3458–3466 *Nucleic Acids Research, 1996, Vol. 24, No. 17*

Hydrogen bond geometry in DNA–minor groove binding drug complexes

Lydia Tabernero*, Jordi Bella and Carlos Alemán[1]

Department of Biological Sciences, Lilly Hall of Life Sciences, Purdue University, West Lafayette, IN 47907, USA and [1]Departament d'Enginyeria Química, ETSEIB, Universitat Politècnica de Catalunya, Diagonal 647, E-08028 Barcelona, Spain

Received February 8, 1996; Revised and Accepted July 2, 1996

ABSTRACT

The geometry of the hydrogen bonding interaction between DNA and minor-groove binding drugs has been analyzed from a sample of 22 crystal structures of DNA–drug complexes, retrieved from the Nucleic Acid Database. Seventy-seven interactions between the drugs and acceptor groups in the nucleotide bases can be classified as hydrogen bonds. Their geometry departs significantly from linearity since, in most instances, the interactions can be described as three-center or multiple hydrogen bonds. Results also show that there is no preference for hydrogen bonds involving positively charged groups in the drugs. Relationships between hydrogen bond geometry and positioning of the drug along the minor groove are also discussed. The information presented may be useful in the design of new specific minor groove binding drugs.

INTRODUCTION

Netropsin, distamycin, Hoechst 33258 or berenil are examples of low molecular weight compounds with antibiotic, antiviral and antitumour activities, which are known to bind specifically to the DNA minor groove. These non-intercalating drugs have a binding preference for stretches of AT-rich over GC sequences (1). Different experimental and theoretical analyses have brought a large amount of information about sequence specificity, binding energies and stability of the several DNA–drug complexes (2–15).

From the very beginning it was postulated that minor groove binding drugs as netropsin could recognize specific DNA sequences by selective hydrogen bonds to the DNA bases (16). X-ray crystallographic studies of oligonucleotide–drug complexes showed soon that the mode of interaction between drugs and DNA could be much more subtle (8). The typical drug in this category is a flat, crescent-shaped molecule, which accommodates itself into the minor groove of the DNA double helix establishing a complex interaction that simultaneously involves electrostatic, hydrogen bonding and van der Waals effects. The relative significance of these terms for the stability of the DNA–drug complexes is still a matter of discussion, but hydrogen bonding between the drug and the DNA bases seems to keep its role as a main responsible for the observed specificity.

Many X-ray structures of different DNA sequences complexed with minor groove binders have been reported in the past years, and hydrogen bonding interactions have been described in all of them (Table 1; see also ref. 17 for a recent review). A comparative analysis of the different observed interactions may provide valuable information for a better understanding of the binding mechanism and specificity.

We present here a detailed analysis of the hydrogen bonding geometries found in the crystal structures of several DNA–minor groove drug complexes, whose coordinates have been retrieved from the Nucleic Acid Database (18). Along with the obvious classification by acceptor types, we have analyzed the geometry of the interaction based on the type of donor group involved. We will discuss the high correlation observed between drug positioning into the minor group and hydrogen bonding geometry. Finally, previous theoretical results (19) seem to suggest that the positively charged donor groups may have a greater stabilizing role than the neutral ones, and therefore our analysis will also extend into the search for a structural correlation with these theoretical results.

MATERIALS AND METHODS

The crystal structures

Coordinates for the crystal structures of oligonucleotide–minor groove binder complexes were retrieved from the Nucleic Acid Database (NDB) [(18), release of December 1994]. Table 1 lists the NDB entries used in this paper.

Classification of hydrogen bonding groups

In all the structures analyzed in this paper, a basic asymmetry arises from the fact that the drug provides the donor hydrogen bonding groups, whereas the oligonucleotide molecule participates with its acceptor groups located at the bottom and walls of the DNA minor groove. To compare hydrogen bonding geometries of charged with non-charged groups, we classified the drug donor groups into three categories: Type A, in which the NH_2 donor group is part of an amidinium or guanidinium moiety; Type B, in which the NH donor group is part of an amide moiety; and Type C, in which the donor NH group is part of a benzimidazole ring. Type A donors, when present, are located at the ends of the drug, with formal charge +1. Type B and C groups are usually internal in the drug molecule and have formal charge zero.

* To whom correspondence should be addressed

Table 1. Crystallographic entries from the NDB (18) used in this work

NDB code	DNA sequence–drug name	R value	Max. res.	Reference
gdl001	CGCGATATCGCG–Netropsin	20.2	2.40	(30)
gdl002	CGCGAATTCGCG–Hoechst 33258	15.7	2.25	(31)
gdl003	CGCAAATTTGCG–Distamycin	20.2	2.20	(25)
gdl004	CGCGATATCGCG–Netropsin	20.1	2.40	(30)
gdlb05	CGCGAATT^{Br}CGCG–Netropsin	21.1	2.21	(8)
gdl006	CGCGAATTCGCG–Hoechst 33258	14.0	2.20	(32)
gdl008	CGCGAATTCGCG–DAPI	21.5	2.40	(33)
gdl009	CGCGAATTCGCG–Berenil	16.7	2.50	(20)
gdl010	CGCGAATTCGCG–Hoechst 33258	15.7	2.00	(26)
gdl011	CGCGAATTCGCG–Hoechst 33258	15.7	2.00	(26)
gdl012	CGCGAATTCGCG–Hoechst 33258	15.2	1.90	(26)
gdl013	CGCGAATTCGCG–Hoechst 33258	14.9	2.00	(26)
gdl014	CGCAAATTTGCG–Netropsin	19.8	2.20	(27)
gdl015	CGCGAATTCGCG–Pentamidine	19.4	2.10	(34)
gdl016	CGCAAATTTGCG–Berenil	18.3	2.00	(35)
gdlb17	CGC$^{(E)}$GAATTCGCG–Netropsin	15.6	2.50	(36)
gdl018	CGCGAATTCGCG–Netropsin	16.4	2.20	(36)
gdlb19	CGC$^{(E)}$GAATTCGCG–Hoechst 33258	14.5	2.50	(37)
gdlb20	CGC$^{(E)}$GAATTCGCG–Hoechst 33342	15.7	2.50	(37)
gdl021	CGCGAATTCGCG–Hoechst 33342	16.8	2.25	(37)
gdl022	CGCGAATTCGCG–Hoechst 33258	17.2	2.00	(37)
gdl023	CGCGAATTCGCG–Propamidine	17.4	2.10	(38)

On the DNA side, we considered two major acceptor types: carbonyl groups and aromatic nitrogen atoms, since they impose different geometry requirements on the hydrogen bonding interaction. We also included less frequent interactions involving O′ atoms from deoxyribose rings, but we did not consider solvent-mediated interactions in this study. Water bridges are not a common feature in most of the structures, although they may be of some importance in helping the binding of drugs like berenil to DNA (20).

Most of the hydrogen bonding groups discussed above are also present in proteins, for which exhaustive hydrogen bonding statistical analysis are available (21), and can be used for comparison. The –N= aromatic groups are genuine of DNA in macromolecules, and their hydrogen bonding geometry has been analyzed in small molecule crystal structures (22).

Metric analysis and selection criteria for hydrogen bonding interactions

Ideal positions for the hydrogen atoms were calculated and built up for each set of coordinates. Pertinent planar geometry was adopted for all three types of donor groups. Occasionally the original coordinates showed significant deviations from planarity on guanidinium, amidinium or amide groups. In these cases hydrogen atoms were built as closely as possible to their ideal position in a planar group. No attempt was made to correct the positions of the non-hydrogen atoms.

Donor–acceptor pairs were first selected from the 22 structures in Table 1, using a distance cutoff of 3.5 Å between the two non-hydrogen atoms. This yielded a total set of 164 putative hydrogen bonding interactions. Of these, only contacts with distances $d(H\cdots A) < 2.9$ Å (A being the acceptor atom), and angles $\alpha(N–H\cdots A) > 90°$ were retained in succeeding geometry calculations. This selection criterion resulted in a total of 77 hydrogen bonds. If classified by acceptor type, 39 cases involve the O2 atom from thymines or cytosines, 26 cases involve the N3 atom from adenines and eight cases are hydrogen bonds to the O4′ or O3′ atoms of sugar rings (the standard nucleotide nomenclature is adopted for all atoms). The remaining four cases correspond to an unusual geometry in which the NH_2 groups of guanine residues are in better disposition as to be considered the hydrogen bonding donors in their interaction with the drug. These interactions will be discussed separately.

Atomic subsets of coordinates were created for every hydrogen bonding interaction by selecting only those atoms that were relevant for describing the geometry of the interaction. Every subset of coordinates was inspected and then reduced to a common reference system using the graphics program CHAIN (23).

As it will be discussed later, many of the interactions turned out to be parts of multiple hydrogen bonds, mainly those in which one hydrogen bonding donor group is shared between two hydrogen bonding acceptor groups. We will designate these interactions as three-center hydrogen bonds, and will use the term ‘bifurcated’ hydrogen bonds for those cases in which one single donor group uses two protons to interact with a single acceptor atom (24).

RESULTS

Score of hydrogen bonding

Table 2 presents the scoring of success (X) and failure (0) for the different interactions in each DNA–drug complex according to the criterion stated above. Those interactions with unusual geometry in which the donor group would correspond to the oligonucleotide part are designated as *. For those drugs with only terminal charged groups (Berenil, Pentamidine, Propamidine), the percentage of putative interactions fulfilled is 87.5% (considering at least one hydrogen bond per donor group). For drugs with only neutral, Type C donor groups (Hoechst), the rate of success is 85%. Finally, on those drugs that have both charged and neutral donor groups (Netropsin, Distamycin, DAPI), the scoring percentage drops to 53% for charged groups and 52% for neutral ones. One of the complexes in this category shows no hydrogen bonding interactions.

Table 2. Score table for hydrogen bond interactions

NDB code	Drug name	G^+/A^+	A_0	A_1	A_2	A_3	A^+
gdl001	Netropsin	*	–	0	X	0	X
gdl002	Hoechst	–	–	X	X	–	–
gdl003	Distamycin	–	X	0	X	X	X
gdl004	Netropsin	X	–	0	X	X	*
gdlb05	Netropsin	0	–	X	0	X	X
gdl006	Hoechst	–	–	X	X	–	–
gdl008	DAPI	X	–	X	–	–	0
gdl009	Berenil	X	–	–	–	–	0
gdl010	Hoechst	–	–	X	0	–	–
gdl011	Hoechst	–	–	X	X	–	–
gdl012	Hoechst	–	–	X	X	–	–
gdl013	Hoechst	–	–	X	X	–	–
gdl014	Netropsin	X	–	0	X	X	X
gdl015	Pentamidine	X	–	–	–	–	X
gdl016	Berenil	X	–	–	–	–	X
gdlb17	Netropsin	0	–	0	0	0	0
gdl018	Netropsin	*	–	X	0	0	X
gdlb19	Hoechst	–	–	X	0	–	–
gdlb20	Hoechst	–	–	X	0	–	–
gdl021	Hoechst	–	–	X	X	–	–
gdl022	Hoechst	–	–	X	X	–	–
gdl023	Propamidine	X	–	–	–	–	X

X indicates that at least one hydrogen bond is formed between the particular donor group in the drug and the DNA molecule; 0 indicates no hydrogen bonds; * corresponds to unusual geometries (see text). Non-charged donor groups along the drug backbone are designated as $A_{1–3}$ (internal groups) or A_0 (distamycin terminal amide group). G^+ and A^+ correspond to guanidinium or amidinium groups at the ends of the drug molecule.

Table 3. Statistics of the hydrogen bond interactions: average values for typical hydrogen bonding parameters

Donor/acceptor	Number	N···Y (Å)	H···Y (Å)	N–H···Y (°)	H···O=C (°)
Type A/O	9	2.84 (0.29)	2.10 (0.37)	128 (15)	145 (31)
Type A/N	11	3.12 (0.26)	2.39 (0.28)	129 (16)	
Type A/O4′, O3′	8	2.92 (0.38)	2.37 (0.28)	115 (24)	
Average Type A	28	2.94 (0.31)	2.26 (0.33)	124 (18)	
Type B/O	8	2.95 (0.32)	2.25 (0.41)	125 (10)	143 (13)
Type B/N	7	3.16 (0.12)	2.50 (0.21)	127 (27)	
Average Type B	15	3.07 (0.26)	2.39 (0.34)	126 (19)	
Type C/O	22	2.96 (0.24)	2.38 (0.30)	116 (14)	144 (18)
Type C/N	8	3.06 (0.19)	2.23 (0.27)	138 (13)	
Average Type C	30	2.99 (0.23)	2.35 (0.29)	122 (19)	
Average ABC/O	39	2.93 (0.26)	2.29 (0.35)	121 (14)	
Average ABC/N	26	3.11 (0.21)	2.37 (0.27)	131 (19)	
Type A/Gua N2*	4	2.96 (0.22)	2.24 (0.37)	130 (23)	135 (19)
Average	77	2.99 (0.27)	2.32 (0.32)	124 (18)	143 (20)

Standard deviations are shown in parentheses.

Hydrogen bonding geometry at the donor side

The two parameters analyzed on the donor part are the distance H···A and the angle N–H···A (A meaning the acceptor atom). Table 3 summarizes the global statistics for these and other hydrogen bonding parameters. Histograms showing the distribution of selected hydrogen bonding parameters are plotted in Figure 1. Figures 2 and 3 show the distribution of these parameters for the interactions included in the above defined criterion, classified by donor and acceptor types.

Overall, hydrogen bonds in this sample deviate significantly from a linear geometry (Table 3), with average values of 2.32 Å for the H···A distance and 124° for the N–H···A angle. Typical, linear, hydrogen bonds show H···A distances into the 1.9–2.0 Å range and N–H···A angles close to 160° (24). This non-linear behaviour seems to be quite independent from the different donor or acceptor types involved (Figs 2 and 3 and Table 3).

These deviations from a typical hydrogen bonding geometry can be related to the formation of three-center or multiple

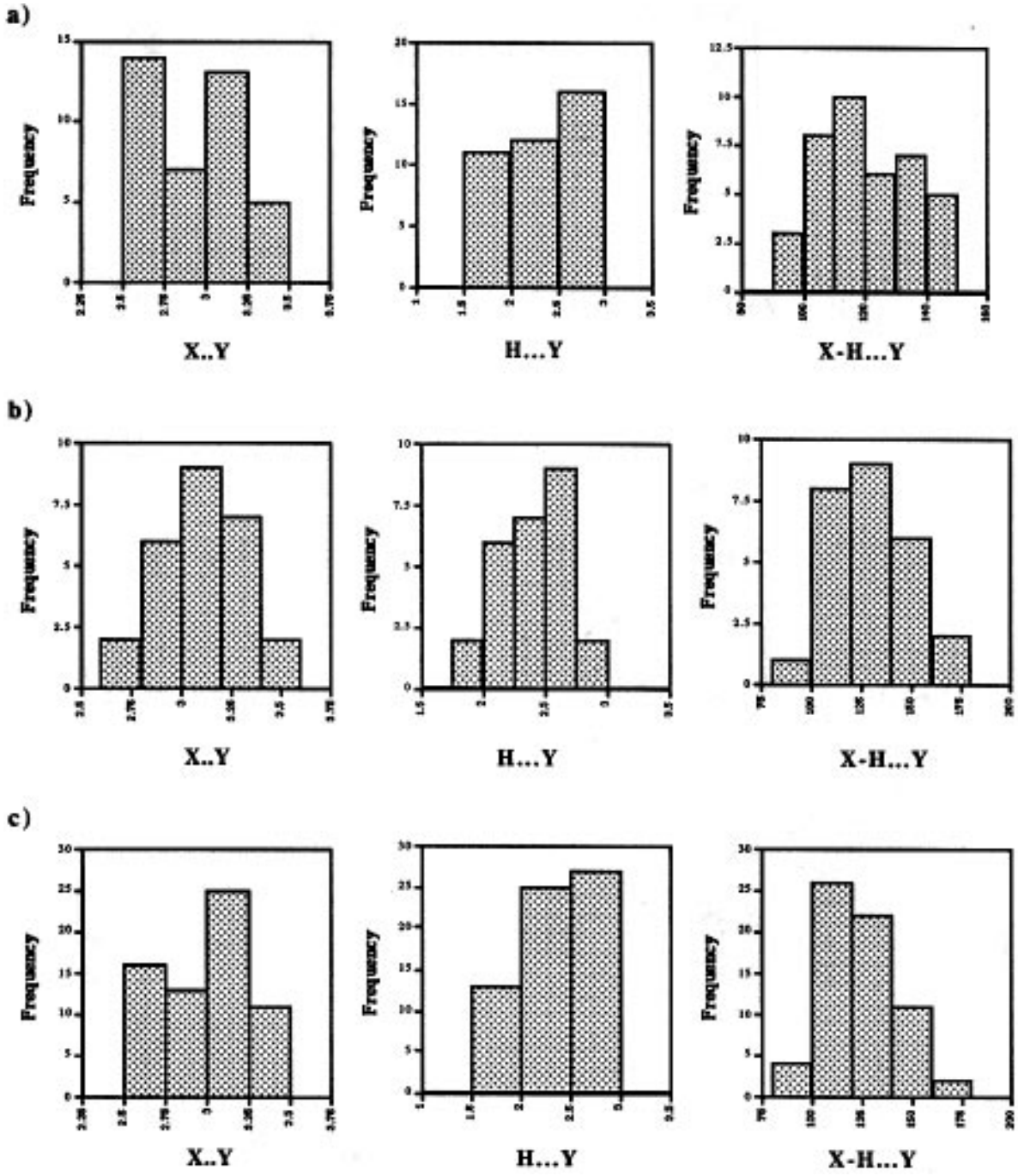

Figure 1. Histograms for the three types of donor groups combined: (**a**) N–H···O=C interactions, (**b**) N–H···N interactions and (**c**) all interactions N–H···A.

hydrogen bonds, with the donor groups typically located halfway between two different acceptor groups. For example, several cases of three-center hydrogen bonds have been originally reported for Type B or C donors (8,25–27), but a close examination of the structures in Table 1 indicates that even those reported as two-center could be classified into the first category. To analyze the geometry of the three-center hydrogen bonds, we have selected all cases of B or C donors that participate in at least one hydrogen bond, as defined throughout this work. Then we have plotted the H···A and N–H···A parameters for their two possible acceptors assuming a three-centered geometry (Fig. 4a), despite the actual values for the 'secondary' hydrogen bond. All the shorter components have been grouped together on the positive side of the N–H···A angle (lower quadrant in Fig. 4a), whereas the longer or secondary components have been assigned negative values of that angle (upper quadrant). The resulting plot suggests that most of the hydrogen bonds can be considered as three-centered, with a slightly unsymmetrical distribution of the acceptor groups between shorter and longer components.

Three-center hydrogen bonds are much more common in biological molecules than previously thought (24). Up to 20% of the common N–H···O=C hydrogen bonds are actually three-

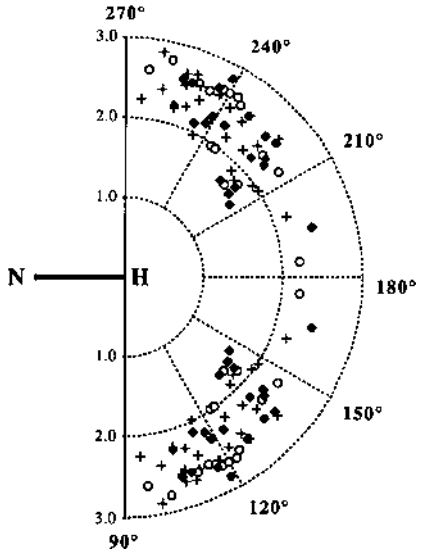

Figure 2. Polar diagram of N–H···A angles versus H···A distances for all N–H···A interactions selected as hydrogen bonds. Labels correspond to different donor types: (◆) type A, (○) type B and (+) type C. Data points are repeated for angle values greater and smaller than 180°.

centered, as shown by a survey on X-ray and neutron crystal structures of biological small molecules (28). Three-center hydrogen bonds can be symmetrical, with comparable values for hydrogen bonding distances H···A, H···A′ and angles N–H···A, N–H···A′. Most often they exhibit unsymmetrical geometry, with a major and a minor component, whose H···A distances can differ by as much as 1.0 Å (24). Therefore, it is not uncommon for the minor components of three-center hydrogen bonds to show H···A′ distances of ≥2.9 Å, and N–H···A′ angles of 90°. This is the situation for many of the Type B and C hydrogen bonds analyzed in this work. Figure 4b shows the average hydrogen bonding parameters for Types B and C three-center hydrogen bonds when grouped according to major and minor components. The average values are consistent with an unsymmetrical, three-center hydrogen bond description of these interactions, even for those cases in which the minor component geometry falls clearly out of the range initially used as a selection criterion for this work.

The attractive character of the three-centered hydrogen bonds is related to the coplanarity of the hydrogen atom with the plane defined by the atom donor N and its two acceptors A, A′. The sum of the angles N–H–A, N–H–A′ and A–H–A′ should be near 360° for a three-center hydrogen bond with good geometry, and the hydrogen atom should be between 0.0 and 0.2 Å from the plane A–N–A′. For the interactions shown in Figure 4a–b the average deviation of the planarity is 20° ± 15°, which corresponds to an out-of-plane component of 0.4 Å for the hydrogen atoms.

Further analysis by donor types does not show significant differences between the geometry of three-center hydrogen bonds involving Type B or Type C groups, suggesting that they behave similarly.

The last class of donors, Type A, have a broader variety of interactions, given their situation at the terminal regions of the drugs. For example, all N–H···O drug–sugar interactions belong to this class. Usually, Type A donors do not participate in three-center hydrogen bonds as those already discussed (a notable exception is the drug berenil in the gdl016 structure), and in several instances they seem to be involved in multiple interactions using both hydrogens from their NH_2 groups. The N–H···A angles for Type A hydrogen bonds are not more linear than those of the three-centered interactions (Table 3). Additionally, a common motif is observed in 10 cases, for which the NH_2 group is placed with its two hydrogen atoms around one acceptor atom. The primary interaction shows an average H···A distance of 2.35 (±0.28) Å, and an N–H···A angle of 121° (±15°). The secondary interaction H′···A has much worse geometry, 3.01 (±0.30) Å and 79° (±8°) respectively, and cannot be considered a hydrogen bond on itself. However it may be responsible for the deviation from linearity of the primary H···A hydrogen bond.

Hydrogen bonding geometry at the acceptor side

Table 3 shows the average values of the H···O=C angle for the drug–DNA hydrogen bonds in this sample. They are very similar to those observed for peptide N–H···O=C hydrogen bonds in

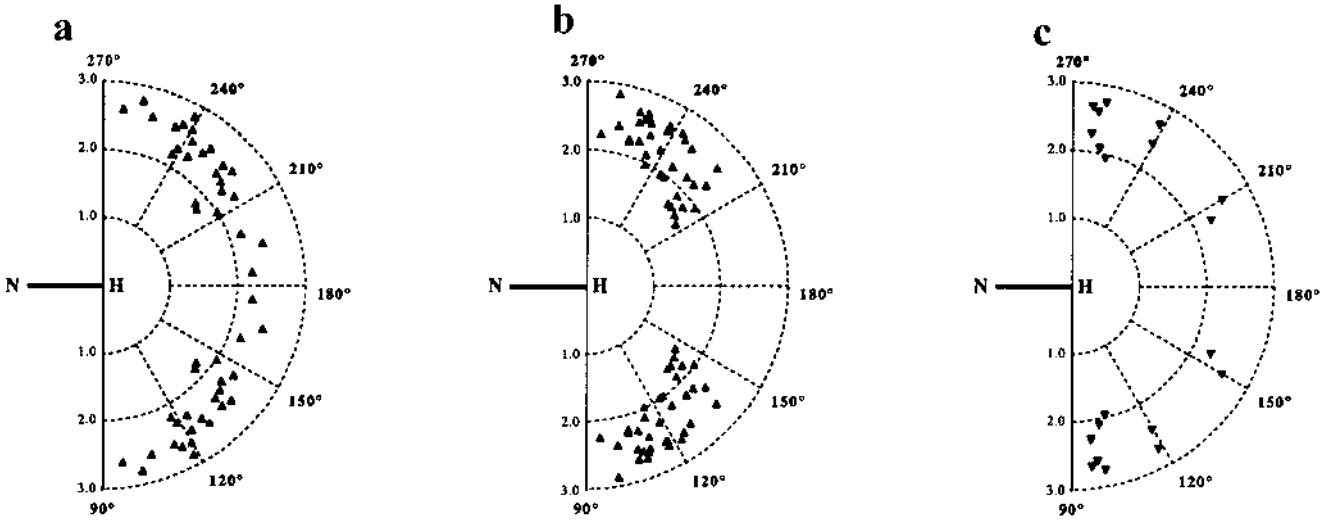

Figure 3. Polar diagrams classified by acceptor type: (**a**) N from purine rings, (**b**) O from carbonyl groups and (**c**) O′ from sugar rings.

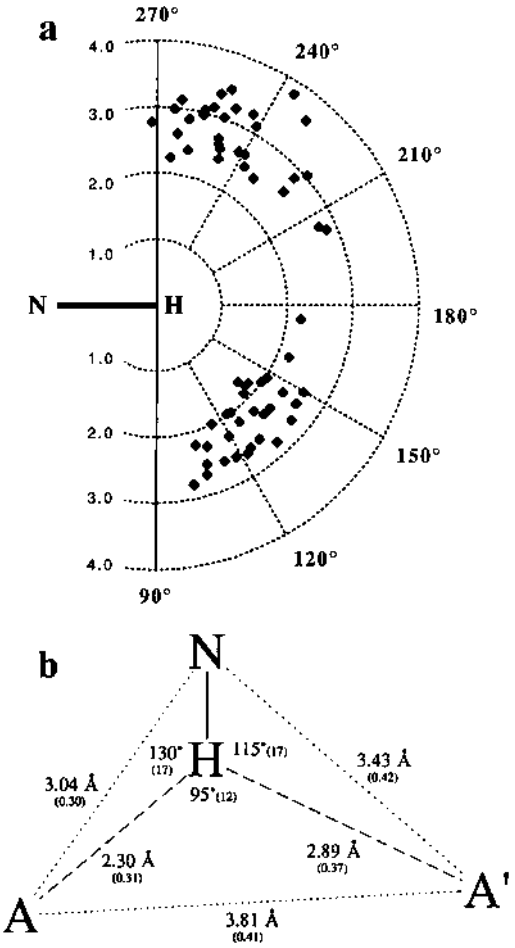

Figure 4. (a) Polar diagram of interatomic distances and angles between type B and C donors and their putative three-center hydrogen bond mates. Closer neighbours are grouped on the 90–180° quadrant, and farther neighbours are plotted on the 180–270° quadrant. **(b)** Average hydrogen bonding parameters for three-centered geometry if classified in shorter (N···A) and longer (N···A') components as stated above.

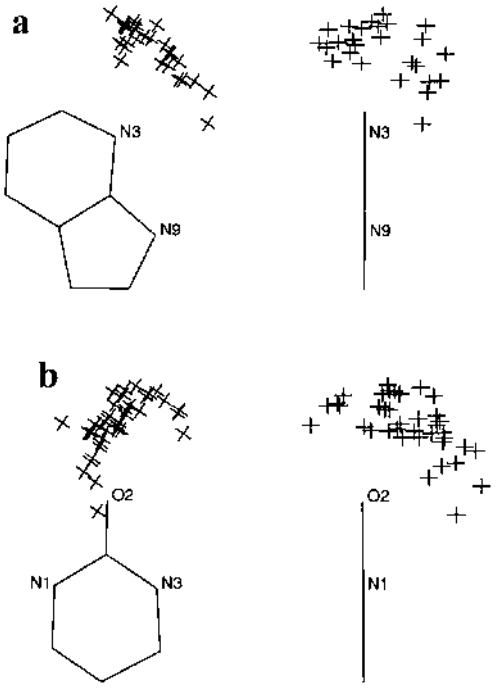

Figure 5. Scatter plots of the donor hydrogen atoms around (**a**) N3 atoms from adenines and guanines; (**b**) O2 atoms from thymines and cytosines. Only the unvariant atoms from the rings are shown.

protein crystal structures: 147° for α-helices or 151° for β-sheets (21). The distribution of donor hydrogen atoms around the C=O group (Fig. 5b) is also similar to that observed around carbonyl groups in proteins. This distribution is very broad and does not follow the ideal orientation expected for the lone pairs on the oxygen atom, which would be coplanar with the aromatic ring, at ±120° from the axis of the C=O bond. Approaching the C=O bond from the plane of the ring is not possible for a minor-groove binding drug without deforming the DNA double helix: either the sugar ring from the same nucleotide or the nitrogenous base from the complementary chain would pose unsurmountable steric obstacles. Instead, the mode of approach for all drugs involved in minor-groove binding is at a certain angle with the plane of the ring (Fig. 5b). This angle can be derived from the dihedral angle H···O2=C2–N1, which should be 0° or 180° for an on-plane approach. For the hydrogen atoms shown in Figure 5b, the average H···O2=C2–N1 dihedral angle is 115°, with a very broad distribution.

The distribution of donor hydrogens around the N(sp^2) acceptor atoms is also very broad and deviates from the predicted position for the lone pair of the N(sp^2) atom (Fig. 5a). This behaviour differs from what is observed for N–H···N(sp^2) hydrogen bonds in crystal structures of small molecules, for which the highest concentration of hydrogen bonding interactions occurs in the lone pair direction (22). As in the H···O=C hydrogen bonds discussed above, the approach of the drug following the N(sp^2) lone pair is precluded by steric obstacles coming from sugar rings at the minor groove wall. Thus, the average value for the H···N···C6 angles is 134° (±11°). An ideal orientation from the acceptor point of view would correspond to an H···N···C6 angle close to 180°. The departure of the hydrogen positions from the plane of the fused rings can be evaluated through the dihedral angle between that plane and the plane defined by the three atoms H···N···C6. For the hydrogen atoms in Figure 5a, that dihedral angle averages 22°, with a very broad distribution.

Unusual geometry: guanine NH_2 groups acting as donors?

An unexpected hydrogen bonding geometry is observed in four cases, all them involving the NH_2 groups from guanine residues in the fourth position and different nitrogen atoms from netropsin molecules: N1 in gdl001, N10 in gdl004, N1 in gdl018 and N3 in gdl018. Nitrogen–nitrogen distances for these pairs are 3.04, 3.03, 2.64 and 3.14 Å respectively, which might suggest a drug–DNA hydrogen bonding interaction. However, when standard hydrogen positions are built for these groups, the resulting geometry suggests in fact that the hydrogen bonds go in the other direction, that is, with the NH_2 groups from guanines acting as donors and nitrogen atoms on guanidinium or amidinium groups acting as acceptors. Only two complexes show this unexpected type of interaction, since gdl001 and gdl004 correspond to the same netropsin–DNA complex with the drug refined in two

different orientations. In other complexes the distances from guanine NH_2 groups to netropsin nitrogens are usually >4 Å, with the NH_2 hydrogen atoms pointing away from the drug.

DISCUSSION

Position of the drugs along the minor groove and hydrogen bonding geometry

The positioning of the drug along the minor groove has some variability among all the structures. Even for the same drug, different complexes show slightly shifted positions. The structures of netropsin bound to three different dodecanucleotide sequences can be grouped in two main categories, as has been recently suggested by Goodsell *et al.* (29). Class I complexes include gdlb05, gdl014 and gdl003 (Distamycin complex). In them, pyrrole rings are opposed to the A–T base pairs and amide moieties are located between two successive base pairs (Fig. 6). With this disposition, most amide nitrogens form three-center hydrogen bonds to adenine or thymine bases on both strands of the DNA fragment, although in some cases the three-center hydrogen bond can be rather unsymmetrical. In class II complexes, gdl001, gdl004, gdlb17 and gdl018, the amide moieties lay in the plane of the base pairs and the pyrroles are positioned midway between two successive steps (Fig. 6). Many of the putative DNA–drug interactions in the class II complexes do not fulfill our selection criterion for a hydrogen bond: either the donor–acceptor distances are >3.5 Å or the X–H···Y angles are too closed (<90°) (Table 2). Additionally, those are the only complexes that show the unusual interaction involving guanine NH_2 groups as donors (Table 2, indicated by the asterisk).

Similarly, complexes with Hoechst drugs can also be classified into two main structural classes that differ in the positioning of the imidazole nitrogen atoms on the DNA minor groove (Fig. 6). In class I, including gdl002, gdl006, gdl011, gdl012, gdl013 and gdl021, the imidazole rings are positioned between successive base pairs, while benzyl rings are opposed to the base pairs. With this arrangement, those nitrogen atoms from imidazole rings that are facing the minor groove, make three-center hydrogen bonds with adenine and thymine bases at adjacent steps on different strands of the oligonucleotide. Class II complexes, gdl010, gdlb19, gdlb20 and gdl022, have their imidazole rings slid up and therefore nitrogen atoms are facing the base pairs whereas benzyl rings are located in between steps. As for netropsin class II complexes, the geometry for the interactions in this set of structures is deficient and most of the hydrogen bonds are not formed (Table 2).

Goodsell *et al.* claim that DNA–netropsin class II complexes may in fact represent poorly refined models and that the position of the drug in these complexes is incorrectly shifted along the minor groove by one-half base pair step (29). These authors adduce lower quality of the experimental data for class II crystal structure determinations, and remark that with this drug positioning the hydrogen bonding capabilities between drug and the DNA minor groove are not fulfilled. The results of our analysis point in that direction for both netropsin or Hoechst class II DNA–drug complexes: the score and overall geometry of the hydrogen bonding interactions are better for class I complexes than for class II ones.

The re-examination of the gdlb05 crystal structure by the same authors (29), suggests that repositioning the netropsin drug in that complex with a class II topology would imply fitting the guanidinium group into the electron density corresponding to a solvent molecule, probably water (see figure 6 in ref. 29). This offers an explanation for the unexpected hydrogen bond arrangements that we have noticed in netropsin–class II complexes, that is with guanine NH_2 groups acting as donors. If we assume that in these complexes the drug is incorrectly shifted, those nitrogen atoms from guanidinium or amidinium groups that seem to act as acceptors would actually correspond to water molecules from the intrinsic hydration shell of the minor groove.

Hydrogen bonding geometry and formal charge of the donor group

We have not observed any significant correlation between nature of the hydrogen bonding donor and goodness of the hydrogen bonding geometry. For all three types of hydrogen bonding donors we observe a significant deviation from the ideal hydrogen bond parameters, which can be at least partially rationalized as a result of the formation of different types of multiple hydrogen bonds.

Charged and non-charged groups do exhibit different strategies because of their positioning along the drug. The arrangement observed in the crystal structures from Table 1 seems to indicate that the terminal, charged groups, have more options at hand to fulfill hydrogen bonding interactions with the DNA molecule (for example with oxygen atoms from sugar rings), whereas the neutral, internal groups are somehow restricted by the more rigid conformation of the pyrrole or benzimidazole rings and the available acceptor groups at the bottom of the minor groove. The intrinsic curvature of long drug molecules like Netropsin or Distamycin does not match perfectly the curvature of the bottom of the minor groove. This precludes the formation of all possible hydrogen bonds for a flat, completely extended Netropsin-like drug. Indeed, all Netropsin or Distamycin molecules in Table 1 complexes show a rotation of the plane from their charged groups with respect to the mean plane of the central amide groups (see for example Fig. 6). Thus, flexible ends in different DNA–drug complexes adopt alternative orientations as to find different hydrogen bond mates, either at the DNA bases or at the sugar rings. Some of these interactions may involve multiple or bifurcated hydrogen bonds, although the observed geometry in the crystal structures of this sample does not provide definitive evidence for the later case. Interestingly, in gdl015 and gdl023 complexes, Propamidine and Pentamidine terminal groups attach themselves to only one of the DNA strands instead of adopting a three-centered geometry. Three out of four of the hydrogen bonding NH_2 groups in these drugs are positioned as to orient both protons towards the acceptor atom in a bifurcated-like geometry.

Amide and benzimidazole groups are often kept farther away from the base pairs and therefore they adopt a different strategy. In the most favourable case these internal donor groups are positioned between successive base pair steps and form three-center hydrogen bonds to both strands of the DNA double helix. These interactions define a characteristic pattern that is repeated for either amide or benzaimidazole donors (Fig. 6 and ref. 29). Deviations of this pattern by sliding the position of the drug along the minor groove result in a worse geometry for the interactions and loss of several hydrogen bonds to the bases. Thus, appropriate spacing between donor groups remains as one of the essential characteristics for a successful minor-binding drug.

It is possible that the inclusion of charged groups at the flanking regions of a minor groove-binding drug may increase its

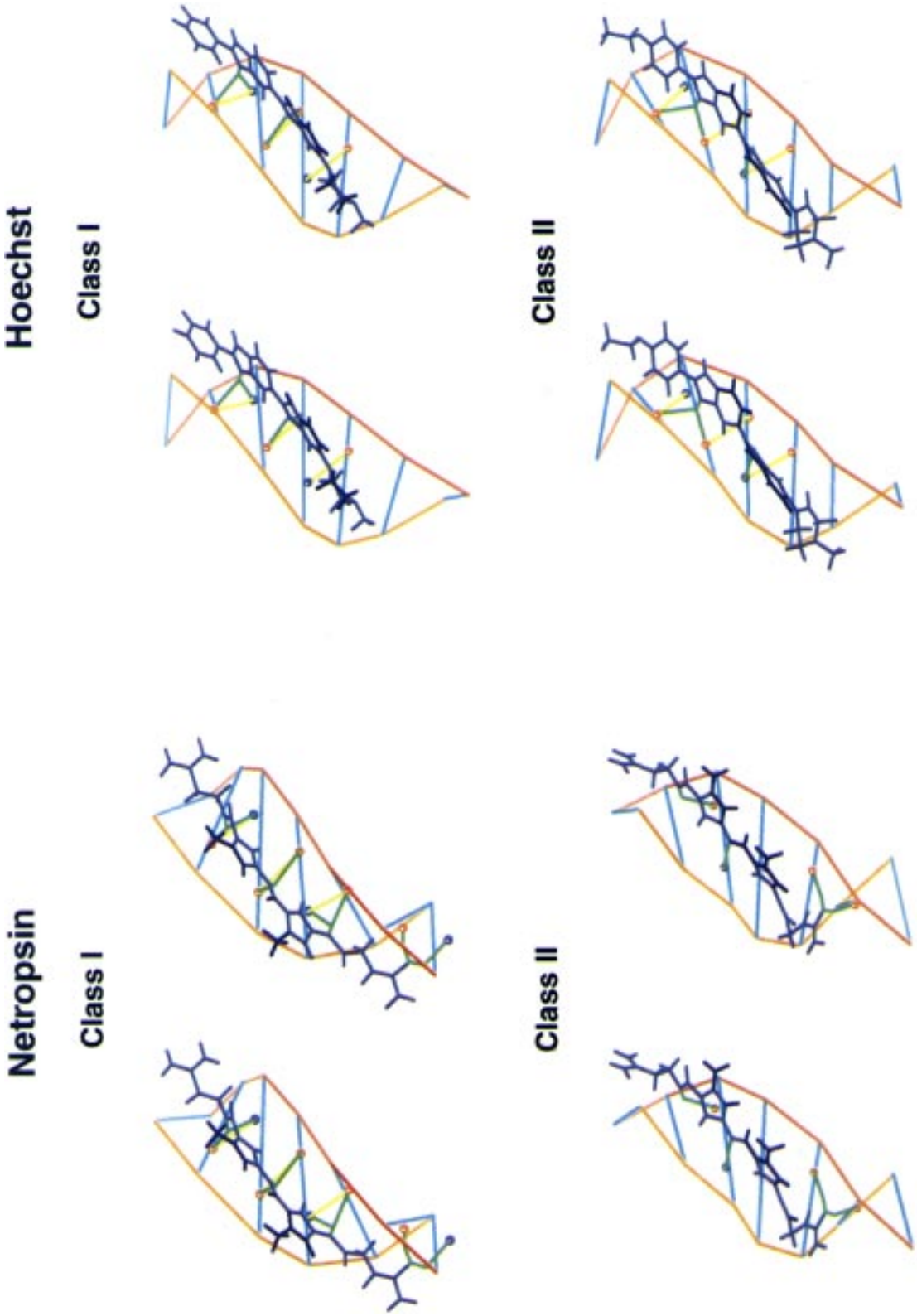

Figure 6. Stereo diagrams showing the binding position for Class I and Class II drugs into the minor groove (see text). Steps represent the base-pairs and the spheres the acceptor atoms on the bases. Hydrogen bonds are shown as green lines.

anchorage capabilities during and after the initial steps of the complex formation, compensating for any lack of interaction at other points in the drug. Charge itself may be critical at the initial steps of the binding process, as the drug must compete with and replace the pre-existing water molecules at the minor groove (8,10). In a related study, we have carried out *ab initio* quantum mechanical calculations of model compounds for the hydrogen bonding interactions between amide and thymine carbonyl groups. These calculations predict a much tighter interaction of charged nitrogen groups with thymine carbonyl acceptors (19). Furthermore, after introduction of solvent effects, the calculated interaction energy between a donor group and a thymine carbonyl acceptor, is negative for positively charged amide groups, and positive for neutral amides (19). This would suggest a leading role for the positively charged moieties at the initial steps of the formation of the DNA–drug complex, whereas interactions between neutral groups would be established later. However, as we mentioned above, no correlation is observed between charged donor groups and better hydrogen bonding geometry in the crystal structures studied. This probably occurs because the crystal structures represent the final stage in the complex formation where a combination of different interactions takes place.

Hydrogen bonding geometry and DNA flexibility

The conformation of the DNA double helix in these complexes is mainly restricted by the local geometry of the specific sequence. A certain degree of flexibility can be achieved by means of propeller twist variations and the capacity of the base pairs to adopt a more favorable disposition in optimizing their hydrogen bonding interactions with the particular drug. Sequences with alternating ATAT steps have a lower degree of propeller twist in complexes with drugs that sequences with successive AATT or AAATTT steps (27). This indicates that tracts of A–T base pairs are able to adapt their local conformation to improve the hydrogen bonding geometry with the internal donor groups from the drugs and therefore increase their specificity.

Final remarks: effect of the resolution of the crystal structures

One of the main drawbacks of the sample of the crystal structures analyzed in this work is their relative low resolution (Table 1), that introduces two undesirable effects. First, the final geometries of the crystallographic model can be rather poor, mainly on the non-covalent interactions, for which very soft or no restraints are applied during the crystallographic refinement. Second, the interpretation of the electronic density maps at the early stages of the refinement can be very ambiguous as to where and how to model the drug molecule on the DNA minor groove (29). Still, meaningful information can be deduced from comparison between the available structures and identification of common, repetitive motifs. Thus, the description of a particular hydrogen bond in one crystal structure can be quite inaccurate, but the average values over several structures will probably represent a better description of the actual interactions taking place between DNA fragments and minor groove binding drugs. We think that the geometric features described in this work provide a useful framework to which future structural determinations of DNA complexes can be compared.

REFERENCES

1 Zimmer, C. and Wahnert, U. (1986) *Proc. Biophys. Mol. Biol.*, **47**, 31–112.
2 Taylor, J.S., Schultz, P.G. and Dervan, P.B. (1984) *Tetrahedron*, **40**, 457–465.
3 Dervan, P.B. (1986) *Science*, **232**, 464–471.
4 Portugal, J. and Waring, M.J. (1987) *Eur. J. Biochem.*, **167**, 281–289.
5 Liquier, J., Mchami, A. and Taillandier, E. (1989) *J. Biomol. Struct. Dyn.*, **7**, 119–126.
6 Patel, D.J. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 6424–6428.
7 Orozco, M., Laughton, C.A., Herzyk, P. and Neidle, S. (1990) *J. Biomol. Struct. Dyn.*, **8**, 359–373.
8 Kopka, M.L., Yoon, C., Goodsell, D., Pjura, P. and Dickerson, R.E. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 1376–1380.
9 Van Dyke, M.W., Hertzberg, R.P. and Dervan, P.B. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 5470–5474.
10 Chalikian, T.V., Plum, G.E., Sarvazyan, A.P. and Breslauer, K.J. (1994) *Biochem.*, **33**, 8629–8640.
11 Adnet, F., Liquier, J., Taillandier, E., Singh, M.P., Rao, K.E. and Lown, J.W. (1992) *J. Biomol. Struc. Dyn.*, **10**, 565–575.
12 Bailly, C., Henichart, J.P., Colson, P. and Houssier, C. (1992) *J. Mol. Recogn*, **5**, 155–171.
13 Jenkins, T.C., Lane, A.N., Neidle, S. and Brown, D.G. (1993) *Eur. J. Biochem.*, **213**, 1175–1184.
14 Singh, S.B., Ajay, Wemmer, D.E. and Kollman, P.A. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 7673–7677.
15 Rhee, Y., Wang, C., Gaffney, B.L. and Jones, R.A. (1993) *J. Am. Chem. Soc.*, **115**, 8742–8746.
16 Berman, H.M., Neidle, S., Zimmer, C. and Thurm, H. (1979) *Biochim. Biophys. Acta*, **561**, 124–131.
17 Westhof, E., and Patel, D (1995) *Curr. Opin. Struct. Biol.*, **5**, 279–343.
18 Berman, H.M., Olson, W.K., Beveridge, D.L., Westbrook, J., Gelbin, A., Demeny, T., Hsieh, S.-H., Srinivasan, A.R. and Schneider, B. (1992) *Biophys. J.*, **63**, 751–759.
19 Aleman, C., Tabernero, L. and Bella, J. (1996) *J. Phys. Chem.*, **100**, 11480–11487.
20 Brown, D.G., Sanderson, M.R., Skelly, J.V., Jenkins, T.C., Brown, T., Garman, E., Stuart, D.I. and Neidle, S. (1990) *EMBO J.*, **9**, 1329–1334.
21 Baker, E.N. and Hubbard, R.E. (1984) *Prog. Biophys. Mol. Biol.*, **44**, 97–179.
22 Llamas-Saiz, A. and Foces-Foces, C. (1990) *J. Mol. Struct.*, **238**, 367–382.
23 Sack, J.S. (1988) *J. Mol. Graphics*, **6**, 224–225.
24 Jeffrey, G.A. and Saenger, W. (1991) *Hydrogen Bonding in Biological Structures*. Springer-Verlag, New York.
25 Coll, M., Frederick, C.A., Wang, A.H.-J. and Rich, A. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 8385–8389.
26 Quintana, J.R., Lipanov, A.A. and Dickerson, R.E. (1991) *Biochemistry*, **30**, 10294–10306.
27 Tabernero, L., Verdaguer, N., Coll, M., Fita, I., van der Marel, G.A., van Boom, J.H., Rich, A. and Aymami, J. (1993) *Biochemistry*, **32**, 8403–8410.
28 Taylor, R., Kennard, O. and Versichel, W. (1984) *Acta Crystallogr.*, **B40**, 280–288.
29 Goodsell, D.S., Kopka, M.L. and Dickerson, R.E. (1995) *Biochemistry*, **34**, 4983–4993.
30 Coll, M., Aymami, J., van der Marel, G.A., van Boom, J.H., Rich, A. and Wang, A.H.-J. (1989) *Biochemistry*, **28**, 310–320.
31 Teng, M., Usman, N., Frederick, C.A. and Wang, A.H.-J. (1988) *Nucleic Acids Res.*, **16**, 2671–2690.
32 Pjura, P.E., Grzeskowiak, K. and Dickerson, R.E. (1987) *J. Mol. Biol.*, **197**, 257–271.
33 Larsen, T.A., Goodsell, D.S., Cascio, D., Grzeskowiak, K. and Dickerson, R.E. (1989) *J. Biomol. Struct. Dyn.*, **7**, 477–491.
34 Edwards, K.J., Jenkins, T.C. and Neidle, S. (1992) *Biochemistry*, **31**, 7104–7109.
35 Brown, D.G., Sanderson, M.R., Garman, E. and Neidle, S. (1992) *J. Mol. Biol.*, **226**, 481–490.
36 Sriram, M., van der Marel, G.A., Roelen, H.L.P.F., van Boom, J.H. and Wang, A.H.J. (1992) *Biochemistry*, **31**, 11823–11834.
37 Sriram, M., van der Marel, G.A., Roelen, H.L.P.F., van Boom, J.H. and Wang, A.H.-J. (1992) *EMBO J.*, **11**, 225–232.
38 Nunn, C.M., Jenkins, T.C. and Neidle, S. (1993) *Biochemistry*, **32**, 13838–13843.