4882–4889 *Nucleic Acids Research, 1996, Vol. 24, No. 24*

Designer DNA-binding drugs: the crystal structure of a *meta*-hydroxy analogue of Hoechst 33258 bound to d(CGCGAATTCGCG)$_2$

George R. Clark[1,2], Christopher J. Squire[2], Emily J. Gray[1], Werner Leupin[3,+] and Stephen Neidle[1,*]

[1]The CRC Biomolecular Structure Unit, Institute of Cancer Research, Sutton, Surrey SM2 5NG, UK, [2]Chemistry Department, University of Auckland, Auckland, New Zealand and [3]F. Hoffman-La Roche Ltd, Preclinical Research Pharma Gene Technologies, CH-4002 Basel, Switzerland

Received September 11, 1996; Revised and Accepted November 5, 1996 DDBJ/EMBL/GenBank accession nos X96024, X96025

ABSTRACT

An analogue of the DNA binding compound Hoechst 33258, which has the *para* hydroxyl group altered to be at the *meta* position, together with the replacement of one benzimidazole group by pyridylimidazole, has been cocrystallized with the dodecanucleotide sequence d(CGCGAATTCGCG)$_2$. The X-ray structure has been determined at 2.2 Å resolution and refined to an R factor of 20.1%. The ligand binds in the minor groove at the sequence 5′-AATTC with the bulky piperazine group extending over the C·G base pair. This binding is stabilised by hydrogen bonding and numerous close van der Waals contacts to the surface of the groove walls. The *meta*-hydroxyl group was found in two distinct orientations, neither of which participates in direct hydrogen bonds to the exocyclic amino group of a guanine base. The conformation of the drug differs from that found previously in other X-ray structures of Hoechst 33258–DNA complexes. There is significant variation between the minor groove widths in the complexes of Hoechst 33258 and the *meta*-hydroxyl derivative as a result of these conformational differences. Reasons are discussed for the inability of this derivative to actively recognise guanine.

INTRODUCTION

It is well established that a variety of small molecules such as netropsin, berenil, pentamidine, distamycin and Hoechst 33258 bind in the minor groove of duplex DNA, with a marked preference for AT-rich regions (1–3). These groove binding species and their derivatives have shown a wide range of biological activity from antitumour to antiviral, antifungal and antimicrobial (for example, refs 4–9). Their modes of action all involve initial groove binding to the DNA possibly followed by interference in the function of DNA topoisomerases (10,11) or with transcription factor binding (12–15). In order to bind effectively in the minor groove these compounds should be able to follow the contours of the minor groove without unfavourable contacts and their charge distribution must be appropriate for the electrostatic potential encountered within the binding site of the groove. The minor groove binders therefore tend to have similar structural properties, such as an inherently concave shape along one long edge and cationic charge, often at the ends of the molecule. They often contain a series of linked aromatic ring systems which are close to coplanar when viewed from the side.

NMR (16,17) and X-ray (18–24) studies of minor groove binding drugs complexed to DNA oligonucleotides have characterised the non-covalent interactions involved in drug–DNA recognition. These studies have shown the drug bound in the narrow minor groove of AT-rich sequences stabilised by an array of van der Waals interactions (25) involving the hydrogen atoms of the walls and floor of the minor groove. Direct hydrogen bonding between the drug and various sites on the DNA is frequently (though not invariably) observed (23). The sequence selectivity of groove binders is therefore likely to be dominated by their ability to recognise variations in DNA groove width, with DNA–drug hydrogen bonding interactions being a secondary factor (26).

Much of the current focus of minor-groove drug design is aimed towards improving the selectivity of the drugs to specific base pair sequences of duplex DNA. A number of minor groove binders have been designed that can bind to both A·T and G·C base pairs in stretches of mixed sequence by taking advantage of the sequence dependence of minor groove width in B-form DNA (27–30). The bis-benzimidazole compound Hoechst 33258 (**1**; Fig. 1) is a primitive example of a GC-tolerating molecule (at one end of its binding site) as a result of the bulk of the saturated piperazine ring on one end of the molecule requiring a wide GC minor groove region.

Hoechst 33258 has been studied extensively by X-ray crystallography (31–37) bound to oligonucleotide sequences d(CGCXXXXXGCG)$_2$ (where XXXXXX is either 5′-GAATTC, 5′-AAATTT or 5′-GATATC) and has shown significant variations in its direct hydrogen bonding patterns to DNA and also some variation in binding site. These vary not only with the DNA sequence involved (eg. 34) but also with the temperature of data collection (33). Hoechst 33258 has been used as the starting point for a number of rational drug design studies aimed at altering the biological activity and/or sequence selectivity of the drug. A number of the structural variants of Hoechst 33258 developed to

*To whom correspondence should be addressed. Tel/Fax: +44 181 643 1675; Email: steve@iris5.icr.ac.uk

+Present address: Department of Physical Chemistry, University of Basel, Klingelbergstrasse 80, CH-4056 Basel, Switzerland

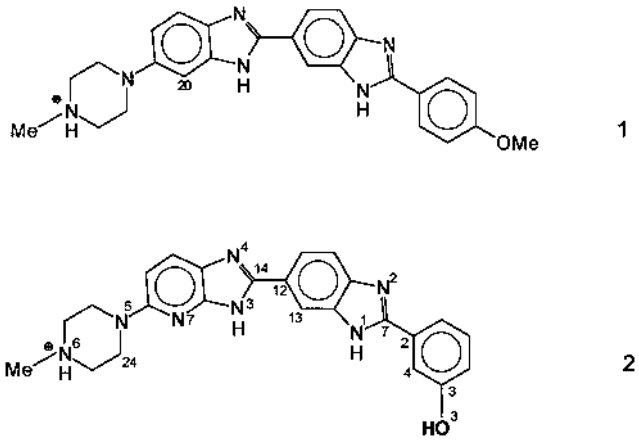

Figure 1. The structures of Hoechst 33258 (**1**) and 3-{5-[5-(4-methyl-piperazin-1-yl)-1H-imidazo(4,5-b)pyridin-2-yl]-benzimidazol-2-yl}-phenol (**2**). Atom N7 and the hydroxyl group HO3 in **2** are highlighted in bold.

date have been aimed at compromising the AT selectivity to allow a degree of GC recognition (38,39). A complementary approach has been to use the AT-directing ability of Hoechst 33258 to bind the molecule strongly to AT-rich tracts but adding additional sequence recognition features to the ends of the molecule. One such compound ('*meta*-Hoechst') has the phenolic OH substituent of Hoechst 33258 on the *meta* rather than the *para* position of the ring (40–42).

Meta-Hoechst has been studied by NMR, molecular modelling, DNA footprinting and other techniques (40–42). These studies have shown that *meta*-Hoechst binds to AT-rich regions of duplex DNA in a very similar site to that of Hoechst 33258 itself. NMR and molecular modelling studies of *meta*-Hoechst complexed to d(CGCGAATTCGCG)$_2$ have predicted that moving the OH from *para* to *meta* should introduce the possibility of additional hydrogen bonding between the *meta*-hydroxyl group and the exocyclic 2-amino group of guanine at position 4 and the 2-oxygen of cytosine 21 in the opposite strand. This model also predicts that the phenolic ring would be held tightly in place with the hydroxyl group pointing directly into the groove, limiting the phenolic ring flipping observed for Hoechst 33258, when it is in the *para* position (43–45). Molecular modelling further predicted that this directional binding of the *meta*-hydroxyl group into the groove may result in a change in dihedral angle between the phenolic ring and the adjacent benzimidazole from that seen in Hoechst 33258, which could affect hydrophobic and π–π interactions (42).

This paper reports the crystal structure of the complex formed between the complementary dodecamer d(CGCGAATTCGCG)$_2$ and an analogue of *meta*-Hoechst in which one benzimidazole group is replaced by a pyridylimidazole moiety (**2**; Fig. 1). This latter substitution in another analogue of Hoechst 33258 has been suggested by NMR to increase the GC tolerance of the drug molecule (46). The present crystal structure enables us to examine the effect of the *para*→*meta* change on the sequence selectivity of the drug, particularly in terms of the resultant drug–DNA hydrogen bonding.

MATERIALS AND METHODS

Synthesis and crystallisation

The DNA dodecamer d(CGCGAATTCGCG)$_2$ was purchased from the Oswel DNA Service (University of Edinburgh) and annealed before use. The *meta*-OH(N) Hoechst compound **2** was prepared as the hydrochloride salt. Its synthesis and biological properties will be reported elsewhere; its DNA-binding properties have been described (42). The complex was grown from hanging drops at 18°C as colourless plates. The crystal used for data collection was grown from a drop containing 4 µl of 35% w/v 2-methylpentane-2,4-diol, 2 µl of 5 mM drug, 2 µl of 200 mM $MgCl_2$ and 3 µl of 2 mM dodecamer, equilibrated against a reservoir containing 1 ml 35% w/v 2-methyl-2,4-pentanediol. All solutions were prepared in 30 mM sodium cacodylate buffer at pH 7.0. An X-ray quality crystal was obtained after ~3 months.

Data collection

The single crystal used for data collection was of approximate dimensions 0.30 × 0.24 × 0.08 mm and was mounted in a 0.5 mm Lindemann glass capillary with a small amount of mother liquor. Diffraction data were collected at 20°C to a maximum resolution of 2.2 Å using a Siemens-Xentronics multiwire area detector and a rotating-anode X-ray generator operated at 3 kW. A crystal-to-detector distance of 10 cm and a swing angle of 15° was used with χ set at 45° to obtain data sets at φ values of 0 and 60°. The crystal was rotated through 100° in ω, collecting 180 s frames every 0.2°. The crystal showed no significant decay during data collection. Data processing was performed using XENGEN version 1.3. After merging, the data comprised 3292 of the possible 3514 unique reflections (i.e. 93.0%) to 2.2 Å, with a merging R value of 4.5%.

Structure refinement

The crystal was found to be in the space group $P2_12_12_1$ with unit cell dimensions of a = 24.62, b = 40.35 and c = 66.23 Å. This cell is isomorphous with that for the native sequence and other drug-dodecamer structures and so the coordinates for the native structure [PDB entry 9BNA (47)] were used for the initial model. The refinement process was carried out using the program X-PLOR version 3.1 (48). Initial rigid-body refinement of the whole duplex model was carried out for data in the range 8.0–4.0 Å, the resolution range of the data being gradually increased to 8.0–3.0 Å. The DNA duplex was then divided into 24 rigid groups comprising the 24 nucleotide units and the upper resolution limit was gradually increased to the maximum of 2.2 Å during further rigid-body refinement. At this stage the R factor was 31.5%. Positional and temperature factor refinement reduced the R factor to 28.3%. Electron density maps were calculated and displayed using the graphics package FRODO/TOM (version 3.1 Alberta/Caltech). The DNA molecule fitted the electron density well in all regions of the model and a continuous volume of electron density was observed in the minor groove of the DNA. The molecular structure of the *meta*-OH(N) Hoechst molecule was built from one of the Hoechst 33258 crystal structures (35) using the computer modelling package Insight II (49). The lobe of continuous electron density in the minor groove showed sufficient detail to allow the drug molecule to be unambiguously fitted to it with a unique orientation and position and with the piperazine group in the stereochemically-acceptable conformation (Fig. 2). The location of the *meta*-hydroxyl group itself was ambiguous

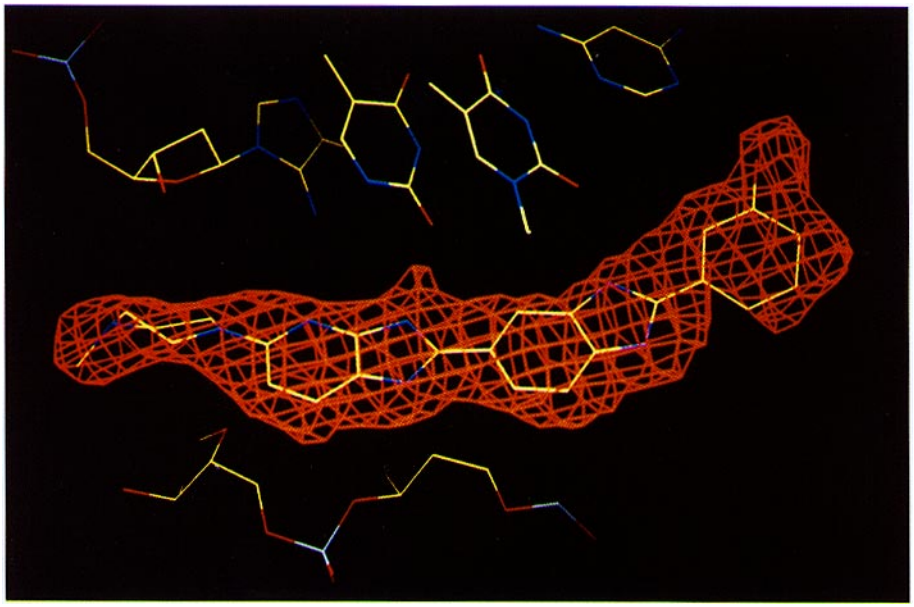

Figure 2. The molecule of **2** superimposed on the electron density calculated from a $2F_o–F_c$ omit map, drawn at a contour level of 1σ.

and two possible structures were refined from this point, one with the hydroxyl pointed into the groove and the other with it pointing out. Partial charges were calculated for the drug molecule by an empirical calculation using the cvff (default) force-field of Insight II. Planar restraints were applied to each individual aromatic group. Positional and temperature refinement was continued with the drug molecule included. During refinement a total of 84 water molecules and a magnesium ion were located in $F_o–F_c$ electron density maps and included in the structure. The final R factor for the 'OH in' conformation was 20.2% for all observed data in the range 8.0–2.2 Å, with rms deviations from the ideal of 0.019 Å and 1.37° for bond lengths and angles, respectively. The X-PLOR refinement and R factor calculations used 3097 reflections with $F > 2\sigma(F)$. For the 'hydroxyl out' structure the final R factor was 20.4%, with 84 water molecules and the magnesium ion. Final refined coordinates for both 'hydroxyl in' and 'hydroxyl out' structures, together with the observed and calculated structure factors have been deposited in the Nucleic Acid Database with the identity codes X96024 and X96025.

RESULTS AND DISCUSSION

Overall features of the complex

The DNA duplex adopts a B-DNA conformation with the ligand lying in the narrow central region of the minor groove (Fig. 3). The asymmetric unit contains one anti-parallel double helix, the two strands being crystallographically distinct. Packing of the helices in the unit cell is analogous to that observed for almost all other dodecanucleotide crystal structures in this space group, with the two terminal G·C base pairs of each helix interacting with a neighbouring duplex via minor groove–minor groove base pair hydrogen bonding and crystal packing interactions. This leaves the central 8 bp free of any direct crystal packing forces and allows a drug molecule to reside in this region of the groove, free of direct crystal packing constraints. There is no evidence of the drug molecule being disordered along the groove, where it adopts a unique position and orientation, similar to that seen for a number of other Hoechst 33258 complexes, even though the conformation of the phenol ring was ambiguous due to possible ring flipping. The drug is located in the sequence at the 5′-AATTC site, with the phenol ring lying towards the 5′ end of the sequence and the bulky piperazine group close to the C9·G16 base pair (the 'pip-down' orientation). The two possible conformations of the phenol ring have the hydroxyl group pointing directly either into or out of the groove. Both resulting structures show good geometry but only the 'hydroxyl in' structure shows any evidence of (weak) direct hydrogen bonding of the hydroxyl group to the DNA and also has a slightly lower R factor (20.2% compared with 20.4%). The former orientation has been predicted by NMR studies. The two possible conformations and the comparatively larger temperature factors for the phenol ring when compared with the rest of the molecule (Table 1) could be evidence of ring flipping occurring although this dynamic effect was not observed in the NMR studies of *meta*-Hoechst (40,41,43). Both the 'hydroxyl in' and 'hydroxyl out' structures, while not showing significant direct hydrogen bonding between hydroxyl group and DNA, contain a number of indirect hydrogen bonds to the DNA via extended water networks. The existence of the 'hydroxyl out' conformation is itself indicative of a weak propensity for the hydroxyl group to interact with DNA. We shall discuss the 'hydroxyl in' conformation in greater detail since this one has been predicted to be capable of directly recognising a guanine base edge.

Table 1. Average group temperature factors (Å^2) in the crystal structure of the *meta*-OH(N) Hoechst-d$(CGCGAATTCGCG)_2$ complex

Phenyl	Benzimidazole	Pyridylimidazole	Piperazine
38	28	22	26

DNA–drug hydrogen bonding and water structure

There are two sets of bifurcated hydrogen bonds between the edges of the bases and the two inner facing nitrogen atoms N1 and N3 of the benzimidazole and pyridylimidazole rings. These are

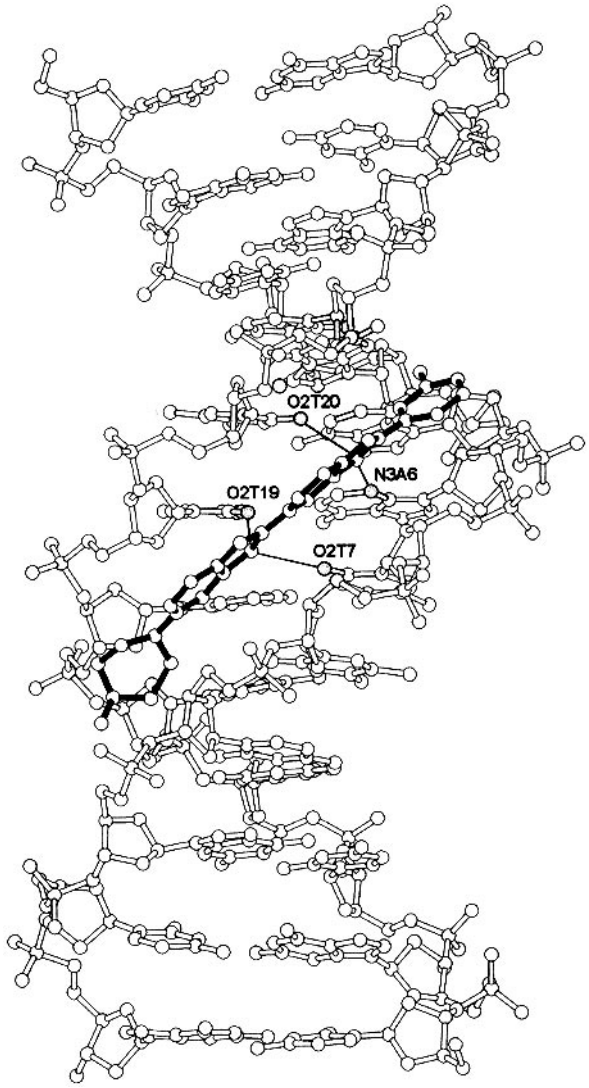

Figure 3. Computer plot of the complex, with the drug shown with filled-in bonds. Drug–DNA hydrogen bonds involving the imidazole groups are shown as thin lines, and the acceptor atoms on the DNA are shaded.

from N1 of the drug to O2 of T20 and N3 of A6, and from N3 of the drug to O2 of T19 and O2 of T7 (Fig. 4). These bifurcated hydrogen bonds result in the ligand being held in a binding site that is overall similar to that found in the various Hoechst 33258 X-ray structures where the same pattern of bifurcated hydrogen bonds is seen although the distances vary considerably (Table 2). The pyridyl nitrogen atom N7 does not take part in any hydrogen bonding with the DNA and so the present structure is highly likely to be similar to that of *meta*-Hoechst itself. N7 is in van der Waals contact (3.5 Å) with atom N3 on the edge of A18, on the floor of the groove. The *meta*-hydroxyl oxygen atom O3 makes two direct, albeit very weak, hydrogen bonds to N3 of A5 and to O2 of C21 (Table 4). O3 does hydrogen bond to two water molecules, W30 and W43, which themselves bridge between base edge atoms and O4′ atoms of a sugar ring (Fig. 4). There is no direct interaction between O3 and the 2-amino group of G4, in contrast with predictions from earlier studies (40–42). The OH ... N2G4 distance is 4.2 Å. The oxygen atom of this hydroxyl group is in a similar position to, and shows the same pattern of hydrogen bonding as the N7 hydrogen bond donor of distamycin in its complex with d(CGCAAATTTGCG)$_2$ (50). The water molecules W30 and W43 are part of an extensive hydrogen-bonded water network in and around the minor groove involving around 30 water molecules. This network, which extends into the major groove and also contains a hydrated magnesium ion, makes numerous contacts with the DNA including contacts to symmetry-related helices. An analogous network also stabilises the 'hydroxyl out' structure and may contribute to the stability of this alternative conformation, as observed in the crystal structure. A well-defined octahedral mass of electron density was assigned as the $Mg(H_2O)_6^{2+}$ ion with each of the co-ordinated waters hydrogen bonded to various sites on two symmetry-related helices. These magnesium co-ordinated waters form hydrogen bonds to N7 or O6 atoms of G2 or C22 on one helix, or to phosphate oxygen atoms of A6 or T7 on the other helix. A magnesium ion was also located at this same site in the low-temperature structures of Hoechst 33258 bound to d(CGCGAATTCGCG)$_2$ (33). At the other end of the minor groove, N6 of the piperazine group forms a bifurcated hydrogen bond to water molecules W62 and W65. These water molecules are part of a group of eight hydrogen-bonded waters which form a total of six hydrogen bonds to various DNA atoms including sugar, base and phosphate oxygen sites. One of the waters twice removed from N6 hydrogen bonds to O2PG10 of a symmetry-related helix. Another group of seven waters are found hydrogen-bonded via W84 to the benzimidazole atom N2. Again, one water twice removed from the ligand is hydrogen-bonded to the DNA at O3′C21 and another water hydrogen bonded to this one, itself hydrogen bonds to a symmetry-related phosphate oxygen. Many of the waters in these networks also fulfil a space-filling role.

Table 2. Drug–DNA hydrogen bond distances (Å) in *meta*-OH(N) Hoechst and selected Hoechst 33258 complexes with d(CGCGAATTCGCG)$_2$

	N1–N3A6	N1–O2T20	N3–O2T7	N3–O2T19
meta-H[a]	3.3	3.2	3.3	2.7
15[b]	3.2	2.8	3.7	3.0
0[c]	2.9	3.1	2.7	2.7
–25[c]	2.9	3.2	2.7	3.1
–100[c]	3.3	3.2	3.7	2.9

[a]Present *meta*-OH(N) Hoechst structure.
[b]Room-temperature structure (32).
[c]Low-temperature structures (33).

Drug conformation

The degree of drug twisting between each of its component groups is quite different from that observed for other Hoechst 33258/d(CGCGAATTCGCG)$_2$ structures (31–33) (Table 3), although the torsion angles between phenyl and benzimidazole rings are similar to those found in the –100°C low temperature (LT) structure (33). They are also close to those found in the NMR study of Hoechst 33258 complexed with the sequence d(GTGGAATTC-CAC)$_2$ (52,53). The LT studies found that lowering the temperature resulted in the two benzimidazole groups becoming progressively more coplanar. It was suggested (33) that at low temperatures the molecule would have insufficient energy to overcome the barrier to rotation between the benzimidazole ring systems, so that they could become more coplanar as a result of the favourable delocalisation of electron density. This coplanarity results in a poor fit to the walls of the minor groove and so at higher temperatures

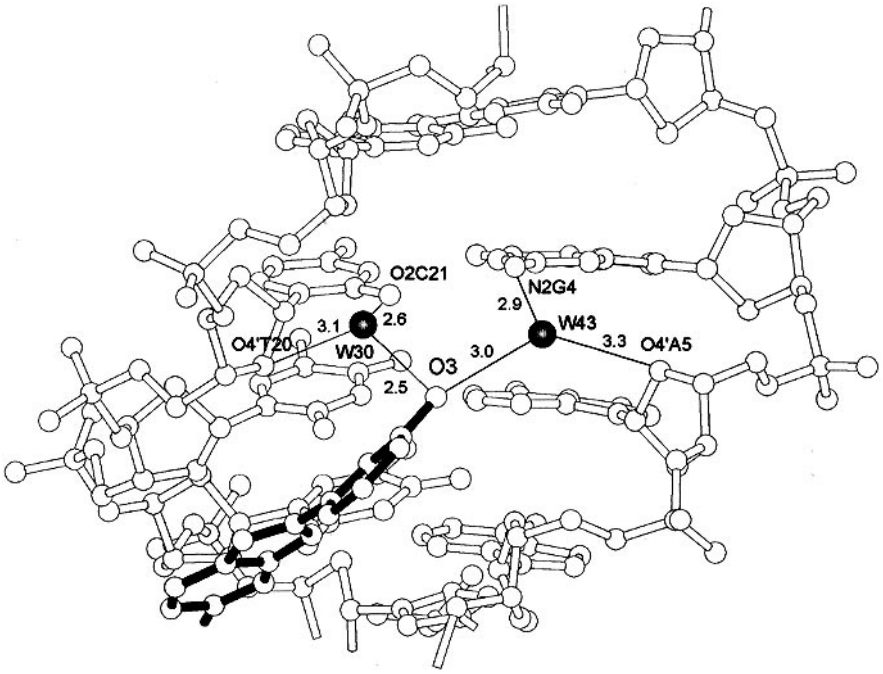

Figure 4. Plot of the immediate environment around the *meta* hydroxyl atom of the drug, with donor atoms on the DNA being shaded. Hydrogen bonds involving water molecules W30 and W43 are shown as thin lines. Distances are in Å.

groove contacts favour a twist at each linker. The cause of the large phenol–benzimidazole torsion angle for the –100°C structure was unclear. It is possible that as the benzimidazole and pyridylimidazole rings become coplanar, groove wall contacts with the phenyl ring become unfavourable and the conjugation between phenyl and benzimidazole is lost. The torsion angles for the present ambient temperature structure can be explained by a different mechanism. Here, in order for a number of interactions to take place simultaneously, i.e. the phenyl hydroxyl to form hydrogen bonds with water molecules and weak contacts made with the DNA, for the benzimidazole and pyridylimidazole hydrogen bond donors to remain in phase with the base acceptor sites, and for the ligand to follow the groove topology, the molecule is forced to adopt the observed conformation. Whereas the delocalisation between phenyl and benzimidazole groups is weakened, that between the benzimidazole and pyridylimidazole moieties is favourable. This type of conformational change between RT Hoechst and *meta*-OH species was predicted previously by the NMR studies of *meta*-Hoechst (40,41).

Table 3. Inter-ring torsion angles (°), defined in terms of the ligand atoms facing into the minor groove, for *meta*-OH(N) Hoechst and selected Hoechst 33258 complexes

	Phenol–Benzimidazole	Benzimidazole–Benzimidazole
meta-H[a]	36	8
15[b]	0	28
0[c]	3	14
–25[c]	–11	12
–100[c]	24	0

[a]Present *meta*-OH(N) Hoechst structure, where the second benzimidazole has been replaced by a pyridylimidazole group.
[b]Room-temperature structure (32).
[c]Low-temperature structures (33).

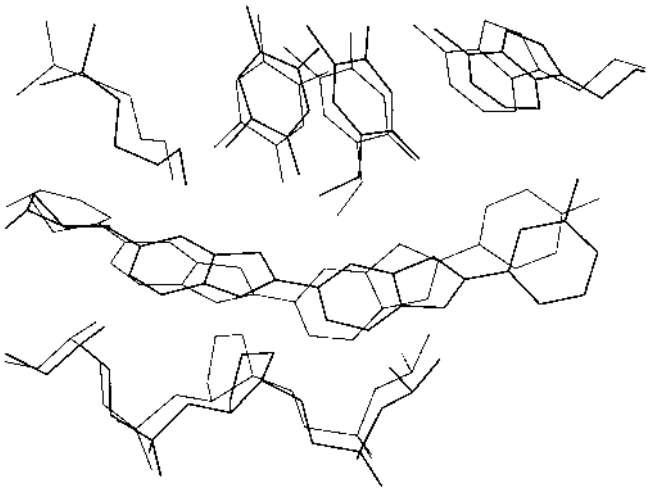

Figure 5. Least squares superposition of the present structure (bold) and the 15°C Hoechst structure, illustrating their conformational differences but overall similarity in binding site.

There are numerous close contacts between the drug molecule and the minor groove (Table 4), which reflect the conformation that the drug has adopted. The drug conformations in the present structure and the analogous ambient temperature Hoechst 33258 structure with the same DNA sequence (32) are compared in Figure 5. The binding site in the present structure is very similar to that of the Hoechst one but with the ligand moved slightly towards the 5′ end of the duplex as a result of the formation of the phenol hydrogen bonds to the DNA. The out-of-plane twist of the phenol ring with respect to the rest of the drug molecule is necessary in order for these hydrogen bonds to form and this twisting and the small movement up the groove alters the

sugar–phenol–sugar π–π interaction seen in the Hoechst structure. The *meta*-hydroxyl group also limits the depth that the phenol ring can protrude into the groove. The imidazole moiety of the adjacent benzimidazole is sited similarly to the Hoechst structure but its phenyl moiety binds deeper in the groove due to the ligand rotating into the groove at the opposite end to the phenol ring. This rotation and the unprotonated N7 atom allow the attached pyridylimidazole group to approach significantly closer to the floor of the groove than the analogous benzimidazole group in the Hoechst complex. This results in the formation of the N3 ... O2T7 hydrogen bond, which was not observed in the 15°C Hoechst structure (32). The small shift of the ligand up the groove is also instrumental in the formation of this hydrogen bond. The first benzimidazole group is twisted towards base A6 more than in the Hoechst structure, which is found closer to T20. The pyridylimidazole group is twisted towards T19 and away from T7. This twisting results in the particular hydrogen bonding pattern and distances discussed above. The piperazine rings in the two structures show similar conformations and positions in the groove although the *meta*-Hoechst piperazine twists somewhat more and has closer approaches to the groove walls.

Table 4. Van der Waals contact distances (Å) between *meta*-OH(N) Hoechst and d(CGCGAATTCGCG)$_2$

Drug atom	DNA atom	Distance
O3	N3A5	3.3*
O3	C2A5	3.4
O3	O2C21	3.4*
C3	O4′A6	3.5
N1	N3A6	3.3*
N1	O2T20	3.2*
N1	O4′T7	3.4
C13	O2T7	3.2
C13	O2T19	3.3
N3	O2T7	3.3*
N3	O2T19	2.7*
N7	N3A18	3.5
N7	C4′T19	3.4
N7	O4′T19	3.1
C15	O4′T19	3.5
C18	C5′C9	3.4
C19	C4′T19	3.4
C23	O4′C9	3.3
C23	O2C9	3.4
C24	O4′C9	3.2
C25	C5′G10	3.2

*Indicates hydrogen bond.

Further evidence of differences between the structures is shown by the minor groove widths (Fig. 6). The drug-induced minor groove widening for *meta*-OH(N) Hoechst shows a trend similar to that found in the 15°C Hoechst structure. The present structure shows significant groove widening at the phenol binding site, greater than that observed for both Hoechst structures (32,33) in part because of the large dihedral angle between phenol and benzimidazole and the resulting groove contacts, and in part on account of the influence of the two water molecules W30 and W43. Over the binding sites of the benzimidazole and pyridylimidazole groups the groove widening is lessened but still remains large and is again indicative of the somewhat poor fit of these coplanar

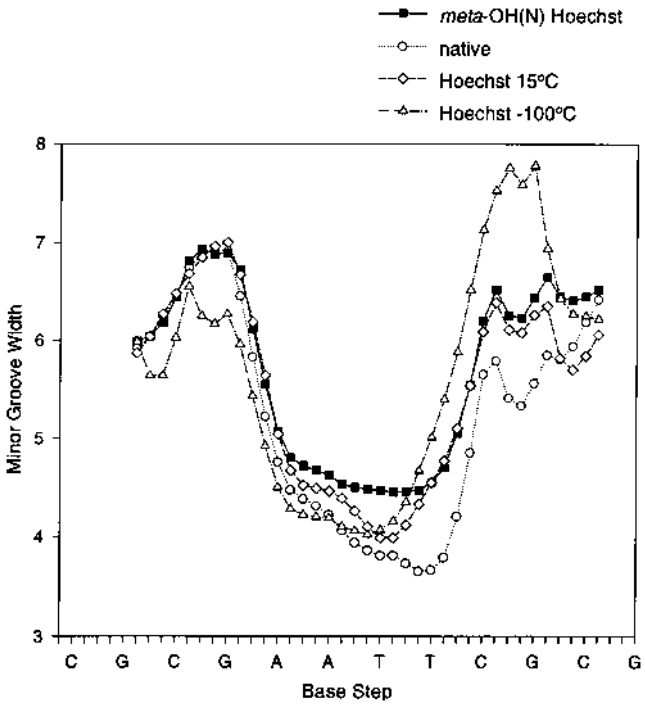

Figure 6. Plot of minor groove widths based on C1′ atoms, for the present, 15°C Hoechst, –100°C Hoechst and native structures.

moieties to the groove curvature. The ligand is kept centred in the groove by this widening. The LT structure, while having a drug conformation similar to that in the present structure, shows a significantly different pattern of groove width variation, which is presumably the result of the lower data collection temperature. All the complexes show very significant groove widening at the piperazine binding site due to the bulk of that ring system. Drug-induced groove widening is also greater at the G12·C13 end of the helix due to the greater freedom in crystal packing here compared with the other end of the helix which is held more tightly in place by neighbouring molecules.

Table 5. Inter-base pair twist (°) and rise (Å) for *meta*-OH(N) Hoechst and selected dodecamer structures

Step	Twist			Rise		
	meta-H[a]	Hoe[b]	nat[c]	*meta*-H	Hoe	nat
C1/G2	35	44	37	3.6	3.7	3.4
G2/C3	37	35	37	3.5	3.7	3.5
C3/G4	29	34	28	3.3	3.3	3.4
G4/A5	39	30	39	3.5	3.5	3.2
A5/A6	**36**	**47**	**36**	**3.4**	**3.5**	**3.4**
A6/T7	**33**	**25**	**33**	**3.2**	**2.9**	**3.3**
T7/T8	**36**	**42**	**36**	**3.5**	**4.1**	**3.2**
T8/C9	**39**	**33**	**40**	**3.5**	**3.1**	**3.5**
C9/G10	**29**	**35**	**31**	**3.2**	**3.2**	**3.2**
G10/C11	41	38	39	3.4	3.6	3.7
C11/G12	39	33	35	3.7	3.7	3.1

[a]Present *meta*-OH(N) Hoechst structure.
[b]Room-temperature Hoechst structure (32).
[c]Native (52) dodecamer structure.
The drug binding site in the present structure is highlighted in bold type.

Helical parameters for this structure, the Hoechst 33258 structures and the native dodecamer were calculated using the program CURVES, version 5 (51). A comparison of these shows that the present structure is more like the native dodecamer than the 15°C Hoechst one. The inter-base pair twist and rise and also the propeller twist over the drug binding site are indicative of the overall weak hydrogen bonding between drug and DNA and the effect of drug binding on DNA structure (Table 5 and Table 6). There are larger propeller twists at G4·C21 and A5·C20 in the present structure compared with the native one. These are indicative of the phenol OH bifurcated hydrogen bonds to the two water molecules W30 and W43. Overall, the groove widening observed for this structure does not appear to have a major effect on the base and base-pair morphological parameters. It is not obvious what particular parameters give rise to the groove widening, rather it is likely that a combination of them is responsible.

Table 6. Propeller twist (°) in *meta*-OH(N) Hoechst and selected structures

Step	Propeller twist		
	meta-H	Hoechst	native
C1–G24	–10	–2	–13
G2–C23	–17	–9	–12
C3–G22	–10	–7	–6
G4–C21	–13	–2	–10
A5–T20	**–23**	**–10**	**–17**
A6–T19	**–20**	**–27**	**–21**
T7–A18	**–18**	**–17**	**–18**
T8–A17	**–22**	**–7**	**–19**
C9–G16	**–10**	**–9**	**–13**
G10–C15	3	–9	0
C11–G14	–29	–21	–16
G12–C13	12	–6	–11

The bases involved in hydrogen bonding to the drug in the *meta*-OH(N) Hoechst structure are highlighted in bold.

Minor groove binding of *meta*-OH(N) Hoechst

Hoechst 33258 and *meta*-OH(N) Hoechst have very similar structures. Both contain a protonated, non-planar saturated heterocyclic ring system at one end of the molecule and a phenolic ring at the other. The benzimidazoles of the former and the benzimidazole/pyridylimidazole rings of the latter bind to the 5′-AATTC site in the minor groove by a combination of non-bonded interactions to the walls and floor of the groove and bifurcated hydrogen bonding to adenine N3 and thymine O2 atoms. The *meta*-OH substitution has little effect on the overall binding site of the drug, with the bulky N-heterocyclic ring requiring a widened groove for binding and the remainder of the drug favouring the non-bonded interactions associated with the narrow AATT region. This similarity in binding site and in the binding interactions overall is in accord with the closely similar DNA binding (and antitumour) properties of the two drugs (41,42). Thus, measurements of the extent of thermal stabilisation of the double helix gave essentially identical T_m values for them (41). The principal difference seen with *meta*-OH(N) Hoechst compared with the other Hoechst 33258 complexes is the (weak) bifurcated hydrogen bonding to N3 of A5 and to O2 of C21. This is distinct from the bifurcated bonding to the 2-amino groupof G4 and O2 of C21 predicted by NMR (40,41). The phenol ring flipping indicated by the X-ray structure is not evident in the NMR studies of *meta*-Hoechst. This poses the question as to whether or not ring flipping is occurring in solution and is not being detected, or whether the ring flipping seen in the X-ray structure is solely an artefact of the crystallisation process. The weakness of the phenol OH hydrogen bonding to the DNA as seen in the present structure may well be an indication that phenol ring flipping can occur in solution although this cannot be automatically assumed from the present solid state crystal structure.

The position of the hydroxyl group being *meta* forces the drug away from the floor of the minor groove at the phenol binding site and rotates the rest of the molecule towards the groove for effective binding. The pyridylimidazole moiety is in closer contact with the groove floor than the second benzimidazole in the parent Hoechst 33258 because of this rotation and the presence of the non-protonated N7 atom. This closer approach allows the formation of the N3 ... O2T7 hydrogen bond, which was not seen in the 15°C Hoechst structure. The replacement of a carbon atom in Hoechst with N7 was designed to allow a greater tolerance of G·C base pairs (42,46). The unprotonated N7 atom could tolerate the increased steric hindrance of the exocyclic amino group at G2, compared with the C2 hydrogen atom on adenine bases, as well as the possibility of N7 ... NH_2G hydrogen bond formation. These interactions do not occur in the present structure.

The drug–DNA hydrogen bonding, while probably being a secondary factor in the binding, has an influence on the drug conformation. The formation of the phenol–DNA hydrogen bonds and the requirement that the molecule follow the groove topology and also keep its hydrogen bond donors in phase with the DNA acceptor sites at the edges of the base pairs, forces the *meta*-OH(N) Hoechst molecule into a conformation quite different from that of other Hoechst structures in 1:1 ligand–DNA complexes. This results in groove widening over the binding sites of the phenol and benzimidazole rings greater than that seen in the 15°C Hoechst structure. The most significant increase in minor groove width is found at the phenol binding site because of the high ring twist and shallow binding depth. The present structure also highlights the importance of water networks, not only in stabilising drug binding but also in stabilising the DNA double helical structure with very extensive hydrogen bonding networks linking symmetry-related helices as well as forming intramolecular bridges.

Sequence specificity of binding

The observation of *meta*-OH(N) Hoechst binding in the same site as Hoechst 33258 itself, i.e. mostly in the narrow groove AT region, provides further support for the hypothesis that binding of such molecules to DNA is largely dominated by groove width factors, with complementarity of drug narrow cross-section and surface of groove walls (24,25,27–30). The piperazine ring is the sole moiety of the Hoechst 33258 and *meta*-OH(N) Hoechst molecules which has a pronounced requirement to bind in the wider minor groove of the GC region in these complexes. The *meta* hydroxyl group cannot achieve GC recognition in the minor groove for two reasons: (i) the phenyl ring has too narrow a cross-section for effective binding to a wide GC-rich minor groove, and most importantly, (ii) the hydroxyl group is positioned insufficiently deeply into the groove for interaction with a guanine exocyclic amino group. The present structure demonstrates that it is almost equivalent in position to the inner-facing nitrogen atom of

a benzimidazole or pyridylimidazole ring, and so its weak bifurcated hydrogen-bonding to N3A5 and to O2C21 is unsurprising. Effective GC recognition would require a substituent at this end of the bis-benzimidazole core that addresses both of the problems posed by *meta*-OH(N) Hoechst, i.e. that has sufficient size to contact the walls of the widened groove, as well as having substituents that reach the groove floor.

ACKNOWLEDGEMENTS

This work was supported by a CRC programme grant to S.N. GRC is grateful to the University of Auckland for sabbatical leave at The Institute of Cancer Research when this work was performed. We are grateful to Christine Nunn, John Trent and Alexis Wood (ICR) for discussions and assistance on crystallographic aspects of this work.

REFERENCES

1 Zimmer, C. and Wähnert, U. (1986) *Prog. Biophys. Mol. Biol.*, **47**, 31–112.
2 Dervan, P.B. (1986) *Science*, **232**, 464–471.
3 Kopka, M.L. and Larsen, T.A. (1992) In Propst, C.L. and Perun, T.J. (eds) *Nucleic Acid Targeted Drug Design*, Marcel Dekker, Inc., New York, pp. 304–374.
4 Gravatt, G.L., Baguley, B.C., Wilson, W.R. and Denny, W.A. (1994) *J. Med. Chem.*, **37**, 4338–4345.
5 Boykin, D.W., Kumar, A., Spychala, J., Zhou, M., Lombardy, R.J., Wilson, W.D., Dykstra, C.C., Jones, S.K., Hall, J.E., Tidwell, R.R., Laughton, C., Nunn, C.M. and Neidle, S. (1995) *J. Med. Chem.*, **38**, 912–916.
6 Cory, M., Tidwell, R.R. and Fairley, T.A. (1992) *J. Med. Chem.*, **35**, 431–438.
7 Tidwell, R.R., Jones, S.K., Geratz, J.D., Ohemeng, K.A., Cory, M. and Hall, J.E. (1990) *J. Med. Chem.*, **33**, 1252–1257.
8 Broggini, M., Erba, E., Ponti, M., Ballinari, D., Geroni, C., Spreafico, F. and D'Incalci, M. (1991) *Cancer Res.*, **51**, 199–204.
9 Prakash, A.S., Valu, K.K., Wakelin, L.P.G., Woodgate, P.D. and Denny, W.A. (1991) *Anti-Cancer Drug Des.*, **6**, 195–206.
10 Beerman, T.A., McHugh, M.M., Sigmund, R., Lown, J.W., Rao, K.E. and Bathini, Y. (1992) *Biochim. Biophys. Acta*, **1131**, 53–61.
11 Chen, A.Y., Chiang, Y., Gatto, B. and Liu, L.F. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 8131–8135.
12 Dorn, A., Affolter, M., Müller, M., Gehring, W.J. and Leupin, W. (1992) *EMBO J.*, **11**, 279–286.
13 Chiang, S.-Y., Welch, J., Rauscher III, F.J. and Beerman, T.A. (1994) *Biochemistry*, **33**, 7033–7040.
14 Welch, J.J., Rauscher III, F.J. and Beerman, T.A. (1994) *J. Biol. Chem.*, **269**, 31051–31058.
15 Henderson, D., and Hurley, L.H. (1995) *Nature Med.*, **1**, 525–527.
16 Jenkins, T.C., Lane, A.N., Neidle, S. and Brown, D.G. (1993) *Eur. J. Biochem.*, **213**, 1175–1184.
17 Blaskó, A., Browne, K.A. and Bruice, T.C. (1994) *J. Am. Chem. Soc.*, **116**, 3726–3737.
18 Kopka, M.L., Yoon, C., Goodsell, D., Pjura, P. and Dickerson, R.E. (1985) *J. Mol. Biol.*, **183**, 553–563.
19 Coll, M., Frederick, C.A., Wang, A.H.-J. and Rich, A. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 8385–8389.
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 Edwards, K.J., Jenkins, T.C. and Neidle, S. (1992) *Biochemistry*, **31**, 7104–7109.
22 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.
23 Nunn, C.M., Jenkins, T.C. and Neidle, S. (1993) *Biochemistry*, **32**, 13838–13843.
24 Wood, A.A., Nunn, C.M., Czarny, A., Boykin, D.W. and Neidle, S. (1995) *Nucleic Acids Res.*, **23**, 3678–3684.
25 Neidle, S. (1992) *FEBS Lett.*, **298**, 97–99.
26 Nunn, C.M. and Neidle, S. (1995) *J. Med. Chem.*, **38**, 2317–2325.
27 Dwyer, T.J., Geierstanger, B.H., Bathini, Y., Lown, J.W. and Wemmer, D.E. (1992) *J. Am. Chem. Soc.*, **114**, 5911–5919.
28 Mrksich, M., Wade, W.S., Dwyer, T.J., Geierstanger, B.H., Wemmer, D.E. and Dervan, P.B. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 7586–7590.
29 Mrksich, M. and Dervan, P.B. (1993) *J. Am. Chem. Soc.*, **115**, 2572–2576.
30 Geierstanger, B.H., Jacobsen, J.P., Mrksich, M., Dervan, P.B. and Wemmer, D.E. (1994) *Biochemistry*, **33**, 3055–3062.
31 Pjura, P.E., Grzeskowiak, K. and Dickerson, R.E. (1987) *J. Mol. Biol.*, **197**, 257–271.
32 Teng, M., Usman, N., Frederick, C.A. and Wang, A.H.-J. (1988) *Nucleic Acids Res.*, **16**, 2671–2690.
33 Quintana, J.R., Lipanov, A.A. and Dickerson, R.E. (1991) *Biochemistry*, **30**, 10294–10306.
34 Carrondo, M.A.A.F. de C.T., Coll, M., Aymami, J., Wang, A.H.-J., van der Marel, G.A., van Boom, J.H., and Rich, A. (1989) *Biochemistry*, **28**, 7849–7859.
35 Spink, N., Brown, D.G., Skelly, J.V. and Neidle, S. (1994) *Nucleic Acids Res.*, **22**, 1607–1612.
36 Vega, M.C., Saez, I.G., Aymami, J., van der Marel, G.A., van Boom, J.H., Rich, A. and Coll, M. (1994) *Eur. J. Biochem.*, **222**, 721–726.
37 Sriram, M., van der Marel, G.A., Roelen, H.L.P., van Boom, J.H. and Wang, A.H.-J. (1992) *EMBO J.*, **11**, 225–232.
38 Bathini, Y., Rao, K.E., Shea, R.G. and Lown, J.W. (1990) *Chem. Res. Toxicol.*, **3**, 268–280.
39 Rao, K.E., Shea, R.G., Yadagiri, B. and Lown, J.W. (1990) *Anti-Cancer Drug Des.*, **5**, 3–20.
40 Parkinson, J.A., Ebrahimi, S.E., McKie, J.H. and Douglas, K.T. (1994) *Biochemistry*, **33**, 8442–8452.
41 Ebrahimi, S.E., Bibby, M.C., Fox, K.R. and Douglas, K.T. (1995) *Anti-Cancer Drug Des.*, **10**, 463–479.
42 Leupin, W., Bur, D., Dorn, A., Ji, Y.-H., Labhardt, A., Fede, A., Billeter, M. and Wüthrich, K. (1994) *Actualité Chim. Thérap.*, **21**, 153–170.
43 Parkinson, J.A., Barber, J., Buckingham, B.A., Douglas, K.T. and Morris, G.A. (1992) *Magn. Reson. Chem.*, **30**, 1064–1069.
44 Searle, M.S. and Embrey, K.J. (1990) *Nucleic Acids Res.*, **18**, 3753–3762.
45 Embrey, K.J., Searle, M.S. and Craik, D.J. (1993) *Eur. J. Biochem.*, **211**, 437–447.
46 Kumar, S., Yadagiri, B., Zimmermann, J., Pon, R.T. and Lown, J.W. (1990) *J. Biomol. Struct. Dyn.*, **8**, 331–357.
47 Westhof, E. (1987) *J. Biomol. Struct. Dyn.*, **5**, 581–600.
48 Brünger, A.T., Kuriyan, J. and Karplus, M. (1987) *Science*, **235**, 458–460.
49 Insight II, Version 2.3.0, Biosym Technologies (1993).
50 Coll, M., Frederick, C.A., Wang, A.H.-J. and Rich, A. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 8385–8389.
51 Lavery, R. and Sklenar, H. (1988) *J. Biomol. Struct. Dyn.*, **6**, 63–91.
52 Fede, A., Labhardt, A., Bannworth, W. and Leupin, W. (1991) *Biochemistry*, **30**, 11377–11388.
53 Fede, A., Billeter, M., Leupin, W. and Wüthrich, K. (1993) *Structure*, **1**, 177–186.