X-ray crystallographic study of DNA duplex cross-linking: simultaneous binding to two d(CGTACG)(2) molecules by a bis(9-aminoacridine-4-carboxamide) derivative

Published online 1 December 2006 *Nucleic Acids Research, 2006, Vol. 34, No. 22* ***6663–6672***
doi:10.1093/nar/gkl930

X-ray crystallographic study of DNA duplex cross-linking: simultaneous binding to two d(CGTACG)$_2$ molecules by a bis(9-aminoacridine-4-carboxamide) derivative

Nicholas H. Hopcroft[1], Anna L. Brogden[1,2], Mark Searcey[2] and Christine J. Cardin[1,*]

[1]School of Chemistry, University of Reading, Whiteknights, Reading, Berkshire RG6 6AD, UK and
[2]Department of Pharmaceutical and Biological Chemistry, School of Pharmacy, University of London, 29-39 Brunswick Square, London WC1N 1AX, UK

Received September 7, 2006; Revised October 16, 2006; Accepted October 18, 2006 NDB code: DD0078 and PDB code: 2GB9

ABSTRACT

Acridine-4-carboxamides form a class of known DNA mono-intercalating agents that exhibit cytotoxic activity against tumour cell lines due to their ability to inhibit topoisomerases. Previous studies of bis-acridine derivatives have yielded equivocal results regarding the minimum length of linker necessary between the two acridine chromophores to allow bis-intercalation of duplex DNA. We report here the 1.7 Å resolution X-ray crystal structure of a six-carbon-linked bis(acridine-4-carboxamide) ligand bound to d(CGTACG)$_2$ molecules by non-covalent duplex cross-linking. The asymmetric unit consists of one DNA duplex containing an intercalated acridine-4-carboxamide chromophore at each of the two CG steps. The other half of each ligand is bound to another DNA molecule in a symmetry-related manner, with the alkyl linker threading through the minor grooves. The two crystallographically independent ligand molecules adopt distinct side chain interactions, forming hydrogen bonds to either O6 or N7 on the major groove face of guanine, in contrast to the semi-disordered state of mono-intercalators bound to the same DNA molecule. The complex described here provides the first structural evidence for the non-covalent cross-linking of DNA by a small molecule ligand and suggests a possible explanation for the inconsistent behaviour of six-carbon linked bis-acridines in previous assays of DNA bis-intercalation.

INTRODUCTION

Since the first evidence of naturally occurring antibiotics binding to DNA by bis-intercalation (1), there have been several studies aimed at the design of synthetic analogues for therapeutic use. Bis-intercalation has the potential to generate both kinetically and thermodynamically strong DNA binding, leading to the inhibition of DNA replication, transcription or topoisomerase activity. Antibiotics such as echinomycin and the triostins, which exhibit anti-tumour as well as anti-microbial activity, bind to DNA by inserting two quinoxaline groups into the helical stack, with 2 bp occluded between them (2,3). This finding is consistent with the 'neighbour exclusion principle' first formulated to describe mono-intercalator binding, which states that intercalation cannot occur at two consecutive base pair steps of a DNA duplex. Much of the work on designing synthetic analogues of these antibiotics has focused on bis-acridine compounds, due to the stronger intercalative binding of the acridine moiety compared with quinoxaline, and there have been several efforts to determine the minimum linker length between the two acridine groups necessary for bis-intercalation. These studies have generally employed hydrodynamic techniques to detect unwinding of supercoiled plasmid DNA or lengthening of linear DNA upon intercalation by the ligand. Although compounds that have linkers of seven or more atoms length have consistently been shown to bis-intercalate (4–8), analogues with shorter linkers have remained the subject of some controversy. Early experiments gave conflicting results concerning whether linkers of five or six atoms length allowed bis-intercalation (4,5), leading to the suggestion that the nature of the linker and any ring substituents may affect the length required (9). At 7.5 and 8.8 Å, respectively (assuming all-*trans* geometry), five or six atom linkers are too short to allow bis-intercalation without violating the neighbour exclusion principle. An NMR study, whilst not resulting in a full structure determination, indicated that a simple bis-acridine with a six carbon atom linker bound mono-intercalatively (7), in contrast to the hydrodynamic data for the same compound (5). However, a recent electrophoretic mobility assay suggested that compound **1** (Figure 1), a derivative with relatively potent antitumour

*To whom correspondence should be addressed. Tel: +44 118 931 8215; Fax: +44 118 931 6331; Email: c.j.cardin@reading.ac.uk

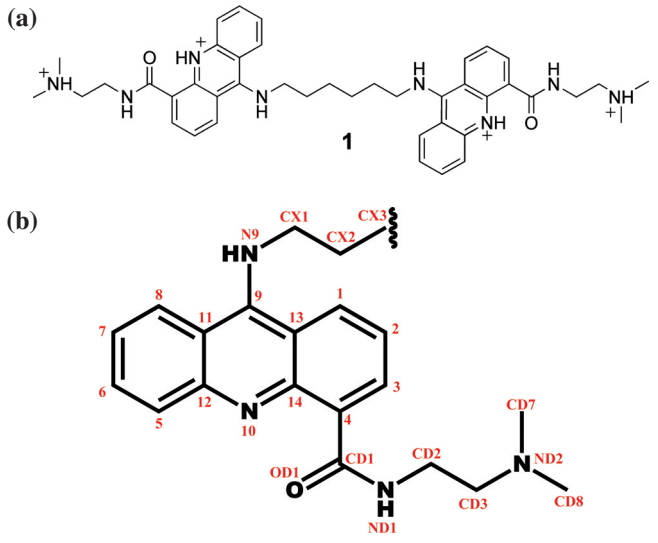


Figure 1. Molecular formula (**a**) and atom numbering scheme (**b**) of 9,9′-(1,6-hexanediyldiimino)bis{*N*-[2-(dimethylamino)ethyl]}acridine-4-acarboxamide, referred to as compound **1** throughout the text and designated A4C in the coordinate files. The numbering scheme is shown for only half of the ligand due to internal symmetry. The two crystallographically independent chromophores are referred to as DAC1 and DAC2, after the common name for the *N*-[2-(dimethylamino)ethyl]acridine-4-carboxamide monomer, DACA.

activity, bis-intercalates in violation of the neighbour exclusion principle (8).

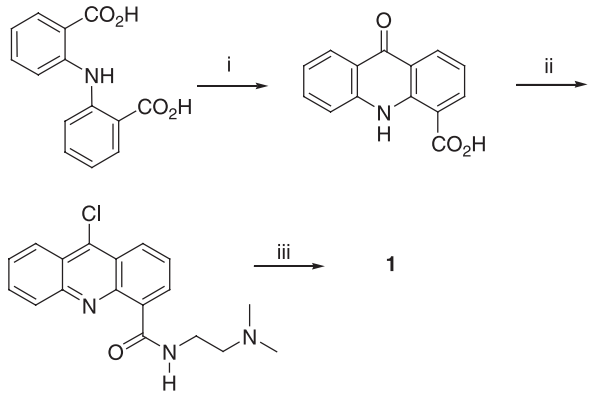
The elucidation of the structures of triostin A and echinomycin bound to d(CGTACG)$_2$ (2,3) has been followed by other crystal structures of synthetic ligands bis-intercalating palindromic DNA duplexes (10–13). However, such structures are relatively elusive, partly due to the multiple chromophore orientations which are accessible when long-flexible linkers are used. In the naturally occurring antibiotics, which possess cyclic linkers, there is a much smaller accessible conformational space and a structure which has evolved to have a single DNA-binding mode. In contrast to the bis-intercalators, many structures of mono-intercalating ligands bound to DNA duplexes have been solved, including several of acridine-4-carboxamides in complex with d(CGTACG)$_2$ (14–17). Some members of this class of compounds are known to have cytotoxic effects due to their ability to inhibit topoisomerases (18). It has been suggested that bis-intercalating derivatives of these, such as compound **1**, the subject of this work, could additionally act as transcription inhibitors if they formed more long-lasting complexes than their mono-intercalating equivalents (8). The only structure of one of this series of bis-functional compounds bound to DNA is that of the 8-carbon-linked analogue of **1** in complex with d(CGTACG) (19). In those crystals, grown in the presence of Co^{2+} ions, the DNA did not exist as a simple duplex, but instead formed a junction-like structure involving four duplexes. Although the ligand linker was long enough to allow classical bis-intercalation, the compound instead bound to a novel quadruplex-like intercalation cavity formed by exchange of bases between different duplexes in the junction. Several similar structures, all determined using crystals grown in the presence of Co^{2+} ions, have been obtained with mono-intercalators bound (20–24). Although the positive charge of the bound ligands may be important in allowing the juxtaposition of phosphate groups, this ability is clearly not specific to bis-intercalators, and the Co^{2+} ions, identifiable in the electron density maps, appear to play a crucial role in stabilizing this structure.

A further mode of bis-intercalator binding, in which the ligand cross-links two duplexes with one chromophore incorporated into the base pair stack of each, has also been suggested (25). Gel electrophoresis experiments using ligands with rigid extended linkers, designed to promote cross-linking over classical bis-intercalation, have provided circumstantial evidence for this mode of binding in the form of knotted and catenated DNA (25,26). The authors of that work noted, however, that formal proof of duplex cross-linking would require structural characterization of the ternary complex formed. Here, we provide the first such structural evidence for a duplex cross-linking mode of bis-intercalator binding, describing the X-ray crystal structure of the six-carbon-linked bis(acridine-4-carboxamide) **1** bound to d(CGTACG)$_2$ molecules. The structure of the mono-intercalator analogue of compound **1** bound to the same DNA was solved previously by multi-wavelength anomalous dispersion (MAD) using brominated derivatives of the ligand and DNA (14,15). In that structure, the DNA crystallized with a 2-fold axis relating the two strands within the duplex, and two symmetry-related ligand molecules intercalating the CG steps at each end. The ligand side chain occupied the major groove of the DNA, with ND2 forming a hydrogen bond with the G2 nucleotide, although the side chain was disordered to a certain extent. In the structure presented here of the bis-functional intercalator **1** in complex with the same DNA, the individual chromophores bind in a manner similar to the mono-intercalator, whilst the linker makes no direct contacts with the DNA. Two crystallographically independent acridine-4-carboxamide moieties are bound to the duplex; in each case the other half of the ligand binds to another duplex in a symmetry-related manner. The subtle differences between the modes of binding of the two ligand molecules in this structure and that of the monomer are discussed, as are the implications of this structure on studies of the effects of linker length.

MATERIALS AND METHODS

Synthesis of the six-carbon-linked bis(9-aminoacridine-4-carboxamide)

The synthesis of compound **1** is outlined in Scheme 1 and is similar to previously described methods (8,27). Commercially available 2,2-iminodibenzoic acid was treated with excess concentrated sulphuric acid to give 9(10H)acridone-4-carboxylic acid, which was converted to 9-chloroacridine-4-carbonyl chloride. The resulting unstable acid chloride was immediately reacted with an excess of *N*,*N*-dimethylethylenediamine in anhydrous dichloromethane buffered with triethylamine and gave a good yield of the 9-chloroacridine-4-carboxamide with no concomitant replacement of the nuclear chlorine. The synthesis was completed by following a modified version of the procedure described previously (8). The 9-chloroacridine-4-carboxamide was dissolved in phenol and allowed to react with half a molar



Scheme 1. Synthesis of **1:** (i) Conc. H_2SO_4, 100°C, 3 h, 99% yield. (ii) (a) $SOCl_2$, DMF (two drops), reflux, 30 min, 85% yield; (b) *N*, *N*-dimethylethylenediamine, CH_2Cl_2, 0°C, 2 h, 67% yield. (iii) (a) Phenol, 120°C, 1 h; (b) 1,6-diaminohexane, phenol, 55°C, 24 h, 46% yield.

equivalent of 1,6-diaminohexane. After heating at 55°C for 24 h, and suitable work up, compound **1** was purified using flash column chromatography. The tetrahydrochloride salt of the compound was prepared by treating the free base in MeOH with 12 M HCl, followed by precipitation with EtOAc. The salt was freeze-dried from water to give a yellow powder.

Crystallization and data collection

The self-complementary DNA oligonucleotide d(CGTACG) was purchased from Eurogentec as a HPLC-purified solid. Crystals containing the ligand **1** were grown by vapour diffusion from sitting drops at 291 K. The drop contained 4 µL 10% (v/v) 2-methyl-2,4-pentanediol (MPD), 40 mM sodium cacodylate, pH 7.0, 12 mM spermine, 80 mM $SrCl_2$; 1 µL 1 mM DNA duplex and 1 µL of ~375 µM **1**. The drop was equilibrated against a 1 mL reservoir of 35% (v/v) MPD. Yellow crystals of octahedral shape and approximate dimensions 0.2 mm × 0.1 mm × 0.1 mm appeared after ~1 week. Initial screening and crystal indexing were carried out at EMBL Hamburg beam line X13 and final data collection was performed at the ESRF BM14. The structure was solved by SAD at the Sr^{2+} K-edge absorption peak, in a modification to a Sr^{2+} MAD procedure to be described in detail elsewhere. X-ray diffraction data were collected at 100 K using radiation of wavelength 0.769 Å and a Mar Research CCD detector. Data reduction and processing were performed using the programs Denzo and Scalepack (28) (Table 1).

Structure solution and refinement

The structure was solved using the SHELXC/D/E programs (29,30), as integrated in the HKL2MAP routines (31). The space group of the crystal was determined as $P4_322$ (with a correlation coefficient of 71.43%, compared with 54.25% for $P4_122$) with two Sr^{2+} ions located in the asymmetric unit, one at half occupancy due to its location on an axis of 2-fold rotational symmetry. The initial model was obtained from the previously determined structure of 6-bromo-9-amino-DACA bound to $d(CG^{5\text{-}Br}UACG)_2$ (14) (NDB code DDF073). Refinement was performed using SHELXL (32) and model rebuilding carried out using Xtalview (33). Five per cent of the reflections were selected randomly to monitor the validity of the refinement and model rebuilding process using the R_{free} factor (34) (Table 1). The final coordinates have been deposited in the NDB (35) (code DD0078) and PDB (36,37) (code 2GB9), and structure factors are available from the PDB. Full conformational analysis of the DNA is available via the NDB, whilst a summary of the deviations from the B-form conformation is provided below. Figures 2a and 4–7 were produced using Molscript (38) and Raster3D (39).

Table 1. X-ray data and refinement statistics

X-ray data	
Unit cell dimensions, Å ($P4_322$ spacegroup)	
a (Å)	37.46
b (Å)	37.46
c (Å)	53.55
α, β, γ (degrees)	90.00
Resolution (Å) (outer shell)	37.2–1.7 (1.79–71.70)
Unique reflections	4509 (626)
Anomalous multiplicity[a]	15.4 (15.3)
Anomalous completeness (%)	100.0 (100.0)
$< I/\sigma(I) >$	4.2 (3.4)
R^b_{merge}	6.4% (20.4%)
Refinement and model correlation	
Resolution range (Å)	8.0–1.7
No. of reflections used in refinement	4204
R-factor[c]	21.2%
No. of reflections used for R_{free}	221
R_{free} [c]	25.5%
No. of DNA atoms	240
No. of ligand atoms	52
No. of water molecules	103
Average B-factors (Å^2)	
DNA	19.1
Ligands	18.7
Water molecules	35.9
R.m.s. deviations from ideal geometry (targets in parentheses)	
Bond lengths (Å)	0.026
1,3 Distances defining angles (Å)	0.021

[a]The average number of observations of the same reflection.
[b]The value of the merging *R*-factor between equivalent measurements of the same reflection, $R_I = \Sigma|I - \langle I \rangle|/\Sigma I$.
[c]Crystallographic *R*-factor (R_{free}) = $\Sigma||F_o| - |F_c||/\Sigma|F_o|$.

RESULTS

Crystal packing and overall structure description

The complex of **1** bound to $d(CGTACG)_2$ crystallized in the space group $P4_322$ in the presence of Sr^{2+} ions, in a packing which has not previously been observed for this DNA molecule or any complex involving DNA-intercalating ligands. The novel packing is due to the non-covalent cross-linking of DNA duplexes by the bis-intercalator (Figure 2). The overall conformation of an individual duplex shows a strong similarity to that already observed for the mono-intercalator (14), allowing the model built *de novo* for that work to be placed unchanged in the SAD electron density map as a starting model for this structure, despite the significant differences in packing. The asymmetric unit contains two d(CGTACG) strands (designated A and B) that interact through Watson–Crick base pairing to form a single duplex. This is in contrast to the structures of the same DNA molecule with bound mono-intercalators, in which a crystallographic axis of 2-fold rotational symmetry relates the two strands within a duplex. The electron-density maps (Figure 3) allowed the

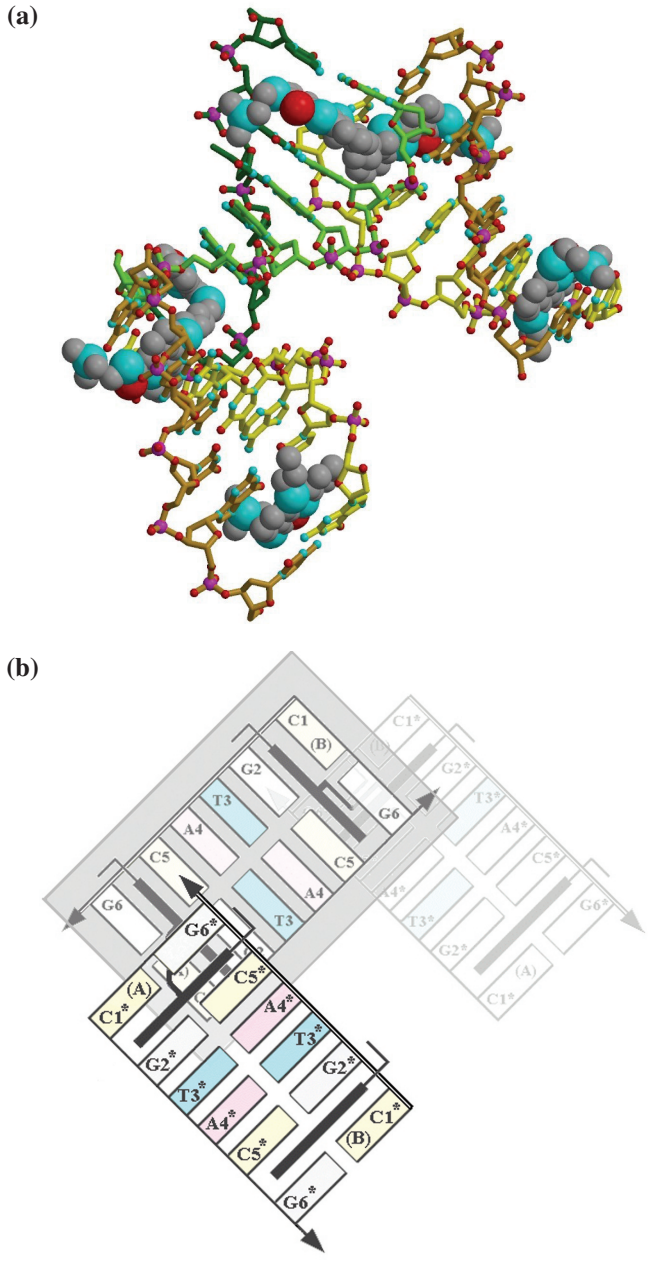


Figure 2. (**a**) Ball-and-stick representation of DNA duplex cross-linking by compound **1** in the crystal. The asymmetric unit contains a single duplex with two mono-intercalated ligands. Each molecule of compound **1** lies on an axis of 2-fold symmetry, with the other chromophore bound to a symmetry-related duplex. The DNA molecule of focus is shown in green whilst symmetry-related molecules are shown in yellow. In each case, strand B is indicated by darker shading than strand A. The ligand is shown in space filling format. (**b**) Schematic representation of the DNA cross-linking depicted in (a). The asymmetric unit is indicated by the shaded box and symmetry-related nucleotides are labelled with an asterisk. The major and minor grooves of the DNA are depicted as hollow and filled arrows, respectively. Alternative stereo views of Figures 2a, 4, 5, 6b and 7 are provided as online Supplementary Data.

complete modelling of two acridine-4-carboxamide chromophores in the asymmetric unit, as well as the DNA duplex (except the 5′ phosphate group of each strand), 103 water molecules and one and a half Sr^{2+} ions, the anomalous signal from which was used to obtain the SAD maps. The two

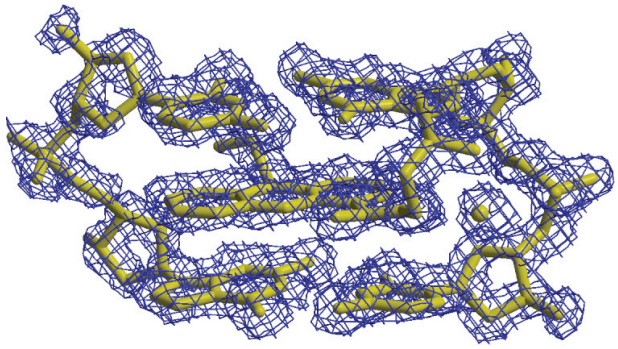


Figure 3. Example of the final electron density map, contoured at 1 σ-level. The region shown corresponds to the intercalation site of the DAC1 chromophore, viewed from the major groove face to highlight the definition of the electron density for the side chain, which is semi-disordered in the mono-intercalator complex (15).

acridine-4-carboxamide moieties intercalate into the DNA at the two distinct CG steps of the duplex. In each case, the linker of the intercalator extends away from the duplex and meets an axis of 2-fold rotational symmetry; thus, the whole ligand is made up of two symmetry-related halves. The bis-intercalators themselves mediate crystal packing by cross-linking DNA molecules. Additionally, direct stacking occurs between base pairs at opposite ends of symmetry-related duplexes, without the involvement of a 'spacer' compound as seen in structures of this DNA molecule in the presence of mono-intercalators (14–17). Cross-linking occurs where these stacks pack against each other, with the helix axes approximately orthogonal, although the two ligand-binding sites are not equivalent in terms of crystal packing. The close juxtaposition of duplexes linked through their minor grooves allows direct hydrogen bonding interaction between the phosphate group of G2(A) and the N2 group of G6(B) of a symmetry-related helix.

Strontium ion interactions

Sr1 is situated in the minor groove of the DNA, at a pseudo-symmetric position at the central TA step, and is therefore remote from the ligand binding sites (Figure 4). The electron density maps allowed the identification of six water molecules coordinating Sr1. Previous studies of Sr^{2+} binding to the DNA Holliday junction (40) have shown the ion to coordinate up to eight water molecules, so the hydration sphere observed here is incomplete. The irregular arrangement of the observed water molecules around this ion supports the expectation that other, less ordered, water molecules are also involved in its coordination. Two of the water molecules coordinated to Sr1 form hydrogen bonds with the N3 groups of A4(A) and A4(B), respectively, whilst a third interacts with the phosphate group of a symmetry-related C5 nucleotide. Sr2 is located on a crystallographic 2-fold axis and four coordinating water molecules could be identified in the electron density maps, consisting of two symmetry-related pairs. As for Sr1, both the number and geometry of these water molecules suggest that they do not form the complete hydration shell for the ion. Sr2 does not lie in either groove of the DNA duplex but is instead located between duplexes, forming bridging interactions between the DNA backbones

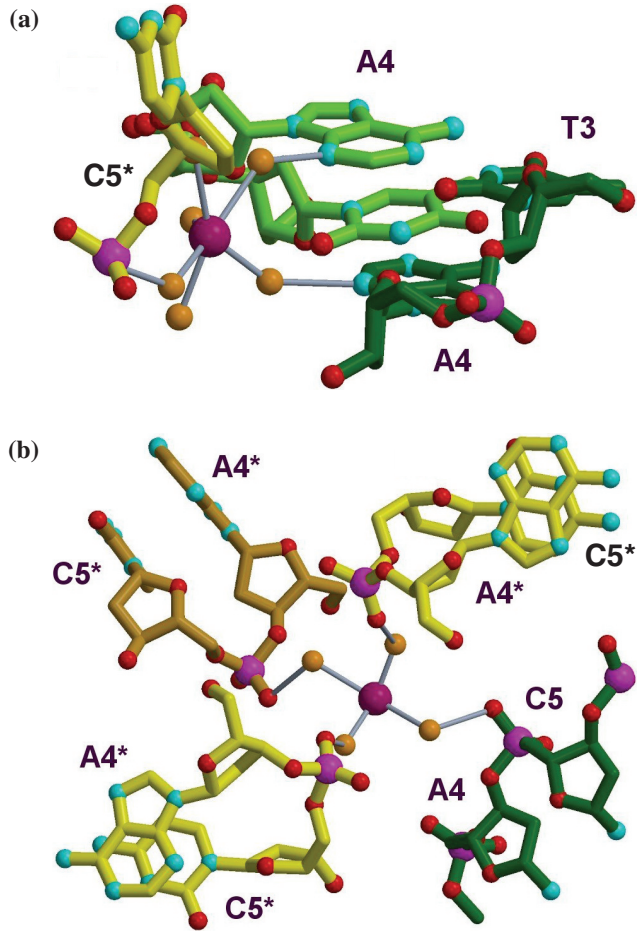


Figure 4. Interactions formed by the two Sr^{2+} ions in the asymmetric unit. (**a**) Sr1 bound in the minor groove of the DNA duplex, at the central TA step. (**b**) Sr2-mediating crystal packing contacts between the backbones of several DNA strands. Sr^{2+} ions are shown in magenta and water molecules in orange. Coordination bonds to the ion and hydrogen bonds (except those involved in standard Watson–Crick base pairing) are represented by grey lines.

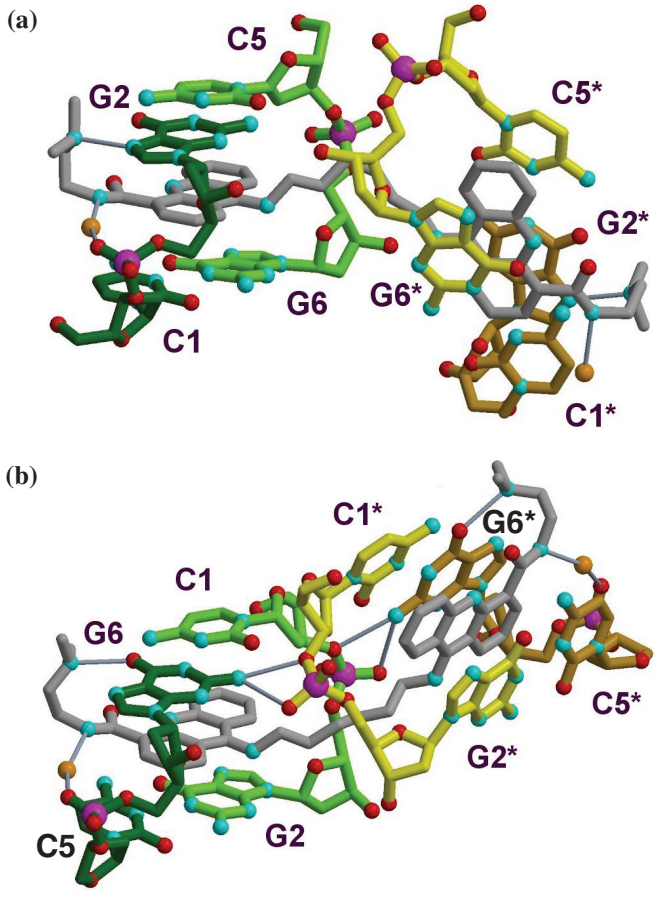


Figure 5. DNA duplex cross-linking by DAC1 (**a**) and DAC2 (**b**). The left-hand side of each ligand is shown in a similar orientation, to highlight the similarities and differences in binding between the two independent observations.

that may be important to crystal packing. One pair of symmetry-related water molecules coordinated to Sr2 form hydrogen bonds with the phosphate groups of their corresponding C5(A) nucleotides, whilst another pair each contact the phosphate groups of C5(B) nucleotides. Thus, a total of four symmetry-related duplexes are linked by Sr2 and its primary hydration shell.

Conformation of the ligand and DNA/ligand interactions

The two bound chromophores have very similar overall orientations within the intercalation cavity, but differently oriented side chains. The first chromophore (DAC1) intercalates at the C1G2(B)/C5G6(A) step of the DNA duplex, with the ring that bears the side chain closest to strand B (Figure 5a). The long axis of the acridine lies closest to that of the G2(B)/C5(A) base pair, such that the intercalator forms extensive π–π stacking interactions with the bases of C5(A), G6(A) and G2(B), but less so with that of C1(B). The side chain protrudes into the major groove of the DNA duplex, adopting an extended conformation in the 3′ direction of strand B. Thus, the protonated ND2 group lies approximately in the plane of the G2(B) base, where it forms a hydrogen bond with N7 at a distance of 2.8 Å. ND1 forms a water-mediated hydrogen bonding interaction with the phosphate group of G2(B) identical to that observed in previously determined structures of acridine-4-carboxamide mono-intercalators bound to this DNA molecule. The CD1 atom of the side chain does not form a planar centre; it lies in the plane of the acridine ring, whereas both OD1 and ND1 are positioned on the same side of this plane as the rest of the side chain. The carbonyl oxygen of the carboxamide group, OD1, forms a hydrogen bond with N10 of the acridine ring, demonstrating that N10 must be protonated under the conditions in the crystal. The linker lies in the minor groove and extends away from the DNA molecule, surrounded by largely disordered solvent.

The second chromophore within the asymmetric unit (DAC2) is situated at the C5G6(B)/C1G2(A) step (Figure 5b). The conformation of this intercalator is similar to that of the other, but with some subtle differences in the contacts it forms with the DNA. The side chain lies in the major groove, extending in the 3′ direction along strand B of the DNA, and the acridine ring is again positioned such that its long axis is almost parallel to that of the base-pair nearest the side chain. Thus the ligand forms extensive π–π stacking interactions with the bases of C1(A), G2(A) and G6(B), but a less extensive

contact with the base of C5(B). The geometry at CD1 is again non-planar, with the OD1 and ND1 groups on the same side of the acridine plane as the rest of the side chain. The OD1 carbonyl group forms a hydrogen bond with N10 of the chromophore, indicating protonation of the ring nitrogen. The most significant difference between the interactions formed by the two chromophores is the nature of the hydrogen bond formed by the ND2 group of the side chain. In this case, ND2 acts as a hydrogen bond donor to O6 of the G6(B) base, at a distance of 2.7 Å. ND1 forms hydrogen bonding interactions with the phosphate groups of G6(B) and C5(B) via one and two water molecules, respectively. The contact with G6 is essentially the same as the conserved interaction described above for DAC1 with G2. However, there is no equivalent of the C5 interaction for DAC1, as the phosphate group of C1 is disordered. The linker of DAC2 extends away from the DNA through the minor groove, but the solvent is more ordered here than in the vicinity of DAC1. Thus, a network of water-mediated hydrogen bond interactions linking N9 of the ligand to O4′ of G2(B) and O2 of C1(B) is observed.

Conformation of the DNA

The two distinctive conformations of the ligand, with the side chain forming a hydrogen bond with either the guanine O6 or N7 atoms, show both similarities and differences in their effect upon the DNA around them (Figure 6). In strand B, the opening of the CG steps to allow intercalation is achieved through a series of subtle changes in backbone torsion angles that are almost identical for the two binding sites. The β angles of the two guanine nucleotides in this strand are changed by an average of +38° and the ε angle of G2 by −37° (there is no ε angle for the terminal G6 nucleotide). The ζ angles of all nucleotides in strand B are higher than the average value for B-DNA. This is particularly the case for the two cytosine nucleotides, where the change is +34°. Similarly, the δ angles are lower than in B-DNA in all strand B nucleotides except the two guanines. This effect is also most pronounced in the cytosine nucleotides, at an average change of −42°. In contrast to the subtle conformational changes in strand B, in strand A specific backbone torsion angles are significantly perturbed from the values expected for B-DNA, and these changes are different for the two ligand binding sites. The α and γ angles of G2 differ from those for B-DNA by −114° and +131°, respectively, whilst the same angles in G6 are changed by +129° and −129° relative to B-DNA. The ζ angles of strand A are generally higher than expected, but the two cytosine nucleotides show very different behaviour to each other. The largest change in ζ is +49° in C1, whereas C5 shows a change of −27°. Overall, the conformation of the DNA backbone at the two ligand-binding sites is very similar in strand B, to which both side chains bind, but significantly different in strand A. This is evidenced by the fact that pairs of equivalent backbone torsion angles at the intercalation sites differ by an average of only 6.3° in strand B, but by 56.2° in strand A.

The DNA structure shows no major perturbations of individual base pair geometry compared to native B-DNA, and no apparent correlations between the binding sites of the two ligands. The most notable feature in this regard is that the C5(A)–G2(B) base pair is flattened with respect to its propeller twist, a feature that is accompanied by a slightly higher buckle than that of the other base pairs. The base pair step parameters, however, show more notable deviations from previously observed norms. As expected, the intercalation sites for DAC1 and DAC2 have similar local rise values of 6.8 and 6.7 Å, respectively, approximately twice the size of the non-intercalated base pair steps. The only shift or slide value > 0.5 Å is a slide (translation along base pair long axis) of 0.75 Å at the intercalation site of DAC1. The intercalation site of DAC2 shows the largest tilt and roll, at 5.2° and 11.0°, respectively. The twist angles at the two intercalation sites,

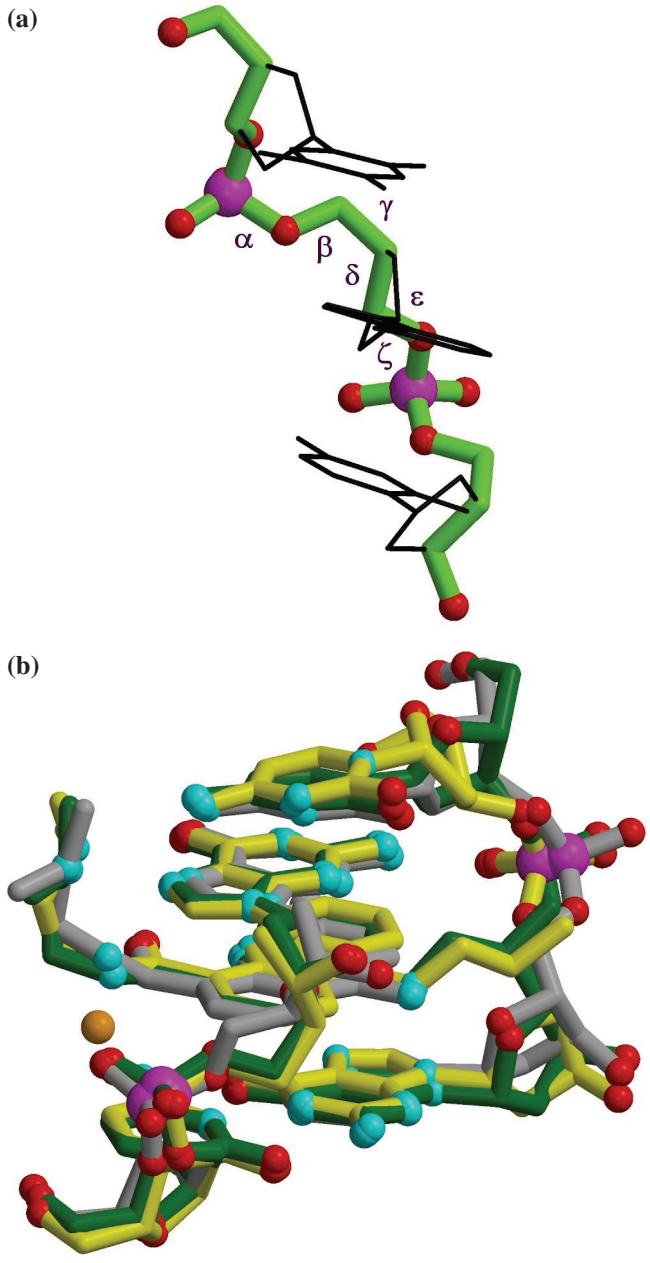


Figure 6. Conformational analysis of the DNA. (**a**) Definition of DNA backbone torsion angles, using the A4(A) nucleotide (which adopts a conformation close to B-form DNA) as an example. For clarity, the ribose rings and bases are shown only as outlines. (**b**) Superposition of the DAC1 (green), DAC2 (yellow) and 9-amino-DACA mono-intercalator (14) (grey) binding sites.

when defined relative to the helix axis, are in very close agreement with each other at 22.6° and 22.4°. However, these angles are 21.9° and 18.8° for DAC1 and DAC2, respectively, when defined locally. The helical twist angles for the non-intercalated steps, in the 5′ to 3′ direction along strand B, are 25.0°, 37.0° and 31.9°, with the locally defined angles very similar.

DISCUSSION

The structure presented here represents the first crystallographic evidence of a small molecule ligand non-covalently cross-linking two DNA duplexes. There are two crystallographically independent observations of this phenomenon in the structure, and comparison of the two with each other, and with complexes of related mono-intercalators, can yield valuable information on the general principles of how these classes of compounds bind to DNA. A notable feature of this structure is that each half of the bis-intercalator binds to the DNA in a fashion similar to that of the equivalent mono-intercalator (Figure 6b). Intercalation is accompanied by helix extension and unwinding, both of which have been used to detect ligand binding in viscometric studies (4–6,8,9). Indeed, compound **1** has previously been classified as a bis-intercalator on the basis that it causes an unwinding of 36° per ligand in plasmid DNA, approximately double that of the monomer (8). The helix unwinding angle suggested by this structure, taking account of the effects on adjacent base pair steps, is 29° per chromophore. Classical bis-intercalation by this ligand would require either violation of the neighbour exclusion principle or significant bending of the DNA, because the linker is too short to span 2 bp in an undistorted duplex. The structure presented here raises the possibility that a certain amount of DNA cross-linking during solution studies may have complicated these experiments, accounting for the inconsistencies in results regarding bis- versus mono-intercalation for compound **1** and other ligands with similar linker lengths (4–9). Here, we show directly that compound **1** can bind to DNA by duplex cross-linking, in the crystalline state at least, whereas the solution data suggesting that **1** may bis-intercalate result from indirect methods. It remains to be determined whether other ligands of interest with six-atom linkers also cross-link DNA, and what the minimum and maximum linker lengths for this mode of binding are. A consideration of the geometry of cross-linking in the crystal suggests no reason why the same mode of binding should not be adopted in solution; the precise arrangement of duplexes seems more likely to be an effect of the ligand cross-linking than a cause of it, especially because the same DNA molecule, when mono-intercalated, forms crystals with very different symmetry. The fact that crystals with identical morphology and very similar unit cell dimensions (data not shown) were grown in the presence of various different metal ions shows that the structure presented here is not influenced significantly by the Sr^{2+} ions used for obtaining phase information. Furthermore, the hexameric DNA helices pack end-to-end to form pseudo-infinite stacks in the crystal, indicating that there would be no steric clashes if the two cross-linked DNA molecules were longer. Although both independent observations of cross-linking involve DNA helices that are at ~90° to each other, this is not an exact relationship forced by crystal symmetry. The precise spatial arrangement of the two helices, and therefore the overall conformation of the ligand, differs between the two observations (Figure 5). In each case, the first three carbon atoms of the linker extend away from the helix in a similar fashion, through the minor groove (Figure 6b). The major conformational difference between the two linkers is found at the central carbon–carbon bond, where DAC1 exhibits a *gauche* torsion angle and DAC2 adopts a *trans* geometry. This, coupled with less dramatic differences in the two torsion angles either side (which are identical to each other due to symmetry), results in the linkers of the two ligands each bending ~90°, in almost opposite directions. Hence, the flexible nature of the linker allows the ligand to cross-link duplexes in an adaptable manner, despite previous suggestions that linker rigidity may promote a cross-linking mode of binding (25,26).

A striking feature of this structure is that the side chains of the two ligand molecules extend in the same direction along the helix, whereas in all of the mono-intercalator complexes the side chains of the two symmetry-related bound ligands have both pointed towards the centre of the hexanucleotide. In the crystal, the DNA helices pack together to form pseudo-infinite stacks, and ligand cross-linking occurs between duplex molecules in neighbouring stacks. Unlike in the structures containing mono-intercalators, there is no 'spacer' ligand molecule stacked between the duplexes, so the extended DNA stacks in the crystal resemble longer DNA except for the breaks in the backbones every 6 bp. The fact that the side chain of DAC2 interacts with G6(B) sterically prevents that of DAC1 interacting with G6(A), because these two guanines are part of adjacent base pairs in the pseudo-infinite stack (Figure 7). The DAC1 chromophore therefore intercalates in the correct orientation for its side chain to form a hydrogen bond with G2(B). Hence the side chains of both ligands contact guanines in strand B rather than strand A. Bearing in mind that the two ends of each hexameric DNA duplex are not equivalent in terms of their packing with molecules in other helical stacks, it is unclear whether a different crystal of the same morphology and packing could contain ligand chromophores that are all rotated by 180° relative to the situation here, with their side chains all extending in the opposite direction along the minor groove and all forming hydrogen bonds with guanines in strand A. There is also no steric reason why, even in the absence of intra-duplex crystallographic symmetry, the two ligands bound to each hexamer should not have their side chains both extending towards the centre of the duplex, as in the structure with bound monomer. Alternatively, the crystal could even contain a mixture of alternative side chain orientations, as long as the two ligands at neighbouring ends of adjacent molecules within a stack adopt the same side chain direction as each other. Hence, the orientations adopted here can be taken to represent a genuine preference of the ligand under these crystalline conditions.

Previous structural work involving the mono-intercalating equivalent of compound **1** has shown that the protonation state of N10 has a significant effect on side chain geometry (14). In the present structure, both ligands are protonated at this site, and therefore have the side chain oriented with the carboxamide oxygen close enough to N10 to act as the acceptor in an intramolecular hydrogen bond. This contrasts

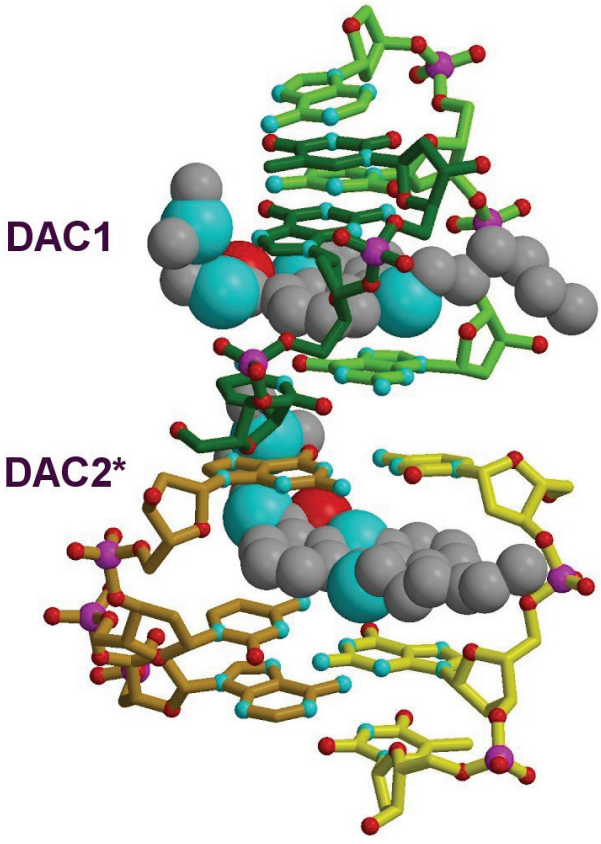


Figure 7. The packing of the end of one DNA duplex with the opposite end of a symmetry-related duplex to form a pseudo-infinite stack. The bound ligands are shown in space filling format to illustrate that the side chain of DAC2 sterically prevents that of DAC1 from interacting with the terminal base pair of the duplex to which it is bound. The symmetry-related chromophore of each ligand has been omitted for clarity.

with the conformation observed if N10 is not protonated, in which the side chain is rotated by ~180° so that the carboxamide nitrogen acts as the donor in a hydrogen bond with N10. Despite the two ligands adopting the same geometry at the carboxamide group, they differ in terms of the hydrogen bond contacts that the side chains make in the major groove of the DNA. In DAC1, the ammonium group of the side chain forms a hydrogen bond with N7 of G2(B), in a manner identical to the major conformation of the monomer (14). The side chain of DAC2, however, forms a hydrogen bond with O6 of G6(B). It has been noted before for the monomer that the dimethylaminoethyl section of the side chain is partially disordered, and that refinement of a single conformation can result in a position in which hydrogen bond formation to both O6 and N7 is plausible (15). In the complex with compound **1** presented here, the two independent observations of the ligand show two distinct conformations, each of which is well-ordered. This suggests that the disorder observed for the mono-intercalator results from a combination of these two preferred alternative conformations. The central conformation reported before (15), in which there are two putative hydrogen bonds, appears to be an average of the two positions observed here which, although it could represent a transitional intermediate, is not necessarily highly populated itself. The reasons behind the greater side chain ordering in the bis-intercalator are not clear, but a possible reason for the distinction between the two observations of the ligand is the difference in electrostatic potential at the two binding sites. At G2, the sugar-phosphate backbone is contiguous, and therefore the positively charged side chain is attracted towards the phosphate group as well as the partial negative potential of the major groove face of the guanine base. The side chain of DAC1 thus lies close to N7, and forms a hydrogen bond there. In contrast, the lack of phosphate negative charge at the interduplex site 3′ to G6 means that the side chain of DAC2 occupies a more central position within the major groove and thus forms a hydrogen bond with O6.

Strands A and B might be expected to show some distinctive structural characteristics due to the fact that both ligand side chains interact with strand B, making it the main focus of phosphate charge neutralization by the compound. The temperature factors for strand B are noticeably higher for the AT base pairs than for the GC ones, but those for strand A increase steadily from 5′ to 3′ (data not shown). This may be because the guanines and cytosines of strand B receive a stabilizing effect from hydrogen bond interactions with the ligand side chains, whereas the only stabilizing hydrogen bond for strand A is a crystal packing contact involving the 5′ end of the DNA backbone. When comparing the effects of ligand binding on various nucleotides in this structure and those of the mono-intercalator complexes it is important to distinguish between nucleotides that are co-ordinated by ligand side chains and those that are not. The region C1–G2 in the mono-intercalator complex is equivalent to C1(B)–G2(B) and C5(B)–G6(B) in the sense of interacting with the ligand side chain, and C5–G6 in the previous structures is equivalent to both C1(A)–G2(A) and C5(A)–G6(A). Hence, it is possible to distinguish specific effects of the intercalated ligand from more general position effects. The coupled changes in the α and γ torsion angles of G2 in the mono-intercalator structure, which have been suggested to be necessary for the opening up of the intercalation site, are found in neither G2(B) nor G6(B), although they are found in the non-equivalent G2(A). The different changes in α and γ associated with G6 in the mono-intercalator structure are found only in G6(A) and not in G2(A) as expected, while in the present structure α and γ remain remarkably unchanged from B-form DNA values in the G2(B) and G6(B) nucleotides that are contacted by the ligand side chains. When assessed in the same way, the previously identified decrease in δ (and associated switch to a C3′-*endo* sugar pucker) for the cytosine in the strand contacted by the ligand side chain appears to be a consistent feature of intercalation. Changes in ζ for both the cytosine and guanine in the side chain coordinating strand also appear to be reasonably consistent.

SUMMARY

The structure presented here is the first showing a ligand noncovalently cross-linking two DNA duplexes. Each half of the symmetrical compound is based on known topoisomerase inhibitors and binds in a way similar to those ligands, through intercalation and hydrogen bond formation with the major groove face of guanine. The structure suggests that linkers

of six-atoms (~8.8 Å) length may promote a cross-linking mode of binding rather than bis-intercalation in violation of the neighbour exclusion principle. Such non-covalent DNA cross-linking could be of practical use in the study of inter-duplex contacts during replication, transcription and topoisomerase activity. The possible role of this mode of binding in the cytotoxicity of the compound studied here requires further investigation.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank the BBSRC for financial support of NHH (grant 19997) and the University of Reading and University of London School of Pharmacy for financial support of ALB We are grateful to the MRC for providing access to the ESRF CRG beam line BM14 and to Hassan Belrhali for his help during data collection. We also acknowledge use of EMBL beam lines at the DORIS storage ring, DESY, Hamburg. The Open Access publication charges for this article were waived by Oxford University Press.

Conflict of interest statement. None declared.

REFERENCES

1. Waring,M.J. and Wakelin,L.P. (1974) Echinomycin: a bifunctional intercalating antibiotic. *Nature*, **252**, 653–657.
2. Wang,A.H., Ughetto,G., Quigley,G.J., Hakoshima,T., van der Marel,G.A., van Boom,J.H. and Rich,A. (1984) The molecular structure of a DNA–triostin A complex. *Science*, **225**, 1115–1121.
3. Ughetto,G., Wang,A.H., Quigley,G.J., van der Marel,G.A., van Boom,J.H. and Rich,A. (1985) A comparison of the structure of echinomycin and triostin A complexed to a DNA fragment. *Nucleic Acids Res*., **13**, 2305–2323.
4. Le Pecq,J.B., Le Bret,M., Barbet,J. and Roques,B. (1975) DNA polyintercalating drugs: DNA binding of diacridine derivatives. *Proc. Natl Acad. Sci. USA*, **72**, 2915–2919.
5. Wakelin,L.P.G., Romanos,M., Chen,T.K., Glaubiger,D., Canellakis,E.S. and Waring,M.J. (1978) Structural limitations on the bifunctional intercalation of diacridines into DNA. *Biochemistry*, **17**, 5057–5063.
6. King,H.D., Wilson,W.D. and Gabbay,E.J. (1982) Interactions of some novel amide-linked bis(acridines) with deoxyribonucleic acid. *Biochemistry*, **21**, 4982–4989.
7. Assa-Munt,N., Denny,W.A., Leupin,W. and Kearns,D.R. (1985) 1H NMR study of the binding of bis(acridines) to $d(AT)_5.d(AT)_5$. 1. Mode of binding. *Biochemistry*, **24**, 1441–1449.
8. Wakelin,L.P.G., Bu,X., Eleftheriou,A., Parmar,A., Hayek,C. and Stewart,B.W. (2003) Bisintercalating threading diacridines: relationship between DNA binding, cytotoxicity and cell cycle arrest. *J. Med. Chem*., **46**, 5790–5802.
9. Wright,R.G.M., Wakelin,L.P.G., Fieldes,A., Acheson,R.M. and Waring,M.J. (1980) Effects of ring substituents and linker chains on the bifunctional intercalation of diacridines into deoxyribonucleic acid. *Biochemistry*, **19**, 5825–5836.
10. Gao,Q., Williams,L.D., Egli,M., Rabinovich,D., Chen,S.L., Quigley,G.J. and Rich,A. (1991) Drug-induced DNA repair: X-ray structure of a DNA–ditercalinium complex. *Proc. Natl Acad. Sci. USA*, **88**, 2422–2426.
11. Peek,M.E., Lipscomb,L.A., Bertrand,J.A., Gao,Q., Roques,B.P., Garbay-Jaureguiberry,C. and Williams,L.D. (1994) DNA distortion in bis-intercalated complexes. *Biochemistry*, **33**, 3794–3800.
12. Hu,G.G., Shui,X., Leng,F., Priebe,W., Chaires,J.B. and Williams,L.D. (1997) Structure of a DNA–bisdaunomycin complex. *Biochemistry*, **36**, 5940–5946.
13. Shui,X., Peek,M.E., Lipscomb,L.A., Gao,M., Ogata,C., Roques,B.P., Garbay-Jaureguiberry,C., Wilkinson,A.P. and Williams,L.D. (2000) Effects of cationic charge on three-dimensional structures of intercalative complexes: structure of a bis-intercalated DNA complex solved by MAD phasing. *Curr. Med. Chem*., **7**, 59–71.
14. Todd,A.K., Adams,A., Thorpe,J.H., Denny,W.A., Wakelin,L.P. and Cardin,C.J. (1999) Major groove binding and 'DNA-induced' fit in the intercalation of a derivative of the mixed topoisomerase I/II poison *N*-(2- (dimethylamino)ethyl)acridine-4-carboxamide (DACA) into DNA: X-ray structure complexed to $d(CG^{5Br}UACG)_2$ at 1.3 Å resolution. *J. Med. Chem*., **42**, 536–540.
15. Adams,A., Guss,J.M., Collyer,C.A., Denny,W.A. and Wakelin,L.P. (1999) Crystal structure of the topoisomerase II poison 9-amino-[*N*-(2-dimethylamino)ethyl]acridine-4-carboxamide bound to the DNA hexanucleotide $d(CGTACG)_2$. *Biochemistry*, **38**, 9221–9233.
16. Adams,A., Guss,J.M., Collyer,C.A., Denny,W.A., Prakash,A.S. and Wakelin,L.P. (2000) Acridinecarboxamide topisomerase poisons: structural and kinetic studies of the DNA complexes of 5-substituted 9-amino-(*N*-(2-dimethylamino)ethyl)acridine-4-carboxamides. *Mol. Pharmacol*., **58**, 649–658.
17. Adams,A., Guss,J.M., Denny,W.A. and Wakelin,L.P. (2002) Crystal structure of 9-amino-*N*-[2-(4-morpholinyl)ethyl]-4-acridinecarboxamide bound to $d(CGTACG)_2$: implications for structure–activity relationships of acridinecarboxamide topoisomerase poisons. *Nucleic Acids Res*., **30**, 719–725.
18. Bridewell,D.J., Finlay,G.J. and Baguley,B.C. (2001) Topoisomerase I/II selectivity among derivatives of *N*-[2-(dimethylamino)ethyl] acridine-4-carboxamide (DACA). *Anticancer Drug Des*., **16**, 317–324.
19. Teixeira,S.C., Thorpe,J.H., Todd,A.K., Powell,H.R., Adams,A., Wakelin,L.P., Denny,W.A. and Cardin,C.J. (2002) Structural characterisation of bisintercalation in higher-order DNA at a junction-like quadruplex. *J. Mol. Biol*., **323**, 167–171.
20. Yang,X.L., Robinson,H., Gao,Y.G. and Wang,A.H. (2000) Binding of a macrocyclic bisacridine and ametantrone to CGTACG involves similar unusual intercalation platforms. *Biochemistry*, **39**, 10950–10957.
21. Adams,A., Guss,J.M., Collyer,C.A., Denny,W.A. and Wakelin,L.P. (2000) A novel form of intercalation involving four DNA duplexes in an acridine-4-carboxamide complex of $d(CGTACG)_2$. *Nucleic Acids Res*., **28**, 4244–4253.
22. Thorpe,J.H., Hobbs,J.R., Todd,A.K., Denny,W.A., Charlton,P. and Cardin,C.J. (2000) Guanine specific binding at a DNA junction formed by $d(CG^{5\text{-}Br}UACG)_2$ with a topoisomerase poison in the presence of Co^{2+} ions. *Biochemistry*, **39**, 15055–15061.
23. Adams,A., Guss,J.M., Denny,W.A. and Wakelin,L.P. (2004) Structure of 9-amino-[*N*-(2-dimethylamino)propyl]acridine-4-carboxamide bound to $d(CGTACG)_2$: a comparison of structures of $d(CGTACG)_2$ complexed with intercalators in the presence of cobalt. *Acta Crystallogr D Biol Crystallogr*., **60**, 823–828.
24. Valls,N., Steiner,R.A., Wright,G., Murshudov,G.N. and Subirana,J.A. (2005) Variable role of ions in two drug intercalation complexes of DNA. *J. Biol. Inorg. Chem*., **10**, 476–482.
25. Mullins,S.T., Annan,N.K., Cook,P.R. and Lowe,G. (1992) Bisintercalators of DNA with a rigid linker in an extended configuration. *Biochemistry*, **31**, 842–849.
26. Annan,N.K., Cook,P.R., Mullins,S.T. and Lowe,G. (1992) Evidence for cross-linking DNA by bis-intercalators with rigid and extended linkers is provided by knotting and catenation. *Nucleic Acids Res*., **20**, 983–990.
27. Atwell,G.J., Cain,B.F., Baguley,B.C., Finlay,G.J. and Denny,W.A. (1984) Potential antitumour agents. 43. Synthesis and biological activity of dibasic 9-aminoacridine-4-carboxamide, a new class of antitumour agent. *J. Med. Chem*., **27**, 1481–1485.
28. Otwinowski,Z. and Minor,W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol*., **276**, 307–326.
29. Scheider,T.R. and Sheldrick,G.M. (2002) Substructure solution with SHELXD. *Acta Crystallogr D Biol Crystallogr*., **58**, 1772–1779.
30. Sheldrick,G.M. (2002) Macromolecular phasing with SHELXE. *Z. Kristallogr*., **217**, 644–650.
31. Pape,T. and Schneider,T.R. (2004) HKL2MAP: a graphical user interface for phasing with SHELX programs. *J. Appl. Cryst*., **37**, 843–844.

32. Sheldrick,G.M. and Schneider,T.R. (1997) SHELX-97: high-resolution refinement. *Methods Enzymol*., **277**, 319–343.
33. McRee,D.E. (1999) XtalView/Xfit—a versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol*., **125**, 156–165.
34. Brunger,A.T. (1992) The free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*, **355**, 472–475.
35. 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) The nucleic acid database: a comprehensive relational database of three-dimensional structures of nucleic acids. *Biophys. J*., **63**, 751–759.
36. Bernstein,F.C., Koetzle,T.F., Williams,G.J., Meyer,E.F., Jr, Brice,M.D., Rodgers,J.R., Kennard,O., Shimanouchi,T. and Tasumi,M. (1977) The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol*., **112**, 535–542.
37. Berman,H., Henrick,K. and Nakamura,H. (2003) Announcing the worldwide Protein Data Bank. *Nature Struct. Biol*., **10**, 980.
38. Kraulis,P.J. (1991) Molscript—a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst*., **24**, 946–950.
39. Merritt,E.A. and Bacon,D.J. (1997) Raster3D: photorealistic molecular graphics. *Methods Enzymol*., **277**, 505–524.
40. Thorpe,J.H., Gale,B.C., Teixeira,S.C. and Cardin,C.J. (2003) Conformational and hydration effects of site-selective sodium, calcium and strontium ion binding to the DNA Holliday junction structure $d(TCGGTACCGA)_4$. *J. Mol. Biol*., **327**, 97–109.