup de Química Macromolecular, CID-CSIC, i Departament d'Enginyeria Química, Universitat Politècnica de Catalunya, Diagonal 647, 08028 Barcelona, Spain, Gorlaeus Laboratory, Leiden State University, 2300 RA Leiden, The Netherlands, and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Received October 26, 1992; Revised Manuscript Received April 12, 1993

ABSTRACT: The molecular structure of the complex between the minor groove binding drug netropsin and the dodecamer d(CGCAAATTTGCG) has been solved and refined by X-ray diffraction analysis to an *R*-factor of 19.8% and 2.2-Å resolution. The drug lies in the narrow minor groove of the B-DNA fragment, covering five of the six A·T base pairs (from A5·T20 to T9·A16). The long six A·T base pair tract allows the drug to bind in a position that optimizes its contacts with the DNA, establishing hydrogen bonds with O_2 of thymines and N_3 of adenines. The DNA molecule shows a high propeller twist only at the A6·T19 step of the A-tract. Two three-centered hydrogen bonds are observed in the major groove at half of the A-tract.

Interactions between DNA and the antibiotics netropsin and distamycin have been widely studied in the past years by multiple techniques like footprinting experiments (Taylor et al., 1984; Dervan et al., 1986; Portugal & Waring, 1987), NMR (Patel, 1982; Klevit et al., 1986), FTIR (Liquier et al., 1989), and single-crystal X-ray diffraction studies (Kopka et al., 1985; Coll et al., 1987, 1989). A general conclusion from these studies is that these drugs have a binding preference for stretches of AT-rich over GC sequences (Zimmer & Wanhert, 1986) and that alternating ATAT regions appear to bind netropsin less efficiently than do continuous runs of A's or T's (Wanhert et al., 1975; Zimmer et al., 1976; Gursky et al., 1977).

These nonintercalative drugs bind in a narrowed minor groove of the B-DNA double helix through a combination of electrostatic interactions with their positive charged ends, hydrogen bonds to N_3 of adenines and O_2 of thymines at the floor of the minor groove, and van der Waals contacts with the sugar–phosphate backbone at the side wall of the groove. Netropsin is known to bind to DNA, covering about four consecutive A·T base pairs. Distamycin, being a larger molecule, needs at least five or six consecutive A·T base pairs. In all DNA–netropsin crystal structures studied up to now the DNA contained a stretch of four A·T base pairs whereas in the DNA–distamycin structure the DNA sequence contained six A·T base pairs. An open question was how one of these drug molecules would be positioned when there are longer AT sequences than strictly needed. Would the drug bind in several alternative positions shifted one base pair along the DNA or would it have a preference for a particular site?

Another relevant question that remains still open is the conformational diversity of DNA and its relation to sequence. A large number of single-crystal X-ray diffraction analyses have been done with different sequences [for a review, see Kennard and Hunter (1989)]. In particular, structural studies on DNA oligomers containing oligo[d(A)]·oligo[d(T)] sequences (or A-tracts) revealed a substantial reduction of the minor groove width and a high degree of propeller twist that favor the formation of three-centered or bifurcated hydrogen bonds in the major groove between N_6 of adenines and O_4 of two adjacent thymines in the opposite strand (Coll et al., 1987; Nelson et al., 1987; Di Gabriele et al., 1989; Aymamí et al., 1990). This kind of interaction can explain some properties of poly[d(A)]·poly[d(T)] which melts at a temperature 6 °C higher than poly[d(AT)], does not switch to an A-DNA conformation upon dehydration (Arnott & Selsing, 1974), and maintains a helical repeat of 10 base pairs per turn in solution whereas random DNA goes to 10.5 base pairs (bp) per turn (Rhodes & Klug, 1981). Furthermore, lengths of about 80–100 bp of poly[d(A)]·poly[d(T)] are partially excluded from nucleosomes (Kunkel & Martinson, 1981), and runs of about 20 bp are positioned at the end of nucleosomes (Prunell, 1982). Shorter lengths can be included into the nucleosome, preferentially at positions of minimum curvature (Satchwell et al., 1986).

Several models have been proposed for poly[d(A)]·poly[d(T)] (Arnott et al., 1983; Alexeev et al., 1987), the most recent ones including bifurcated hydrogen bonds and a narrow minor groove (Aymami et al., 1989; Lipanov et al., 1990).

Alternating AT sequences display a minimum propeller twist in two different dodecamer structures, even though the minor groove remains narrow (Coll et al., 1989; Yoon et al., 1988). On the other hand, the groove is wide in the central region of the d(CGATATATCG) decamer structure (Yuan et al., 1992). In the latter, close-packing interactions in that region might stabilize a wide groove conformation. It is unclear however what makes the minor groove narrow. Is it related

† This work was supported by CICYT Grants SAL90/824 to M.C. and PB90/605 to J. A. Subirana. A.R. was supported by the American Cancer Society. L.T. and N.V. acknowledge a MEC fellowship.

* To whom correspondence should be addressed.
‡ Universitat Politècnica de Catalunya.
§ Leiden State University.
⊥ Massachusetts Institute of Technology.

0006-2960/93/0432-8403$04.00/0

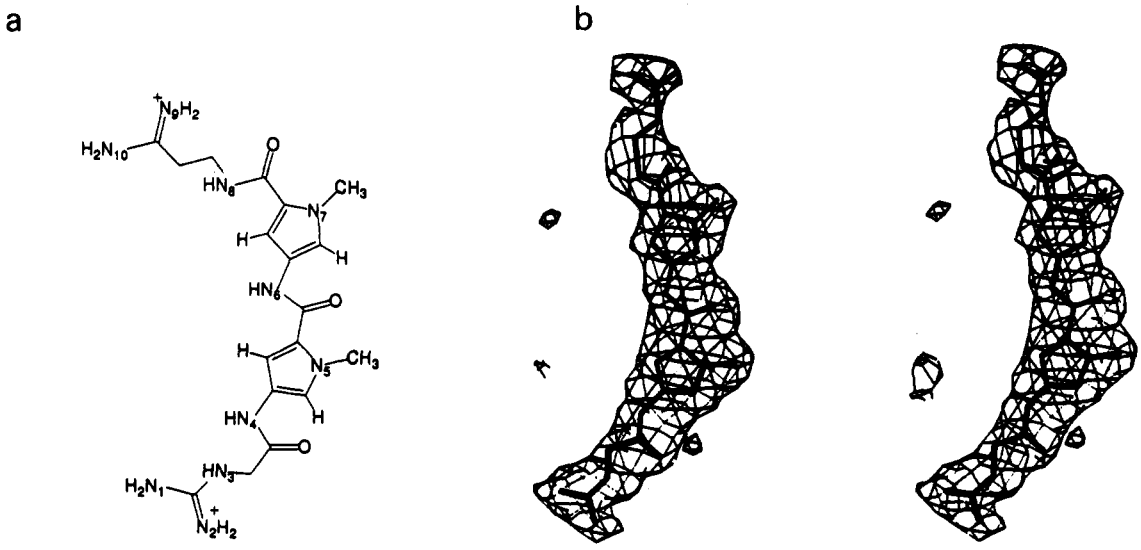

FIGURE 1: (a) Numbering scheme of netropsin. (b) Stereo difference Fourier omit map of the area where netropsin binds in the DNA minor groove calculated with coefficients $F_o - F_c$ and α_c after removing the drug molecule and refining five cycles with NUCLSQ. The drug molecule is shown fitted in the electron density.

to the propeller twist or to the presence of A·T base pairs itself? Furthermore, the variability in the propeller twist and three-centered hydrogen bonds inside the A-tract makes it difficult to establish cause–effect relationships for these particular stretches. The few known structures, most of them having similar crystal packing interactions, are clearly insufficient to determine general rules unambiguously.

We have approached these problems by solving the crystal structure of the complex between netropsin and the DNA dodecamer d(CGCAAATTTGCG), which contains a six-base pair A-tract, two more than needed for netropsin to bind. We compare this structure with the previously solved complexes of DNA with netropsin (Kopka et al., 1985; Coll et al., 1989) and distamycin (Coll et al., 1987). A detailed analysis of the DNA itself compared with related sequences is also provided.

MATERIALS AND METHODS

The DNA dodecamer d(CGCAAATTTGCG) was synthesized by an improved phosphodiester method as described previously (van der Marel et al., 1981). Crystallization of the complex was achieved in a few days from a solution containing 0.9 mM DNA dodecamer (single-strand concentration), 40 mM sodium cacodylate buffer at pH 6.5, 9 mM magnesium chloride, 1.0 mM netropsin, 0.9 mM spermine, and 8.3% 2-methyl-2,4-pentanediol (MPD), equilibrating by vapor diffusion against 50% MPD at room temperature. The complex crystallizes in the orthorhombic space group $P2_12_12_1$ with unit cell dimensions a = 25.65 Å, b = 42.03 Å, and c = 65.33 Å. A crystal measuring 0.4 × 0.5 × 1.2 mm was mounted in a sealed glass capillary with a droplet of mother liquor. Diffraction data were collected at room temperature on a CAD4 Turbo diffractometer (Enraf-Nonius) with an Elliot GX21 rotating anode X-ray generator, using an ω-scan mode. A total of 4097 reflections were collected up to 2-Å resolution. After the symmetry-related reflections were averaged with the program SHELXS-86 (Sheldrick, 1985), 2348 reflections between 15 and 2.1 Å were considered to be observable at the $2.0\sigma(F)$ level above background. Empirical absorption, Lorentz-polarization (Lp), and decay corrections were applied.

Being isomorphous with other related dodecamers, DNA atomic coordinates from the distamycin–d(CGCAAATTTGCG) complex (Coll et al., 1987) were used as a starting model. This model was first refined as a rigid body using the XPLOR package (Bruenger, 1987), and the R-factor dropped from 35.2% to 29.3%. A combined energetic and X-ray refinement followed with XPLOR, leading to an R-factor of 25%. The XPLOR parameter file was modified to strengthen the planarity of the bases by increasing the force constants for impropers and dihedrals. A first Fourier map calculated with coefficients $2F_o - F_c$ clearly showed positive electron density corresponding to the drug lying in the minor groove. After the addition of 30 water molecules outside the minor groove and a cycle of B-factor refinement, the R-factor dropped to 23%. At this stage, a difference Fourier map ($F_o - F_c$) was displayed on a Cyber 910 graphic system and the netropsin molecule was fitted into the residual electron density in the minor groove using the program TOM (Cambillau & Horjales, 1987), a modified version of FRODO (Jones, 1978).

Due to the quasi 2-fold symmetry of the netropsin molecule, there are two possible orientations for the drug in which it may interact with the DNA double helix (Coll et al., 1989). In one of them the drug seems to fit better than in the other. However, because of the resolution of the data and the possible partial occupancy of the drug molecule, the second orientation cannot be ruled out easily. Therefore, both models were refined independently, the difference in the final R-factor being only 0.2%. In the orientation with a lower final R-factor, the NH groups of netropsin made more satisfactory hydrogen bonds to the bases at the floor of the minor groove and the guanidinium end made a hydrogen bond to a nearby base. This model is the one reported in this paper. There was no evidence of disorder in the position of the netropsin molecules along the A·T stretch.

The final XPLOR refined model, including 72 solvent molecules, has an R-factor of 19.8% and root-mean-square deviations from target values of 0.021 Å for the single bond lengths and 4.1° for bond angles. No metal ions or spermine molecules could be identified.

An alternative refinement of the final stages were performed with the program NUCLSQ starting with the resulting coordinates from XPLOR and initially using data between 8- and 2.5-Å resolution. The resolution was progressively increased. At the final round, with thermal parameters refined, the R-factor was 19.4% for data between 8 and 2.1 Å ($F > 3\sigma(F)$). The final NUCLSQ model has root-mean-square deviations from target values of 0.011 Å for single bond lengths

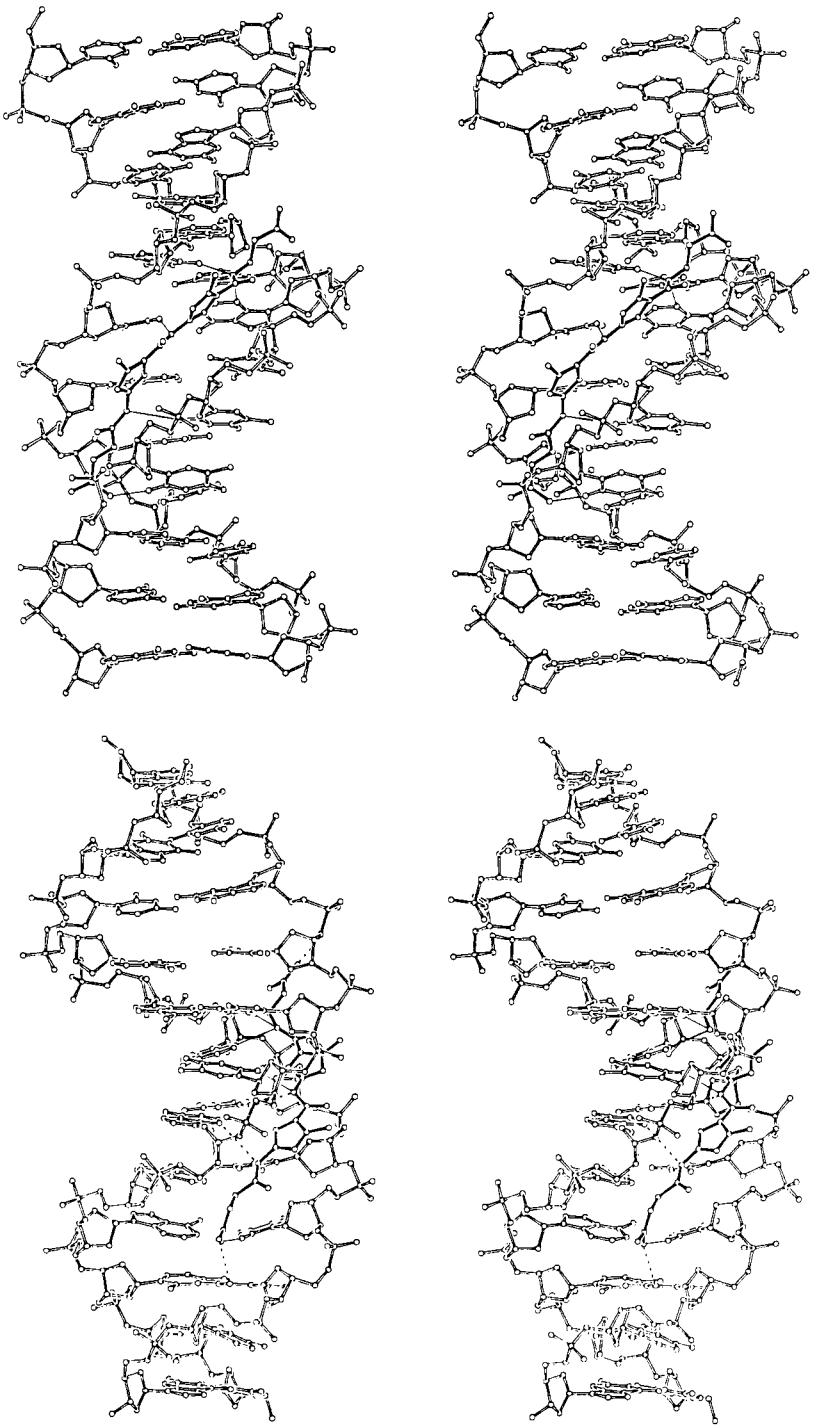

FIGURE 2: Skeletal stereo drawings of the d(CGCAAATTTGCG)–netropsin complex, with the DNA in open bonds and netropsin in filled bonds. The two views in (a, top) and (b, bottom) are 90° apart about the DNA helix axis. Thin lines indicate the H-bonds between netropsin and DNA. Dashed lines indicate weak H-bond interactions.

and 0.024 Å for one to three distances (bond angles). The coordinates from the NUCLSQ model were analyzed and compared with the XPLOR coordinates, the only significant difference being at thymine T9 where a slightly higher inclination of this base was observed in the NUCLSQ model. No differences were found in the three-center hydrogen bonds at the major groove of the DNA or in the hydrogen bonds between the drug molecule and the DNA. The final atomic coordinates have been deposited with the Brookhaven Protein Data Bank.

RESULTS AND DISCUSSION

Netropsin–DNA Interactions. Figure 1a shows the numbering scheme of netropsin. Figure 1b shows a difference Fourier omit map, with the netropsin molecule fitted in the electron density. The structure of the complex is shown in a skeletal stereo drawing in Figure 2. The general mode of interaction is similar to what has been observed in other complexes with minor groove binders. The netropsin molecule fits tightly in the narrow minor groove at the central region, covering five of the six A·T base pairs from A5·T20 to T9·A16. The drug is slightly displaced from the center of the dodecamer. The two pyrrole rings of netropsin are slightly twisted so that each ring can sit parallel to the enclosing walls of the DNA minor groove.

Amide NH groups of netropsin form bifurcated or three-centered hydrogen bonds to adenine N_3 and thymine O_2 atoms

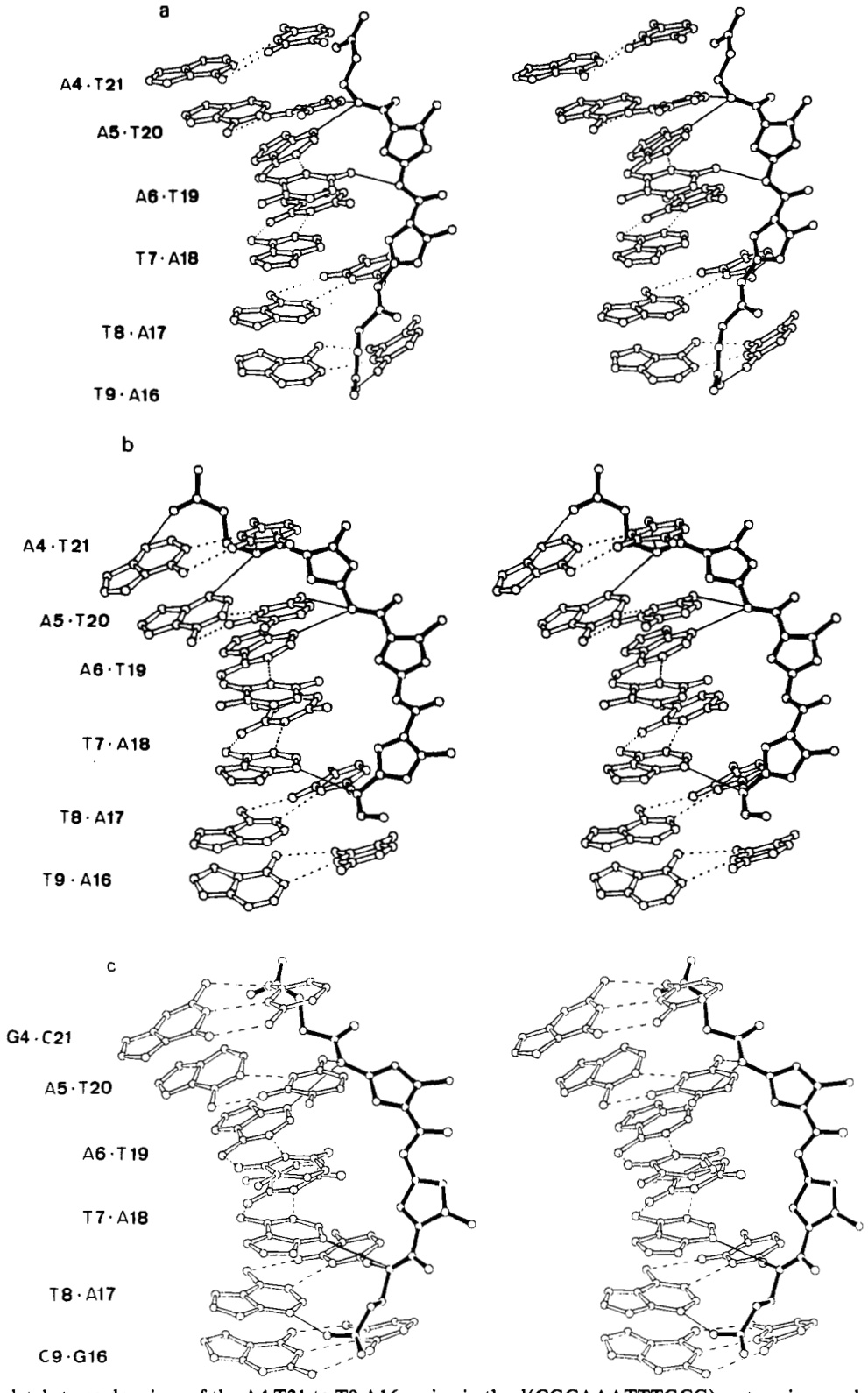

FIGURE 3: (a) Skeletal stereo drawings of the A4·T21 to T9·A16 region in the d(CGCAAATTTGCG)–netropsin complex. Only the DNA bases are shown for clarity. Hydrogen bonds (<3.5 Å) between the drug and the DNA are shown as continuous thin lines. (b) Same region in the d(CGCAAATTTGCG)–distamycin complex. (c) Same region in the d(CGCGAATTCGCG)–netropsin complex.

at the floor of the minor groove (Figure 2 and 3a) as the water molecules of the spine of hydration do (Kopka et al., 1985). Hydrogen bond distances between N_6 of netropsin and O_2 (T7) and between N4 of netropsin and N_3 (A18) emerged longer than 3.5 Å after refinement (Table I). Although having the right geometry for being part of three-centered H-bond systems, these interactions are weak. They are indicated by broken lines in Figure 2. At the guanidinium end nitrogen N_1 forms a hydrogen bond to N_3 of G10 and O_2 of T9. At the amidinium end nitrogen N9 contacts oxygen $O_{3'}$ of thymine T21 (3.2-Å distance). Thus, the amidinium group is displaced away from the floor of the minor groove, not interacting with the bases. The electron density is weak at this end of the drug, probably due to some disorder in the position of the amidinium moiety. Selected distances between DNA and netropsin atoms are listed in Table I. As observed in other

Table I: Hydrogen Bond Distances of d(CGCAAATTTGCG)–Netropsin

neotropsin atom	DNA atom	distance (Å)	neotropsin atom	DNA atom	distance (Å)
N_1	O_2 (T9)	2.5	N_8	O_2 (T20)	2.7
N_1	N_3 (G10)	3.2	N_8	N_3 (A6)	2.9
N_4	O_2 (T8)	3.4	N_3	$O_{4'}$ (A18)	3.3
N_4	N_3 (A18)	3.5	N_9	$O_{3'}$ (T21)	3.2
N_6	O_2 (T19)	2.7	N_{10}	$O_{3'}$ (T6)	3.5
N_6	O_2 (T7)	3.7			

complexes, there are close contacts between the drug atoms and the adenine C_2 atoms. If those A·T base pairs were replaced by G·C base pairs, there would be severe clashes between the N_2 amino groups of guanine and the netropsin.

In addition to hydrogen-bonding interactions, another predominant force in stabilizing the drug–DNA complex is provided by the van der Waals contacts between the netropsin backbone atoms and the $O_{4'}$ atoms of the sugars and other atoms of the sugar–phosphate backbone forming the walls of the narrow minor groove.

When the structure of the complex d(CGCAAATTTGCG)–netropsin is compared with previously reported complexes of the same drug (Kopka et al., 1985; Coll et al., 1989), several differences are apparent. The addition of two A·T base pairs in the binding region might favor the better disposition of the netropsin molecule, exploiting the structural complementarity between the crescent shape of the drug and the natural curvature of the DNA helix. Although the netropsin molecule could occupy several positions along the six A·T base pairs, no sign of positional disorder is observed in the electron density maps. The drug chooses a central position which must be one of the possible structures with lower energy. For example, in this low-energy complex almost all possible hydrogen bonds between the drug and the DNA are satisfied. This is achieved not only by the positioning of the drug but also by the capability of the A·T bases in adopting a high propeller twist conformation, as occurs at base pair A6·T19, where a propeller twist of −25.7° is observed. Compared with distamycin complexed to the same sequence, netropsin occupies a more central position in the A-tract (Figure 2b). The position of the drug is also central in the d(CGCGAATTCGCG)–netropsin complex (Kopka et al., 1985) although the orientation of the drug molecule is reversed as can be seen when parts a and c of Figure 3 are compared.

DNA Conformation. The overall feature of the DNA dodecamer d(CGCAAATTTGCG) in this complex is similar to that of the previously reported complex with distamycin (Coll et al., 1987). Figure 4 shows a superposition of both structures. They differ by a *rms* of 0.43 Å. Several DNA dodecamers containing a central AT-rich region and crystallized isomorphously share the same characteristics (Drew et al., 1980; Coll et al., 1987; Nelson et al., 1987; Di Gabriele et al., 1989; Yoon et al., 1988; Edwards et al., 1992). All of them belong to the B-DNA family, in contrast to a recent dodecamer structure, d(CCCCCGCGGGGG), crystallized in a different lattice and found to be in the A-DNA conformation (Verdaguer et al., 1991). Most B-type dodecamers have a narrow minor groove in their central A·T segments. Among these oligonucleotide structures, the ones having continuous runs of adenines—or A-tracts—display a high degree of propeller twist in the A·T base pairs. This large propeller twist is sometimes associated with the formation of three-centered or bifurcated hydrogen bonds in the major groove.

Table II shows a comparison of propeller twists in different dodecamers. In the present structure one of the central A·T base pairs, A6·T19, has the highest propeller of all dodecamers, −25.7°. Compared with the same sequence in the distamycin complex, the propeller twist in this base pair is about 5° higher. However, the propeller twist is low in all other base pairs of the A-tract. The mean propeller twist in the AT region is −13°, a value that is lower than in all other A-tract dodecamers, the variability being higher in the present structure. This can be rationalized in terms of a flexible adaptation of the DNA and of the drug molecule in order to optimize its interactions. In general an A-tract will display a high propeller twist, but the precise value of this parameter can vary according to different environmental factors, including the nature of adjacent sequences (Yanagi et al., 1991).

When looking at Figures 3a and 5, one can see the very high propeller twist of the A6·T19 base pair. Hydrogen bonds of N_3 (A6) and O_2 (T19) to different amide nitrogens of netropsin clearly stabilize this conformation. Here the drug is pulling apart the two bases of the same base pair. Additionally, a three-centered H-bond from O_4 of T19 to N_6 of A6 and N_6 of A5 in the major groove contributes to stabilize the twisting. Another three-centered H-bond, adjacent to the one previously mentioned, between O_4 of T20 and N_6 of A5 and N_6 of A4 is observed in the major groove. The propeller twist is −13.2° at base pair A5·T20, whereas it is only −1.6° at base pair A4·T21. However, A4 and T21 are not in the same plane because of their pronounced stagger. This effect places N_6 of A4 close enough to O_4 of T20 to establish a H-bonding interaction. Therefore, propeller twisting is not the only deformation of the base pairing leading to bifurcated hydrogen bonds in the major groove.

Two crystal structures with the d(CGCAAATTTGCG) sequence have been reported recently. One is the redetermination of the native dodecamer (Edwards et al., 1992) which shows only one possible three-centered hydrogen bond in the DNA major groove at nitrogen N_6 of A5. Major groove cross distances between N_6 of adenines and O_4 of thymines on the next base pair were already observed to be longer in the d(CGCAAATTTGCG) structure than in its complex with distamycin (Coll et al., 1987). Another reported structure is the complex with the drug berenil (Brown et al., 1992). In this case, the DNA is very similar to the native dodecamer and the three-centered hydrogen bond practically disappears.

With four different structures of the d(CGCAAATTTGCG) dodecamer it is possible now to analyze the influence of drug binding in this DNA sequence. It seems that netropsin and distamycin have the effect of increasing the propeller twist at some of the A·T base pairs, bringing the N_6 atom of adenine closer to the O_4 of thymine on the next base pair. The effect is not so strong with berenil, a drug that only makes hydrogen bonds to the DNA with its two amidinium ends. This is probably a softer binding mode than in the case of netropsin and distamycin where, in addition to the amide, amidinium, and guanidinium hydrogen bonds along the drug length, the pyrrole rings may contribute to the propeller by pushing apart the A and T bases of the same base pair.

Another important deformation from canonical base pairing is observed at base pair T9·A16 having a buckle of 21.9° (Figures 3a and 5). Here again the drug molecule could be responsible for the unusual conformation since a strong H-bond (2.6 Å) is present. However, we have to caution that T9 was the only base having a slightly different orientation in the NUCLSQ and the XPLOR refinements. Even though in both models the buckle was very high, some degree of disorder

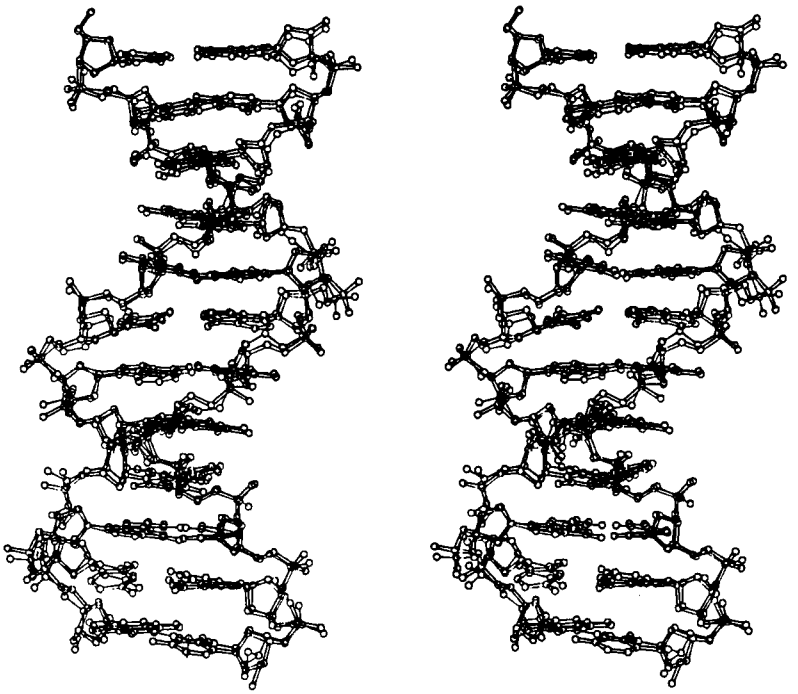

FIGURE 4: Skeletal stereo drawings showing the superposition of the d(CGCAAATTTGCG) dodecamers from the complex with netropsin (open bonds) and the complex with distamycin (thin bonds).

Table II [a]

	propeller twist (deg)				
base pair	A_3T_3 net[b]	A_3T_3 dis[c]	A_2T_2 net[d]	A_2T_2[e]	ATAT net[f]
1·24	−7.8	−7.2	−16.0	−13.0	−5.1
2·23	−22.1	−13.2	−14.1	−10.8	−13.2
3·22	−4.9	−18.0	−7.4	−3.7	−7.1
4·21	**−1.6**	**−13.8**	−6.7	−10.4	−8.6
5·20	**−13.1**	**−16.8**	**−20.0**	**−16.2**	**−1.9**
6·19	**−25.6**	**−20.1**	**−19.1**	**−17.5**	**−14.0**
7·18	**−11.1**	**−19.4**	**−23.0**	**−17.0**	**−15.6**
8·17	**−13.2**	**−16.7**	**−21.7**	**−17.0**	**−7.2**
9·16	**−16.0**	**−23.6**	−12.7	−16.0	−11.7
10·15	−14.9	−3.7	−5.1	−4.7	−10.9
11·14	−21.0	−10.4	−20.0	−16.9	−15.7
12·13	−0.1	−13.6	−6.5	−2.4	−1.9
mean A.T	−13.2	−18.4	−21.0	−16.9	−9.7
mean C.G	−11.8	−11.0	−11.1	−9.7	−9.3

[a] All parameters calculated with NEWHEL90 written by R. Dickerson. [b] d(CGCAAATTTGCG)–netropsin (this work). [c] d(CGCAAATTTGCG)–distamycin (Coll et al., 1987). [d] d(CGCGAATTCGCG)–netropsin (Kopka et al., 1985). [e] d(CGCGAATTCGCG) (Drew et al., 1980). [f] d(CGCGATATCGCG)–netropsin (Coll et al., 1988).

Table III: Helix Parameters for the d(CGCAAATTTGCG)–Netropsin Complex[a]

base step	roll (deg)	tilt (deg)	twist (Å)	rise (Å)	pr tw (deg)	buckle (deg)
C1·G24	1.7	−0.3	40.6	3.2	−7.8	7.2
C2·G23	1.4	−1.0	37.9	3.2	−22.1	5.1
C3·G22	5.3	6.8	28.3	4.1	−4.9	2.9
A4·T21	3.7	−5.4	32.5	3.1	−1.6	2.3
A5·T20	−4.9	−6.9	40.5	3.3	−13.1	8.3
A6·T19	−1.4	1.4	31.7	3.0	−25.6	−14.1
T7·A18	3.1	1.2	37.5	3.1	−11.1	2.6
T8·A17	0.5	2.6	39.4	3.2	−13.2	13.4
T9·A16	−6.6	−1.6	34.0	3.3	−16.0	21.9
G10·C15	−7.5	2.1	34.3	2.8	−14.9	6.3
C11·G14	0.4	3.2	41.2	3.9	−21.0	−5.1
C12·G13					−0.1	−16.6

[a] All parameters were calculated with NEWHEL90 written by R. Dickerson.

cannot be excluded since the O_2 atom had very weak density in the $2F_o - F_c$ map. In the d(CGCAAATTTGCG)–distamycin complex (Coll et al., 1987) the buckle at this level is 12.5°. The difference in conformation with the netropsin complex is apparent when parts a and b of Figure 2 are compared. In the distamycin complex the drug does not cover base pair T9·A16; although it is a longer molecule, it is displaced from the center of the dodecamer. Instead a water molecule was observed in this location, a remainder of the spine of hydration in the minor groove. Furthermore, this base pair shows a propeller twist of −23.6° in the distamycin complex but only −14.9° in the present structure. It seems as if in the latter N_1 of netropsin is "pulling" O_2 of T9, causing the original propeller to decrease in favor of a higher buckle.

Figure 6 shows a polar projection plot of the DNA backbone. Minimum distances across the minor groove between backbone atoms of both strands are indicated. This is a finer measure of the groove width than phosphate–phosphate distances. The general trend is a narrower minor groove in the AT segment compared with the CG segment. Because of the helical nature of DNA, shortest distances across the minor groove are not between atoms of paired residues but between atoms of a residue in one strand and atoms of residues on the other strand two or three base pairs away in the 3′ direction. For example, two residues of the CG segment at the 3′ end of each strand (G10, C11 and G22, C23) have shortest distances across the groove with A17, A18 and A5, A4. There is a short distance of 7.2 Å between $O_{4'}$ of C11 and $O_{4'}$ of A17. Therefore, the narrow minor groove in the central region extends further at the lower end of the double helix in Figure 6. Although phosphate C11 "opens" the minor groove, the sugar moiety of C11 and phosphate G12 "close" the groove. This peculiarity was observed in other dodecamer structures, i.e., when the P–P distances across the minor groove were plotted. It seems appropriate to note here the breakdown of the 2-fold symmetry of the double helix even though the DNA sequence is palindromic.

Conclusions. About 100 DNA oligonucleotide structures have been determined up to now. Many of the structures adopting the B-DNA form are dodecamers, packing in an

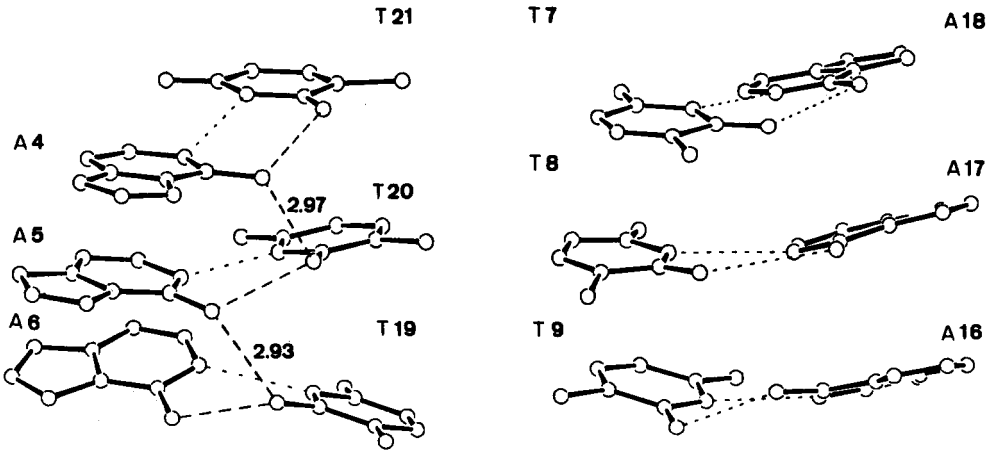

FIGURE 5: Diagrams illustrating base pairing in the A-tract region of the dodecamer. Three-centered or bifurcated H-bonds are observed only in three base pairs.

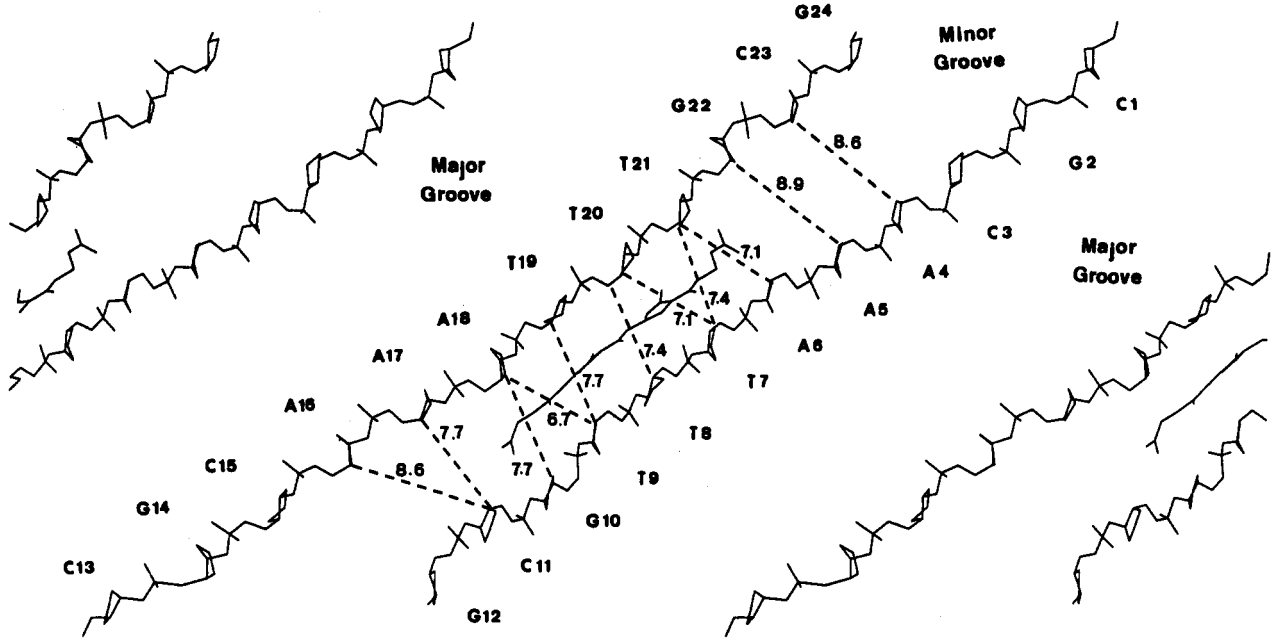

FIGURE 6: Cylindrical projection of the d(CGCAAATTTGCG)–netropsin complex. DNA bases are not plotted. The shortest distances between atoms across the minor groove are indicated by dotted lines and marked in angstroms. At the central A-tract, where netropsin lies, the groove is clearly narrower.

orthorhombic lattice with two C·G base pair overlaps. All these dodecamers, like the one described here, exhibit a similar trend in their conformation. Part of it might be attributed to crystal packing forces, in particular at both ends of the oligonucleotide. Is the narrow minor groove in the midst of the dodecamers also due to a packing effect? The fact is that all B-DNA dodecamers, but one, have similar sequences, with CG's at the ends and AT's in the middle. The exception is the oligonucleotide d(CGCAAGCTGGCG) which crystallizes in the same lattice not having a narrow minor groove in the middle (Webster et al., 1990). Therefore, DNA dodecamers without a narrow minor groove are allowed to pack in this lattice.

Even though having a similar general trend, the various dodecamer sequences show many dissimilarities in the details. The most remarkable one is the high propeller twist of the A·T base pairs when there is an A-tract in the middle of the dodecamer, compared with alternating AT-tracts which tolerate only low propeller twist values. The present structure indicates that this is indeed a possible characteristic of A·T base pairs in the A-tract but not a strict general rule. Out of six A·T base pairs only one shows a high propeller.

The base pair propeller twist appears to be enhanced by the binding of netropsin in step A6·T19, but diminished at step T9·A16. It is interesting however that the mean propeller twist in the A-tract is −13.3°, unlike the other A-tract oligonucleotide structures which have higher values. It seems that A·T base pairs are quite plastic when they are part of an A-tract. They tolerate a high degree of propeller twist and buckle. When the propeller twist is high enough in consecutive base pairs, three-centered hydrogen bonds occur in the major groove between consecutive base pairs, although this is not the only way to achieve these kinds of bonds.

Long runs of adenines most probably cooperatively accentuate some of these characteristics. A model for poly[d(A)]·poly[d(T)] based on fiber diffraction patterns proposes a propeller twist of −29° and continuous three-centered hydrogen bonds in the major groove (Aymami et al., 1989). We have seen that both features vary considerably from one base pair to another inside the A-tract of the dodecamer structure described here.

The plasticity of the A·T base pair in the A-tract in terms of the inclination of the individual base planes (both propeller and buckle) might be exploited by DNA binding molecules

such as drugs and proteins, in order to optimize their interactions. This is the case in the DNA–phage 434 repressor complex where large values of propeller twist and buckle have been described (Aggarwal et al., 1988). The netropsin molecule takes advantage of this plasticity as well. Because the six A·T run is long enough, the drug can find a position where interactions with high propeller twisted and buckled bases are optimized.

ACKNOWLEDGMENT

We thank Dr. Jordi Bella for helping with the plotting programs and Prof. J. A. Subirana for his generous support.

REFERENCES

Alexeev, D. G., Lipanov, A. A., & Skuratovskii, I. Ya. (1987) *Nature 325*, 821–823.

Aggarwal, A. K., Rodgers, D. W., Drottar, M., Ptashne, M., & Harrison, S. C. (1988) *Science 242*, 899–907.

Arnott, S., & Selsing, E. (1974) *J. Mol. Biol. 88*, 509–521.

Arnott, S., Chandrasekaran, R., Holl, I. H., & Puigjaner, L. C. (1983) *Nucleic Acids Res. 11*, 4141–4155.

Aymami, J., Coll. M., van der Marel, G. A., van Boom, J. H., Wang, A. H-J., & Rich, A. (1989) *Nucleic Acids Res. 17*, 3229–3245.

Aymami, J., Coll, M., Frederick, C., Wang, A. H-J., & Rich, A. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 2526–2530.

Brown, D. G., Sanderson, M. R., Garman, E., & Neidle, S. (1992) *J. Mol. Biol. 226*, 481–490.

Berman, H. M., Neidle, S., Zimmer, C., & Thrum, H. (1979) *Biochem. Biophys. Acta 561*, 124–131.

Bruenger, A. T. (1987) *XPLOR. A System for X-ray Crystallography and NMR*, Yale University Press, New Haven, CT.

Cambillau, C., & Horjales, E. (1987) *J. Mol. Graphics 5*, 174–177.

Carrondo, M. A., Coll, M., Aymamí, J., Wang, A. H-J., van der Marel, G. A., van Boom, J. H., & Rich, A. (1989) *Biochemistry 28*, 7849–7859.

Coll, M., Frederick, C. A., Wang, A. H-J., & Rich, A. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8385–8389.

Coll, M., Aymamí, 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., Kopka, M. L., & Pjura, D. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 7099–7103.

DiGabriele, A. D., Sanderson, M. R., & Steitz, T. A. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 1816–1820.

Drew, H., Takan, T., Tanaka, S., Takura, K. K., & Dickerson, R. E. (1980) *Nature 286*, 567–573.

Edwards, K. J., Brown, D. G., Spink, N., Skelly, J. V., & Neidle, S. (1992) *J. Mol. Biol. 226*, 1161–1173.

Gursky, G. V., Tumanyan, V. G., Zasedatelev, A. S., Zhuge, A. L., Grokhovsky, S. L., & Gottikh, B. P. (1977) in *Nucleic Acid-Protein Recognition* (Vogel, H. J., Ed.) pp 189–217, Academic Press, New York.

Hendrickson, W. A., & Konnert, J. H. (1981) in *Biomolecular Structure, Function and Evolution* (Srinivasan, R., Ed.) Vol. 1, pp 43–57, Pergamon, Oxford.

Jones, T. A. (1978) *J. Appl. Crystallogr. 11*, 268–272.

Kennard, O., & Hunter, W. N. (1989) *Q. Rev. Biophys. 22*, 327–379.

Klevit, R. E., Wemmer, D. E., & Reid, B. R. (1986) *Biochemistry 25*, 3296–3303.

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

Kunkel, G. R., & Martinson, H. G. (1981) *Nucleic Acids Res. 9*, 6869–6888.

Lipanov, A. A., Chuprina, V. P., Alexeev, D. G., & Skuratovskii, I. Ya. (1990) *J. Biomol. Struct. Dyn. 7*, 811–826.

Liquier, J., Mchami, A., & Taillandier, E. (1989) *J. Biomol. Struct. Dyn. 7*, 119–126.

Nelson, H. C. M., Finch, J. T., Luisi, B. F., & Klug, A. (1987) *Nature 330*, 221–226.

Patel, D. J. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 6424–6428.

Portugal, J., & Waring, M. J. (1987) *Eur. J. Biochem. 167*, 281–289.

Prunell, A. (1982) *EMBO J. 1*, 173–179.

Rhodes, D., & Klug, A. (1981) *Nature 292*, 378–380.

Satchwell, S. C., Drew, H. R., & Travers, A. A. (1986) *J. Mol. Biol. 191*, 659–675.

Sheldrick, G. M. (1985) in *Crystallographic Computing 3* (Sheldrick, G. M., Krueger, C., & Goddard, R., Eds.) pp 175–189, Oxford University Press, Oxford.

Taylor, J. S., Schultz, P. G., & Dervan, P. B. (1984) *Tetrahedron 40*, 457–465.

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

Van der Marel, G. A., van Boeckel, C. A. A., Willie, G., & van Boom, J. H. (1981) *Tetrahedron Lett. 22*, 3887–3890.

Verdaguer, N., Aymamí, J., Fernández-Forner, D., Fita, I., Coll, M., Huynh-Dinh, T., Igolen, J., & Subirana, J. (1991) *J. Mol. Biol. 221*, 623–635.

Wähnert, V., Zimmer, C., Luck, G., & Pitra, C. (1975) *Nucleic Acids Res. 2*, 391–404.

Webster, G. D., Sanderson, M. R., Skelly, J. V., Neidle, S., Swann, P. F., Li, B. F., & Tickle, I. J. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 6693–6697.

Yanagi, K., Privé, G. G., Dickerson, R. E. (1991) *J. Mol. Biol. 217*, 201–214.

Yoon, C., Prive, G. G., Goodsell, D. S., & Dickerson, R. E. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 6332–6336.

Yuan, H., Quintana, J., & Dickerson, R. E. (1992) *Biochemistry 31*, 8009–8021.

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

Zimmer, C., Luck, G., & Fric, I. (1976) *Nucleic Acids Res. 3*, 1521–1532.