Crystal Structure of a Covalent DNA−Drug Adduct: Anthramycin Bound to C-C-A-A-C-G-T-T-G-G and a Molecular Explanation of Specificity[†,‡]

Mary L. Kopka,[*] David S. Goodsell, Igor Baikalov, Kazimierz Grzeskowiak, Duilio Cascio, and Richard E. Dickerson

Molecular Biology Institute, Department of Chemistry and Biochemistry, and Institute of Geophysics and Planetary Physics, University of California, Los Angeles, California 90024

Received May 3, 1994; Revised Manuscript Received August 2, 1994[⊗]

ABSTRACT: A 2.3-Å X-ray crystal structure analysis has been carried out on the antitumor drug anthramycin, covalently bound to a ten base pair DNA double helix of sequence C-C-A-A-C-G-T-T-G-G. One drug molecule sits within the minor groove at each end of the helix, covalently bound through its C11 position to the N2 amine of the penultimate guanine of the chain. The stereochemical conformation is C11*S*, C11a*S*. The natural twist of the anthramycin molecule in the C11a*S* conformation matches the twist of the minor groove as it winds along the helix; a C11a*R* drug would only fit into a left-handed helix. The C11*S* attachment is roughly equatorial to the overall plane of the molecule, whereas a C11*R* attachment would be axial and would obstruct the fitting of the drug into the groove. The six-membered ring of anthramycin points toward the 3′ end of the chain to which it is covalently attached or toward the end of the helix. The acrylamide tail attached to the five-membered ring extends back along the minor groove toward the center of the helix, binding in a manner reminiscent of netropsin or distamycin. The drug−DNA complex is stabilized by hydrogen bonds from C9-OH, N10, and the end of the acrylamide tail to base pair edges on the floor of the minor groove. The origin of anthramycin specificity for three successive purines arises not from specific hydrogen bonds but from the low twist angles adopted by purine−purine steps in a B-DNA helix. Binding of anthramycin induces a low twist at T-G in the T-G-G sequence of this DNA−drug complex, by comparison with the structure of the free DNA. The origin of anthramycin's preference for adenines flanking the alkylated guanine arises from a netropsin-like fitting of the acrylamide tail into the minor groove.

Among the drugs known to interact with DNA, two classes of binding have been studied extensively by crystallographers—noncovalent AT-specific minor groove binding drugs such as netropsin and berenil (Kopka et al., 1985a−c; Coll et al., 1987; Larsen et al., 1989; Quintana et al., 1991; Brown et al., 1990; Edwards et al., 1992) and drugs such as daunomycin whose aromatic rings intercalate between the base pairs of DNA (Ughetto et al., 1985; Frederick et al., 1990; Gao et al., 1990; Kamitori & Takusagawa, 1992). A third class of groove-binding drugs that has not been investigated structurally is organic molecules that make covalent bonds with DNA. One such covalently-binding drug molecule is anthramycin, a member of the pyrrolo[1,4]-benzodiazepines (Figure 1a), which forms a linkage to guanine. The present crystal structure analysis of this guanine-specific drug offers a look at GC recognition at the molecular level and a first look at a naturally occurring covalent complex of an antibiotic with DNA.

Anthramycin is an antitumor antibiotic from a thermophilic actinomycete, *Streptomyces refuineus*, originally found growing in a compost heap and subsequently isolated from a fermentation broth. It originally was called "refuin" from the Hebrew word for health, at a time when it was believed to be a protein (Tendler & Korman, 1963; Korman & Tendler, 1965). Later it was renamed "anthramycin" when its structure was deduced (Leimgruber et al., 1965a,b) to reflect the anthranilic acid component of the left half of the molecule as drawn in Figure 1a. Investigators in other parts of the world also separated other members of this benzodiazepine family from various strains of *Streptomyces*. Sibiromycin (from Siberia; Gause & Dudnik, 1971) was discovered in Russia, and tomaymycin (from the Tomei region between Tokyo and Nagoya, Figure 1b; Arima et al., 1972) and the neothramycins A and B (Takeuchi et al., 1976) were isolated in Japan, making these drugs a truly international family.

Leimgruber and co-workers (1965a,b, 1968) isolated, characterized, and eventually synthesized anthramycin. This initial work was followed by a series of papers (Kohn & Spears, 1970; Glaubiger et al., 1974; Kohn et al., 1974) establishing the following: (1) Anthramycin binds to double-stranded DNA but not to RNA. (2) A covalent bond is formed between the drug and DNA. This was demonstrated by showing that anthramycin remains bound to one strand of denatured DNA and that standard methods of removing the drug fail. (3) Anthramycin cannot be an intercalator, because the plane of its rings is oriented approximately 36° to the helix axis. (4) The drug forms a bond with guanine residues, because only guanine-containing polydeoxynucleotides react with anthramycin. (5) The N2 of guanine is

[†] This work was carried out with the support of NIH Program Project Grant GM-31299 and American Cancer Society Grant DNP-22G.

[‡] Final coordinates have been deposited with the Brookhaven Protein Data Bank (access code 274D).

* To whom correspondence should be addressed at the Molecular Biology Institute, UCLA.

[⊗] Abstract published in *Advance ACS Abstracts,* October 15, 1994.

0006-2960/94/0433-13593$04.50/0

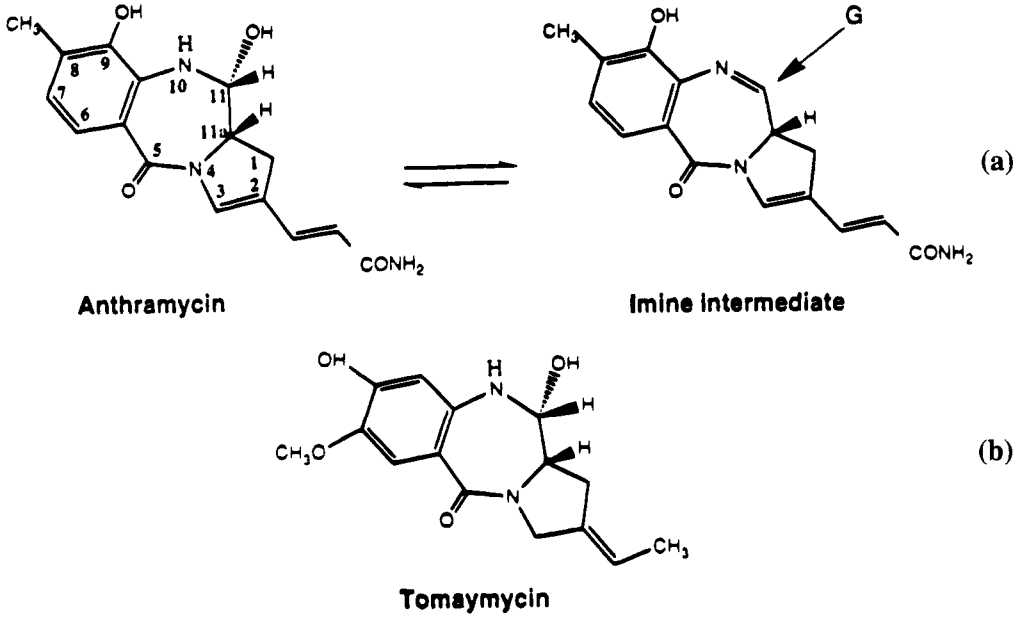

FIGURE 1: Pyrrolo[1,4]benzodiazepine antibiotics (a) anthramycin and (b) tomaymycin. Each drug consists of three rings; a six-membered substituted phenol ring, a diazepine ring in the boat conformation, and a pyrrole ring of a degree of saturation varying from one antibiotic to the other. Anthramycin forms a covalent bond between atom C11 and the N2 amine of guanine. The reaction is postulated to go through an imine intermediate (right) in which the original stereochemistry at C11 is lost.

implicated in bond formation because anthramycin does not react with poly(dI)·poly(dC), which lacks an amine group at the 2-position.

Nuclear magnetic resonance experiments confirmed the aminal linkage between the N2 of guanine and the C11 of anthramycin (Graves et al., 1984). For anthramycin, the carbinolamine function of the drug, N10−C11-OH, is necessary for binding, as is the acrylamide tail (Horowitz et al., 1971; Horowitz, 1971). One of the most interesting biological results of the research by Horowitz is that anthramycin inhibits the DNA-dependent RNA and DNA polymerase reactions by binding to the DNA template, thus explaining how the drug works in the cell.

X-ray crystal structure analyses of anthramycin methyl ether, with $-OCH_3$ in place of $-OH$ on carbon C11 (Mostad, 1978; Arora, 1979), and tomaymycin methyl ether (Arora, 1981) have demonstrated that, in the free drug, the stereochemistry at the two chiral carbons, C11 and C11a, is *R* and *S*, respectively, as drawn in Figure 1a.

In contrast to other groove-binding drugs that interact with DNA in seconds or milliseconds, the reaction of anthramycin with DNA takes about 1 h to complete. Mechanisms were proposed by Hurley (Hurley, 1977; Hurley et al., 1977) and by Lown and Joshua (1979). Hurley originally favored a nucleophilic attack via an Sn_1ca mechanism, while Lown believed complex formation occurred via an imine as shown at the right of Figure 1a. At present, the imine mechanism is the accepted explanation. The imine intermediate loses chirality at atom C11, but as we shall see from the X-ray analysis, only one stereoisomer is re-formed in the covalent DNA complex, probably because of the remaining chirality at atom C11a and of constraints involved in fitting the drug into the minor groove.

MATERIALS AND METHODS

Preparation of DNA Adducts. The C-C-A-A-C-G-T-T-G-G decamer (Privé et al., 1991) was synthesized in our laboratory by the solid-phase phosphoramidite method and purified on a DE52 (Whatman) column using a KCl gradient. Anthramycin methyl ether was kindly donated by Hoffman-La Roche. Anthramycin methyl ether, dissolved in absolute methanol (Fisher Scientific) at a concentration of 5 mg/mL, was mixed with the DNA in a 4:1 drug to decamer helix ratio and allowed to shake on a LABQUAKE test-tube rocking platform for 48 h at 4 °C. This reaction procedure was based on the sample preparation used for NMR analysis of anthramycin bound to a hexamer (Graves et al., 1985). An analytical size sample (~0.1 OD) then was injected into a Beckman HPLC system using a reversed-phase Beckman Ultrasphere ODS column. A gradient, consisting of 0−50% Fisher HPLC-grade acetonitrile in 5 mM ethylenediamine−acetate buffer (EDAA), was used to elute the sample. Figure 2 shows the profile obtained. If running time was increased, improved resolution of the peaks could be obtained. Therefore, a semipreparative size column was substituted for the analytical one, and larger samples were injected (~75 OD), using an increased running time of 75 min rather than 40 min. The individual peaks were collected, UV spectra were taken, and the 321/349-nm ratio (Kohn et al., 1974) was calculated to characterize each peak in Figure 2.

Peaks b and c in Figure 2 represent covalent adducts of anthramycin with DNA, which will be designated as Ant-1 and Ant-2, respectively. Each adduct was purified a second time on HPLC, and adduct Ant-2 was given a third purification (Table 1). Ant-1 did not go through a third purification because of a lack of material. At the end of purification, Ant-1 and Ant-2 were dialyzed against distilled water to remove any EDAA. Each adduct then was lyophilized to a pale yellow "cottony" solid. The adducts were finally dissolved in Millipore-filtered distilled water to a concentration of ~10 mg/mL.

Crystallization and Data Collection. Crystals could be obtained at all phases of purification, including the unseparated mixture as well as the individual adducts, and unit cell and space group information is given in Table 1. However, the crystal quality, as judged by resolution of the diffraction

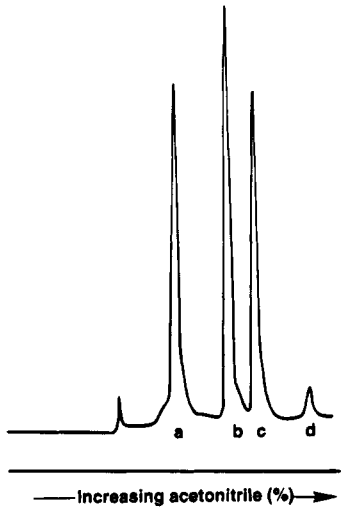

FIGURE 2: HPLC of the anthramycin–CCAACGTTGG reaction mixture after 48 h of reaction time. The running time for HPLC was 40 min. Increasing the running time to 75 min led to even sharper separation of individual peaks. Peaks: a, DNA ($OD_{321}/OD_{349} = 0$; DNA does not absorb at these wavelengths); b, adduct 1 ($OD_{321}/OD_{349} = 0.693$); c, adduct 2 ($OD_{321}/OD_{349} = 0.720$); d, unreacted anthramycin ($OD_{321}/OD_{349} = 1.113$). The OD_{321}/OD_{349} ratio is expected to be 1.134 for free anthramycin and 0.697 for bound drug.

pattern, increased dramatically during purification. Since the highest resolution was obtained with three times purified Ant-2, this form was used for structure analysis. Crystals of Ant-2 were grown under high-salt conditions by vapor diffusion at 4 °C in sitting drops. The crystallizing medium contained 0.45 mM decamer–anthramycin adduct, 317 mM $MgCl_2$, and 9% v/v absolute methanol with 19.5% v/v 2-methyl-2,4-pentanediol (MPD) as the precipitating agent. The methanol was added to help keep the DNA–adduct in solution. The external reservoir contained 45% MPD and was increased by either 5% MPD or the addition of $MgCl_2$ every 2 weeks (e.g., first change was 45% MPD + 0.48 M $MgCl_2$). Under the very high salt conditions needed for crystal formation, crystals would not form without salt in the reservoir as well as MPD. Two yellow-gold crystal polymorphs—hexagonal plates and hexagonal rods—appeared at 60% MPD + 0.48 M $MgCl_2$. Both crystal forms gave identical diffraction patterns. The crystal habit appeared to be dependent on the solution environment. A plate-shaped crystal was used for X-ray survey photography, and a rod of 0.3 mm × 0.2 mm × 0.1 mm was used for data collection on a Rigaku R-Axis II image plate. The crystals were very sensitive to X-rays. On a Rigaku AFC-5R automatic diffractometer at 4 °C, intensities would remain strong until ca. 8 h of exposure from a rotating anode source and then suffer a rapid and catastrophic decay, making a full data set impossible to obtain. But a rapid data set could be collected within 5 h on the image plate without serious decay.

Structure Solution. (A) Ant-1 vs Ant-2. For both Ant-1 and Ant-2, the space group was initially undecided between the two enantiomers $P3_121$ and $P3_221$. A choice between enantiomers could not be made until each structure had been solved: Ant-1 to a low resolution that was not pursued further and Ant-2 as described below. Ant-1 ultimately proved to adopt space group $P3_121$ with ten base pairs or one double helix per asymmetric unit (asu), and Ant-2, space group $P3_221$ with five base pairs or one strand per asu. Statistics of data collection and refinement for Ant-2 are given in Table 2.

(B) Rotation Search. The structure was solved by molecular replacement using X-PLOR version 3.0 (Brunger et al., 1987). The starting model was one strand of an idealized B-DNA decamer helix (Chandresekaran & Arnott, 1989) with the helix axis parallel to z. The solution with the highest Patterson correlation coefficient, $(\theta_1, \theta_2, \theta_3) = (0°, 23.4°, 342°)$, placed the model helix axis perpendicular to the x axis of the unit cell and rotated 23.4° down from the z direction around that axis. Rotation $\theta_3 = 342°$ about the helix axis itself generated a duplex with the minor groove facing in the positive x direction. A solution with the minor axis facing the negative x direction, or with $\theta_3 = 162°$, gave a markedly lower Patterson correlation coefficient. The value of 23.4° for helix inclination from z was corroborated by a cluster of strong 3.4-Å reflections in the survey X-ray photographs, as is encountered frequently in crystals of B-DNA oligomers.

(C) Translation Search. Space group symmetry and the molecular dyad through the helix require the center of the helix to lie at either $z = 1/3$ or $1/6$ for the two possible enantiomers, $P3_121$ or $P3_221$, respectively, but the location along the x axis is undetermined. To find the correct x position, a translational search was carried out at 1.0-Å intervals along the entire 24-Å cell length, for both $z = 1/3$ and $z = 1/6$. The solution with highest correlation coefficient was $x = 16.0$ Å (0.668), $z = 14.82$ Å ($1/6$), and space group $P3_221$.

(D) Simultaneous Rotation/Translation Search. Space group symmetry requires that the perpendicular dyad of the double helix lie along a crystallographic 2-fold, which can be defined as the x axis. This determines two of the three rotations and two of the three translations, leaving only two unknown variables: rotation around and displacement along the x axis. Figure 3 shows a two-dimensional rotation/translation search in space group $P3_221$ using the correlation coefficient between I_o and I_c. The single 6σ peak of height 0.34 yields a clear answer: rotation of 24° from the z direction and an x coordinate of 0.68. Reversing the helix to face the minor groove in the opposite direction, or choosing the opposite $P3_121$ enantiomer, produced maximum correlation peak heights of less than 0.20. Hence, the two trials gave the same solution within limits of convergence of subsequent refinement.

Structure Refinement and Fitting of Anthramycin. X-PLOR rigid body refinement using data initially from 8 to 3 Å was carried out in five steps: (a) the entire strand as one rigid unit; (b) five rigid bodies, nucleotides 1–2, 3–4, 5–6, 7–8, and 9–10; (c) 10 individual rigid nucleotides; (d) nucleosides and phosphates separated; and (e) bases, sugars, and phosphates uncoupled. Data were extended from 3.0 to 2.7 Å in the final rigid body step. The initial R-factor of 50.3% dropped to 36.6% during this overall process.

NUCLSQ refinement (Hendrickson & Konnert, 1980) followed, with gradual addition of data to 2.3 Å. A difference map at this stage indicated substantial positive electron density within the minor groove near guanine 9 (and of course in the symmetry-equivalent position near guanine 19 at the opposite end of the helix). Before building a drug molecule into this density, phases were improved by adding seven water molecules in locations far from the minor groove. For a peak to be interpreted as a solvent molecule,

Table 1: Increased Resolution with Purification

purity stage	resolution (Å)	unit cell dimensions[a] $a = b$ (Å)	c (Å)	space group	data collection	base pairs per asu
mixture of adducts	7.0	33.14	103.7	$P3_121$ or $P3_221$?	X-ray photos	?
1 × HPLC						
Ant-1	3.4	33.50	107.1	$P3_121$	diffractometer	10
Ant-2	3.2	23.95	89.6	$P3_221$	X-ray photos	5
2 × HPLC						
Ant-1	3.0	33.45	107.95	$P3_121$	image plate	10
Ant-2	3.0	23.43	89.95	$P3_221$	diffractometer	5
3 × HPLC						
Ant-2	2.3	23.95	88.89	$P3_221$	image plate	5

[a] All unit cell angles: $\alpha = 90°$, $\beta = 90°$, and $\gamma = 120°$.

Table 2: Statistics of Data Collection and Refinement at 2.3 Å for the 3 × HPLC Ant-2 Crystals of Table 1

no. of unique reflections over $2\sigma(F)$	1212
completeness of data (%)	90
2.4−2.3-Å outermost shell (%)	58
R-merge (on I) (%)	6.38
final F-factor (%)	21.2
no. of DNA atoms (per strand or asymmetric unit)	202
no. of anthramycin atoms (per strand or asu)	22
no. of bound water molecules (per strand or asu)	21
stereochemical refinement parameters[a]	
bond distances (Å)	
bases and sugars	0.030 (0.030)
phosphates	0.032 (0.030)
angle distances (Å)	
bases and sugars	0.058 (0.040)
phosphates	0.061 (0.040)
average isotropic B-values (Å^2)	
overall B	34.26
phosphates	43.96
sugars	34.51
bases	28.17
drug	29.12
waters	47.27

[a] Numbers in stereochemical parameters are rms deviations from ideal values in the final model, followed in parentheses by target variance used in refinement.

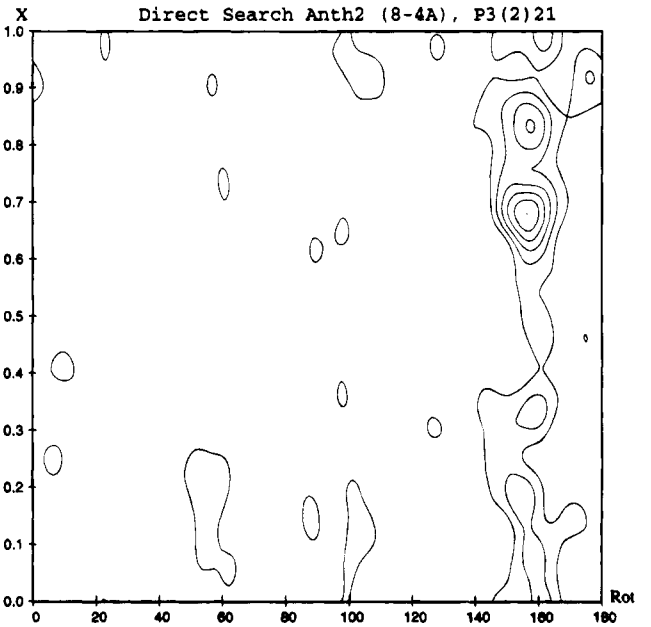

FIGURE 3: Two-variable rotation−translation Patterson correlation function plot. The helix sits on the x axis with its internal 2-fold symmetry axis along x. The vertical axis measures the distance of the helix from the origin along x, and the horizontal axis measures the tip of the helix axis away from being parallel to the z axis. See text for details.

it had to rise simultaneously above the 1σ level in the $(2F_o - F_c)$ map, rise above 2.5σ for the $(F_o - F_c)$ map, and lie within good hydrogen-bonding distances of acceptor or donor atoms. The R-factor just prior to addition of anthramycin was 26.2%. The helix axis of the refined dodecamer without drug is inclined 22° to the c axis.

Coordinates for the anthramycin molecule itself were available from two independent X-ray structure analyses (Mostad et al., 1978; Arora, 1979). The drug was fitted into density within the minor groove (Figure 4) on an Evans and Sutherland PS390 picture system using FRODO (Jones, 1978). However, since the covalent bond from the DNA is made to a chiral C11 atom on the drug, four isomers were possible: two stereoisomers, R and S, and two geometric isomers with the six-membered ring facing toward either the 3′ or 5′ end of the alkylated DNA strand. (For reference, the S isomer has the attachment from guanine oriented as with the C11−H bond in Figure 1a or on the same side of the drug molecule as the C11a hydrogen. Attachment from the direction of the C11−OH bond produces the R isomer.) These four geometries will be termed the 11S-3′, 11R-3′, 11S-5′, and 11R-5′ conformations, following the nomenclature of Boyd et al. (1990b).

Anthramycin molecules were fitted into the observed density in all four conformations. The 11R-5′ conformer was rejected at the outset for failing to fit the observed electron density. Models 11S-5′, 11R-3′, and 11S-3′ were refined by X-PLOR, treating the entire drug molecule as one rigid body, and at this point gave R-factors of 31.9%, 30.3%, and 29.5%, respectively. These differences were suggestive but hardly conclusive. Refinement was continued with one cycle of positional refinement and then one cycle of B refinement using NUCLSQ. Observation of difference maps indicated that 11S-3′ was slightly favored over 11R-3′, and both were much better than 11S-5′ with the molecule turned in the opposite direction. Upon continued NUCLSQ refinement, the 11R-3′ model inverted spontaneously to 11S-3′. NUCLSQ constraints themselves could not have produced the stereochemical interconversion. In the geometry dictionary for NUCLSQ as set up for this problem, the guanine base, anthramycin phenyl ring, and amide tail were constrained to be planar, and only bond lengths and angles were restrained. No torsion constraints or preferences for R vs S stereochemistry were used. Hence, the spontaneous inversion at C11 under refinement can be considered evidence for the correctness of that isomer.

The final R-factor at 2.3-Å resolution with 21 waters added is 21.2%. Both intensity data and final coordinates have been

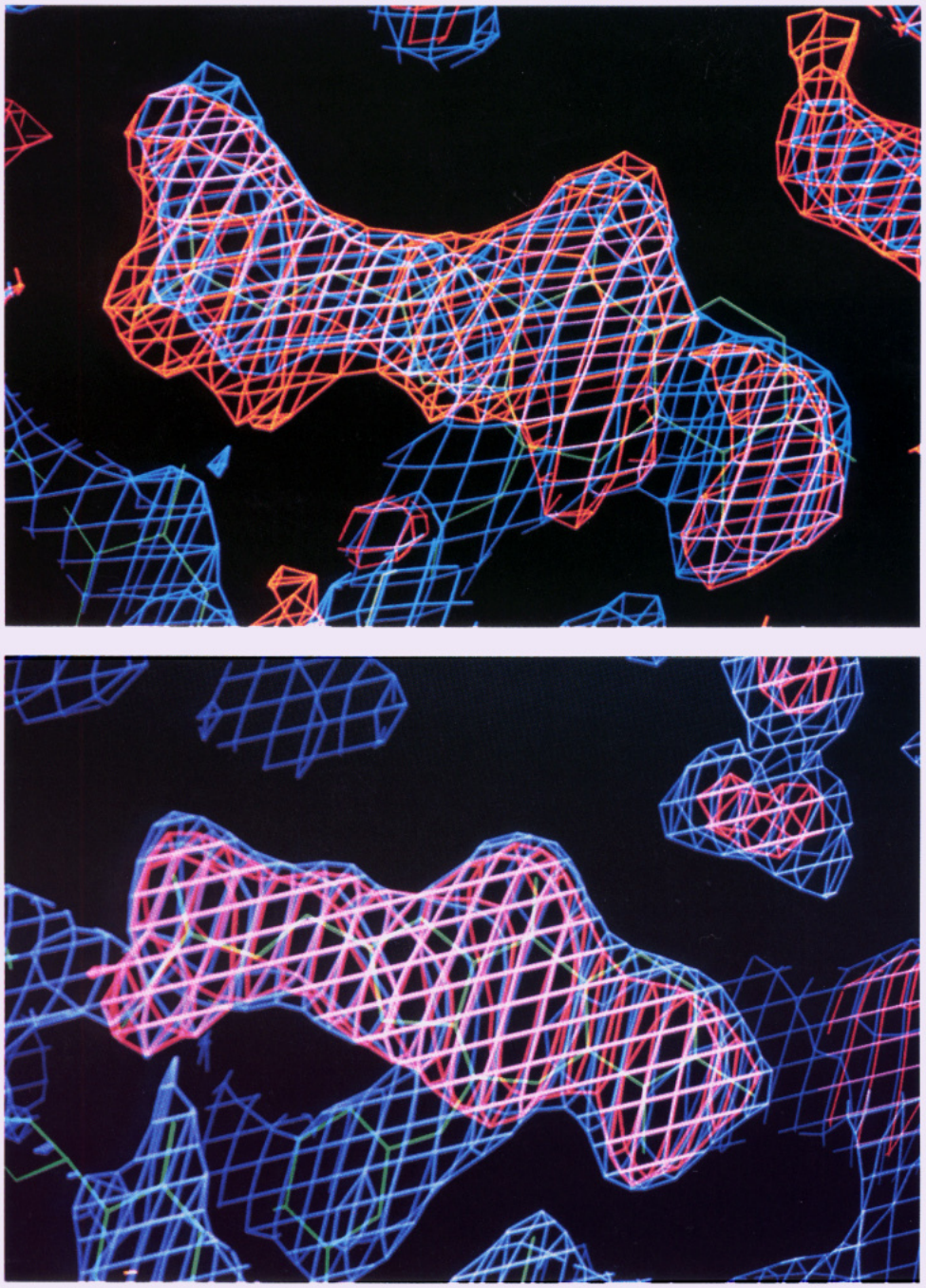

FIGURE 4: $2F_o - F_c$ density (blue) and $F_o - F_c$ difference density (orange or pink) maps for anthramycin positioned in the minor groove of C-C-A-A-C-G-T-T-G-G. (a, top) Initial drug image before anthramycin had been incorporated in the refinement. This was the density image used for initial positioning of the drug. (b, bottom) Final drug image at the conclusion of refinement. This map is calculated by deleting the refined drug molecule from the coordinate list for calculation of phases.

deposited with the Brookhaven Protein Data Bank and are available for immediate release.

RESULTS

Binding of Anthramycin to DNA. The structure of the Ant-2 complex with C-C-A-A-C-G-T-T-G-G is shown in Figure 5. One anthramycin molecule sits within the minor groove at each end of the double helix, with its six-membered ring directed toward the end of the helix and its acrylamide tail pointing backward toward the middle of the helix. The two drug molecules and two ends of the helix are related by a crystallographic symmetry axis; whatever is observed at base n on one strand is also seen at base $n \pm 10$ on the other strand.

The primary attachment of drug to DNA occurs by a covalent bond between the C11 atom of the anthramycin seven-membered ring and the N2 of guanine 9. The complex is stabilized further by hydrogen bonds. The O2 atom of cytosine 2 accepts a bond from both anthramycin atoms N10 (3.4 Å) and C9−OH (3.1 Å). Longer interactions extend from anthramycin N10−H to the N3 of guanine 20 (3.7 Å) and from anthramycin C9−OH to the N2 of guanine 20 (3.9 Å). At the other end of the drug molecule, the amide nitrogen of the acrylamide tail donates a H-bond to the O2 of thymine 18 (2.9 Å) and the O4′ of sugar 19 (3.3 Å).

Some of these interactions were predicted by model building (Hurley & Petrosek, 1979; Arora, 1981), molecular mechanics (Rao et al., 1986; Remers et al., 1986), and NMR

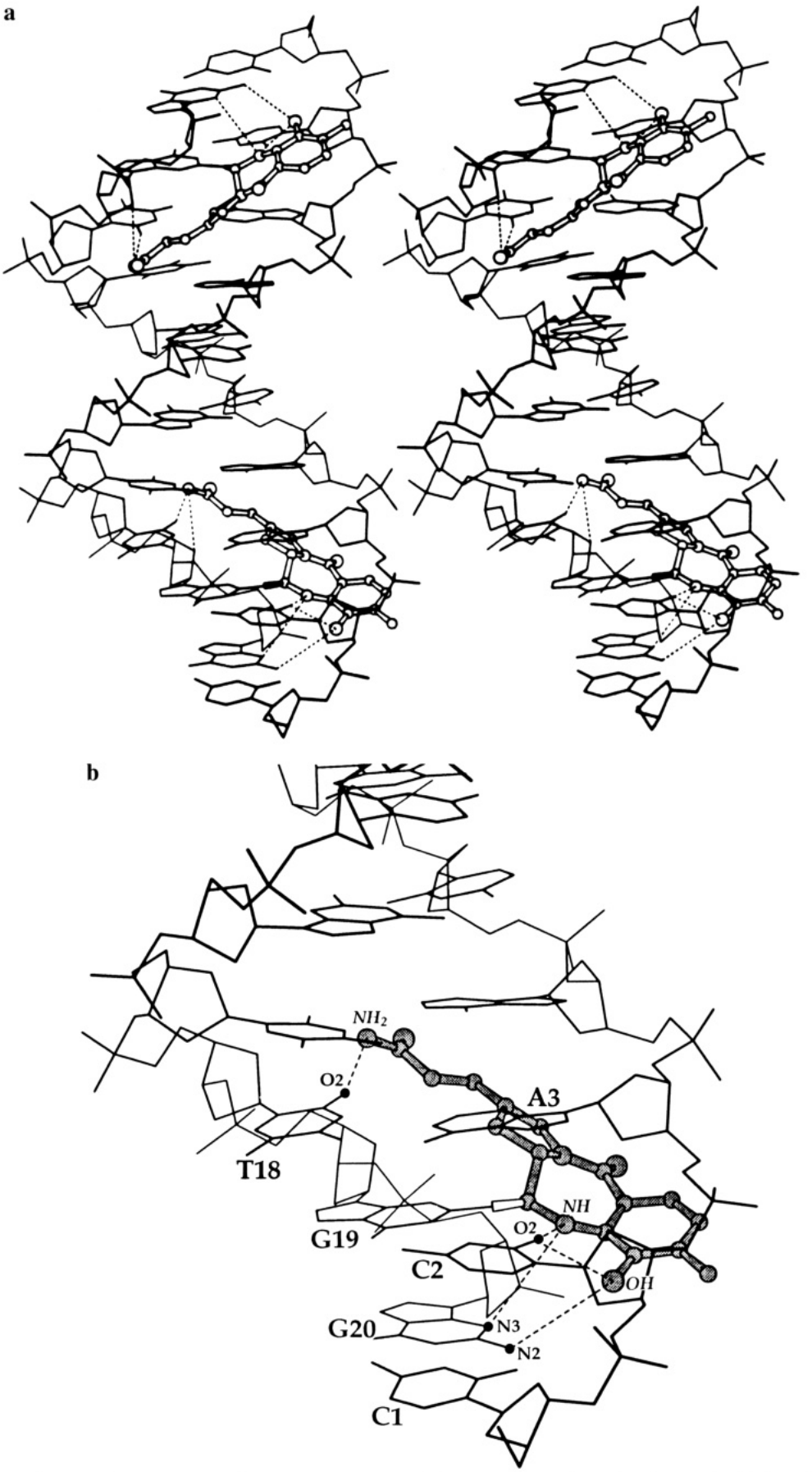

FIGURE 5: Structure of the B-DNA decamer C-C-A-A-C-G-T-T-G-G with two anthramycin molecules covalently bound within the minor groove to the N2 of guanine on the second base pair from each end of the helix. DNA to drug H-bonds are dotted. (a) View into the minor groove at top and major groove at bottom, displaying anthramycin binding. (b) Close-up of the bottom half of (a), showing details of interactions involving base pairs C1•G20, C2•G19, and G3•C18. Three of the six distances drawn are between 2.8 and 3.5 Å; all six are below 4.0 Å. The open bond linkage is the covalent bond to guanine 19.

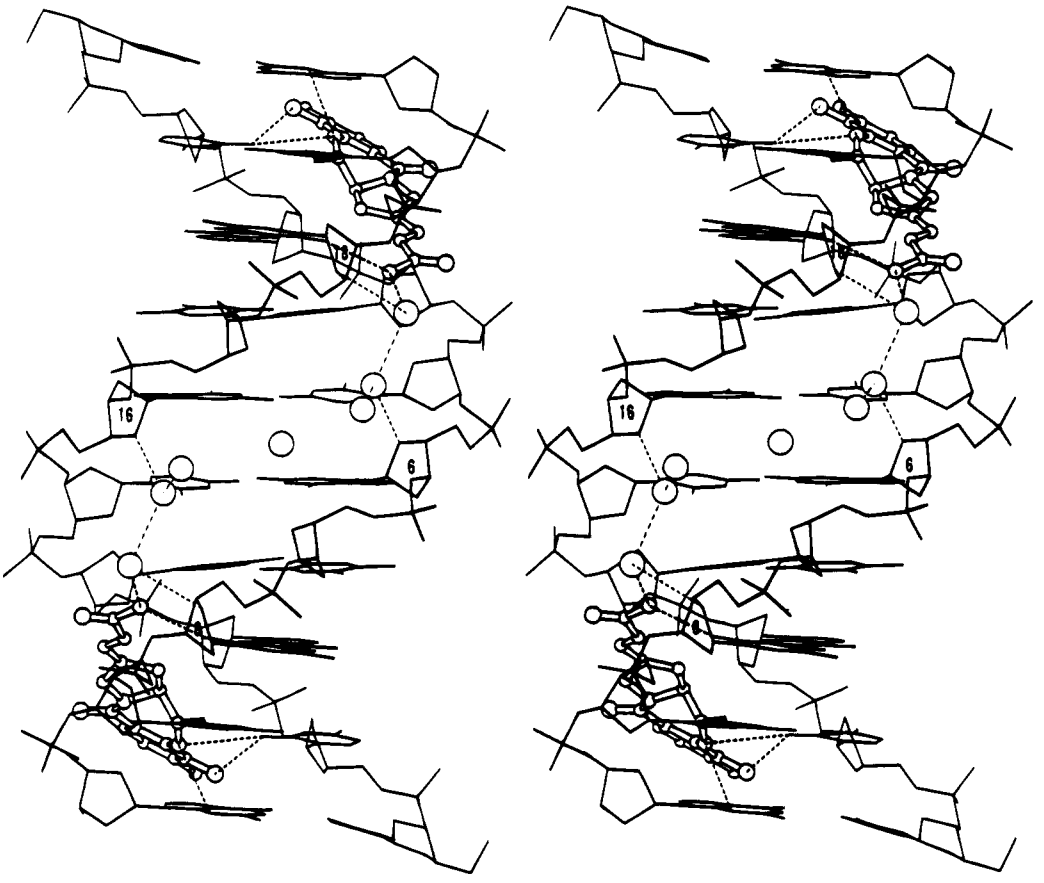

FIGURE 6: View into the minor groove at center, with localized water molecules shown as large spheres. These form a zigzag spine of hydration connecting the tails of the two anthramycin molecules. Sugar rings 6, 8, 16, and 18, which form water bridges via O4′ atoms, are numbered.

(Boyd et al., 1990b). In the model-building studies of Remers et al., it was proposed that the nitrogen on the amide tail could bond to thymine O2 in the sequence T-G-C, just as is observed in our T-G-G complex. NMR also revealed a close contact between the acrylamide amide nitrogen and thymine in the sequence T-G-C (Graves et al., 1985).

Hence, for anthramycin, both the N10 imide and the C9 hydroxyl contribute stabilizing hydrogen bonds to the DNA–drug complex. Tomaymycin can form the first of these bonds, but not the second, since its hydroxyl group is shifted on the six-membered ring from position 9 to position 8 (Figure 1). The bond from anthramycin C9–OH to the O2 of the cytosine paired with the covalently bonded guanine was predicted by three research groups using different techniques: Hurley and Petrosek (1979) from CPK model building, Arora (1981) from computer modeling, and Rao et al. (1986) and Remers et al. (1986) from molecular mechanics calculations.

The observed binding site for this C-C-A-A-C-G-T-T-G-G sequence is the three base pair T-G-G at each end of the double helix. Nuclear magnetic resonance and footprint protection experiments (Graves et al., 1985; Hertzberg et al., 1986) had indicated previously that the anthramycin binding site should be two or three base pairs, with a strong preference for R-G-R sites over all others (R = purine; Y = pyrimidine). The order of preference is R-G-R > Y-G-R = R-G-Y ≫ Y-G-Y. Hence, in our sequence, T-G-G should be preferred over the central C-G-T, as the crystallographic results confirm. Indeed, because of the internal crystallographic 2-fold symmetry of this space group, a DNA–drug complex with anthramycin bound at the central C-G-T could only adopt this crystal packing mode if helices were randomly disordered end for end. It is more probable that a helix with centrally bound drug would have adopted a different crystal packing mode entirely. This could be the reason for the different Ant-1 crystal form. But because the Ant-1 crystals only diffract to 3 Å, the drug molecule could not be located with certainty from difference density; recurring features down the minor groove could be ascribed alternatively to a drug molecule, a spine of hydration, or experimental error from the lower resolution.

Minor Groove Hydration. Figure 6 shows the hydration pattern in the minor groove of the anthramycin complex. As usual in narrow regions of the minor groove, the central region is occupied by a zigzag string of solvent peaks, in this case leading from one anthramycin tail to the other. On the 2-fold symmetry axis at the center of the helix sits a lone solvent peak, with somewhat too large distances to neighboring water molecules (4.0 Å) for good hydrogen bonding by a water molecule. This could merely reflect uncertainty in water definition at 2.3-Å resolution. It is reasonable to conclude that the minor groove between the two anthramycin tails is filled with a continuous spine of hydration. The remaining localized solvent peaks—waters and ions—surrounding the helix act as bridges from one helix to a neighbor. A network of bridging peaks interlaces between phosphates and base atoms on adjacent helices and helps maintain the integrity of the crystal.

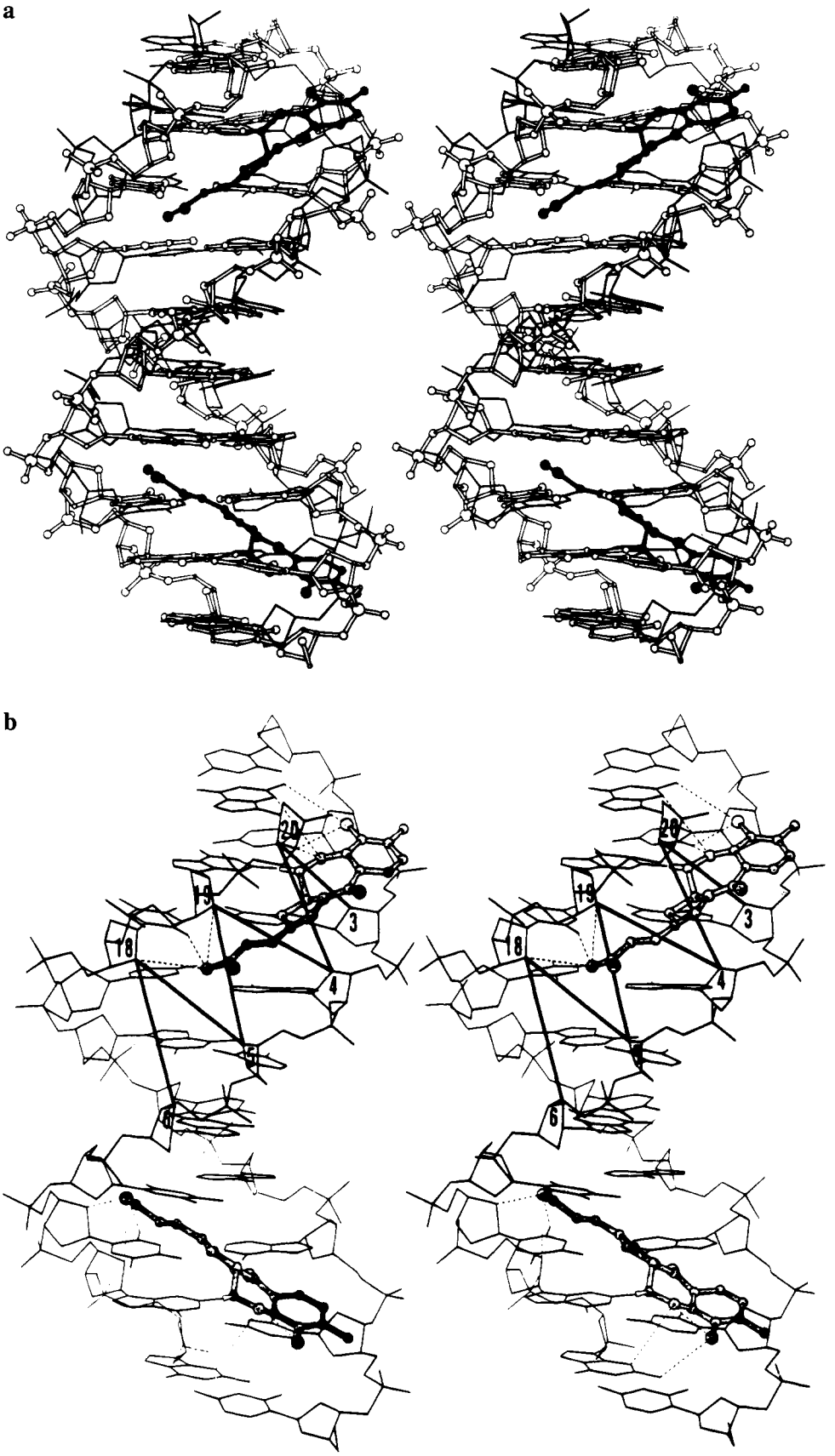

FIGURE 7: (a) Superposition of the structure of C-C-A-A-C-G-T-T-G-G alone (Privé et al., 1991) and its complex with anthramycin (this work), obtained by minimizing the root mean square difference in positions of the central four base pairs. DNA alone is shown in thin bonds and the DNA−drug complex in open bonds. The root mean square difference between the central four base pairs is 0.888 Å, while that between all atoms is 1.965 Å. (b) Same view of the DNA−drug complex, illustrating the O4′−O4′ vectors that are used to measure the width of the minor groove. This is a truer measure of groove width than P−P vectors, because O4′ atoms of sugars are nested deep in the groove near the drug, whereas phosphate groups sit high on the rim of the groove. These distances are plotted in the following figure.

Minor Groove Narrowing upon Drug Binding. As Figure 7a shows, the minor groove narrows considerably at the anthramycin binding sites, compared with the same sequence in the absence of drug. In the latter case (Privé et al., 1991) the minor groove is wide at either end and narrow in the middle, whereas the present anthramycin complex exhibits essentially constant width all down the helix.

Different measures of minor groove width have been tried in the past: originally the shortest P−P distances across the groove (Kopka et al., 1985a,c) and more recently a combination of these distances and the O4′−O4′ distances parallel to them (Privé et al., 1991). For a decamer, these P−P vectors are 20−5, 19−6, 18−7, etc. and for O4′−O4′, 20−4, 19−5, 18−6, etc. In principle, P−P distances less 5.8 Å for two phosphate group radii and O4′−O4′ distances less 2.8 Å for two oxygen radii should combine to yield a smooth curve that measures the width of the opening between walls of the groove. But in isolated cases this merging of reduced phosphorus and oxygen separations fails. The combined P−P and O−O groove width curve was smooth for monoclinic C-C-A-A-C-G-T-T-G-G (Privé et al., 1991), for orthorhombic C-G-A-T-C-G-A-T-C-G (Grzeskowiak et al., 1991), C-G-A-T-T-A-A-T-C-G (Quintana et al., 1992), and C-G-A-T-A-T-A-T-C-G (Yuan et al., 1992), and for both monoclinic and trigonal C-C-A-A-C-I-T-T-G-G (Lipanov et al., 1993). But the groove width plot for trigonal C-G-A-T-C-G-^{6Me}A-T-C-G (Bailkalov et al., 1993) exhibited saw-toothed alternations, with reduced P−P distances consistently 1.0−1.5 Å higher than adjacent and parallel reduced O4′−O4′ distances. The effect was as though the phosphate groups running along the crest of the two walls of the minor groove had been rotated outward, giving the crest of the groove an artificial widening.

This same effect is found in the present anthramycin complex: reduced P−P distances are ca. 1 Å wider than adjacent O−O distances. An explanation for this anomaly is suggested by Figure 7b. Phosphate groups sit high atop the two sides of the minor groove, like crenallations on a castle wall. Sugar O4′ atoms, by contrast, lie deep within the groove, in the region occupied by the drug molecule itself. O4′ atoms are in van der Waals contact with drug; phosphate groups are not. Individual phosphates could be rocked to one side of the minor groove wall or the other because of crystal packing or other factors, yielding apparent perturbations in the minor groove width plot that are structurally meaningless. Where drug packing is the issue, it seems most sensible to measure minor groove width via O4′ atoms.

Figure 7b shows that O4′ atoms occur in zigzag alternation across the minor groove, and the best vectors with which to measure groove width are distances between O4′ atoms of residues 3−20, 20−4, 4−19, 19−5, 5−18, etc. These distances are plotted in Figure 8 for the anthramycin−DNA complex (Ant) and for three DNA decamers without drug (CG, CICa, and CIMg). (See Figure 8 caption for sequences.) The crystal structures of C-C-A-A-C-G-T-T-G-G with and without anthramycin (Ant and CG) involve quite different local helix packing, so it is not immediately clear whether observed differences arise from drug binding or from crystal packing. Fortunately, C-C-A-A-C-I-T-T-G-G has been studied in both monoclinic and trigonal space groups (CICa and CIMg). Its structures provide a bridge between the other two and help unscramble the relative significance of drug binding vs crystal packing.

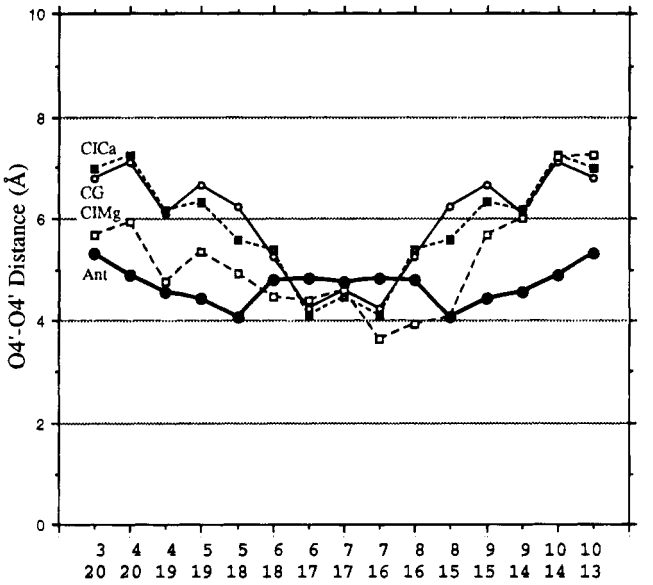

FIGURE 8: Minor groove width for four B-DNA decamer helices, measured between sugar O4′ atoms as shown in the previous figure. Actual distances between atom centers have been decreased by two oxygen radii of 1.4 Å each, to show the width of the gap between walls. O4′ atoms connected are identified along the horizontal axis. Symbols: black dots and thick lines = Ant = C-C-A-A-C-G-T-T-G-G with anthramycin (this work); open circles and thin lines = CG = C-C-A-A-C-G-T-T-G-G without drug (Privé et al., 1991); black squares and short dashes = CICa = C-C-A-A-C-I-T-T-G-G in same monoclinic space group as C-C-A-A-C-G-T-T-G-G without drug (Lipanov et al., 1993); open squares and long dashes = CIMg = C-C-A-A-C-I-T-T-G-G in same trigonal space group as anthramycin/DNA but with quite different crystal packing (Lipanov et al., 1993).

Figure 8 shows that both of the monoclinic structures (CG, CICa) have a wide minor groove at the ends and a narrow groove at the center. When the C-C-A-A-C-I-T-T-G-G sequence is induced to adopt a trigonal packing mode via a change of cation (CIMg), some of this groove widening disappears, but not all. The left end as plotted in Figure 8 still remains wider than the center, and the right end is scarcely affected at all. To this extent the wider minor groove at the ends is indeed a consequence of base sequence and not simply a crystal packing artifact. Now add anthramycin, and the minor groove regularizes: both ends become narrower and the center becomes slightly wider. The groove width now varies only ±0.5 Å from a mean of 4.5 Å, as measured on the O4′−O4′ diagonals, or ca. 95% of this, 4.3 Å, measured perpendicular to the groove. This narrowing can only be an effect of drug binding. But without the evidence from the two C-C-A-A-C-I-T-T-G-G structures, attribution of cause and effect would have been less clear.

Stereochemistry of the Covalent DNA−Drug Bond. The anthramycin molecule contains two chiral carbons in its seven-membered ring: C11 and C11a. Crystal structure analyses of the methyl ethers of both anthramycin and tomaymycin had demonstrated an *R* conformation at C11 and an *S* conformation at C11a, as drawn in Figure 1a (Mostad, 1978; Arora, 1979, 1981). The C11a conformation remains unchanged during DNA binding. But the stereochemistry at C11 is reversed upon binding to DNA; the DNA−drug complex has stereochemistry C11*S*, C11a*S*. Chirality at C11 would be lost in going through the imine reaction intermediate (Figure 1a, right). Hence, the observed

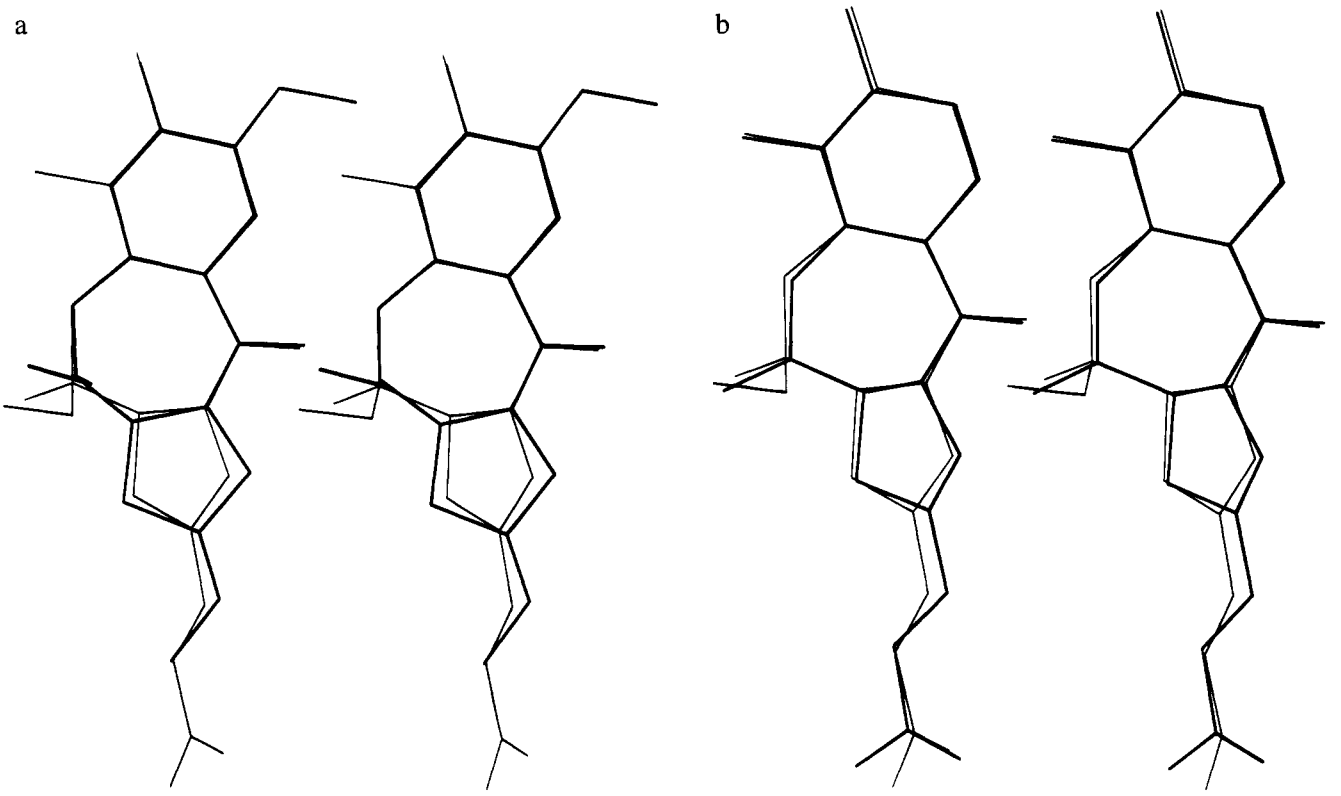

FIGURE 9: (a) Stereo superposition of crystal structures of anthramycin methyl ether (light bonds; Mostad et al., 1978) and tomaymycin methyl ether (dark bonds; Arora, 1981). The six-membered rings of the two molecules were used for superposition. Note that the five-membered ring of anthramycin is twisted out of plane more than with tomaymycin and that the attachments to C11 are shifted as a consequence. But for both drugs the C11−O bond rises axially from the plane of the molecule, whereas the C11−H bond (which will become the attachment to DNA) extends equatorially to the left. (Bond to hydrogen not drawn for tomaymycin.) The angle between normal vectors to the six- and five-membered rings is 35.4° for anthramycin and 9.1° for tomaymycin. (b) Superposition of the anthramycin molecule from its methyl ether crystal structure (light bonds; Mostad et al., 1978) and its DNA complex (dark bonds; this work). The entire molecule was used for superposition, with the exception of substituents on the C11 atom. The root mean square deviation of equivalent atoms is 0.26 Å. The angle between normal vectors to the six- and five-membered rings is 35.4° for free anthramycin and 40.9° for its DNA complex.

C11*S* conformation in the final DNA complex can only result from the fitting of the drug to the minor groove.

For reference, the crystal structures of isolated methyl esters of anthramycin and tomaymycin are compared in Figure 9a, and free anthramycin methyl ester is compared with its DNA complex in Figure 9b. The *S* conformation at C11a gives the three-ring drug molecule a right-handed twist, which makes it fit particularly well within the twisting course of the minor groove around the B-DNA helix. An anthramycin isomer with an *R* configuration at C11a would be twisted in the wrong sense, would clash with the walls of the groove of a right-handed helix, and would not bind well to B-DNA. If left-handed B-DNA existed, then anthramycin with the *R* configuration at C11a might conceivably be a natural substrate. But in a world with only right-handed B-DNA, it is not surprising that only the C11a*S* enantiomer possesses biological activity (Hurley et al., 1988).

The *R* conformation at C11, where the drug is bound to guanine, is equally unfavorable for DNA binding. Because of the *S* configuration at C11a, a bond in the *S* configuration at the neighboring C11 extends in an equatorial direction, only 1° off from the best mean plane through the six- and five-membered drug rings (Figure 9). The three-ring drug molecule can slip into a narrow groove and extend an *S* bond toward the bottom of the groove. In contrast, an *R* configuration at C11 would extend the bond in an axial direction almost perpendicular to the plane of the molecule, like the $-OCH_3$ in free anthramycin and tomaymycin methyl ethers. Either the drug would have to sit partially crossways in a widened groove or the seven-membered ring would have to be repuckered drastically to bring the *R* bond toward the equator, and a major shift of the receiving guanine would be called for. Indeed, in our attempts to fit the C11*R* conformation in place as a starting point for refinement, we found that guanine had to be pushed back as shown in Figure 10, to a position where the guanine N2 to cytosine O2 bond was scarcely credible. As discussed earlier under Materials and Methods, this initial C11*R* conformation refined spontaneously to the C11*S* conformation. Rao et al. (1986) deduced from AMBER molecular modeling calculations that the *R* configuration was unfavorable because if forced apart this same C-G base pair hydrogen bond.

The natural twist of the free anthramycin molecule predisposes it to fit into the groove of a right-handed helix, but this twist is accentuated somewhat when the drug binds to DNA. The six-membered ring of anthramycin is planar, and the five-membered ring is approximately so. Normal vectors to the two rings make an angle of 35.4° to one another in the crystal structure of free anthramycin (Mostad et al., 1978; Arora, 1979) and 40.9° in the DNA complex. If one defines a long axis as extending from atom C8 to atom C2 in Figure 1, then the nonplanarity of the drug molecule can be described in the language that one uses for an individual base pair of DNA. *Propeller* is the relative

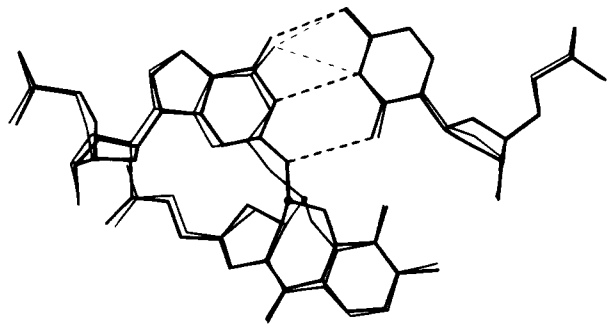

FIGURE 10: Adjustments required to attach *R* and *S* isomers of anthramycin to guanine after the body of the drug has been fitted into the density and difference density shown in Figure 4a. Thick line = fitting of the *S* isomer. Thin line = fitting of the *R* isomer. With the *R* isomer, in order to keep the three-ring drug molecule within the observed difference electron density, the guanine base must be pushed back and rotated until the N2 hydrogen bond with its paired cytosine is broken. Furthermore, the guanine O6 atom moves into a position intermediate between the cytosine N4 and N3 atoms. None of this deformation is required with the *S* isomer of anthramycin.

twist of the two rings about their long axis, and *buckle* describes the bending of the molecule about a short axis at right angles to it, roughly a line from atom C5 to the midpoint of the N10−C11 bond in Figure 1. In the crystal structure of free anthramycin the drug molecule has a propeller twist of 34.4° and a buckle of 8.2°. After it binds to DNA, the propeller increases slightly to 37.1°, and the buckle increases to 17.6°. As we shall see in the following section, this increased bending or buckling of the drug molecule probably arises mainly because of stacking contacts of anthramycin six-membered rings across a junction between helices.

Nuclear magnetic resonance studies have corroborated the 11a*S*, 11*S*-3′ conformation of anthramycin bound to DNA, with the six-membered ring pointing toward the 3′ end of the strand to which the drug is bound. Krugh et al. (1989) demonstrated this conformation with anthramycin bound to the B-helical hexamer A-T-G-C-A-T. Molecular mechanics calculations indicate that the 3′ orientation of the six-membered ring is favored over the 5′ orientation by somewhere between 5.6 kcal/mol (Rao et al., 1986; Boyd et al., 1990b) and 9.3 kcal/mol (Remers et al., 1986).

In contrast, NMR experiments with tomaymycin and this same A-T-G-C-A-T hexamer may illustrate a difference in behavior of the two drugs. The axial twist of the tomamycin molecule in the crystal structure is substantially smaller than that of anthramycin (Arora, 1981). The angle between normals to the six- and five-membered rings is only 9.1° instead of 35.4°, and this can be factored as before into a propeller of 3.3° and a buckle of 8.5°. The flatter tomamycin molecule may be considered as the basis for differential reactivity at C11, but Figure 9 demonstrates that, in the crystal structures at least, the C11−O bond direction in the *R* conformation is fully as axial for tomaymycin as for anthramycin.

Barkley et al. (1986) showed by fluorescence studies that two forms of tomaymcin exist in equilibrium in protic solvents and that these probably are the 11*R* and 11*S* isomers, equilibrating via the imine. Fluorescence spectra and proton NMR implied two ground-state forms of the tomamycin adduct with calf thymus DNA (Barkley et al., 1986) and with A-T-G-C-A-T (Cheatham et al., 1988). The authors proposed that these could be the 11*S*-3′ and 11*R*-5′ isomers, differing simultaneously by inversion of bond geometry at C11 and by rotation of the molecule so it points in the opposite direction along the minor groove. But this conclusion rests heavily on model-building studies; the authors acknowledge that "The alternative explanation would be that the heterogeneous emission arises from a single diastereomer bound in different environments on DNA" (Barkley et al., 1986).

The problem with molecular mechanics model building is that the believability of the result depends heavily on adherence to reasonable energy terms for allowed distortions. The DNA and drug geometries are reasonable in stereo Figure 6 of Barkley et al. (1986), but the drug conformations depicted there are 11*S*-3′ and 11*S*-5′, not 11*S*-3′ and 11*R*-5′ as claimed. In stereo Figure 9 of Cheatham et al. (1988) the two conformations are indeed 11*S*-3′ and 11*R*-5′, but the seven-membered drug ring is drastically repuckered and the DNA helix geometry is distorted beyond belief. The issue can only be settled by direct experimental measurements. Another NMR study with tomaymycin bound to C-I-C-G-A-A-T-T-C-I-C-G found only a single isomer, 11*S*-3′ (Boyd et al., 1990a). In summary, the flatter tomaymycin molecule may possibly adopt more than one conformation within the groove, but both X-ray diffraction and nuclear magnetic resonance indicate only the 11*S*-3′ conformation for anthramycin.

Tilted Stacking of Helices within the Crystal. The arrangement of helices within the trigonal crystals is unlike anything that has been encountered before in DNA or DNA−drug structures. Because 10 base pairs constitute one turn of helix, it is a common pattern for B-DNA decamer helices to stack atop one another in semicontinuous rods that closely simulate continuous helices. The interhelix junction from one decamer to the next is no different from internal helix steps even though connecting phosphates are absent, emphasizing the importance of base pair upon base pair interactions in determining helix structure. The common monoclinic, orthogonal, rhombohedral, and trigonal crystal forms differ chiefly in the manner in which these infinite helical rods pack together. In monoclinic and orthorhombic crystals the rods all are parallel. In all previously known trigonal forms, such as the C-C-A-A-C-I-T-T-G-G used earlier for comparisons, the rods lie at 120° intervals, parallel to the *ab* plane and perpendicular to the 3-fold screw axis around *c* (Bailakov et al., 1993; Lipanov et al., 1993; Goodsell et al., 1994).

Although the same $P3_221$ space group is used for the present anthramycin−DNA complex, the packing of helices is completely different. Individual helices do not lie in the *ab* plane; they are inclined 22° to the *c* axis and spiral up it in a 3-fold screw as shown in Figure 11. With this 22° inclination the helices cannot stack atop one another to simulate a continuous helix. Even without this inclination the 120° rotation of the 3-fold screw would be enough of itself to interrupt regular base pair stacking. The sum of the nine internal helix twist angles in this helix is 333°. With normal stacking of repeated helices, the interhelix step would have a twist of 27°, but the 3-fold axis changes this angle to 27° − 120° = −93°. The terminal base pairs on two successive helices are crossed at right angles, rather than helically stacked. Simple interhelix base pair stacking obviously is not the organizing principle of this crystal form.

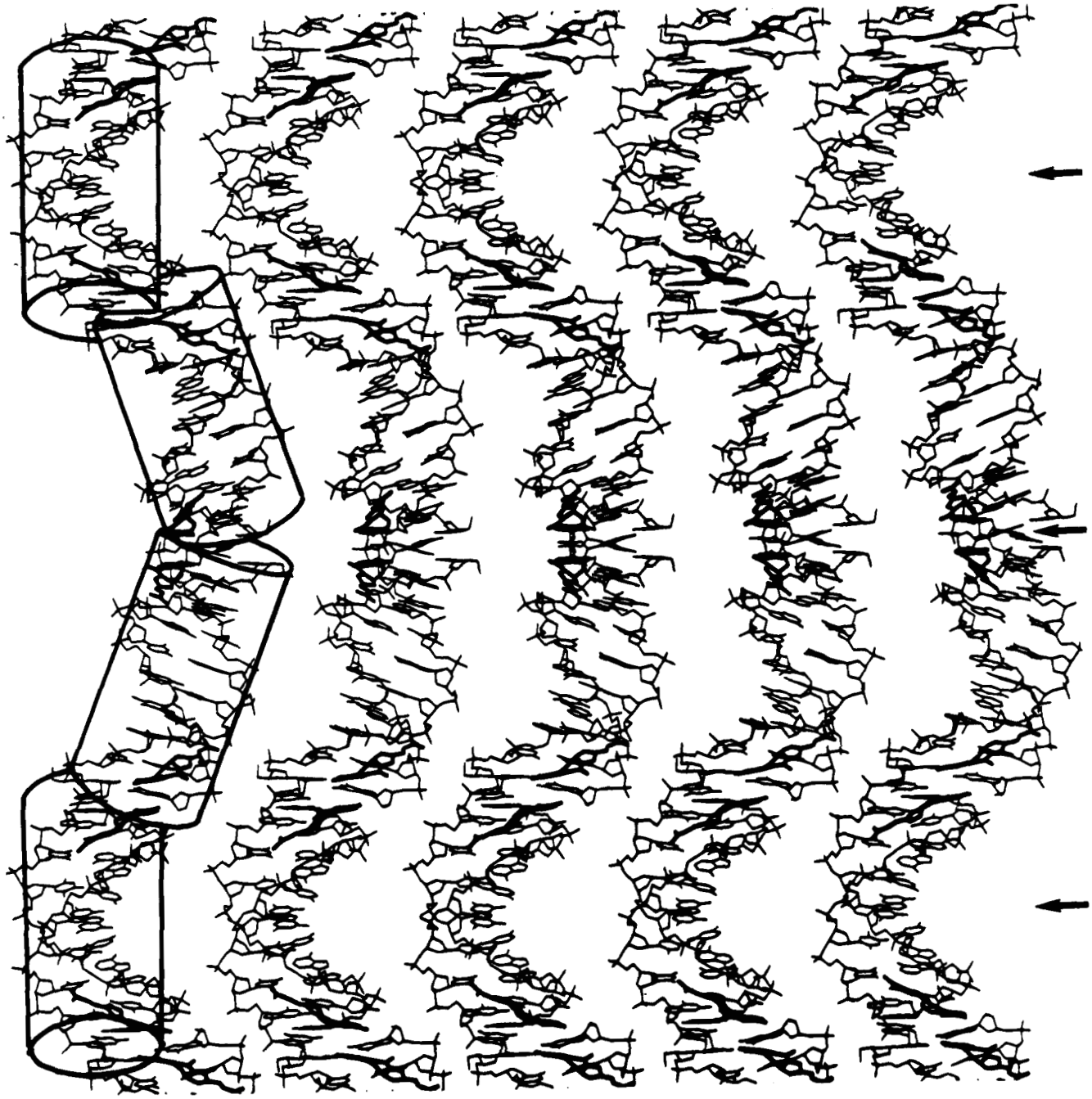

FIGURE 11: Adjacent columns of decamer helices winding around the 3-fold screw axis along *c* (vertical). The repeating unit along *c* is three successive decamers. The *a* axis is the horizontal distance from one column to the next. This figure is a stereogram, with adjacent columns rotated by 6° about the *c* axis. Relax your eyes as if you are looking beyond the page, until six columns are seen instead of five. The central four columns then will appear in three dimensions. Note that a stereoview reveals that this sheet of columns has pronounced horizontal ridges and troughs. The sheet of columns as drawn here are stacked atop one another with sufficient horizontal displacement that the columns of one layer nest against the intercolumn grooves of the next layer. A crystallographic 2-fold axis runs horizontally through the middle of this drawing, and another 2-fold axis runs through the center of a helix 15 base pairs above and below the middle, as marked by the three arrows at the right.

The junction between helices is shown in Figure 12a. The primary interactions between molecules are stacking of the six-membered rings of two opposed anthramycins and stacking of cytosines from the terminal base pairs of the two helices. A hydrogen bond extends from each guanine N2 amine to the cytosine O2 on the adjacent stacked helix. Figure 12b is a view down the *c* axis, showing only the two terminal base pairs at a helix junction and the two anthramycin molecules. Now the cytosine−cytosine and anthramycin ring stacking is evident, as is the stacking of each terminal guanine against the cytosine sugar ring of the opposed helix. The result is a hydrophobic pocket that shields the drug molecules on all but one edge.

Local Helix Parameters. Local helix parameters were calculated with the NEWHEL94 program, a copy of which can be obtained from the Brookhaven Protein Data Bank, or by E-mail from R. E. Dickerson at red@uclaue.mbi.ucla.edu. Figure 13 shows the behavior of three local helix parameters: propeller, helical twist, and slide. The first of these parameters refers to individual base pairs, and the last two refer to steps from one base pair to the next. As before, four decamer helices are compared: C-C-A-A-C-G-T-T-G-G

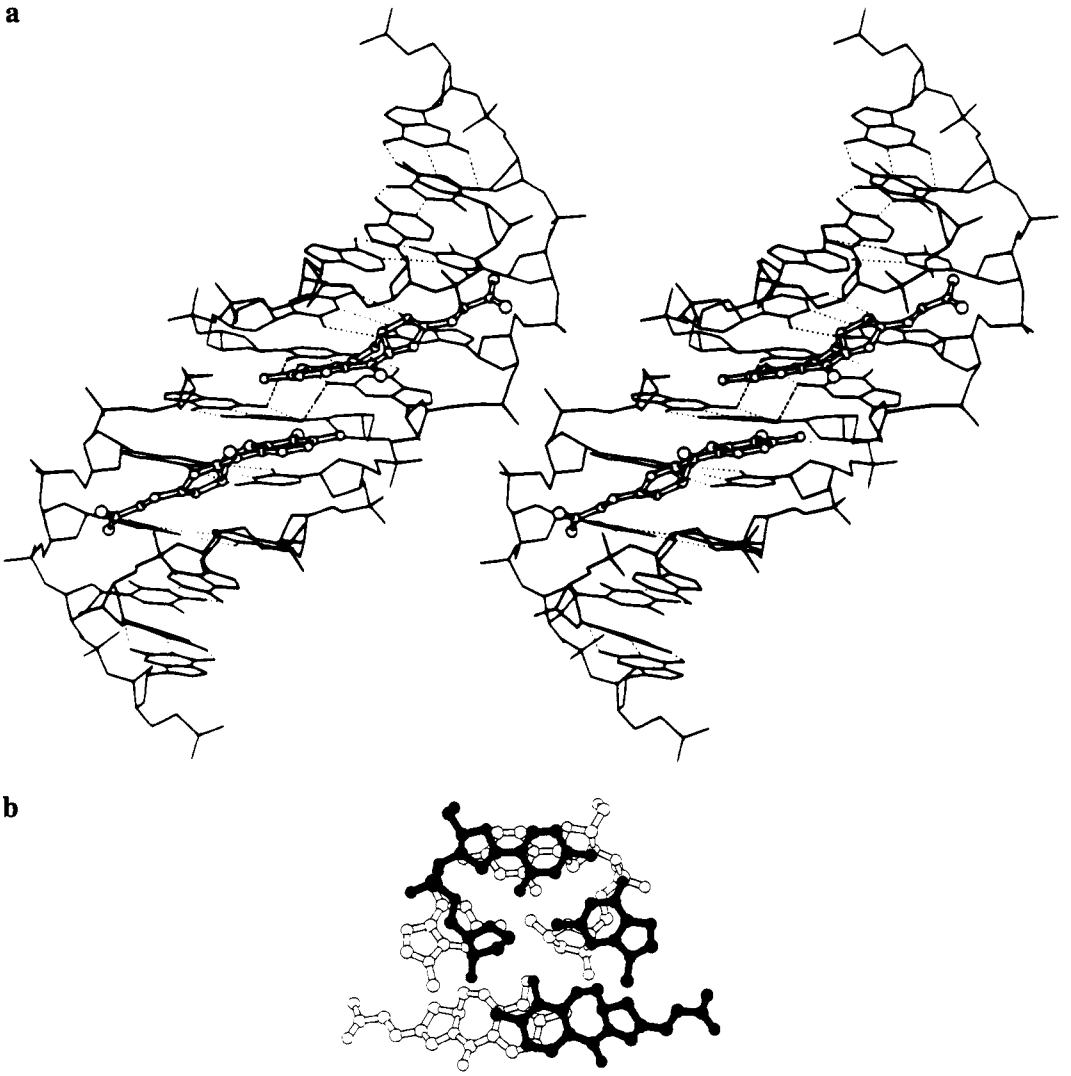

FIGURE 12: (a) Stereo close-up of the interface or junction between two stacked helices. The view is approximately that into Figure 14 at the level marked b. Note the close face-to-face contacts between anthramycin six-membered rings and the stacking of one cytosine over another in the background. Guanines are stacked approximately over opposing cytosine sugars. Hydrogen bonds from guanine N2 to cytosine O2 between helices are dotted. (b) Drawings showing the terminal base pair and anthramycin molecule for two decamers spiraling around the 3-fold screw axis, viewed directly down the *c* axis. Note the stacking of cytosine upon cytosine (top), guanine upon sugar (middle), and anthramycin six-membered rings over one another (bottom). A crystallographic 2-fold symmetry axis runs vertically in the plane of the page.

with and without anthramycin (Ant, CG) and C-C-A-A-C-I-T-T-G-G in two crystal forms to bridge the transition between monoclinic (CICa) and triclinic (CIMg) packing.

(A) Propeller Twist. Propeller in the anthramycin complex (Figure 13a) hovers around a typical −10° region, with the conspicuous exception that the outermost base pairs are forced by the peculiar helix−helix stacking seen in Figure 12a to exhibit a reversed (positive) propeller. The normal propeller sense in these outermost base pairs would break the cytosine O2 to guanine N2 hydrogen bonds that hold two helices together.

(B) Helical Twist. Helical twist changes upon binding of drug (Figure 13b). In the two drug-free monoclinic decamers (CG, CICa) the C-A step 2 exhibits an identically large twist of more than 50°, and the central C-G or C-I step 5 is almost as large. Base steps of the type Y-R (pyrimidine−purine) are inherently unstable because of poor overlap of base pairs and tend in some cases to exhibit bimodal twist behavior (Yanagi et al., 1991). Both steps 2 and 5 are much less twisted in the trigonal CIMg decamer and still less in the Ant complex. In contrast (and perhaps in compensation), the intervening steps 3 and 4 show the opposite behavior, increasing in the order CICa, CG, CIMg, and Ant. It seems probable that the bound drug constrains the first three base pairs into a compact unit with a modest 35° twist at steps 1 and 2, encouraging the helix to "fracture" and exhibit a higher compensatory twist at step 3 just outside the drug-binding zone. Indeed, as elaborated later in the Discussion, this twist behavior at a bimodal T-G (=C-A) step may lie at the heart of the observed base specificity of anthramycin.

(C) Slide. This quantity (Figure 13c) measures the relative motion of two adjacent base pairs along their long axes. In the drug-free monoclinic helices (CG, CICa), the large twist angle at C-A steps is combined with a very large sliding of the second base pair over the third. This behavior is damped down in CIMg and is completely absent in Ant. Figure 7a illustrates an interesting correlation between slide and minor groove width that was unnoticed in the original report of the drug-free CG decamer structure. A large lengthwise slide between the second and third base pairs facilitates stacking of helices in the monoclinic space group adopted by this structure, as shown by Figure 14 of Privé et al. (1991). But

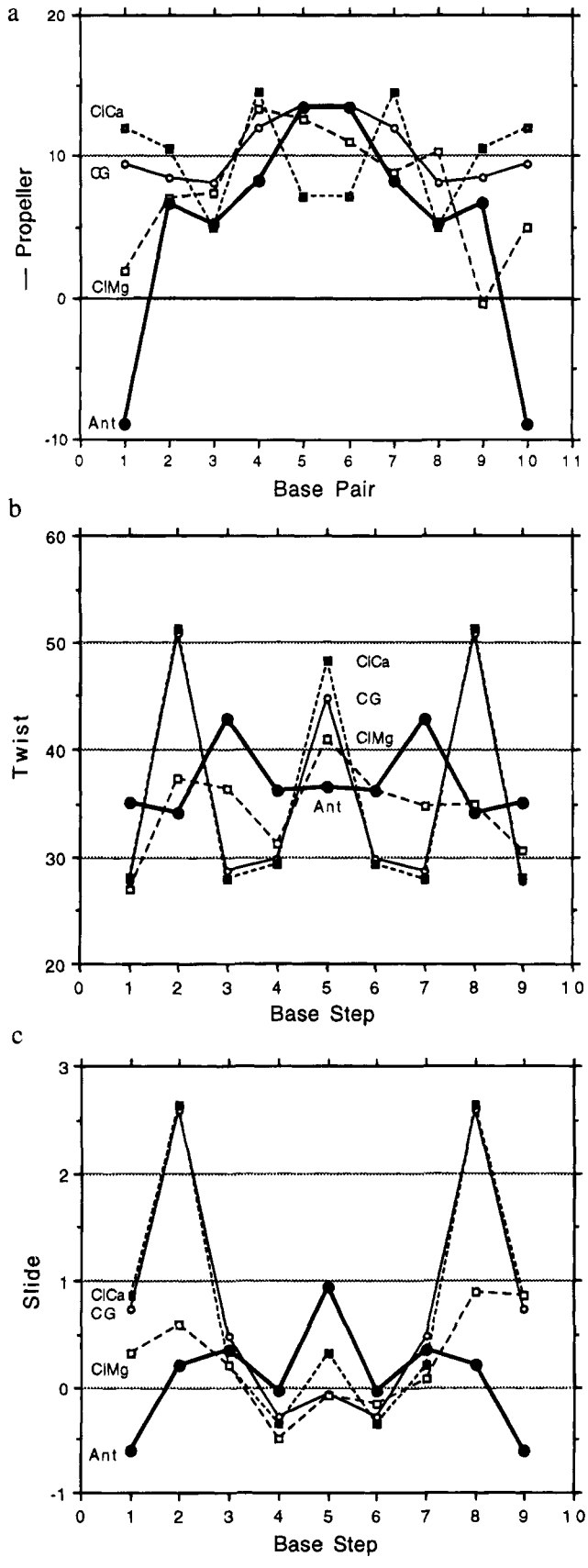

FIGURE 13: Plots of local helix parameters in four decamer structures, identified as in the Figure 8 caption: (a) propeller twist of base pair; (b) helical twist from one base pair to the next; (c) slide along the long axis from one base pair to the next. See text for further definitions and discussion.

as base step T8•G9 of Figure 8 of that paper illustrates particularly well, such a slide pushes phosphates farther apart

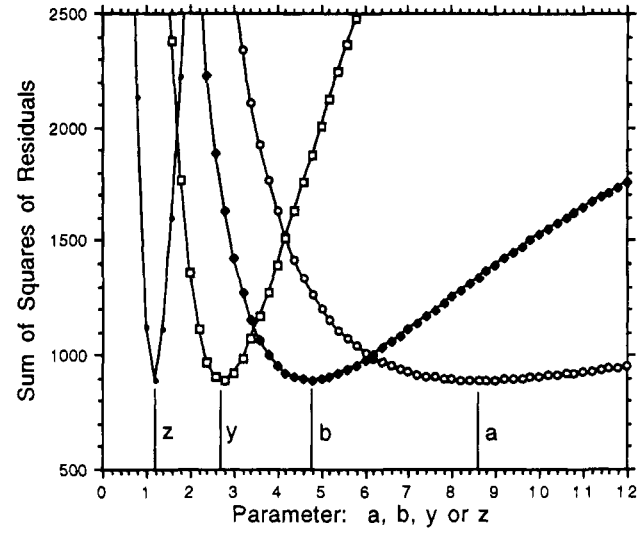

FIGURE 14: Profiles of least squares error sums for the fitting of parameters *a*, *b*, *y*, and *z* to the binding probability data of Table 3a. The global minimum is at *a*, *b*, *y*, *z* = 8.6, 4.8, 2.7, 1.2. Each of the four curves is a plot of error sums across the full range of values for that parameter, holding the other three parameters at their minimum values. The minimum for each curve (at error sum = 892) is marked by a short vertical line at the bottom. See text for further discussion.

at the two ends of the base step and expands the minor groove near each end of the helix. In the present complex of anthramycin with this same sequence, the crystal packing is different, a large slide is not present, and the minor groove has more nearly its normal width (Figure 7a).

DISCUSSION

Because anthramycin can be cardiotoxic, much of the work on structure–activity relationships (Horowitz et al., 1971; Hurley & Thurston, 1984; Hurley et al., 1988; Jones et al., 1990) has concentrated on varying the substituents on the five-, six-, and seven-membered rings of the 1,4-benzodiazepines, hoping that by changing these substituents a better drug could be found. Indeed, some answers have come from such work. But a fundamental structure–activity question remains unanswered: What is the molecular basis for sequence specificity in the benzodiazepines?

Solution-Based Model for Sequence Specificity. DNA footprinting studies with methidiumpropyl-EDTA and an exonuclease III stop assay have demonstrated that both anthramycin and tomaymycin bind best to regions of three stacked purines (R-G-R), with one flanking pyridine (R-G-Y or Y-G-R) being a second choice, and two flanking pyrimidines (Y-G-Y), a poor third (Hertzberg et al., 1986; Hurley et al., 1988). Footprinting of a 117-bp fragment (Hurley et al., 1988) quantitated this preference: R-G-R sites of all types were observed to be occupied by bound anthramycin 85% of the time, 22% of all R-G-Y sites were bound by drug, 18% of all Y-G-R, and only 9% of Y-G-Y sites.

Pierce et al. (1993) quantitated this binding site preference even further, from a combination of exonuclease III digestion, λ exonuclease digestion, and incision analysis with combined UvrA, UvrB, and UvrC enzymes. Using three DNA oligomers totaling 614 base pairs, they undertook an exhaustive analysis of the relative reactivities of all 16 base triplets having a central guanine. Their most sensitive analytical tool was cleavage by combined UvrABC enzymes, which

Table 3: Relative Preferences of Anthramycin for $^{5'}$X-G-X$^{3'}$ Binding Sites: Comparison of Observed and Calculated Values[a]

(a) Observed

first base	third base: A	G	T	C
A	122 (40) [5]	112 (30) [6]	20 (6) [5]	30 (15) [8]
G	58 (20) [5]	26 (11) [4]	13 (8) [4]	1 (1) [12]
T	25 (10) [6]	0 [3]	7 (5) [3]	11 (5) [8]
C	2 (1) [8]	1 (0.5) [10]	0 [4]	3 (3) [11]

(b) Calculated: $a = 8.6$, $b = 4.8$, $y = 2.7$, $z = 1.2$, $K = 0.95$

first base	third base: A	G	T	C
A	*abyz* 127	*aby* 106	*ayz* 26	*ay* 22
G	*abz* 47	*ab* 39	*az* 10	*a* 8
T	*byz* 15	*by* 12	*yz* 3	*y* 3
C	*bz* 6	*b* 5	*z* 1	1 1

[a] Observed data are taken from Table 2 of Pierce et al. (1993). Primary numbers are the relative intensities of UvrABC incision bands from three DNA oligomers of 247, 238, and 129 base pairs. Values in parentheses are standard deviations. Numbers in square brackets are the number of examples of each sequence in the oligomers analyzed. For explanation of part b, see text.

are observed to make dual incisions 6−8 base pairs upstream and 4 base pairs downstream from an anthramycin or tomaymycin binding site. The relative intensities of gel bands, interpreted as the relative probabilities of drug binding to base triplets, are presented in Table 3a. These numbers can be interpreted as relative anthramycin binding probabilities. They indicate a clear ranking of binding preferences among R-G-R triplets as follows: A-G-A > A-G-G > G-G-A > G-G-G. The second best Y-G-R and R-G-Y also favor A over G as the purine.

A simple computational model can be constructed to account for these experimental values and to examine the relative significance of two factors that will be important in anthramycin specificity: purine−purine base steps and A·T base pairs. For each X_1-G-X_3 triplet, construct a four-factor product as follows:

The first factor is a if X_1-G is a purine−purine step and 1 otherwise.

The second factor is b if G-X_3 is a purine−purine step and 1 otherwise.

The third factor is y if X_1 is an A·T base pair and 1 otherwise.

The fourth factor is z if X_3 is an A·T base pair and 1 otherwise.

For certain aspects of recognition as discussed below, the existence of an A·T base pair at a given position is significant, but its end-for-end orientation, A·T vs T·A, is not. The quantities a, b, y, z are to be determined by least squares fitting against the observed relative binding probabilities in Table 3a. The entry for triplet A-G-A is *abyz*, since both steps are purine−purine and both outer base pairs are A·T, and this corresponds to a relative binding probability of 122. The entry for triplet G-G-T is *az*, since only the first step is purine−purine and only the last base is A·T, and this is to be compared with binding probability 13. Other entries are given in the matrix of Table 3b.

The four parameters, a, b, y, z, and an overall scale factor, K, can be fitted to the drug binding frequency data of Table 3a by minimizing the error sum or the sum of squares of residuals over all 16 entries in the matrix. That is:

$$\text{error sum} = (abyz - 122)^2 + (aby - 112)^2 + (ayz - 20)^2 + (ay - 30)^2 + (abz - 58)^2 \ldots$$

The error sum has a single global minimum at $a = 8.6$, $b = 4.8$, $y = 2.7$, and $z = 1.2$, yielding the calculated binding probabilities given in Table 3b. Figure 14 shows four profiles through the full five-dimensional error sum vs parameter plot. For each profile, the error sum is calculated over the entire range of one parameter, fixing the other three parameters at their least squares error value.

Parameters a and b are larger than y and z at the error sum minimum, indicating that purine−purine steps *per se* are more important than are A·T base pairs on the outer positions. Parameter y is larger than x, indicating that A·T base pairs are more conducive to binding at the first base pair rather than the third. Profiles through y and z reveal relatively sharp minima, making the just-enunciated conclusion about A·T base pairs meaningful. Parameter a is nearly twice as large as parameter b, but Figure 14 shows that the minimum is quite shallow for both parameters, a in particular. Indeed, if we try deleting the T-G-G term from the error sum entirely, on the grounds that the zero for that entry in Table 3a must be in error in view of the binding of drug to T-G-G in the present X-ray crystal structure, then curve a of Figure 14 needs to tilt only very slightly to shift its minimum from 8.6 to 5.7, all other parameters remaining within 0.1 of their former values. Hence, one cannot use the data in Table 3a to decide with assurance between the relative importance of purine−purine steps at the 5′ vs 3′ ends of the recognition triplet.

It would not do to carry this analysis too far, especially in view of the large standard deviations of the binding probabilities (parentheses in Table 3a) and the frequently small numbers of examples in the experimental analysis (square brackets in Table 3a). But one can propose four rules or generalizations:

(1) Purine−purine steps favor anthramycin binding and are more important than are A·T base pairs.

(2) Purine−purine steps are of roughly comparable importance at the 5′ and 3′ ends of an X-G-X triplet, within limits of error of the binding data.

(3) A·T base pairs are preferred at both the first and last positions of an X-G-X triplet.

(4) The preference for A·T base pairs is stronger at the first position than the third.

Hence A-G-A is much more favorable for anthramycin binding than T-G-T (122 vs 7) by rule 1, A-G-T and T-G-A are roughly equivalent by rule 2, A-G-A is favored over G-G-A or A-G-G by rule 3, and A-G-G is preferred over G-G-A by rule 4. Having two purine−purine steps in G-G-G is roughly as favorable as having only one purine−purine step but two A·T base pairs in A-G-T and T-G-A. To what extent can these empirical observations be explained by the crystal structure?

Structural Basis for Specificity. (A) Specificity from Hydrogen Bonding. The oldest model for recognition of DNA sequence by a drug or protein involves specific hydrogen bonds between the recognition molecule and O, N, or $-NH_2$ groups along the floor of the major or minor groove (Seeman et al., 1976). In the minor groove a G·C base pair has two hydrogen bond acceptors, N and O, flanking a central donor, $-NH_2$. In contrast, an A·T pair

has the two N and O acceptors but empty space where the donor would otherwise be. Hydrogen-bonding groups probing the minor groove can easily distinguish G·C from A·T base pairs, but they cannot detect a base pair reversal, C·G for G·C or T·A for A·T, because the receptor atoms occupy virtually equivalent positions in space after reversal (Kopka et al., 1985a–c). Hence hydrogen bonding can be at the root of a preference for A over G, but it cannot be invoked to distinguish base pair reversals or to explain a preference for A or G over T or C. *Hydrogen bonding per se cannot be responsible for the preference of anthramycin for a purine–G–purine binding site.* The origin of this specificity must be sought elsewhere.

(B) A·T vs G·C Base Pairs. Minor groove binding drugs such as netropsin and distamycin demand a binding site consisting of several successive A·T base pairs. The length of the binding site depends on the length of the drug molecule itself. Although some differences have been reported in binding of such drugs to homopolymer $A_n \cdot T_n$ sites vs sites of mixed A·T base pair orientation, this is a secondary effect. The primary factor defining a suitable minor groove binding site is the presence of A·T base pairs in either orientation, rather than G·C base pairs. The reason for this sequence specificity was shown by the X-ray crystal structure analysis of the netropsin–DNA complex (Kopka et al., 1995a–c): The absence of an N2 amine on adenine makes the minor groove significantly deeper and better able to accommodate the drug molecule, whereas guanine N2 amine groups push the drug away from the floor of the groove. A generally higher propeller twist in A·T base pairs may also contribute by making the minor groove narrower. Hence, one must be aware of the possibility that the binding of any drug molecule within the minor groove could be facilitated by regions of A·T base pairs, and this is why that factor was included in the analysis of Table 3. But this line of argument once again fails to address the central issue of anthramycin's preference for two successive purine–purine steps.

(C) Twist Angles at Purine–Purine Steps. The most likely explanation of anthramycin's preference for purine–G–purine is suggested by single-crystal X-ray structure analyses of oligonucleotides containing purine–purine steps (Yanagi et al., 1991; Grzeskowiak et al., 1991; Quintana et al., 1992; Yuan et al., 1992; Leonard & Hunter, 1993). In general, A-A, A-G, and G-G steps have smaller than average twist angles, 24°–37°, whereas G-A steps are somewhat larger, 31°–42°. When two purines follow one another in sequence, their double rings stack particularly well, and this stacking is favored by lowering the twist angle between successive base pairs. Even in those examples with a larger twist, the two base pairs are observed to pivot over their purines, so the rings remain stacked even though the pyrimidines at the other end are completely destacked. [See Figure 8 of Privé et al. (1991), Figure 12 of Grzeskowiak et al. (1991), or Figure 10 of Quintana et al. (1992).] The stability of purine–purine ring stacking favors a strategy that enhances ring overlap: either this type of pivoting over the purine rings or a smaller than normal helix twist angle.

Results of the X-ray Crystal Structure Analysis. (A) Purine–Purine Specificity. The primary specificity of anthramycin for R-G-R binding sites probably arises because a low twist angle at R-G and G-R steps creates a good fit with the drug molecule. Both steps within the T-G-G binding site of the present anthramycin–DNA structure have a low molecule twist, 34° and 35°, whereas the same sequence without drug has a very high 51° twist at the T-G step:

	T — T	T — G	G — G
DNA-drug	43°	34°	35°
DNA alone	29°	51°	28°

Single-crystal structure analyses of DNA oligomers have shown consistently that the T-G (=C-A) step is inherently variable, with a twist angle that depends upon local environment (Quintana et al., 1992). In the present case, the change in local environment is the presence or absence of anthramycin bound to the DNA helix. In the drug-free DNA structure the flanking steps, T-T and G-G, are low twist in order to compensate for the central 51° T-G step (Figure 13b). When this large step is brought down to 34° in the DNA–drug complex, the compensating high twist is transferred to the neighboring T-T step.

T-G-G is not the optimum sequence for an anthramycin site, but the binding of drug apparently induces it to behave like an R-G-R sequence, in the sense of exhibiting low twist at both steps. Hence, the local DNA conformation after drug binding is an example of *induced fit.* In the present crystal structure analysis, binding of anthramycin causes the T-G-G sequence to behave like A-G-G and adopt a low twist angle at both steps. Anthramycin recognizes the low twist angles associated with three successive purines or at least a sequence such as T-G-G that can be induced to behave like a triple purine.

(B) A·T Specificity. The preference for A·T base pairs at one or both ends of the X-G-X recognition triplet recalls the binding specificity of netropsin mentioned earlier. Figure 5 shows that, in the 11*S*-3′ conformation of the crystal structure, the 5′ position of the binding site is occupied by the acrylamide tail of the drug, which extends along the floor of the groove like a "pseudo-netropsin". An A·T pair is preferred over G·C at that position for the same reason as with netropsin: absence of the guanine N2 amine makes the minor groove deeper and provides more room for the drug chain. It is tempting to think that the secondary preference for adenine at the 3′ base pair could imply a second, 11*S*-5′, binding mode, with the drug molecule inverted by 180° about the drug–guanine bond so the acrylamide tail nests against the 3′ base pair. This inverted orientation is disfavored in the present crystal structure because the third base is G·C rather than A·T. But with a sequence such as A-G-A, a 11*S*-5′ conformation might be nearly as acceptable as 11*S*-3′. Table 3 suggests that the 11*S*-3′ A-G-G conformation is more favored than the 11*S*-5′ G-G-A (112 ± 30 vs 58 ± 20). The sequence G-G-G is disfavored because it makes no accommodation for the acrylamide tail in either direction. The sequence A-G-A, which as Pierce et al. (1993) and Kizu et al. (1993) have noted is most favored of all, might be so because it allows drug binding in either orientation. In summary, the preference for purines flanking the covalently bonded guanine arises from base stacking and the inherently small helix twist at purine–purine steps. The preference for A·T base pairs at the ends of the binding site arises from a netropsin-like fitting of the acrylamide tail into the minor groove.

(C) Anthramycin vs Tomaymycin. Although hydrogen bonds cannot account for the triple purine preference of anthramycin, they can give anthramycin a binding advantage over tomaymycin. Figure 5 shows that the anthramycin–DNA complex is stabilized by hydrogen bonds to DNA from the C9–OH and the $-NH_2$ on the long acrylamide tail. In tomaymycin (Figure 1b) the –OH group is moved to the C8 position where it cannot interact with DNA in the manner observed for anthramycin, and the hydrogen-bonding ability of the tail is eliminated. Tomaymycin also is presumably hampered in fitting into the minor groove of right-handed B-DNA because its internal molecular twist between six- and five-membered rings is only 9° instead of the 34° of anthramycin. Hence, one would expect anthramycin to bind more strongly than tomaymycin, and it does.

Drug Design and Modeling. This anthramycin structure provides a look at a minor groove binding drug exhibiting GC base specificity. From the beginning of our structure work on AT-specific groove binders such as netropsin, Hoechst 33258, and DAPI, a primary goal has been to learn how to modulate the specificity selectively to read GC-containing sequences. To this end, modification of the pyrrole in netropsin and distamycin to produce a GC-reading element produced "lexitropsins" that turned out to be more GC-permissive than GC-requiring. [For a review of lexitropsins, see Kopka and Larsen (1992).] Connecting a true GC-reading element with an AT-reading drug molecule by a covalent "tether" or linker so each could bind to its adjacent preferred region is a different strategy that has occurred to several groups. Gurski and co-workers (Khorlin et al., 1980; Leinsoo et al., 1989) have synthesized and examined bis-linked analogues of the netropsin/distamycin family, as have Lown and co-workers (Lown et al., 1989; Beerman et al., 1991; Wang & Lown, 1992) and Schultz and Dervan (Schultz & Dervan, 1983; Dervan, 1986). Two anthramycin molecules also have been cross-linked and tested for drug binding (Farmer et al., 1991; Bose et al., 1992; Wang et al., 1992). But in these cases the design of the linker or tether has been empirical. As Walker et al. (1994) illustrate, it now is possible, having the crystal structures of anthramycin and netropsin, to design linkers of optimal size and geometry to ensure that in the tethered-drug complex with DNA, each half of the bis complex is in optimal orientation to the DNA to which it binds. This should lead to stronger binding and more useful DNA-reading drug analogues.

ACKNOWLEDGMENT

We thank Andrei Lipanov for help with X-PLOR refinement. Special thanks to Dr. S. N. Rao for his model coordinates of anthramycin bound to G•C-rich DNA.

REFERENCES

Arima, K., Kohsaka, M., Tamura, G., Imanaka, H., & Sakai, H. (1972) *J. Antibiot. 25*, 437–444.

Arora, S. K. (1979) *Acta Crystallogr. B35*, 2945–2948.

Arora, S. K. (1981) *J. Antibiot. 34*, 462–464.

Baikalov, I., Grzeskowiak, K., Yanagi, K., Quintana, J., & Dickerson, R. E. (1993) *J. Mol. Biol. 231*, 768–784.

Barkley, M. D., Cheatham, S., Thurston, D. W., & Hurley, L. H. (1986) *Biochemistry 25*, 3021–3031.

Beerman, T. A., Woynarowski, J. M., Sigmund, R. D., Gawron, L. S., Rao, K. E., & Lown, J. W. (1991) *Biochim. Biophys. Acta 1090*, 52–60.

Bose, D. S., Thompson, A. S., Ching, J., Hartley, J. A., Berardini, M. D., Jenkins, T. C., Neidle, S., Hurley, L. H., & Thurston, D. E. (1992) *J. Am. Chem. Soc. 114*, 4939–4941.

Boyd, F. L., Stewart, D., Remers, W. A., Barkley, M. D., & Hurley, L. H. (1990a) *Biochemistry 29*, 2387–2403.

Boyd, F. L., Cheatham, S. F., Remers, W., Hill, G. C., & Hurley, L. H. (1990b) *J. Am. Chem. Soc. 112*, 3279–3289.

Brown, D. G., Sanderson, M. R., Skelly, J. V., Jenkins, T. C., Brown, T., Garman, E., Stuart, D. I., & Neidle, S. (1990) *J. Mol. Biol. 211*, 189–210.

Brunger, A. T., Kuriyan, J., & Karplus, M. (1987) *Science 235*, 458–460.

Chandrasekaran, R., & Arnott, S. (1989) in *Landolt-Bornstein, New Series, Group VII* (Saenger, W., Ed.) Vol. 1b, pp 130–170, Springer-Verlag, Berlin.

Cheatham, S., Kook, A., Hurley, L. H., Barkley, M. D., & Remers, W. (1988) *J. Med. Chem. 31*, 583–590.

Coll, M., Frederick, C. A., Wang, A. H.-J., & Rich, A. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8385–8389.

Dervan, P. B. (1986) *Science 232*, 464–471.

Edwards, K. J., Jenkins, T. C., & Neidle, S. (1992) *Biochemistry 31*, 7104–7109.

Farmer, J. D. J., Gustafson, G. R., Conti, A., Zimmt, M. B., & Suggs, J. W. (1991) *Nucleic Acids Res. 19*, 899–903.

Frederick, C. A., Williams, L. D., Ughetto, G., van der Marel, G. A., van Boom, J. H., Rich, A., & Wang, A. H.-J. (1990) *Biochemistry 29*, 2538–2549.

Gao, Y.-G., Liaw, Y.-C., Robinson, H., & Wang, A. H.-J. (1990) *Biochemistry 29*, 10317–10316.

Gause, G. F., & Dudnik, Y. V. (1971) *Prog. Mol. Subcell. Biol. 2*, 33–39.

Glaubiger, D., Kohn, K. W., & Charney, E. (1974) *Biochim. Biophys. Acta 361*, 303–311.

Goodsell, D. S., Grzeskowiak, K., & Dickerson, R. E. (1994) *J. Mol. Biol. 239*, 79–96.

Graves, D. E., Pattaroni, C., Krishnan, B., Ostrander, J. M., Hurley, L. H., & Krugh, R. T. (1984) *J. Biol. Chem. 259*, 8202–8209.

Graves, D. E., Stone, M. P., & Krugh, T. R. (1985) *Biochemistry 24*, 7573–7581.

Grzeskowiak, K., Yanagi, K., Privé, G. G., & Dickerson, R. E. (1991) *J. Biol. Chem. 266*, 8861–8883.

Hendrickson, W. A., & Konnert, J. H. (1980) in *Computing in Crystallography* (Diamond, R., Ramaseshan, S., & Venkatesan, K., Eds.) pp 13.01–13.23, Indian Academy of Sciences, Bangalore.

Hertzberg, R. P., Hecht, S. M., Reynolds, V. L., Molineaux, I. J., & Hurley, L. H. (1986) *Biochemistry 25*, 1249–1258.

Horowitz, S. B. (1971) *Prog. Mol. Subcell. Biol. 2*, 40–47.

Horowitz, S. B., Chang, S. C., Grollman, A. P., & Borkovec, A. B. (1971) *Science 174*, 159–161.

Hurley, L. H. (1977) *J. Antibiot. 30*, 349–370.

Hurley, L. H., & Petrusek, R. (1979) *Nature 282*, 529–531.

Hurley, L. H., & Thurston, D. E. (1984) *Pharm. Res. 1*, 52–59.

Hurley, L. H., Gairola, C., & Zmijewski, M. (1977) *Biochim. Biophys. Acta 475*, 521–535.

Hurley, L. H., Reck, T., Thurston, D. E., & Langley, D. R. (1988) *Chem. Res. Toxicol. 1*, 258–268.

Jones, G. B., Davey, C. L., Jenkins, T. C., Kamal, A., Kneale, G. G., Neidle, S., Webster, G. D., & Thurston, D. E. (1990) *Anti-Cancer Drug Des. 5*, 249–264.

Jones, T. A. (1978) *J. Appl. Crystallogr. 11*, 268−272.
Kamitori, S., & Takusagawa, F. (1992) *J. Mol. Biol. 225*, 445−456.
Khorlin, A. A., Krylov, A. S., Grokhovsky, S. L., Zhuze, A. L., Zasedatelev, A. S., Gursky, G. V., & Gottikh, B. P. (1980) *FEBS Lett. 118*, 311−314.
Kizu, R., Draves, P. H., & Hurley, L. H. (1993) *Biochemistry 32*, 8712−8722.
Kohn, K. W., & Spears, C. L. (1970) *J. Mol. Biol. 51*, 551−572.
Kohn, K. W., Glaubiger, D., & Spears, C. L. (1974) *Biochim. Biophys. Acta 361*, 288−302.
Kopka, M. L., & Larsen, T. A. (1992) in *Nucleic Acid Targeted Drug Design* (Propst, C. L., & Perun, T. J., Eds.) pp 303−374, Marcel Dekker, Inc., New York.
Kopka, M. L., Pjura, P., Yoon, C., Goodsell, D., & Dickerson, R. E. (1985a) in *Structure & Motion: Membranes, Nucleic Acids & Proteins* (Clementi, E., Corongiu, G., Sarma, M. H., & Sarma, R. H., Eds.) pp 461−483, Adenine Press, New York.
Kopka, M. L., Yoon, C., Goodsell, D., Pjura, P., & Dickerson, R. E. (1985b) *Proc. Natl. Acad. Sci. U.S.A. 82*, 1376−1380.
Kopka, M. L., Yoon, C., Goodsell, D., Pjura, P., & Dickerson, R. E. (1985c) *J. Mol. Biol. 183*, 553−563.
Korman, S., & Tendler, M. D. (1965) *J. New Drugs 5*, 275−285.
Krugh, T. R., Graves, D. E., & Stone, M. P. (1989) *Biochemistry 28*, 9988−9994.
Larsen, T. A., Goodsell, D. S., Cascio, D., Grzeskowiak, K., & Dickerson, R. E. (1989) *J. Biomol. Struct. Dyn. 7*, 477−491.
Leimgruber, W., Stefanovic, V., Karr, A., & Berger, J. (1965a) *J. Am. Chem. Soc. 87*, 5791−5793.
Leimgruber, W., Batcho, A. D., & Shencker, F. (1965b) *J. Am. Chem. Soc. 87*, 5793−5795.
Leimgruber, W., Batcho, A. D., & Czajkowski (1968) *J. Am. Chem. Soc. 90*, 5641−5643.
Leinsoo, T. A., Nikolaev, V. A., Grokhovskii, S. L., Surovaya, A. M., Sidorova, N. Yu., Strel'tsov, S. A., Zasedatelev, A. S., Zhuze, A. L., & Gurskii, G. V. (1989) *Mol. Biol. (USSR) 23*, 1616−1637.
Leonard, G. A., & Hunter, W. N. (1993) *J. Mol. Biol. 234*, 198−208.
Lipanov, A., Kopka, M. L., Kaczor-Grzeskowiak, M., Quintana, J., & Dickerson, R. E. (1993) *Biochemistry 32*, 1373−1375.
Lown, J. W., & Joshua, A. V. (1979) *Biochem. Pharmacol. 28*, 2017−2026.
Lown, J. W., Krowicki, K., Balzarini, J., Newman, R. A., & De Clerq, E. (1989) *J. Med. Chem. 32*, 2368−2375.
Mostad, A., Romming, C., & Storm, B. (1978) *Acta Chem. Scand. B32*, 639−645.
Pierce, J. R., Nazimiec, M., & Tang, M.-S. (1993) *Biochemistry 32*, 7069−7078.
Privé, G. G., Yanagi, K., & Dickerson, R. E. (1991) *J. Mol. Biol. 217*, 177−199.
Quintana, J. R., Lipanov, A. A., & Dickerson, R. E. (1991) *J. Mol. Biol. 221*, 919−940.
Quintana, J. R., Grzeskowiak, K., Yanagi, K., & Dickerson, R. E. (1992) *J. Mol. Biol. 225*, 379−395.
Rao, S. N., Singh, U. C., & Kollman, P. A. (1986) *J. Med. Chem. 29*, 2484−2492.
Remers, W. A., Mabilia, M., & Hopfinger, A. J. (1986) *J. Med. Chem. 29*, 2492−2503.
Schultz, P. G., & Dervan, P. B. (1983) *J. Am. Chem. Soc. 105*, 7748−7750.
Seeman, N. C., Rosenberg, J. M., & Rich, A. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 804−808.
Takeuchi, T., Miyamoto, M., Ishizuka, M., Naganawa, H., Kondo, S., Hamada, M., & Umezawa, H. (1976) *J. Antibiot. 29*, 93−96.
Tendler, M. D., & Korman, S. (1963) *Nature 199*, 501.
Ughetto, G., Wang, A. H.-J., Quigley, G. J., van der Marel, G. A., van Boom, J. H., & Rich, A. (1985) *Nucleic Acids Res. 13*, 2305−2323.
Walker, W. L., Kopka, M. L., Filipowsky, M. E., Dickerson, R. E., & Goodsell, D. S. (1994) *Biopolymers* (submitted for publication).
Wang, J.-J., Hill, G. C., & Hurley, L. H. (1992) *J. Med. Chem. 35*, 2995−3002.
Wang, W., & Lown, J. W. (1992) *J. Med. Chem. 35*, 2890−2897.
Yanagi, K., Privé, G. G., & Dickerson, R. E. (1991) *J. Mol. Biol. 217*, 201−214.
Yuan, H., Quintana, J., & Dickerson, R. E. (1992) *Biochemistry 31*, 8009−8021.