Acta Cryst. (1993). D**49**, 458–467

Anthracycline–DNA Interactions at Unfavourable Base-Pair Triplet-Binding Sites: Structures of d(CGGCCG)/Daunomycin and d(TGGCCA)/Adriamycin Complexes

BY GORDON A. LEONARD

Department of Chemistry, University of Manchester, Oxford Road, Manchester M13 9PL, England

TREVOR W. HAMBLEY

School of Chemistry, University of Sydney, 2006 Sydney, New South Wales, Australia

KATHERINE MCAULEY-HECHT AND TOM BROWN

Department of Chemistry, University of Edinburgh, King's Buildings, West Mains Road, Edinburgh EH9 3JJ, Scotland

AND WILLIAM N. HUNTER*

Department of Chemistry, University of Manchester, Oxford Road, Manchester M13 9PL, England

(*Received* 25 *February* 1992; *accepted* 24 *May* 1993)

Abstract

The structures of two hexanucleotide–anthracycline complexes d(CGGCCG)/daunomycin and d(TGGCCA)/adriamycin have been determined using single-crystal X-ray diffraction techniques. In both cases the anthracycline molecule is bound to non-preferred d(YGG) base-pair triplet sites. For both complexes the crystals are tetragonal and belong to the space group $P4_12_12$. Unit-cell dimensions are $a = 28.07$ (2), $c = 53.35$ (1) and $a = 28.01$ (1), $c = 52.99$ (1) Å, respectively, and the asymmetric unit of both structures consists of one strand of DNA, one drug molecule and approximately 50 water molecules. For the d(CGGCCG) complex the refinement converged with an R factor of 0.21 for 1108 reflections with $F \geq 2\sigma(F)$ in the resolution range 7.0–1.9 Å. For the complex involving d(TGGCCA) the final R value was 0.22 for 1475 reflections in the range 7.0–1.7 Å with the same criterion for observed data. Both complexes are essentially isomorphous with related structures but differ in terms of the number of favourable van der Waals interactions of the amino sugars of the drug molecules with the DNA duplexes and the formation in the minor groove of heterodromic pentagonal arrangements of hydrogen bonds involving water molecules which link the amino sugars to the DNA. These differences in structure are used to rationalize the lack of affinity of daunomycin-type anthracyclines for d(YGG) and d(YGC) sites.

* To whom correspondence should be addressed.

Introduction

Anthracyclines are potent antibiotics and in some cases also serve as antitumour agents. This class of compound binds to DNA and the DNA-binding protein topoisomerase II and the biological activities are generally ascribed as being primarily as a result of their interactions with DNA (Valentini, Nicollela, Vannini, Menozzi, Penco & Arcamone, 1985). Although they have been used clinically in the treatment of certain disorders, they possess serious side effects, notably cardiotoxicity. It is therefore important to characterize the structures of DNA–anthracycline complexes in order to understand fully the nature of their interactions with DNA which might then allow the rational design of more selective DNA-binding agents.

The most studied of the anthracyclines are daunomycin, adriamycin and 4′-epiadriamycin (Fig. 1). Various methods including theoretical studies (Chen, Grish & Pullman, 1985; Pullman, 1987*a*,*b*), DNA-footprinting studies (Chaires, Herrera & Waring, 1990; Skavobogaty, White, Phillips & Reiss, 1988*a*,*b*) and ultraviolet melting (Leonard, Brown & Hunter, 1992) have been used to study the interactions of anthracyclines with DNA and have resulted in evidence that suggests preferential binding to d(CGA), d(CGT), d(TGA) and d(TGT) base-pair triplets, and that d(YGG) and d(YGC) base-pair triplets are particularly unfavourable sites. Recently, reports of the structures of a number of DNA–anthracycline complexes have appeared (Wang, Ughetto, Quigley & Rich, 1987; Moore, Hunter,

ISSN 0907-4449

Langlois d'Estaintot & Kennard, 1989; Frederick, Williams, Ughetto, van der Marel, van Boom, Rich & Wang, 1990; Williams, Frederick, Ughetto & Rich, 1990; Nunn, van Meervelt, Zhang, Moore & Kennard, 1991; Langlois d'Estaintot, Gallois, Brown & Hunter, 1992; Leonard, Brown & Hunter, 1992). These reveal that upon binding to DNA the aglycone chromophore of anthracycline molecules intercalates itself at the first pyrimidine–purine step of the base-pair triplets and that the amino sugar interacts, either directly or *via* solvent-mediated hydrogen bonds, with the minor groove of the last base pair in the triplet. The orientation of the amino sugar with respect to the aglycone moiety is flexible (Langlois d'Estaintot, Gallois, Brown & Hunter, 1992; Leonard, Brown & Hunter, 1992) and it has been suggested that this enables the drug molecules to enhance their binding to the variety of preferred sequences mentioned above. To date, all the structures of the DNA–anthracycline complexes reported concern drug binding to what we term preferential sequences. No structures have appeared in which an anthracycline moiety is complexed to a DNA duplex containing a 'non-preferred' binding site. From theoretical calculations (Chen, Grish & Pullman, 1985), it has been suggested that the lack of affinity of anthracycline molecules for d(YGG) or d(YGC) base-pair triplets is a result of the repulsion of the amino sugar by the presence of the exocyclic N2 amino group on the minor groove side of the third base pair (G·C) in the triplets. This amino group is, of course, not present in A·T or T·A base pairs and the effect of this repulsion would be to force the amino sugar away from the minor groove resulting in less tight binding to sequences with a G·C or C·G base pair as the third base pair in the binding site. There has, as yet, been no structural evidence to support or undermine this hypothesis and hence, as part of an ongoing research project designed to further our understanding of the interactions of anthracycline antibiotics with different DNA sequences, we have co-crystallized the synthetic hexadeoxynucleotides d(CGGCCG) and d(TGGCCA) with, respectively, daunomycin and adriamycin, and we report here the three-dimensional structures of the two complexes as determined by means of single-crystal X-ray diffraction techniques using data to resolutions of 1.9 and 1.7 Å, respectively.

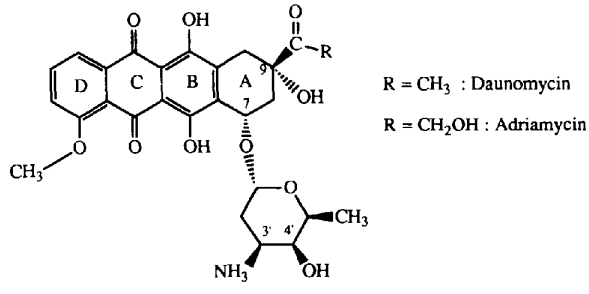


Fig. 1. The molecular formulae of daunomycin and adriamycin. The difference between adriamycin and 4′-epiadriamycin is an inversion of the stereochemistry at the 4′-position of the amino sugar. The aglycone ring consists of three unsaturated rings (*B*, *C*, *D*) and one semi-saturated ring (*A*).

Experimental

Both d(CGGCCG) and d(TGGCCA) were synthesized on an Applied Biosystems 380B DNA synthesizer using phosphoramidite methodology (McBride & Caruthers, 1983) and purified by reversed-phase high-pressure liquid chromatography (Brown & Brown, 1991). Crystals of both complexes were grown by vapour diffusion from drops sitting in Corning glass plates (McPherson, 1982). In both cases, moderately sized red blocks grew within 14 days from 30 μl drops that contained the hexanucleotide (2.3 m*M*), the anthracycline (5.0 m*M*), sodium cacodylate, pH 6.6 (33 m*M*), magnesium chloride (5.0 m*M*), spermine tetrahydrochloride (2.0 m*M*) and 2-methyl-2,4-pentanediol (MPD) [8.3% (*v*/*v*)], which were equilibrated against an external reservoir containing 100 μl of 100% MPD. The crystals of the d(CGGCCG)/daunomycin complex were grown at 277 K and those of the d(TGGCCA)/adriamycin complex at 285 K. Single crystals of approximate dimensions 0.5 × 0.25 × 0.1 mm were sealed in glass capillaries along with a drop of mother liquor and were used for X-ray data collection. We note that the crystals obtained for these complexes are significantly smaller than those we have obtained for other hexanucleotide–anthracycline complexes (Langlois d'Estaintot, Gallois, Brown & Hunter, 1992; Leonard, Brown & Hunter, 1992). Precise details of the data collection for each complex are given in Table 1. In summary, the data were collected on a Rigaku AFC-5*R* four-circle diffractometer using a Rigaku RU200 rotating anode, operating at 50 kV, 140 mA and producing Cu $K\alpha$ radiation (graphite monochromator, λ = 1.5418 Å) with a focal spot of 0.5 mm. The crystal-to-detector distance was 400 mm and a continuously evacuated beam tunnel was used to reduce the absorption of the diffracted X-rays by air. Accurate unit-cell dimensions were obtained from a least-squares fit of 25 reflections in the range $13 < 2\theta < 28°$. Intensities were measured with ω scans of 1° (scan speed 4° min^{-1}) and those reflections with $F < 10\sigma(F)$ were measured in triplicate to improve counting statistics. All data were corrected for Lorentz and polarization effects and an empirical absorption correction was applied (North, Phillips & Mathews, 1968). Three standard reflections (measured every 150 reflections) were used to moni-

tor and correct for crystal decay. Data collection and processing were carried out using the *TEXSAN* program (Molecular Structure Corporation, 1989).

The unit-cell dimensions and space groups (Table 1) for both crystals indicated isomorphism with a series of hexanucleotide–anthracycline complexes and hence the starting models used for the structure refinement of the d(CGGCCG)/daunomycin complex were the refined coordinates for the d(CGTACG)/daunomycin complex (Frederick, Williams, Ughetto, van der Marel, van Boom, Rich & Wang, 1990) and those for the d(TGGCCA)/adriamycin complex were the refined coordinates for the d(CGATCG)/daunomycin complex (Wang, Ughetto, Quigley & Rich, 1987). The refinement procedure adopted (Table 1) was similar to that used for other DNA–drug complexes (Langlois d'Estaintot, Gallois, Brown & Hunter, 1992; Leonard, Brown & Hunter, 1992). Briefly, in the first part of the refinement the starting model was refined as a rigid body using a modified version of *SHELX* (Sheldrick, 1976). Initially, all data with $F > 0$ in the resolution range 10.0–7.0 Å were used and this was increased in 1 Å segments until all reflections in the resolution range 10.0–3.0 Å were included. The rigid-body refinement was followed by positional refinement using restrained least-squares methods (Hendrickson & Konnert, 1981) with the program *NUCLSQ* (Westhof, Dumas & Moras, 1985) and data in the range 7.0–2.5 Å with $F \geq 2\sigma(F)$. Electron-density ($2F_o - F_c$) and difference ($F_o - F_c$) maps were calculated and displayed on an Evans and Sutherland graphics workstation using *FRODO* (Jones, 1978; P. R. Evans, personal communication). At this stage the sequence of the DNA was altered to the correct sequence (all changes being consistent with the electron-density maps) and modifications to the sugar/phosphate backbone and the drug molecules were made to improve the fit of the model to the density. The refinement was continued using *NUCLSQ* with the extension of the data to the resolution limit available, the inclusion of restrained individual isotropic temperature factors and the progressive addition of solvent molecules. Both electron-density and difference maps were examined at regular intervals and used to guide modifications to both the DNA and the drug molecules. Solvent molecules were included only if they appeared with approximately spherical density on both electron-density and difference maps and had good hydrogen-bonding geometry with respect to other atoms already included in the model. This part of the refinement was followed by a stage using the *X-PLOR* software (Brünger, 1990; Brünger, Karplus & Petsko, 1989) which consisted of energy minimization of the structure, simulated annealing using a slow-cooling protocol (Table 1) and one round of Powell minimization. For both crystals this resulted in a significant improvement in the electron-density maps and allowed a reappraisal of both the conformations of the drug molecules and the positions of the solvent

Table 1. *Crystal data, data-collection statistics and refinement details for d(CGGCCG)/daunomycin and d(TGGCCA)/adriamycin complexes*

	d(CGGCCG)/ daunomycin	d(TGGCCA)/ adriamycin
Crystal data		
Chemical formula	$C_{84}H_{100}N_{22}O_{44}P_5.49H_2O$	$C_{83}H_{101}N_{24}O_{44}P_5.43H_2O$
Molecular weight	2319 (DNA/daunomycin only)	2317 (DNA/adriamycin only)
Crystal system	Tetragonal	Tetragonal
Space group	$P4_12_12$	$P4_12_12$
$a = b$ (Å)	28.07 (2)	28.01 (1)
c (Å)	53.35 (1)	52.99 (1)
V (Å^3)	42035.8	41573.8
Z	8	8
Radiation type	Cu $K\alpha$	Cu $K\alpha$
Wavelength (Å)	1.54178	1.54178
No. of reflections for cell parameters	25	25
θ range for cell parameters (°)	6.5–14	6.5–14
Temperature	Ambient	Ambient
Crystal colour	Red	Red
Crystal form	Prism	Prism
Crystal size (mm)	0.5 × 0.25 × 0.1	0.5 × 0.25 × 0.1
Data collection		
Diffractometer	Rigaku AFC-5*R*	Rigaku AFC-5*R*
Data-collection method	ω scans	ω scans
Absorption correction	Empirical	Empirical
T_{min}	0.60	0.73
T_{av}	0.75	0.82
No. of measured reflections	3252 (excluding standards)	4435 (excluding standards)
No. of independent reflections	1742	2433
No. of observed reflections	1108	1475
Observation criterion	$F \geq 2\sigma(F)$	$F \geq 2\sigma(F)$
R_{merge}* [$0.5\sigma(I)$ data only]	0.07	0.07
θ_{max} (°)	24	27
Range of h, k, l	$0 \rightarrow h \rightarrow 14$ $0 \rightarrow k \rightarrow 9$ $0 \rightarrow l \rightarrow 26$	$0 \rightarrow h \rightarrow 16$ $0 \rightarrow k \rightarrow 10$ $0 \rightarrow l \rightarrow 31$
No. of standard reflections	3	3
Frequency of standard reflections	150	150
Intensity variation (%)	−7.0	−5.0
Refinement		
Refinement on	F	F
Final R factor	0.208	0.216
Resolution range (Å)	7.0–1.9	7.0–1.7
Completeness of data (%)	59 (for whole range) 47 (2.2–2.0 Å) 21 (2.0–1.9 Å)	56 (for whole range) 35 (1.96–1.82 Å) 17 (1.82–1.7 Å)
No. of reflections used in refinement	1108	1475
No. of parameters used	828	808
Refinement method	Restrained least squares (Hendrickson & Konnert, 1981); *X-PLOR* (Brünger, 1990; Brünger, Karplus & Petsko, 1989)	Restrained least squares (Hendrickson & Konnert, 1981); *X-PLOR* (Brünger, 1990; Brünger, Karplus & Petsko, 1989)
Source of atomic scattering factors	*International Tables for X-ray Crystallography* (1974, Vol. IV)	*International Tables for X-ray Crystallography* (1974, Vol. IV)
Minimum isotropic B value (Å^2)	2.0	2.0
Slow-cooling protocol		
Initial temperature (K)	1000	2000
Final temperature (K)	300	300
Temperature decrement (K)	25	25
Time step (fs)	0.2	0.2

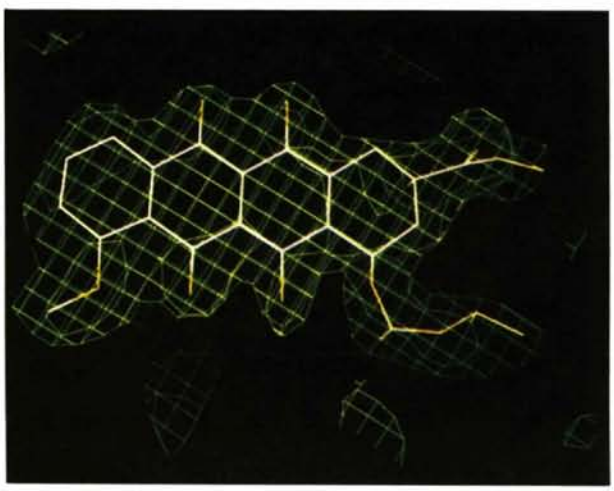
* $R_{merge} = \sum|I(k) - \langle I\rangle|/\sum|I(k)|$ where $I(k)$ and $\langle I\rangle$ are the intensity values for the individual measurements and the corresponding mean value.

molecules. This was followed by a final round of *NUCLSQ* refinement to improve the geometry of the models. For the d(CGGCCG)/daunomycin complex the R factor at the end of the refinement was 0.208 for 1108 reflections with $F \geq 2\sigma(F)$ in the range 7.0–1.9 Å. The model consisted of the DNA (120 atoms), one drug molecule (38 atoms) and 49 solvent molecules all of which were modelled as water O atoms. R is defined as $\sum|F_o - F_c|/\sum|F_o|$ for the d(TGGCCA)/adriamycin complex the final R was 0.216 for 1523 reflections $\geq 2\sigma(F)$ in the range 7.0–1.7 Å. The model consisted of the DNA (120 atoms), an adriamycin molecule (39 atoms) and 47 solvent molecules. Although higher resolution data were included in both of the refinements we judge the effective resolution for each of the structure analyses to be between 2.2 and 2.0 Å. For both models the fit to the electron density is good (Fig. 2) as is the final geometry (Table 2). Coordinates and structure factors for both complexes have been deposited with the Brookhaven Protein Data Bank* (Abola, Bernstein, Bryant, Koetzle & Weng, 1987). The identifying codes are 11D0D and R11D0DSF for the d(CGGCCG)/daunomycin complex; 1DA9 and R1DA9SF for the d(TGGCCA)/adriamycin complex.

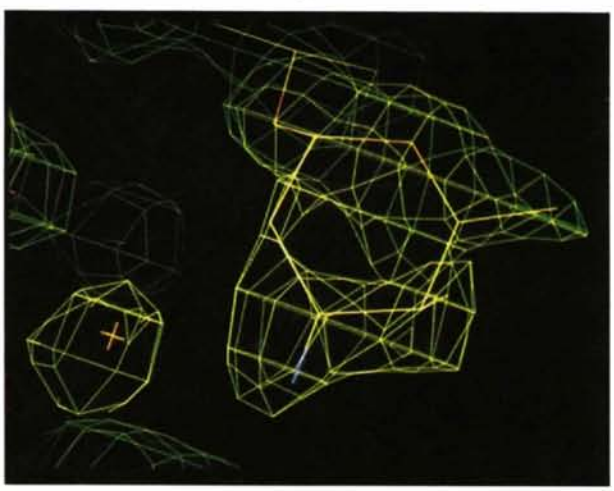
Results and discussion

For both structures described in this paper, the asymmetric unit consists of a single strand of DNA, one anthracycline molecule and a number of ordered solvent molecules. Two asymmetric units, related by a crystallographic dyad, form a distorted B-DNA duplex six Watson–Crick base pairs in length. In each structure the nucleotides are labelled from 1 to 6 on strand 1 and 7 to 12 on strand 2 in the 5′ to 3′ direction. The drug molecules are labelled *D*13 and *D*14. The overall structures of both the

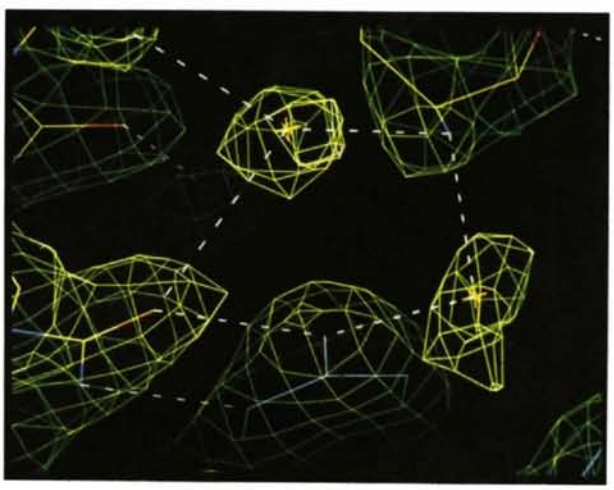
* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (References: 11D0D, R11D0DSF for d(CGGCCG)/daunomycin and 1DA9 and R1DA9SF for d(TGGCCA)/adriamycin). Free copies may be obtained through The Technical Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Supplementary Publication No. SUP 37088). A list of deposited data is given at the end of this issue.



(*a*)



(*b*)



(*c*)

Fig. 2. Examples of the final electron density (green chicken wire) for the d(TGGCCA)/adriamycin complex. The maps were calculated with the coefficients $(2F_o - F_c)$, α_{calc} and are contoured at approximately the r.m.s. deviation of the electron density in the unit cell. Density is shown for (*a*) the aglycone chromophore, (*b*) the amino sugar and (*c*) the solvent arrangement which links the amino sugar to the minor groove of the DNA duplex. Extremely similar figures could be shown for the final electron density for the d(CGGCCG)/daunomycin complex.

Table 2. *Statistics from the final cycles of refinement for the d(CGGCCG)/daunomycin and the d(TGGCCA)/adriamycin complexes*

Restraint groups and parameters	Δ r.m.s.	Standard deviations	$\sum w\Delta^2$*
d(CGGCCG)/daunomycin			
Distance restraints (Å)			
Sugar/base/drug bond lengths	0.022	0.017	
Sugar/base/drug bond angles	0.054	0.030	0.17×10^4
Phosphate bond lengths	0.040	0.025	
Phosphate bond angles	0.062	0.030	
127 distances deviate from ideality by more than 2σ			
Planar restraints			
Deviation from the plane	0.016	0.015	0.89×10^2
Non-bonded contacts			
Single torsions	0.069	0.063	
Multiple torsions	0.135	0.063	0.89
Hydrogen bonds		0.063	
Isotropic thermal parameters (Å^2)			
Sugar/base/drug bonds	9.9	8.0	
Sugar/base/drug angles	12.4	12.0	
Phosphate bonds	12.6	12.0	0.74×10^3
Phosphate angles	14.2	12.0	
Unrestrained chiral volumes	1.3		
d(TGGCCA)/adriamycin			
Distance restraints (Å)			
Sugar/base/drug bond lengths	0.020	0.015	
Sugar/base/drug bond angles	0.046	0.025	0.21×10^4
Phosphate bond lengths	0.046	0.017	
Phosphate bond angles	0.052	0.025	
144 distances deviate from ideality by more than 2σ			
Planar restraints			
Deviation from the plane	0.024	0.015	0.20×10^3
Non-bonded contacts			
Single torsions		0.063	
Multiple torsions	0.200	0.063	0.64×10
Hydrogen bonds		0.063	
Isotropic thermal parameters (Å^2)			
Sugar/base/drug bonds	7.6	8.0	
Sugar/base/drug angles	8.6	12.0	
Phosphate bonds	9.9	12.0	0.40×10^3
Phosphate angles	10.0	12.0	
Unrestrained chiral volumes	1.4		

* $w = 1/\sigma^2$ and Δ = deviation from ideal value.

d(CGGCCG)/daunomycin and d(TGGCCA)/adriamycin complexes are similar to those for all other DNA–anthracycline complexes for which the structures are known (Wang, Ughetto, Quigley & Rich, 1987; Moore, Hunter, Langlois d'Estaintot & Kennard, 1989; Frederick, Williams, Ughetto, van der Marel, van Boom, Rich & Wang, 1990; Williams, Frederick, Ughetto & Rich, 1990, Nunn, van Meervelt, Zhang, Moore & Kennard, 1991; Langlois d'Estaintot, Gallois, Brown & Hunter, 1992; Leonard, Brown & Hunter, 1992). In both structures (Fig. 3) the aglycone chromophore of the anthracycline molecule intercalates itself at approximately right angles to the base step that defines the intercalation site while the amino sugar protrudes into the minor groove where it can interact with the DNA either directly or *via* solvent-mediated hydrogen bonds. The *D* ring of the aglycone chromophore protrudes into the major groove of the double helix, the *C* and *B* rings are stacked into the double helix and the semi-saturated *A* ring is located on the minor-groove side of the duplex. Various parameters describing the geometry of individual base pairs and base steps in both the d(CGGCCG) and d(TGGCCA) duplexes are given in Table 3 and, as can be seen, they are extremely similar to those found in structures where the drug molecules are bound to preferred sequences. The two major distortions that occur on drug binding are the unwinding of the double helix at the second base step – presumably to allow the amino-sugar moiety to interact with the minor groove of the duplex – and the severe buckling of the two base pairs that comprise the actual intercalation site. The degree of buckling observed is indicative that the first base pair of the intercalation site adopts a concave conformation with respect to position of the aglycone chromophore of the drug molecule and that the second base pair is convex in shape. This accommodates the aglycone chromophore into the double helix with no need for extreme adjustment of any of the sugar/phosphate backbone torsion angles, all of which, for both complexes, are well within the expected range for B-DNA (Saenger, 1984) (Table 4).

The planar chromophore of the drug molecule is anchored to the DNA duplex by a combination of van der Waals interactions, aromatic group stacking and direct and solvent-mediated hydrogen bonds (Fig. 4). For the two structures described here the number of solvent-mediated interactions is less than that previously observed in other similar structures. However, as each structure contains a different subset of the interactions found previously in the other structures of DNA–anthracycline complexes, this is liable to be a result of the disorder of the solvent molecules more than a significant deviation from the usual situation.

Significantly, the interactions of the 9-hydroxyl group of the anthracycline molecules with the minor-groove guanine (G8) of the second base pair of the intercalation site are the same in both the d(CGGCCG)/daunomycin complex and its d(TGGCCA)/adriamycin counterpart as they are in other, similar, complexes. We cannot directly elucidate the positions of H atoms at the resolutions of our structure determinations of the two complexes; hence possible hydrogen bonds are evaluated purely on the basis of the interatomic distances of the heavier atoms involved. Based on this criterion, it would seem that the 9-hydroxyl group forms, simultaneously, hydrogen bonds to both the 2-amino group of G(8) and the N3 atom of the same base. It has been suggested that this forms the basis for the molecular recognition of guanine bases by hydroxyl groups. However, as we have already pointed out (Leonard, Brown & Hunter, 1992) in all the structures of similar DNA–anthracycline complexes the N2—O9—N3 angle is extremely acute and it is

Table 3. *Geometrical properties of base-pair steps and base pairs for the d(CGGCCG)/daunomycin and d(TGGCCA)/adriamycin complexes and for some other selected DNA–anthracycline complexes*

All parameters were calculated using the *NEWHEL92* program distributed by R. E. Dickerson and available through the Brookhaven Protein Data Bank (Abola, Bernstein, Bryant, Koetzle & Weng, 1987).

Base pair	Step	Roll (°)	Slide (Å)	Twist (°)	Rise (Å)	Base-pair tilt (°)	Buckle (°)	Propeller twist (°)
d(CGGCCG)/daunomycin								
C1·G12							−8.3	2.6
	1	−6.0	0.9	37.5	5.5	6.8		
G2·C11							12.7	6.0
	2	−2.5	1.0	28.7	3.5	−0.5		
G3·C10							0.4	−7.0
	3	−11.2	0.0	34.5	3.0	0.0		
d(TGGCCA)/adriamycin								
T1·A12							−11.8	0.7
	1	−5.8	0.7	34.9	5.2	6.9		
G2·C11							15.5	4.9
	2	−4.0	0.9	30.5	3.4	−0.6		
G3·C10							2.4	−2.0
	3	−9.4	−0.3	33.3	3.0	0.0		
d(TGTACA)/4′-epiadriamycin (Leonard *et al.*, 1992)								
T1·A12							−9.2	0.0
	1	−9.0	1.0	37.9	5.1	6.2		
G2·C11							18.2	3.2
	2	−8.6	0.1	28.4	3.3	−0.7		
G3·C10							5.7	−6.5
	3	5.8	0.3	35.8	4.2	0.0		
d(CGATCG)/adriamycin (Williams *et al.*, 1990)								
C1·G12							−9.4	1.4
	1	−5.3	1.0	35.3	5.1	8.0		
G2·C11							16.2	−0.1
	2	−6.9	0.4	31.6	3.5	0.1		
A3·T10							6.3	−3.1
	3	−6.6	−0.6	32.0	3.4	0.0		

Table 4. *Sugar/phosphate backbone torsion angles* (°) *and distances* (Å) *between adjacent* P *atoms for d(CGGCCG)/daunomycin and d(TGGCCA)/adriamycin*

Main-chain torsion angles are defined by O3′—P—$\overset{\alpha}{O}$5′—$\overset{\beta}{C}$5′—$\overset{\gamma}{C}$4′—$\overset{\delta}{C}$3′—$\overset{\varepsilon}{O}$3′—$\overset{\zeta}{P}$—O5′. The glycosyl torsion angle χ is defined by O4′—C1′—N1—C2 for pyrimidines; O4′—C1′—N9—C4 for purines.

	χ	α	β	γ	δ	ε	ζ	$P_i - P_{i+1}$
d(CGGCCG)/daunomycin								
C1	−153			66	86	219	−62	
G2	−92	−81	−175	47	142	216	−207	6.8
G3	−99	−6	−250	24	144	200	−99	6.3
C4	−132	−76	−204	64	106	214	−88	6.4
C5	−106	−83	−206	71	130	264	−185	6.6
G6	−108	−111	−158	49	81			
d(TGGCCA)/adriamycin								
T1	−150			43	122	227	−150	
G2	−99	−87	−184	66	122	219	−178	6.6
G3	−112	−55	−218	49	139	187	−112	6.8
C4	−128	−83	−191	66	92	183	−72	6.4
C5	−72	−72	−186	45	146	278	−194	6.6
A6	−91	−72	−190	42	146			

unlikely that both hydrogen bonds are formed, especially given that the O7 atom which forms the ether linkage of the aglycone chromophore and the amino sugar of the anthracycline molecules can form a hydrogen bond to the 2-amino group of G(8) with improved geometry. The situation is no different in the two structures described here and for this reason Fig. 4 shows one hydrogen-bonding interaction for the 9-hydroxyl group and one for the ether O7.

An important aspect of the structures of both the d(CGGCCG)/daunomycin and the d(TGGCCA)/adriamycin complexes is that the amino sugar is not forced away from the floor of the minor groove (Fig. 5*a*) as was suggested by theoretical studies (Chen, Grish & Pullman, 1985). However, the overall conformation of the anthracycline molecules is different to that usually observed in similar hexanucleotide–anthracycline complexes. The orientation of the amino sugar with respect to the aglycone chromophore has changed by rotation about the ether linkage (Fig. 5*b*). This movement probably occurs to alleviate extremely close contacts between the amino sugar and the exocyclic N2 amino group of the (G·C) base pair. Nevertheless, a destabilizing contact between the amino sugar and the guanine N2 is apparent (Table 5). This is counteracted by

Table 5. *Close contacts* (≤ 4 Å) *between the anthracycline amino-sugar moiety and the minor groove of the DNA duplex in the structures of the d(CGGCCG)/daunomycin, d(TGGCCA)/adriamycin, d(TGTACA)/4′-epiadriamycin and d(CGTACG)/daunomycin complexes*

d(CGGCCG)/daunomycin		Distance (Å)
N3′(*D*13)	N2(G8)	3.7
N3′(*D*13)	N2(G9)	2.8
N3′(*D*13)	C2(G9)	3.5
N3′(*D*13)	N3(G9)	3.8
C3′(*D*13)	N2(G8)	3.6
C2′(*D*13)	O2(C5)	3.5
C1′(*D*13)	O4′(G6)	3.6
O5′(*D*13)	O4′(G6)	3.8

d(TGGCCA)/adriamycin		Distance (Å)
N3′(*D*13)	N2(G9)	3.7
C3′(*D*13)	N2(G9)	4.0
C2′(*D*13)	O2(C5)	3.8
C2′(*D*13)	N2(G8)	3.8
C1′(*D*13)	O4′(G6)	3.6
O5′(*D*13)	O4′(G6)	3.8

d(TGTACA)/4′-epiadriamycin		Distance (Å)
N3′(*D*13)	O2(T9)	3.7
C3′(*D*13)	O2(T9)	3.5
C3′(*D*13)	N2(G8)	3.8
C2′(*D*13)	O2(C5)	3.7
C1′(*D*13)	C5′(A6)	3.9
C1′(*D*13)	C4′(A6)	3.8
C1′(*D*13)	O4′(A6)	3.4
O5′(*D*13)	O4′(A6)	3.9
O4′(*D*13)	O2(T9)	3.6
O4′(*D*13)	C1′(T9)	3.7
O4′(*D*13)	C2′(T9)	3.9
O4′(*D*13)	C5′(T9)	4.0

d(CGTACG)/daunomycin		Distance (Å)
N3′(*D*13)	N3(A4)	3.4*
N3′(*D*13)	C2(A4)	3.5
N3′(*D*13)	O4′(C5)	3.4*
N3′(*D*13)	O2(C5)	3.3*
C3′(*D*13)	O2(C5)	3.2
C3′(*D*13)	N2(G8)	3.8
C3′(*D*13)	O2(T9)	3.6
C2′(*D*13)	O2(C5)	3.2
C2′(*D*13)	C1′(C5)	3.9
C2′(*D*13)	O4′(C5)	3.8
C2′(*D*13)	C5′(G6)	4.0
C1′(*D*13)	C5′(G6)	4.0
C1′(*D*13)	C4′(G6)	4.0
C1′(*D*13)	O4′(G6)	3.6
O5′(*D*13)	O2(T9)	3.9
O5′(*D*13)	O2(T9)	3.4

* Possible hydrogen bond.

numerous stabilizing van der Waals contacts to the floor of the DNA minor groove. A balance between these stabilizing and repulsive forces serves to orient the amino sugar close to the DNA.

We have considered the possibility that the change in the orientation of the amino sugar in the minor groove may result in a conformation of the drug molecules which is considerably less than is optimum and that this may be the cause of the lack of affinity of anthracyclines for d(YGG) or d(YGC) base-pair triplets. However, studies on the conformational flexibility of anthracyclines have shown that the optimum conformation occurs when the C8—C7—O7—C1′ and C7—O7—C1′—O5′ torsion angles in the aglycone–amino sugar linkage are in the region of 90 and 290° (Islam & Neidle, 1983) or 90 and 200° (Nakata & Hopfinger, 1980), respectively. In the two structures under discussion here the values for the relevant torsion angles are 87 and 304° [d(CGGCCG)] and 80 and 324° [d(TGGCCA)]. These conformations are reasonably close to one set of the ideal values and it is therefore unlikely that an unstable conformation of the drug is the reason that d(YGC) and d(YGG) base-pair triplets are unfavourable sites for anthracycline binding.

The alteration in the position of the amino sugar in the minor groove of the duplex when the antibiotic is bound to a d(YGG) triplet results in a decrease in the number of favourable van der Waals interactions between the amino sugar and the minor groove of the DNA when compared to other DNA–anthracycline complexes. Table 5 details interatomic distances between the amino sugar and the DNA of less than 4 Å that are observed in the two structures described here as well as for the d(TGTACA)/4′-epiadriamycin and d(CGTACG)/daunomycin complexes. There are approximately twice as many close contacts between the amino sugar and the DNA when the anthracyclines are bound to d(TGTACA) or d(CGTACG) as when they are complexed to d(CGGCCG) or d(TGGCCA). As the majority of these contacts are in the critical 3.4–4.0 Å region it is

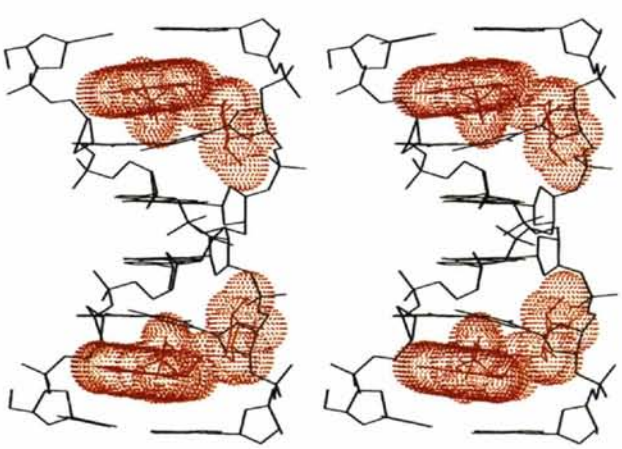


Fig. 3. A stereoview of the d(TGGCCA)/adriamycin complex shown to illustrate the general features of hexanucleotide–anthracycline complexes. The antibiotic molecule is shown with a dot representation of the van der Waals surface, the DNA is depicted as a stick drawing.

likely than an increased number of contacts will stabilize complex formation and this may help explain the apparent preference of daunomycin-type anthracyclines for d(YGT) base-pair triplets over d(YGG) triplets. Although the positions of the amino sugars in the minor grooves of the d(CGGCCG)/daunomycin and the d(TGGCCA)/adriamycin complexes are markedly different to that found in the d(CGTACG)/daunomycin complex, they are similar to that observed in the structure of the d(TGTACA)/4′-epiadriamycin complex (Fig. 5*b*). One might expect, therefore, that in this latter structure there would also be a decrease in the number of favourable van der Waals contacts between the amino sugar and the DNA. As we have already shown (Table 5), this is not the case. This orientation of the amino sugar is not observed in the structures of 4′-epiadriamycin when it is complexed with either d(CGATCG) (Williams, Frederick, Ughetto & Rich, 1990) or d(TGATCG) (Langlois d'Estaintot, Gallois, Brown & Hunter, 1992) and it is likely that in the d(TGTACA)/4′-epiadriamycin complex the movement occurs to avoid close contacts between the O4′ hydroxyl group of the drug and the minor-groove thymine O2 atom which would otherwise result from an inversion of the third base pair in the triplet from A·T to T·A. However, the inversion of the stereochemistry at the O4′ position of 4′-epiadriamycin means that the amino sugar can still form a significant number of favourable van der Waals interactions with the minor groove of the DNA.

A significant consequence of the position of the amino sugar of anthracycline molecules in the minor grooves of the d(CGGCCG)/daunomycin, the d(TGGCCA)/adriamycin and the d(TGTAGA)/4′-epiadriamycin complexes is that this moiety can no longer interact with the DNA *via* direct hydrogen bonds involving the N3′ ammonium group as it does in the structures of related complexes (Fig. 6). Rather it interacts *via* solvent-mediated hydrogen bonds. This does not seem to result in a destabiliza-

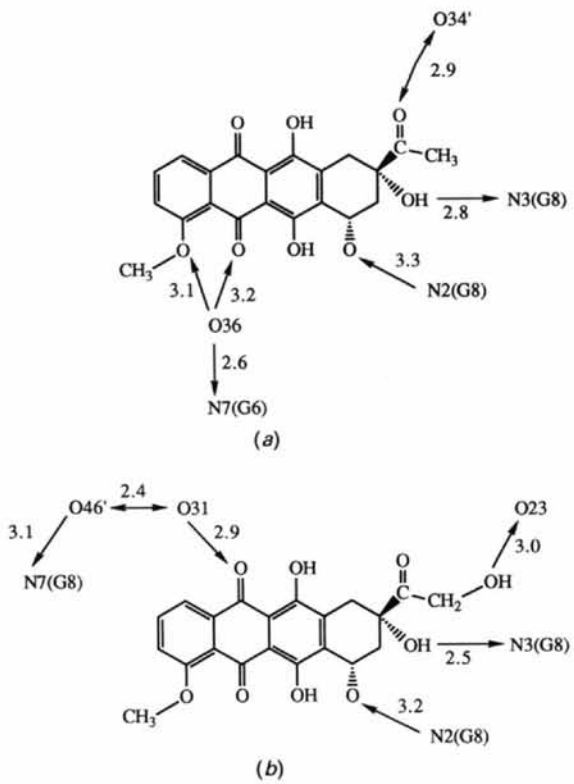


Fig. 4. Schematic diagrams showing the direct and solvent-mediated interactions of the aglycone chromophore with the DNA. (*a*) d(CGGCCG)/daunomycin. (*b*) d(TGGCCA)/adriamycin. An arrow signifies the direction of a probable hydrogen bond.

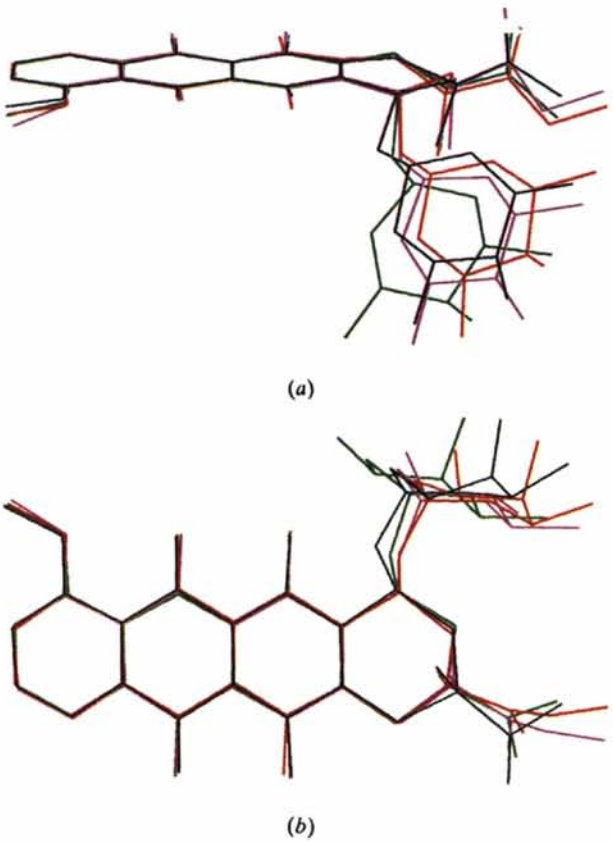


Fig. 5. The molecular superposition of anthracycline molecules in DNA–drug complexes. The colour scheme is as follows. Green: daunomycin as found when complexed to d(CGTACG) (Wang, Ughetto, Quigley & Rich, 1987). Black: daunomycin when complexed to d(CGGCCG). Red: adriamycin in the d(TGGCCA)/adriamycin complex. Pink: 4′-epiadriamycin as found when complexed to d(TGTACA) (Leonard, Brown & Hunter, 1992). In (*a*) the flexibility of the position of the amino sugar with respect to the aglycone chromophore is shown while (*b*) shows that in the d(CGGCCG)/daunomycin and the d(TGGCCA)/adriamycin complexes the amino sugar is not forced away from the floor of the minor groove.

tion in the case of the d(TGTACA)/4′-epiadriamycin complex as evidenced by ultraviolet melting studies (Leonard, Brown & Hunter, 1992).

An important feature of the hydration of the d(TGGCCA)/adriamycin complex is the presence, in the minor groove, of a pentagonal arrangement of hydrogen bonds that links the amino-sugar moiety of the adriamycin molecule to the DNA (Figs. 2*c*, 7*a*). The corners of this pentagon are the N3′ ammonium group of the drug molecule, two water molecules (O20 and O30) and the two minor-groove functional groups of the C(4)·G(9) base pair. The distribution of functional groups in this pentagon allows us to deduce the direction of all the hydrogen bonds and the arrangement defines itself as heterodromic (Saenger, 1984). A similar arrangement can be found in the structure of the d(CGGCCG)/daunomycin complex although some of the interatomic distances are rather longer than one might expect for strong hydrogen-bond formation (Fig. 7*b*). These particular pentagonal arrangements of hydrogen bonds which link the amino sugar can only occur in the interaction of anthracyclines with DNA base-pair triplets having a G·C or C·G base pair at the third position as only these base pairs have the minor-groove interbase hydrogen bond that completes the pentagon. Similar pentagonal arrangements of hydrogen bonds have been found in the structures of α-cyclodextrins (Leysing & Saenger, 1981), the structure of the complex formed between proflavin and the dinucleotide d(CpG) (Neidle, Berman & Shieh, 1980) and the structure of the A-DNA duplex d(GGTATACC) (Kennard, Cruse, Nachman, Prange, Shakked & Rabinovitch, 1986), and are thought to be extremely stable.

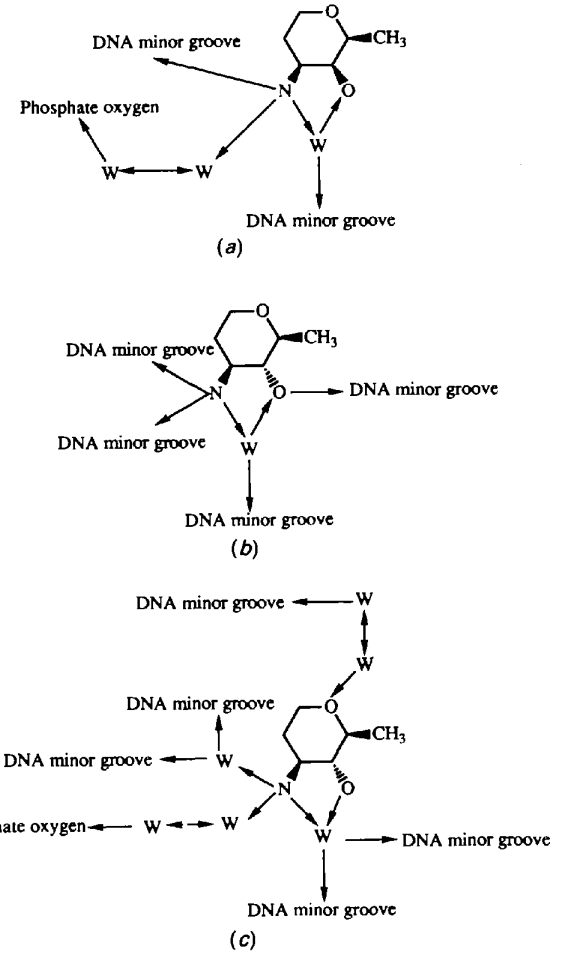


Fig. 6. Different modes of interaction of the amino sugar of anthracyclines with the minor groove. (*a*) The interactions as found in the d(CGTACG)/daunomycin complex (Wang, Ughetto, Quigley & Rich, 1987); (*b*) the interactions as found in the d(CGATCG)/4′-epiadriamycin complex (Williams, Frederick, Ughetto & Rich, 1990); (*c*) the interactions as found in the d(TGTACA)/4′-epiadriamycin complex (Leonard, Brown & Hunter, 1992). Note that in the latter diagram the amino-sugar moiety (shown in pink in Fig. 5) can only interact with the DNA *via* solvent-mediated hydrogen bonds. The convention used regarding the arrows in Fig. 4 also applies here.

Concluding remarks

The single-crystal analyses of the two complexes d(CGGCCG)/daunomycin and d(TGGCCA)/adriamycin have shown that these complexes differ from those of anthracyclines bound to preferred DNA sequences in two important aspects both of which concern the interaction of the amino-sugar moiety of the drug molecules with the DNA to which they are bound. Firstly, the number of favourable van der Waals interactions with the minor groove of the DNA is significantly less in the structures of the two complexes described here than when anthracycline molecules interact with preferred d(YGT) base-pair triplets. Secondly, we observe the presence of stable pentagonal arrangements of hydrogen bonds which link the amino sugars of the drug molecules to the

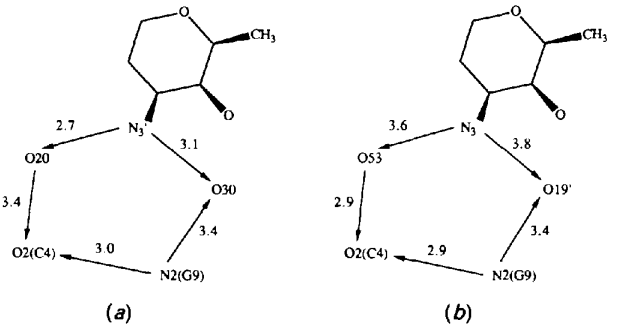


Fig. 7. The pentagonal arrangements, with interatomic distances (Å), of hydrogen bonds found in the minor grooves of (*a*) the d(TGGCCA)/adriamycin and (*b*) d(CGGCCG)/daunomycin complex.

DNA and note that these can only be formed when anthracyclines are bound to d(YGG) or d(YGC) base-pair triplets. In all other aspects the structures of the d(CGGCCG)/daunomycin and the d(TGGCCA)/adriamycin complexes are extremely similar to the complexes formed between anthracyclines and d(YGT) or d(YGA) sequences.

DNA-footprinting studies (Chaires, Herrera & Waring, 1990; Skavobogaty, White, Phillips & Reiss, 1988*a*,*b*) reveal that anthracycline molecules show a marked lack of affinity for DNA base-pair triplets of the type d(YGG) or d(YGC) and prefer, instead, to bind to base-pair triplets of the type d(YGT) or d(YGA). Given that pentagonal arrangements of hydrogen bonds serve to stabilize structures and that the amino-sugar moieties of the drug molecules in the d(CGGCCG)/daunomycin and the d(TGGCCA)/adriamycin complexes are not forced away from the floor of the minor groove by the presence of the exocyclic N2 amino group, it would seem, therefore, that the decrease in the number of favourable van der Waals interactions observed in the structures of d(CGGCCG)/daunomycin and d(TGGCCA)/adriamycin relative to those found in similar structures is responsible for the lack of affinity of anthracyclines for d(CGG) and d(TGG) sites. This may seem a rather tenuous basis for the sequence specificity of daunomycin-type anthracyclines as our structure determinations show that they can form stable complexes with apparently non-preferred sequences. However, it must be remembered that in studies with longer stretches of DNA, the anthracyclines are presented with a variety of possible binding sites rather than the necessarily artificial situation of the complexes used for structure analyses and that even small increases in stability may result in drug molecules binding to some sequences in preference to others.

We thank the Wellcome Trust and the United Kingdom Science and Engineering Research Council for support. TWH gratefully acknowledges the Royal Society and the Australian Academies of Science and Technological Sciences for enabling his stay at the University of Manchester.

References

ABOLA, E. E., BERNSTEIN, F. C., BRYANT, S. H., KOETZLE, T. F. & WENG, J. (1987). *Crystallographic Databases – Information Content, Software Systems, Scientific Applications*, edited by F. H. ALLEN, G. BERGERHOFF & R. SEIVERS, pp. 107–132. Cambridge: Data Commission of the International Union of Crystallography.

BROWN, T. & BROWN, D. J. S. (1991). *Oligonucleotides and Analogues, a Practical Approach*, edited by F. ECKSTEIN, pp. 1–24. Oxford Univ. Press.

BRÜNGER, A. T. (1990). *X-PLOR Manual*. Version 2.2. Howard Hughes Medical Institute, Yale Univ., New Haven, USA.

BRÜNGER, A. T., KARPLUS, M. & PETSKO, G. A. (1989). *Acta Cryst.* A**45**, 50–51.

CHAIRES, J. B., HERRERA, J. E. & WARING, M. J. (1990). *Biochemistry*, **29**, 6145–6153.

CHEN, K.-Y., GRISH, N. & PULLMAN, B. (1985). *J. Biomol. Struct. Dyn.* **3**, 445–466.

FREDERICK, C. A., WILLIAMS, L. D., UGHETTO, G., VAN DER MAREL, G. A., VAN BOOM, J. H., RICH, A. & WANG, A. H.-J. (1990). *Biochemistry*, **29**, 2538–2549.

HENDRICKSON, W. A. & KONNERT, J. H. (1981). *Biomolecular Structure, Conformation, Function and Evolution*, Vol. 1, edited by R. SRINIVASAN, pp. 43–57. Oxford: Pergamon.

ISLAM, S. A. & NEIDLE, S. (1983). *Acta Cryst.* B**39**, 114–119.

JONES, T. A. (1978). *J. Appl. Cryst.* **11**, 268–272.

KENNARD, O., CRUSE, W. B. T., NACHMAN, J., PRANGE, T., SHAKKED, Z. & RABINOVITCH, D. (1986). *J. Biomol. Struct. Dyn.* **3**, 623–647.

LANGLOIS D'ESTAINTOT, B., GALLOIS, B., BROWN, T. & HUNTER, W. N. (1992). *Nucleic Acids Res.* **20**, 3561–3566.

LEONARD, G. A., BROWN, T. & HUNTER, W. N. (1992). *Eur. J. Biochem.* **204**, 69–74.

LEYSING, B. & SAENGER, W. (1981). *Biochim. Biophys. Acta*, **678**, 408–413.

MCBRIDE, L. J. & CARUTHERS, M. H. (1983). *Tetrahedron Lett.* **24**, 245–248.

MCPHERSON, A. (1982). *Preparation and Analysis of Protein Crystals*. San Francisco: Wiley.

Molecular Structure Corporation (1989). *TEXSAN. TEXRAY Structure Analysis Package*. Version 5. Molecular Structure Corporation, 3200 Research Forest Drive, The Woodlands, TX 77381, USA.

MOORE, M. H., HUNTER, W. N., LANGLOIS D'ESTAINTOT, B. & KENNARD, O. (1989). *J. Mol. Biol.* **206**, 693–705.

NAKATA, V. & HOPFINGER, A. J. (1980). *FEBS Lett.* **117**, 259–264.

NEIDLE, S., BERMAN, H. M. & SHIEH, H. S. (1980). *Nature (London)*, **279**, 343–344.

NORTH, A. C. T., PHILLIPS, D. C. & MATHEWS, F. S. (1968). *Acta Cryst.* A**24**, 351–359.

NUNN, C. N., VAN MEERVELT, L., ZHANG, S., MOORE, M. H. & KENNARD, O. (1991). *J. Mol. Biol.* **222**, 167–177.

PULLMAN, B. (1987*a*). *Structure and Expression*, Vol. 2, edited by R. H. SARMA & M. H. SARMA, pp. 237–249. New York: Adenine Press.

PULLMAN, B. (1987*b*). *Anticancer Drug Des.* **7**, 95–105.

SAENGER, W. (1984). *Principles of Nucleic Acid Structure*. New York: Springer-Verlag.

SHELDRICK, G. M. (1976). *SHELX*76. Program for crystal structure determination. Univ. of Cambridge, England.

SKAVOBOGATY, A., WHITE, R. J., PHILLIPS, D. R. & REISS, J. A. (1988*a*). *FEBS Lett.* **227**, 103–106.

SKAVOBOGATY, A., WHITE, R. J., PHILLIPS, D. R. & REISS, J. A. (1988*b*). *Drug Des. Deliv.* **3**, 125–152.

VALENTINI, I., NICOLLELA, V., VANNINI, E., MENOZZI, M., PENCO, S. & ARCAMONE, F. (1985). *Farmaco (Pavia)*, **40**, 377–383.

WANG, A. H.-J., UGHETTO, G., QUIGLEY, G. J. & RICH, A. (1987). *Biochemistry*, **26**, 1152–1163.

WESTHOF, E., DUMAS, P. & MORAS, D. (1985). *J. Mol. Biol.* **184**, 119–145.

WILLIAMS, L. D., FREDERICK, C. A., UGHETTO, G. & RICH, A. (1990). *Nucleic Acids Res.* **18**, 5533–5541.