Crystal and solution structures of 7-amino-actinomycin D complexes with d(TTAGBrUT), d(TTAGTT) and d(TTTAGTTT)

**Eftichia Alexopoulos,[a]
Elizabeth A. Jares-Erijman,[b]
Thomas M. Jovin,[c] Reinhard
Klement,[c] Reinhard Machinek,[d]
George M. Sheldrick[a] and Isabel
Usón[e]***

[a]Department of Structural Chemistry, University of Göttingen, Tammannstrasse 4, 37077 Göttingen, Germany, [b]Departamento de Química Orgánica, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, 1428 Buenos Aires, Argentina, [c]Department of Molecular Biology, Max-Planck-Institute for Biophysical Chemistry, 37077 Göttingen, Germany, [d]Institute of Organic and Biomolecular Chemistry, University of Göttingen, Tammannstrasse 2, 37077 Göttingen, Germany, and [e]ICREA at Instituto de Biología Molecular de Barcelona (IBMB–CSIC), Jordi Girona 18-26, 08034 Barcelona, Spain

Correspondence e-mail: uson@ibmb.csic.es

The formation of the complex of 7-amino-actinomycin D with potentially single-stranded DNA has been studied by X-ray crystallography in the solid state, by NMR in solution and by molecular modelling. The crystal structures of the complex with 5′-TTAG[Br^{5}U]T-3′ provide interesting examples of MAD phasing in which the dispersive component of the MAD signal was almost certainly enhanced by radiation damage. The trigonal and orthorhombic crystal modifications both contain antibiotic molecules and DNA strands in the form of a 2:4 complex: in the orthorhombic form there is one such complex in the asymmetric unit, while in the trigonal structure there are four. In both structures the phenoxazone ring of the first drug intercalates between a BrU–G (analogous to T–G) wobble pair and a G–T pair where the T is part of a symmetry-related molecule. The chromophore of the second actinomycin intercalates between the BrU–G and G–BrU wobble pairs of the partially paired third and fourth strands. The base stacking also involves (A*T)*T triplets and Watson–Crick A–T pairs and leads to similar complex three-dimensional networks in both structures, with looping-out of unpaired bases. Although the available NOE constraints of a solution containing the antibiotic and d(TTAGTTT) strands in the ratio 1:1 are insufficient to determine the structure of the complex from the NMR data alone, they are consistent with the intercalation geometry observed in the crystal structure. Molecular-dynamics (MD) trajectories starting from the 1:2 complexes observed in the crystal showed that although the thymines flanking the d(AGT) core are rather flexible and the G–T pairing is not permanently preserved, both strands remain bound to the actinomycin by strong interactions between it and the guanines between which it is sandwiched. Similar strong binding (hemi-intercalation) of the actinomycin to a single guanine was observed in the MD trajectories of a 1:1 complex. The dominant interaction is between the antibiotic and guanine, but the complexes are stabilized further by promiscuous base-pairing.

1. Introduction

Actinomycin D (AMD) is an anticancer antibiotic frequently used as a component of the clinical VAC therapy (vincristine, actinomycin, cyclophosphamide). It consists of two cyclo-peptide rings connected to each other *via* a phenoxazone moiety. The pharmacological action of AMD has been ratio-nalized by its interaction with the double-stranded DNA helix: the chromophore intercalates between the base pairs of DNA and the cyclopeptides interact with the surface of the minor groove by hydrogen bonding (Waring, 1981). As a result, DNA transcription and replication is inhibited by blocking polymerase translocation along the helix.

The binding of AMD to double-stranded DNA (dsDNA) has been investigated extensively by X-ray crystallography (Kamitori & Takusagawa, 1994; Robinson *et al.*, 2001; Hou *et al.*, 2002), NMR (Brown *et al.*, 1994; Chen *et al.*, 1996; Lian *et al.*, 1996; Chou *et al.*, 2002; Chin *et al.*, 2003) and other spectroscopic techniques and is the subject of numerous thermodynamic studies (Wilson *et al.*, 1986; Chen, 1988; Chen & Sha, 2002; Chen *et al.*, 2003; Qu *et al.*, 2003). In general, the major binding site involves 5′-GpC-3′ sequences. The crystal structures of AMD or AMD derivatives complexed with dsDNA show the phenoxazone ring intercalating into a GpC step, with the cyclic pentapeptides anchored on the helix. Four hydrogen bonds between threonines in AMD and guanine from the DNA and another from the chromophore to the backbone stabilize the complex. Additionally, hydrophobic contacts strengthen the drug–DNA interaction (Jain & Sobell, 1972).

As the drug spans four base pairs, the flanking residues to the guanines also influence complex formation (Chen, 1988). Binding and kinetic studies with oligomers containing self-complementary and non-self-complementary –*XGCY*– sequences indicate differences in binding affinities depending on the nature of the adjacent bases. However, the DNA does not always adopt a helical conformation, *e.g.* the AMD–d(CGATCGATCG) complex consists of a slipped duplex with the ApT dinucleotides looped out (Robinson *et al.*, 2001). The motif of bases looping out of the helix was also described for 5′-G*XC*/C*YG*-3′ sequences (where *X*/*Y* are G/C or T/A) interacting with AMD with the *X*/*Y* bases positioned perpendicular to the stacked bases/chromophore (Chou *et al.*, 2002).

Three NMR (Lian *et al.*, 1996; Chou *et al.*, 2002; Chin *et al.*, 2003) and one X-ray (Hou *et al.*, 2002) crystal structures of non-complementary sequences bound to AMD have been reported so far. In all cases base pairing between the nucleotides seems to be forced. This results in mismatched base pairs, which can be accompanied by strand slippage or formation of hairpin loops. In the case of the sequence d(GATGCTTC)$_2$ (Lian *et al.*, 1996) investigated by NMR methods, the non-complementarity was overcome by formation of T–T mismatches, resulting in a 2:1 DNA strand–drug complex, with the chromophore intercalating between the central GpC/CpG base pairs. The sequences used in two more recent NMR studies (Chou *et al.*, 2002; Chin *et al.*, 2003) formed hairpin loops leading to pairing between the bases of the same strand and an overall strand–drug stochiometry of 1:1. In the only X-ray structure reported in the literature (Hou *et al.*, 2002) the DNA, which contains a CTG triplet sequence, formed slipped duplexes involving mismatched T–T base pairs.

AMD has also been shown to bind tightly and specifically to single-stranded DNA sequences (Wadkins & Jovin, 1991; Wadkins *et al.*, 1996, 2000; Yoo & Rill, 2001; Chen *et al.*, 2003, 2004). Fluorescence studies of 7-amino-actinomycin D (7-AAMD; a fluorescent derivative of AMD shown to bind dsDNA by Graves & Wadkins, 1989) with ssDNA indicate a high sequence-dependence of the binding. In particular, there seems to be an absolute requirement for guanine residues,

although not every guanine constitutes a potential binding site. Thus, sequences containing a 5′-AGT-3′ block were shown to possess a high affinity for the drug. AMD binding to ssDNA has been shown to inhibit HIV reverse transcriptase and other DNA polymerases (Rill & Hecker, 1996; Davis *et al.*, 1998; Imamichi *et al.*, 2003), highlighting the pharmacological and biochemical relevance of these investigations.

In a previous study, we proposed a model for the recognition of ssDNA sequences by actinomycin featuring the hemi-intercalation of the chromophore between the two purine residues (Wadkins *et al.*, 1996). These and subsequent not entirely consistent studies prompted the present investigation of the interactions of 7-AAMD with non-complementary 5′-T$_n$TAGTT$_n$-3′ (*n* = 2,3) by X-ray crystallography, augmented by NMR data and molecular modelling. The crystallographic studies were performed on the 7-AAMD–5′-TTAG[Br^{5}U]T-3′ system. ^{1}H NMR measurements were made on the closely related 5′-TTTAGTTT-3′ sequence, which unfortunately did not yield crystals. Both the 2:1 and 1:1 strand–drug complexes were investigated by molecular-dynamics simulations.

2. Materials and methods

2.1. Crystallization

The DNA oligomers 5′-TTAGTT-3′ and 5′-TTAG[Br^{5}U]T-3′ were obtained from Biotez. Stock solutions were prepared in water without further purification. 7-Amino-actinomycin D was purchased from Sigma–Aldrich Chemie GmbH (Steinheim, Germany) and dissolved in water as a stock solution. The concentrations of both DNA and 7-AAMD were determined by UV-absorption measurements.

Red hexagonally shaped crystals were obtained *via* the hanging-drop method at 293 K from solutions containing 2.3 m*M* 7-AAMD, 2.3 m*M* 5′-TTAGTT-3′, 2.3 m*M* gadolinium nitrate, 2.7 *M* ammonium sulfate, 0.05 *M* sodium/potassium tartrate and 0.1 *M* sodium citrate buffer pH 5.6. Their diffraction pattern at 100 K showed streaks in one direction, so the crystals were unsuitable for data collection, but a unit cell could be determined [*a* = *b* = 71.5 (2), *c* = 109.3 (3) Å, α = β = 90, γ = 120°]. Seeds of these crystals were transferred *via* cat whiskers to drops containing 2.5 m*M* 7-AAMD, 2.5 m*M* 5′-TTAG[Br^{5}U]T-3′, 2.5 m*M* gadolinium nitrate, 2.7 *M* ammonium sulfate, 0.05 *M* sodium/potassium tartrate and 0.1 *M* sodium citrate buffer pH 5.6, which were equilibrated at 293 K. Crystals appeared within a period of 2–3 d and could be frozen without breaking. Long rectangular plates of the orthorhombic form were obtained when seeding experiments were performed under the same conditions but without gadolinium nitrate in either solution, but it is possible that the seeding was incidental to their formation. It is interesting to note that no gadolinium sites were determined in the crystal structure of the hexagonal form once solved. This may be a consequence of the high sulfate concentration in the crystallization buffer.

Table 1
Data-collection statistics for both crystal forms.

Values in parentheses refer to the outer 0.1 Å resolution shell (except for the unit-cell parameters and occupancies, where they are standard uncertainties).

	Orthorhombic form		Trigonal form			
Wavelength (Å)	1.5418	0.8110	0.8463	0.9196	0.9204	0.9050
X-ray source	Rotating anode	X11	BW7B	BW7A	BW7A	BW7A
Resolution (Å)	2.4	2.0	2.5	2.4	2.4	2.5
Independent reflections	3039	5138	11493	12988	12964	11527
Space group	$C222_1$	$C222_1$	$P3_221$	$P3_221$	$P3_221$	$P3_221$
Unit-cell parameters						
a (Å)	51.56 (5)	51.59 (3)	71.48 (7)	71.22 (6)	71.17 (6)	71.22 (7)
b (Å)	70.89 (7)	70.96 (4)	71.48 (7)	71.22 (6)	71.17 (6)	71.22 (7)
c (Å)	39.66 (4)	39.45 (3)	109.32 (13)	108.41 (12)	108.35 (12)	108.40 (13)
Data statistics						
$\langle I/\sigma(I)\rangle$	53.8 (13.3)	19.9 (11.2)	26.4 (8.5)	16.8 (4.4)	16.5 (3.5)	21.3 (5.3)
R_{merge} (%)	7.5 (32)	3.2 (12.3)	7.2 (28)	8.8 (73)	7.5 (70)	8.0 (48)
Completeness (%)	99.5 (99.9)	99.6 (98.8)	99.0 (94.3)	98.9 (92.1)	98.8 (92.1)	99.0 (93.3)
Redundancy	41.5 (29.3)	13.9 (13.5)	10.9 (9.4)	30.6 (3.1)	30.2 (2.9)	30.9 (2.8)
f'† (electrons)	−0.77	−1.14	−1.50	−6.14	−8.50	−3.33
f''† (electrons)	1.28	3.05	3.20	4.34	3.80	3.70
Br occupancy	0.85 (11)	0.57 (4)	0.69 (2)	0.59 (3)	0.50(3)	0.49 (3)

† From Sasaki (1989).

Since a preliminary in-house room-temperature data set (not reported here) of the original trigonal 5′-(TTAGTT)-3′ complex was consistent with the final refined structure of the 5′-(TTAG[Br^{5}U]T)-3′ complex, it is reasonable to assume that the replacement of thymine by bromouracil did not significantly change the structure.

2.2. Data collection and processing

Diffraction data were collected at 100 K for both crystal forms of the 7-AAMD complex with brominated DNA (Table 1). Four synchrotron data sets were collected at EMBL, c/o DESY, Hamburg at different wavelengths from the same trigonal crystal (dimensions 0.3 × 0.3 × 0.01 mm) using MAR345 image-plate (BW7B) and MAR165 CCD (BW7A) detectors and processed with *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997). The space group was assigned as $P3_121$ or $P3_221$ with difficulty because the reflections with l odd were systematically weak. A fluorescence spectrum was recorded before the MAD measurements in order to locate the bromine K edge accurately. As can be seen from the statistics in Table 1, the R_{merge} values for the 2.5–2.4 Å shell collected and processed in two of the data sets are dramatically higher than those for the 2.6–2.5 Å shell and were eventually excluded. The very same crystal had also diffracted to 2.5 Å on our in-house source, so that the stronger source improved the signal-to-noise ratio but not the resolution of diffraction from this crystal.

Two data sets were collected at 100 K from the same orthorhombic crystal of dimensions 0.2 × 0.2 × 0.1 mm. The first data set was collected with a Bruker rotating-anode generator and SMART6000 CCD detector in-house to 2.4 Å resolution using Cu $K\alpha$ radiation. These data were integrated and scaled using the programs *PROTEUM* and *SADABS* (Bruker AXS). The systematic absences indicated a C-centred lattice and space group $C222_1$. A further data set from the same crystal was collected to 2.0 Å on the X11 beamline at

EMBL c/o DESY using a MAR165 CCD detector. However, after processing with *DENZO*, *SCALEPACK* and *XPREP* (Bruker AXS), the space group appeared to be $P2_12_12_1$ with the same unit-cell parameters. The superlattice reflections with $h + k \neq 2n$ were however very weak (about 1/5 of the intensity of the reflections with $h + k = 2n$). The C-centred cell was used for the successful structure solution and refinement reported here.

2.3. ^{1}H NMR spectroscopy

^{1}H NMR spectra were acquired at 291 K on a Varian Inova 500 (500.17 MHz) spectrometer at the Department of Organic Chemistry, University of Göttingen. AMD and ^{2}H$_2$O (99.996% deuterium) were purchased from Sigma.
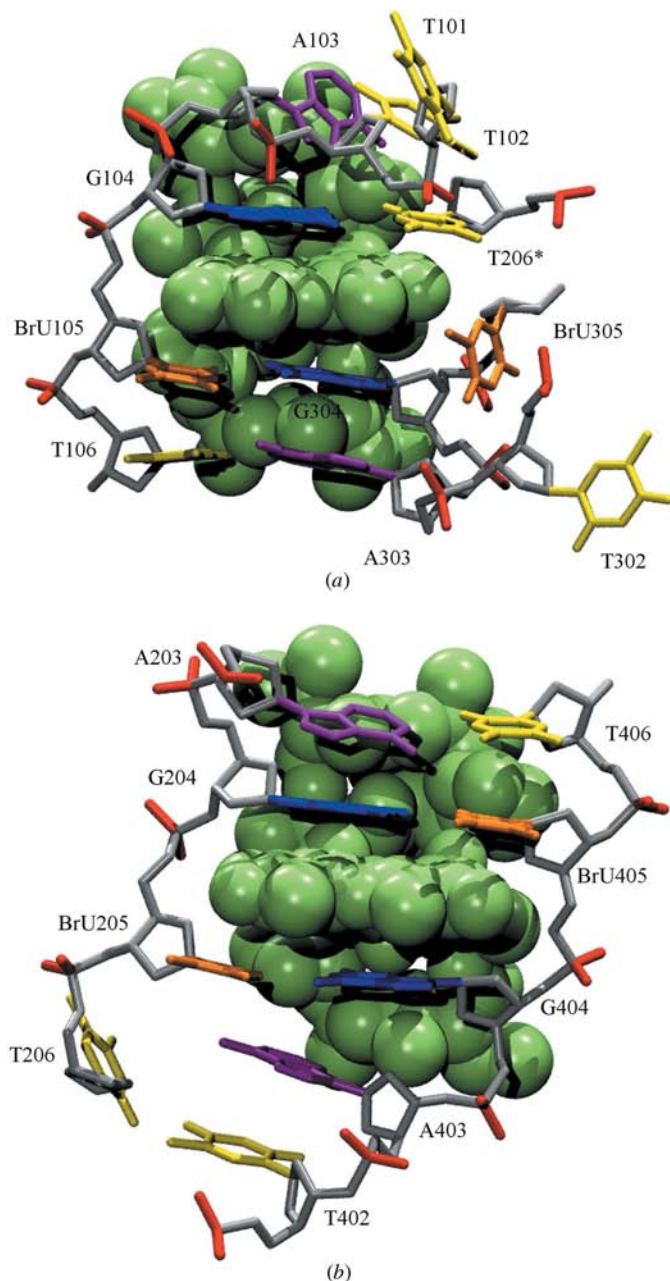
The assignments of AMD in ^{2}H$_2$O were based on literature data (Angerman *et al.*, 1972; Brown *et al.*, 1994) and confirmed by DQF-COSY, TOCSY and NOESY (mixing time 350 ms). Assignments of aromatic protons of T$_3$AGT$_3$ were based on comparison with literature data for mononucleotides, dinucleotides (Bovey, 1972) and a number of duplex DNA molecules and were confirmed by the DQF-COSY, TOCSY and NOESY experiments. NOESY spectra were collected with mixing times of 150, 250, 350, 400 and 500 ms.

3. Results and discussion

3.1. Structure determination of the orthorhombic form

The orthorhombic form was solved first. After a number of inconclusive attempts to solve the structure by SAD (single-wavelength anomalous diffraction) in both $P2_12_12_1$ and $C222_1$, the structure was solved in $C222_1$ by treating the two wavelengths, although neither was close to the absorption edge, as a pseudo-MAD (multiple-wavelength anomalous diffraction) experiment using *XPREP* to generate the F_A values. *SHELXD* (Sheldrick *et al.*, 2001; Schneider & Sheldrick, 2002)

found four heavy-atom sites with high correlation coefficients (CC all/weak = 51.7/31.8). After phasing and density modification with *SHELXE* (Sheldrick, 2002) using a solvent content of 45%, further density modification was performed using *DM* (Cowtan & Main, 1996) with a solvent content of 30%. The contrast and connectivity output by *SHELXE* were a little better for the original hand (0.390 and 0.867, respectively) than for the inverted structure (0.335 and 0.846). A 2:1 DNA–AMD complex was traced by hand in the *DM* map. *A posteriori* the map correlation coefficient against the final

Table 2
Final refinement statistics for the orthorhombic and trigonal structures.

	Orthorhombic	Trigonal
Space group	$C222_1$	$P3_221$
Resolution range (Å)	20–2.0	20–2.5
R_{work}	0.236	0.258
R_{free}	0.279	0.296
Water molecules	13	14
Data/restraints/parameters	5127/2493/2356	11503/17002/8996
R.m.s. deviations from idealized geometry		
Bond lengths (Å)	0.006	0.005
1,3-distances (Å)	0.019	0.015
Non-zero chiral volumes (Å³)	0.004	0.003
Distances from restrained planes (Å)	0.205	0.159
Mean B factors (Å²)		
DNA atoms	21.9	30.1
7-AAMD atoms	15.2	49.8
PDB code	1unm	1unj

refined structure was a little higher after *SHELXE* (0.794) than *DM* (0.775).

The model was subjected to least-squares refinement against F^2 of the synchrotron data set with *SHELXL97* (Sheldrick & Schneider, 1997). 5% of the reflections were selected in thin shells as an R_{free} set for cross-validation (Brünger, 1992). In the first ten refinement steps, the complex was built by hand using *XFIT* (McRee, 1999) with a stepwise increase of the resolution starting from 3 Å. Geometrical 1,2- and 1,3-distance restraints were taken from Parkinson (1996) and additional 1,2- and 1,3-distance restraints were generated with *SHELXPRO* (Sheldrick & Schneider, 1997) from the structures of actinomycin D (Schäfer *et al.*, 1998) and bromouracil (Sternglanz & Bugg, 1975). Planarity, chiral volume and antibumping restraints and a Babinet solvent model (Moews & Kretsinger, 1975) were employed, giving an initial R_{work} of 0.29 (R_{free} = 0.32). A few water molecules were added manually, selecting from the highest difference density peaks those that were approximately spherical and made reasonable contacts, and 12-parameter overall anisotropic scaling (Usón *et al.*, 1999) was applied. During the refinement, it was noticed that the B values of the Br atoms were higher than those of the other atoms of the corresponding base, so a common occupancy was refined for them, giving values of 0.85 (11) for the in-house data and 0.57 (4) for the synchrotron data that were collected later. Although a little bromouracil may have been substituted by thymine by the use of seed crystals, this is clear evidence of bromine loss as a result of radiation damage (Ennifar *et al.*, 2002). The final refinement statistics are summarized in Table 2.

3.2. Structure determination of the trigonal form

The trigonal form was first solved by molecular replacement by a rather indirect route that involved an intermediate lowering of the space-group symmetry using a model from the orthorhombic structure and the program *COMO* (Jogl *et al.*, 2001). Shortly afterwards, the four-wavelength bromine-MAD analysis led to a very clear independent solution of the structure. F_A values were derived by *XPREP* with refinement of the f' and f'' values. *SHELXD* found 16 heavy-atom sites

Figure 1
View of (*a*) the upper part of the complex (unit 1), consisting of strands 1 and 3 interacting with actinomycin molecule 1, and (*b*) the lower part (unit 2), consisting of strands 2 and 4 interacting with actinomycin molecule 2. Colour coding: guanine, blue; thymine, yellow; adenine, pink; bromouracil, orange; sugar, grey; phosphate, red.

with high correlation coefficients (all/weak = 59.2/49.3). 30 cycles of *SHELXE* density modification assuming 30% solvent were performed for both enantiomorphs ($P3_121$ and $P3_221$). The contrast and the connectivity were lower for $P3_121$ (0.321 and 0.843, respectively) than for $P3_221$ (0.382 and 0.885, respectively) confirming the latter space group; the map correlation coefficient against the refined structure was 0.832 (after recalculation with an improved version of *SHELXE*). Four copies of the core of the orthorhombic structure could be fitted into the resulting map and further bases could be traced by hand.

The model was refined in the same way as the orthorhombic form, except that fourfold non-crystallographic symmetry (NCS) restraints were also employed. Although the high-energy remote synchrotron data set was used for the refinement reported in Table 2, a common bromine occupancy was refined for each of the four wavelengths. The occupancies, shown in Table 1 in the order in which the data sets were collected, again strongly indicate that the Br atoms were lost from their original sites during data collection as a result of radiation damage. For both crystal forms, the data sets at wavelengths for which f' of bromine is least negative were measured first. Thus, the dispersive MAD differences and the pseudo-SIR differences caused by loss of bromine during irradiation will have the same phase and reinforce one another, fortuitously strengthening the MAD experiment and explaining the high-quality maps obtained. This is especially true for the orthorhombic structure, where the very modest MAD dispersive difference of −0.37 electrons is enhanced by an average of −9.8 electrons per site (calculated from the change in refined occupancy) caused by loss of bromine! For this reason, it is strongly advisable to collect the wavelength with the most negative f' (*i.e.* the inflection point) last in a MAD experiment.

3.3. Overall crystal structure description

Both crystal modifications of the 7-AAMD–5′-TTAG[Br^{5}U]T-3′ complex have very similar core structures, so the orthorhombic form that was determined to the highest resolution is discussed here. The drug-to-strand stoichiometry is 2:4. The four conformationally different single strands are partly paired, forming two antiparallel double strands with extensive base stacking. In

each one of these double strands an actinomycin molecule intercalates with the phenoxazone ring inserted between two guanines (Fig. 1).

The first double strand, shown in Fig. 1(*a*), is formed by the single strands 1 and 3, which are connected by two stacked base pairs, T106–A303 and BrU105–G304. The latter is an unusual G–T wobble pair (Fig. 2*a*) in which N1 and O6 of the guanine are hydrogen bonded to O3 and N3, respectively, of the pyrimidine. The stacking continues with the phenoxazone ring of the actinomycin and guanine G104, which pairs with thymine T206 of a symmetry-related molecule, again in wobble mode (Fig. 2*b*). In strand 1 the phosphate backbone performs a sharp turn after G104 with a torsion angle δ_{G104} of 137.7°. Adenine A103 and the thymines T102 and T101 do not participate in the stacking; the angles of their planes to that of the guanine G104 are 65.0, 27.8 and 60.7°, respectively. In this way, they can interact with symmetry-related strands. T101 and T102 participate in (A*T)*T triplets (Fig. 2*c*), whereas adenine A303 forms an A–T pair. The remaining two bases of
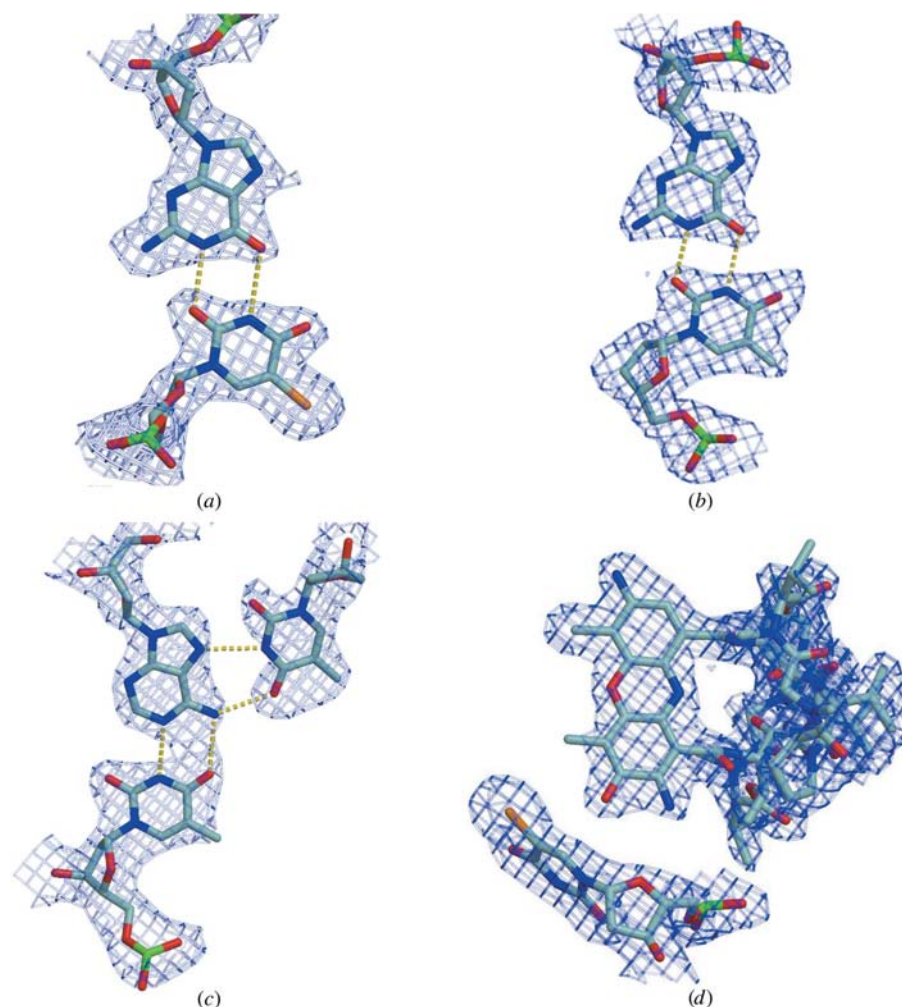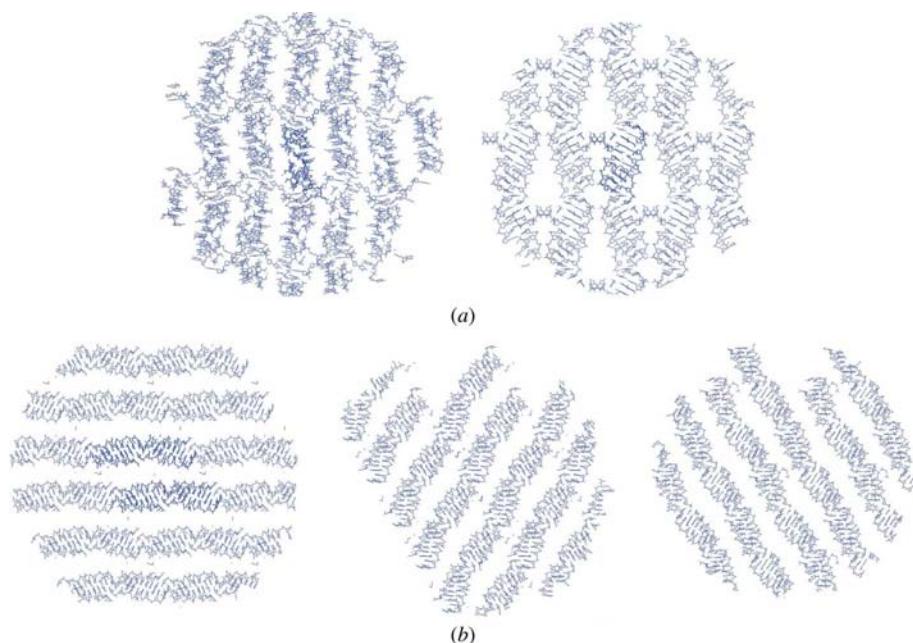

Figure 2
(*a*) The mismatched wobble base pair G304–BrU105 with hydrogen bonding between the N1 and the O6 of the guanine with the O2 and the N3 of the bromouracil, respectively; (*b*) the mismatched base pair G104–T206* showing similar interactions; (*c*) in a T*A*T triplet, A403 makes Watson–Crick interactions with T102* and Hoogsten interactions with T101*; (*d*) π-stacking of BrU305 to the phenoxazone ring of actinomycin molecule 1 (* indicates symmetry equivalent).

strand 3 are not involved in pairing. Bromouracil BrU305 is positioned perpendicularly to the phenoxazone ring (Fig. 2d) in a looped-out fashion with an angle of 98.1°. The Br atom points towards the methyl group C15 of the chromophore. Looped-out bases have also been observed in double-stranded DNA complexes with actinomycin (Chou et al., 2002). The thymine T302 stacks with the symmetry-equivalent base T302* of a neighbouring molecule. The first and last thymine of 3 are missing from the experimental density.

The second part of the complex, consisting of strands 2 and 4, is shown in Fig. 1(b). These strands are connected by three hydrogen-bonded base pairs. A Watson–Crick pair A203–T406 that stacks onto T106–A303 of the first part of the complex is followed by two G–BrU wobble pairs that sandwich the chromophore of the second actinomycin. The remaining bases of strand 4 interact with symmetry-related strands. The complex pattern of pairing and stacking of bases in different strands leads to approximately parallel columns that pack differently in the two crystal forms (Fig. 3). Two thymines of 2 and one thymine of 4 are missing from the experimental density. The missing (disordered) bases in all four strands always flank thymines that are not involved in pairing.

3.4. Analysis of the NMR data in the light of the crystal structure

Owing to the lack of a reasonable number of AMD–DNA correlations, a complete determination of the structure of the complex was not possible based on NOE constraints alone. The observed correlations were useful, however, to assess whether the crystallographic data would fit a solution structure. Most distance constraints from the NOESY spectra in D_2O were classified as strong, medium or weak based on their

Table 3
Comparison of key mean interproton distances from the orthorhombic crystal structure and molecular-dynamics simulations of the 1:1 complex corresponding to observed NOEs in the NMR spectra.

		NOE	Distance (X-ray) (Å)	Distance (MD) (Å)
Guanine H1′	Thr CH$_3$	Strong	3.6	3.8
Thymine H2′1	Thr CH$_3$	Weak	7.4	7.6
Pro H$^\alpha$	Thr CH$_3$	Weak	7.3	7.0

relative intensities at 150 ms mixing time and assigned generous distance bounds of 2.0–4.0, 3.0–5.0 or 4.0–6.0 Å, respectively. The most relevant correlations, not observed for AMD alone but present in the complex, are shown in Table 3; it can be seen that these distance estimates from NMR and the distances observed in the crystal structure are quite similar. The most important intermolecular NOE of the G(H1′)–Thr(CH$_3$) is also present in recent NMR studies of AMD and DNA hairpins (see the supplementary material for Chin et al., 2003).

4. Molecular-dynamics simulations of 1:1 and 2:1 d(TTTAGTTT)–AMD complexes

Molecular-dynamics calculations were carried out on a 1:1 d(TTTAGTTT)–AMD complex and the two 2:1 binding units of ssDNA–AMD found in the crystal structure, with unit 1 d(TAG[Br^{5}U]T)–AMD–d(TAG[Br^{5}U]) (Fig. 1a) and unit 2 d(TAG[Br^{5}U]T)–AMD–d(AG[Br^{5}U]T) (Fig. 1b). Bases of symmetry equivalents were not taken into account. The force-field parameters and charges of the Br^{5}U base were generated with the standard procedures described in the AMBER7 package (Case et al., 2002).

The starting structures for the ssDNA–AMD complex were modelled using the programs NAMOT (Tung & Carter, 1994), VMD (Humphrey et al., 1996), AMBER7 and NAMD (Kalé et al., 1999). For the relative orientation of the ssDNA and the AMD molecule, the two observed NOEs between the AMD and the ssDNA (Table 3) were taken into account: between the threonine methyl and H1′ of guanine (strong) and between the proline H$^\alpha$ and H2′1 of a thymine (weak). This ssDNA–AMD arrangement is also present in the crystal structure. The AMD forms a hemi-intercalation between the 5′-GpT-3′ and the 5′-GpBr^{5}U-3′ steps.

The current model differs from our previously proposed hemi-intercalation in which the AMD chromophore was stacked between the 5′-ApG-3′ step (Wadkins et al., 1996). That model was based on ^{1}H NMR spectra and molecular modelling. No distance data were

Figure 3
Crystal packing of the molecules in (a) the orthorhombic form (two types of layer) and (b) the trigonal form (three layers).

available at that time. The chemical shift calculations of the model structure did not include electrostatic and solvent effects, which are significant, but the results from the calculations could fit the majority of the experimental chemical shift data. The programs available were (and are) basically designed to calculate protein chemical shifts. Parameters for DNA and ligands are not available with the same accuracy as for amino acids. Based on the AMD–ssDNA NOEs and the crystal data now available, we have revised our model of the hemi-intercalation site by displacing the chromophore from the 5′ to the 3′ side of the guanine.
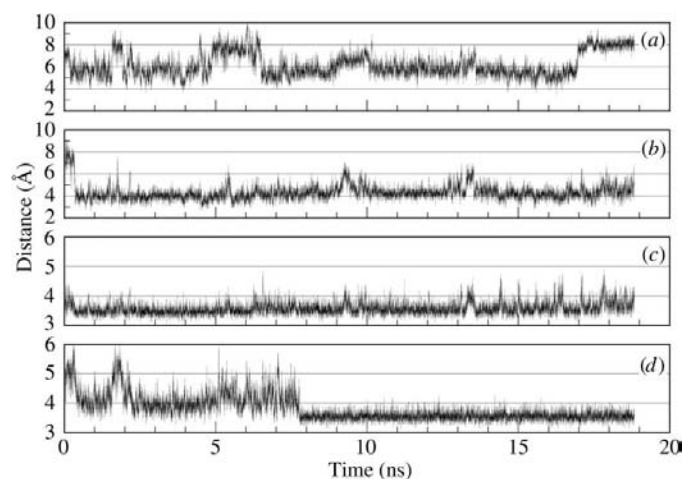
The residues in the ssDNA–AMD model system are numbered 1–11 for the AMD and 12–19 for the DNA. Seven different starting structures were generated by docking the AMD molecule onto the 3′-side of the guanine and the 5′-side of the adjacent thymine. Either the quinoid or the benzenoid ring of the AMD chromophore pointed towards the DNA backbone.

Two starting models of a 2:1 ssDNA–AMD complex were taken from the crystal data, representing the two different binding modes of AMD and DNA (Fig. 1). The top part with strands 1 and 3 is named unit 1 and the lower part with strands 2 and 4 is named unit 2.
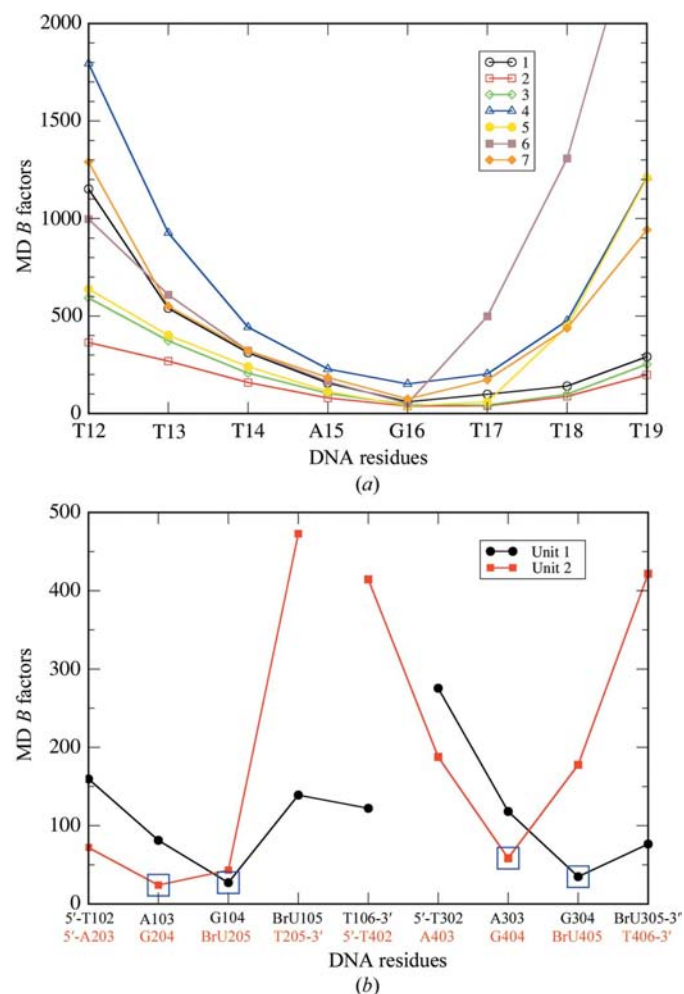
For both 1:1 and 2:1 complexes, the entire AMD–DNA system was solvated in a water box with boundaries 15 Å from the closest solute atom. Ions were added to the system to give a bulk salt concentration of 100 mM NaCl. The force-field parameters were taken from the $AMBER$7 program suite and the MD calculations were carried out with the $NAMD$ v.2.5 program on a parallel Linux cluster. The whole system was equilibrated during a period of 60 ps by slowly increasing the temperature from 0 to 300 K and by releasing the positional constraints of the solute atoms. All MD calculations were carried out with full electrostatic interactions and a Particle

Mesh Ewald (PME) summation technique. The production runs were performed in an NPT ensemble by fixing the pressure at 101.3 kPa. No NMR constraints were imposed on the DNA and AMD structures during these simulations.

MD simulation times for the seven 1:1 ssDNA–AMD starting structures ranged from 5 to 19 ns. The overall development of the trajectories was basically the same for all models. The positioning of the quinoid or the benzenoid ring with respect to the DNA did not influence the behaviour of the trajectories. In all simulations, the AMD chromophore remains tightly stacked on the 3′ side of the guanine G16 with an average distance of about 3.4 Å (Fig. 4c). The thymines flanking the d(AGT) core are rather flexible (Fig. 5a). This is also reflected in the motion of the thymine on the 3′ side of the guanine, which shows strong fluctuations during the first 7.5 ns of the 18.5 ns simulation, after which it stacks on top of the AMD chromophore (Fig. 4d). In one case the 3′-side thymine

Figure 5
B factors for DNA backbone atoms calculated from MD simulations. (a) For all seven simulations of a 1:1 AMD–ssDNA complex the guanine base fluctuates least. The strong fluctuations of the thymine bases 17–19 in model 6 is a consequence of destacking between the AMD chromophore and the adjacent T17. (b) Flexibility of the DNA backbone in MD simulations of the two binding modes of AMD in the crystal. As in the 1:1 model, the guanine bases fluctuate least and remain stacked on the AMD chromophore.

Figure 4
Variation of NOE distances during the MD simulation (model 1): (a) H$^\alpha$(DPR4) with the H2′1(T17) (weak), (b) methyl group of DTH4 and the H1′ of G16 (strong), (c) average distance between the AMD chromophore and the base G16, (d) average distance between the AMD chromophore and the base T17. G16 remains stacked on the chromophore over the entire range of the simulation, whereas T17 at first strongly fluctuates and then locks in into a stacked position.

T17 (Fig. 5*a*, model 6) loses the stacking interaction with the AMD chromophore and does not regain this interaction during the rest of the simulation. This high flexibility of the T-tracks is also the reason for the absence of experimental densities of the terminal thymines in the crystal structure. The conservation of the guanine–chromophore stacking is also seen in the calculated proton distances (Table 4*b*) between the H1′ of guanine and the methyl group of threonine; the average distance during the 18.5 ns trajectory of 3.7 Å accounts for the strong NOE signal. The weakness of the proline Hα–thymine H2′1 NOE arises from the fact that the protons are further apart and short distances are populated less often during the trajectory (Fig. 4*a*). The reproducibility of the G16:H1′–DPR4:Hα NOE distance in the MD simulations is remarkable since there were no distance restraints on the ssDNA–AMD complex.

The 20 ns MD simulations of unit 1 and unit 2 were not intended to reproduce the crystal environment but rather a possible intermediate between the 1:1 ssDNA–AMD complex in solution and the 4:2 ssDNA–AMD crystal packing (see Fig. 6).

For both trajectories, a similar tight stacking of the guanines onto the AMD chromophore was observed (Fig. 5*b*). During the course of the simulations, base pairings observed in the crystal (unit 1, BrU105–G304, T106–A303; unit 2, A203–T406, G204–BrU405, BrU205–G404) are lost owing to the higher flexibility in a solution environment. The flanking DNA residues form van der Waals contacts with either the AMD polypeptide chains or the chromophore. After about 4 ns all four Br⁵U bases orientate in a looped-out and perpendicular manner with respect to the chromophore, similar to the base orientations reported in an NMR study of AMD and DNA

Table 4
Comparison of key mean interproton distances from the orthorhombic crystal structure and molecular-dynamics (MD) simulations of the 1:2 complex corresponding to observed NOEs in the NMR spectra.

		NOE	Distance (X-ray) (Å)	Distance (MD) (Å)
Guanine H1′	Thr CH₃	Strong	3.6	3.8
Thymine H2′1	D-Pro Hα	Weak	7.9	7.6
D-Val Hα	D-Val CH₃$^{\gamma,1,2}$	Strong	3.2	3.1
Thr Hβ	Me-Val CH₃$^{\gamma}$	Medium	5.7	5.6
Thr Hβ	D-Val CH₃$^{\gamma,1}$	Strong	4.5	3.9
Sar Hα	Me-Val NCH₃	Strong	3.0	3.2

hairpins (Chou *et al.*, 2002). Inter-strand stacking of bases is found in unit 2 between T206, T402 and A403. The T406 base stacks onto the D-Thr and Me-Val residues on the back of the AMD. During the time of the MD simulation the DNA strands and the AMD form a stable complex, despite their short lengths of just four and five residues.

5. Conclusions

Combining the X-ray, NMR and the MD results, it is clear that the most important DNA–AMD binding mode is the tight stacking of the guanine base onto the AMD chromophore. This is also seen in other experiments such as the intercalation of AMD in dsDNA between the GpC step or single-stranded sequences forming mini-hairpins including a CG base pair (Chin *et al.*, 2003). In the crystal structure and the MD simulations of the 1:2 AMD ssDNA complex, the chromophore is firmly sandwiched between two guanines. The thymines (or closely related bromouracils) that make wobble pairs with
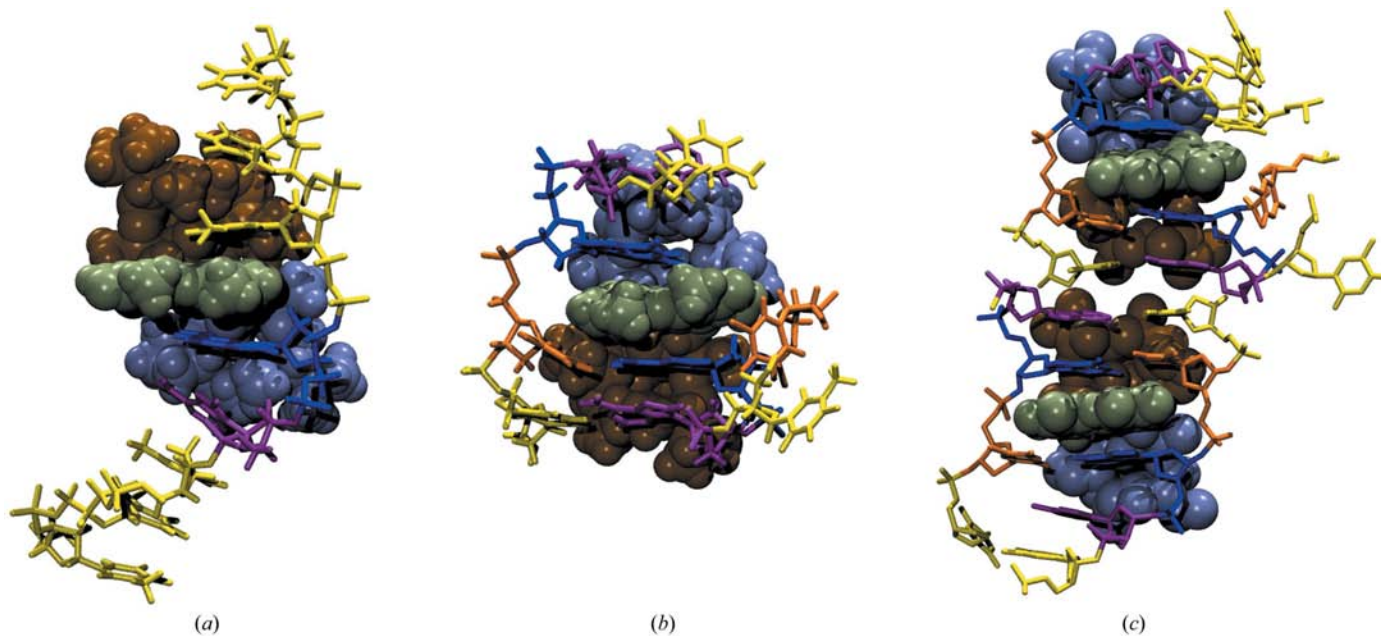
Figure 6
(*a*) Hemi-intercalation MD model of the AMD bound to the d(TTTAGTTT) single-stranded DNA. The AMD chromophore stacks on the 3′ side of the guanine base (G16). (*b*) MD model (unit 1) corresponding to the interaction of AMD molecule 1 with strands 1 and 3 in the crystal. (*c*) The 2:4 complex in the crystal. AMD molecule 1 and DNA strands 1 and 3 are on top and molecule 2 and strands 2 and 4 are at the bottom.

these guanines are fluxional and can become detached within the timescale of the MD simulations; they clearly play at most a minor role in the binding of the AMD molecule. The experimental data coupled with molecular modelling leads us to this new and specific hemi-intercalation model for the complex, which could represent an intermediate structure in the stepwise interaction of the drug with single-stranded DNA. The combination of crystallography, NMR spectroscopy and MD simulations suggest a hierarchy of association of DNA–drug complexes. Starting from a 1:1 drug:DNA stoichiometry at low AMD and DNA concentrations in solution, this complex can acquire a second ssDNA with 1:2 stoichiometry, which finally forms 2:4 aggregates with quasi-antiparallel double-stranded DNA in the crystal structure.

References

Angerman, N. S., Victor, T. A., Bell, C. L. & Danyluk, S. S. (1972). *Biochemistry*, **11**, 2402–2411.

Bovey, F. A. (1972). *High Resolution NMR of Macromolecules*. New York: Academic Press.

Brown, D. R., Kurz, M., Kearns, D. R. & Hsu, V. L. (1994). *Biochemistry*, **33**, 651–664.

Brünger, A. T. (1992). *Nature (London)*, **355**, 472–475.

Case, D. A., Pearlman, D. A., Caldwell, J. W., Cheatham, T. E., Ross, W. S., Simmerling, C. L., Dardeb, T. A., Merz, K. M., Stanton, A. L., Cheng, A. L., Vincent, J. J., Crowley, M., Ferguson, D. M., Radmer, R. J., Seibel, G. L., Singh, U. C., Weiner, P. K. & Kollman, P. A. (2002). *AMBER*7 computer program. University of California, San Francisco, USA.

Chen, F.-M. (1988). *Biochemistry*, **26**, 4323–4331.

Chen, F.-M. & Sha, F. (2002). *Biochemistry*, **41**, 5043–5049.

Chen, F.-M., Sha, F., Chin, K.-H. & Chou, S.-H. (2003). *Biophys. J.* **84**, 432–439.

Chen, F.-M., Sha, F., Chin, K.-H. & Chou, S.-H. (2004). *Nucleic Acids Res.* **32**, 271–277.

Chen, H., Liu, X. & Patel, D. J. (1996). *J. Mol. Biol.* **258**, 457–479.

Chin, K.-H., Chen, F.-M. & Chou, S.-H. (2003). *Nucleic Acids Res.* **31**, 2622–2629.

Chou, S.-H., Chin, K.-H. & Chen, F.-M. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 6625–6630.

Cowtan, K. & Main, P. (1996). *Acta Cryst.* D**52**, 43–48.

Davis, W. R., Gabbara, S., Hupe, D. & Peliska, J. A. (1998). *Biochemistry*, **37**, 14213–14221.

Ennifar, E., Carpentier, P., Ferrer, J.-L., Walter, P. & Dumas, P. (2002). *Acta Cryst.* D**58**, 1262–1268.

Graves, D. E. & Wadkins, R. M. (1989). *J. Biol. Chem.* **264**, 7262–7266.

Hou, M.-H., Robinson, H., Gao, Y.-G. & Wang, A. H.-J. (2002). *Nucleic Acids Res.* **30**, 4910–4917.

Humphrey, W., Dalke, A. & Schulten, K. (1996). *J. Mol. Graph.* **14**, 33–38.

Imamichi, T., Murphy, M. A., Adelsberger, J. W., Yang, J., Watkins, C. M., Berg, S. C., Baseler, M. W., Lempicki, R. A., Guo, J., Levin, J. G. & Lane, H. C. (2003). *J. Virol.* **77**, 1011–1020.

Jain, S. C. & Sobell, H. M. (1972). *J. Mol. Biol.* **68**, 1–20.

Jogl, G., Tao, X., Xu, Y. & Tong, L. (2001). *Acta Cryst.* D**57**, 1127–1134.

Kalé, L., Skeel, R., Bhandarkar, M., Brunner, R., Gursoy, A., Krawetz, N., Phillips, J., Shinozaki, A., Varadarajan, K. & Schulten, K. (1999). *J. Comput. Phys.* **151**, 283–312.

Kamitori, S. & Takusagawa, F. (1994). *J. Am. Chem. Soc.* **116**, 4154–4165.

Lian, C., Robinson, H. & Wang, A. H.-J. (1996). *J. Am. Chem. Soc.* **118**, 8791–8801.

McRee, D. (1999). *J. Struct. Biol.* **125**, 156–165.

Moews, P. C. & Kretsinger, R. H. (1975). *J. Mol. Biol.* **91**, 201–225.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Parkinson, G. (1996). *Acta Cryst.* D**52**, 57–64.

Qu, X., Ren, J., Riccelli, P. V., Benight, A. S. & Chaires, J. B. (2003). *Biochemistry*, **42**, 11960–11967.

Rill, R. L. & Hecker, K. H. (1996). *Biochemistry*, **35**, 3525–3533.

Robinson, H., Gao, Y.-G., Yang, X.-L., Sanishvili, R., Jaochimiak, A. & Wang, A. H.-J. (2001). *Biochemistry*, **40**, 5587–5592.

Sasaki, S. (1989). KEK Report 88-14, pp. 1–136.

Schäfer, M., Sheldrick, G. M., Bahner, I. & Lackner, H. (1998). *Angew. Chem. Int. Ed.* **37**, 2381–2384.

Schneider, T. R. & Sheldrick, G. M. (2002). *Acta Cryst.* D**58**, 1772–1779.

Sheldrick, G. M. (2002). *Z. Kristallogr.* **217**, 644–650.

Sheldrick, G. M., Hauptmann, H. A., Weeks, C. M., Miller, R. & Usón, I. (2001). *International Tables for Crystallography*, Vol. F, edited by E. Arnold & M. Rossmann, pp. 333–351. Dordrecht: Kluwer Academic Publishers.

Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol.* **277**, 319–343.

Sternglanz, H. & Bugg, C. E. (1975). *Biochim. Biophys. Acta*, **378**, 1–11.

Tung, C. S. & Carter, E. S. (1994). *Comput. Appl. Biosci.* **10**, 427–433.

Usón, I., Pohl, E., Schneider, T. R., Dauter, Z., Schmidt, A., Fritz, H.-J. & Sheldrick, G. M. (1999). *Acta Cryst.* D**55**, 1158–1167.

Wadkins, R. M., Jares-Erijman, E. A., Klement, R., Ruediger, A. & Jovin, T. J. (1996). *J. Mol. Biol.* **262**, 53–68.

Wadkins, R. M. & Jovin, T. M. (1991). *Biochemistry*, **30**, 9469–9478.

Wadkins, R. M., Tung, C. S., Vallone, P. M. & Benight, A. S. (2000). *Arch. Biochem. Biophys.* **384**, 199–203.

Waring, M. J. (1981). *Annu. Rev. Biochem.* **50**, 159–192.

Wilson, W. D., Jones, R. L., Zon, G., Scott, E. V., Banville, D. L. & Marzille, L. G. (1986). *J. Am. Chem. Soc.* **108**, 7113–7114.

Yoo, H. & Rill, R. L. (2001). *J. Mol. Recog.* **14**, 145–150.