Articles

Binding of the Antitumor Drug Nogalamycin and Its Derivatives to DNA: Structural Comparison[†,‡]

Yi-Gui Gao, Yen-Chywan Liaw, Howard Robinson, and Andrew H.-J. Wang*

Department of Physiology and Biophysics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801

Received April 16, 1990; Revised Manuscript Received July 18, 1990

ABSTRACT: The three-dimensional molecular structures of the complexes between a novel antitumor drug nogalamycin and its derivative U-58872 with a modified DNA hexamer d[m^{5}CGT(pS)Am^{5}CG] have been determined at 1.7- and 1.8-Å resolution, respectively, by X-ray diffraction analyses. Both structures (in space group $P6_1$) have been refined with constrained refinement procedure to final R factors of 0.208 (3386 reflections) and 0.196 (2143 reflections). In both complexes, two nogalamycins bind to the DNA hexamer double helix in a 2:1 ratio with the elongated aglycon chromophore intercalated between the CpG steps at both ends of the helix. The aglycon chromophore spans across the GC Watson–Crick base pairs with its nogalose lying in the minor groove and the aminoglucose lying in the major groove of the distorted B-DNA double helix. Most of the sugars remain in the C2′-endo pucker family, except three deoxycytidine residues (terminal C1, C7, and internal C5). All nucleotides are in the anti conformation. Specific hydrogen bonds are found in the complex between the drug and guanine–cytosine bases in both grooves of the helix. One hydroxyl group of the aminoglucose donates a hydrogen bond to the N7 of guanine, while the other receives a hydrogen bond from the N4 amino group of cytosine. The orientation of these two hydrogen bonds suggests that nogalamycin prefers a GC base pair with its aglycon chromophore intercalating at the 5′-side of a guanine (between NpG), or at the 3′-side of a cytosine (between CpN) with the sugars pointing toward the GC base pair. The binding of nogalamycin to DNA requires that the base pairs in DNA open up transiently to allow the bulky sugars to go through, suggesting that nogalamycin prefers GC sequences embedded in a stretch of AT sequences.

Nogalamycin (Figure 1) is an interesting antitumor anthracycline antibiotic derived from *Streptomyces*. It is active against a number of tumor cell lines (Bhuyan & Reusser, 1970; Wiley, 1979). Although the drug has not been investigated further for clinical use since the mid-1970s due to the difficulty in administering the drug, it has been continually studied owing to its unique DNA binding properties. Nogalamycin, in contrast to other anthracycline antibiotics (e.g., daunomycin and adriamycin) which have a single amino sugar linked to the C7 position in ring A of the aglycon chromophore, contains two sugar moieties (nogalose and aminoglucose) attached to rings A and D, respectively. Many natural and semisynthetic derivatives of nogalamycin, e.g., disnogalamycin and U-58872 (Figure 1), have been studied to attempt to identify better agents for therapeutic purposes (Wiley, 1979; Wiley et al., 1982). One of them, menogaril, is currently under phase II clinical trial (Adams et al., 1989).

These anthracycline antibiotics bind to the DNA double helix by intercalation (Crooke & Reich, 1980; Wang, 1987). Their DNA binding affinity and sequence specificity are likely to be closely related to their biological activities. To better correlate the structure–function relationship and to design better agents on the basis of these correlations, it is useful to have a detailed view of how drug molecules interact with their target DNA molecules by different biochemical and biophysical methods. Toward this goal, we have recently determined the three-dimensional structure of molecular complexes between daunomycin (and adriamycin) and several DNA hexamers by high-resolution X-ray diffraction analysis (Wang et al., 1987; Frederick et al., 1990; Williams et al., 1990b). These studies have provided valuable information regarding the role of various functional groups of the drug molecules. For example, it was found that the elongated aglycon chromophore intercalated between two CG base pairs with the amino sugar lying in the minor groove. In addition, the essential O9 hydroxyl group (for the biological activity) is in a position to form hydrogen bonds with the guanine base adjacent to the aglycon ring.

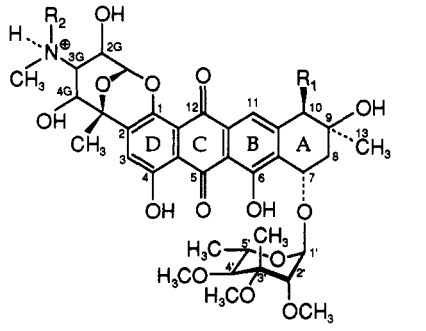
Nogalamycin, with bulky sugars attached at both ends of the chromophore, poses an intriguing question with respect to the ways in which it inserts itself between the base pairs. In the normal right-handed B-DNA double helix, it is difficult to stretch the sugar–phosphate backbones to such an extent that these bulky sugars can go through between the base pairs. Therefore, it has been suggested that the drug binds only to premelted DNA regions (Fox & Waring, 1984). On the basis of model building, it seemed that nogalamycin binds to DNA double helix with these two sugars lying separately in the major and minor grooves (Arora, 1983), although this was not firmly shown experimentally. Furthermore, it was not easy to distinguish in which groove the nogalose or aminoglucose sugar resides, since both orientations seem equally plausible (Collier et al., 1984). Another important issue related to nogalamycin binding to DNA is its nucleotide sequence specificity. DNase

[†] This work was supported by NSF and NIH (A.H.-J.W.). Y.-C.L. acknowledges support from the Institute of Molecular Biology, Taiwan (ROC).

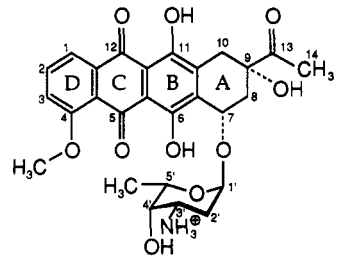
[‡] Atomic coordinates of two crystal structures have been submitted to the Brookhaven Protein Data Bank.

* To whom correspondence should be addressed.

0006-2960/90/0429-10307$02.50/0



Nogalamycin : $R_1 = COOCH_3$; $R_2 = CH_3$
Disnogalamycin : $R_1 = H$; $R_2 = CH_3$
U-58872 : $R_1 = COOCH_3$; $R_2 = CHO$



Daunomycin

FIGURE 1: Molecular formula of two anthracycline antibiotics, nogalamycin and its derivatives (top) and daunomycin (bottom). Both contain an aglycon chromophore with four fused rings (A–D). Rings B–D are unsaturated with exocyclic oxygen atoms, whereas ring A is semisaturated. Nogalamycin and its derivatives have two sugars attached to the aglycon with nogalose at C7 and a positively charged α-D-3,6-dideoxy-3-(dimethylamino)glucose (abbreviated aminoglucose in text) at C1/C2 positions. Daunomycin has an amino sugar attached at the C7 position. The hydroxyl group at the C9 position has an opposite chirality in the two drugs.

I footprinting experiments of nogalamycin on several DNA restriction fragments suggested that the drug binds to alternating purine–pyrimidine sequences such as TpG and GpT, though some ambiguities still remain (Fox & Waring, 1986).

Recently, the interactions between nogalamycin and a DNA hexamer d(GCATGC), which contains the putative preferred binding sites of nogalamycin, have been studied by nuclear magnetic resonance spectroscopy (Searle et al., 1988). That work suggested that in the 2:1 complex two nogalamycin aglycon chromophores intercalated between the CpA (and its complement TpG) steps, with the aminoglucose and nogalose lying in the major and minor grooves, respectively.

In order to visualize the exact mode of interaction between nogalamycin and DNA, we have carried out high-resolution X-ray diffraction analysis of nogalamycin and its derivatives complexed to DNA oligonucleotides. In this paper we describe the three-dimensional structures of the hexagonal ($P6_1$) form of the 2:1 complexes of nogalamycin (and its derivative U-58872) with a methylated DNA hexamer, d[m^{5}CGT(pS)m^{5}CG], determined at 1.7-Å (and 1.8-Å) resolution. The structure of a related hexagonal ($P6_122$) form of the same nogalamycin complex at 2.0-Å resolution has been published recently (Williams et al., 1990a). We compare these structures with that of the orthorhombic form of nogalamycin complexed to d[CGT(pS)ACG] (Liaw et al., 1989) at 1.3-Å resolution and those of the daunomycin (and adriamycin) complexed to related DNA hexamers, d(CGTACG) and d(CGATCG) (Wang et al., 1987; Moore et al., 1989; Frederick et al., 1990).

Materials and Methods

The thiophospho-containing oligonucleotides d[CGT(pS)ACG] and d[m^{5}CGT(pS)Am^{5}CG], where pS in an internucleotidic phosphorothioate linkage in the *R* configuration, were synthesized according to a procedure published earlier (Marrugg et al., 1986). Nogalamycin and its derivative U-58872 were gifts from Dr. Paul Aristoff of the Upjohn Co.; they were dissolved in methanol as stock solutions for crystallization. For the three nogalamcyin–DNA complexes, a typical crystallization mixture contained 0.5 mM of oligonucleotide, 8 mM $MgCl_2$, 30 mM sodium cacodylate (pH 5), 0.5 mM nogalamycin, and 2% (w/v) poly(ethylene glycol) 400 (PEG 400). The solution was equilibrated with 30% PEG 400 at room temperature (~25 °C) by vapor diffusion. The relevant crystal data are shown in Table I. Diffraction intensity data were collected on a Rigaku AFC-5R rotating anode X-ray diffractometer at 25 °C using the ω-scan mode with graphite monochromated CuKα radiation to different resolutions shown in Table I. For the two isomorphous hexagonal forms, the diffraction patterns showed pseudo higher symmetry of space group $P6_122$ with the *R* factor between *F*(*hkl*) and *F*(*khl*) being 0.16 and 0.17 for the nogalamycin and U-58872 complexes, respectively. Lp, empirical absorption, and decay corrections were applied for all three data sets.

Table I: Crystal Data of Nogalamycin–DNA Complexes

	nogalamycin + d[CGT-(pS)ACG]	nogalamycin + d[m^{5}CGT-(pS)Am^{5}CG]	U-58872 + d[m^{5}CGT-(pS)Am^{5}CG]
unit cell dimensions (Å)	a = 22.98 b = 47.27 c = 64.44	a = 26.31 b = 26.31 c = 100.26	a = 26.27 b = 26.27 c = 100.23
crystal system	orthorhombic	hexagonal	hexagonal
space group	$C222_1$	$P6_1$	$P6_1$
Z	8	6	6
V (Å^3)	70000	60104	59903
V_{asym} (Å^3)	8750	10017	9983
unique data	7580 5983 [$>3\sigma(F_o)$]	5276 3386 [$>2\sigma(F_o)$]	4210 2143 [$>2\sigma(F_o)$]
resolution (Å)	1.3	1.7	1.8
final R values	0.192	0.208	0.196

The Patterson maps of the two $P6_1$ crystal forms showed clearly the planar base stacking direction is along the *a*-axis (and the symmetry-related *b*-axis) direction which has a 26.3-Å repeat. This repeating dimensions are multiples (8 times) of 3.3 Å, suggesting that the complex is a hexamer duplex with two nogalamycins intercalated in it. The structure of the $P6_1$ form of nogalamycin complex was first solved by the molecular replacement method using the program ULTIMA (Rabinovich & Shakked, 1984) using a model derived from the molecular structure of the daunomycin–d(CGTACG) complex (Wang et al., 1987). Since the diffraction pattern has a pronounced pseudo $P6_122$ symmetry, we surmised that the molecular 2-fold rotation axis of the drug–DNA complex should nearly coincide with the unit cell (100) axis which would be a crystallographic 2-fold axis if the space group were $P6_122$. The complex was placed on and translated along this axis to search for the correct solution by using the conventional *R* factor (of the data between 25 to 8 Å) as an indicator for the correctness of the position. It was found that when the molecular center of gravity is located at (0.112, 0.223, 0.375), the *R* factor is at a minimum of 0.37. The model derived from this location was then initially refined by using a rigid body refinement method to an *R* factor of 35.0% to 6.0-Å resolution and further refined by using the Konnert–Hendrickson constrained refinement procedure (Hendrickson & Konnert, 1979) to an *R* factor of ~30% to 3.0-Å resolution. At this stage we were quite confident that a correct structure was in hand, and the refinement was extended to higher resolution. A series of Fourier maps were calculated to locate the solvent water

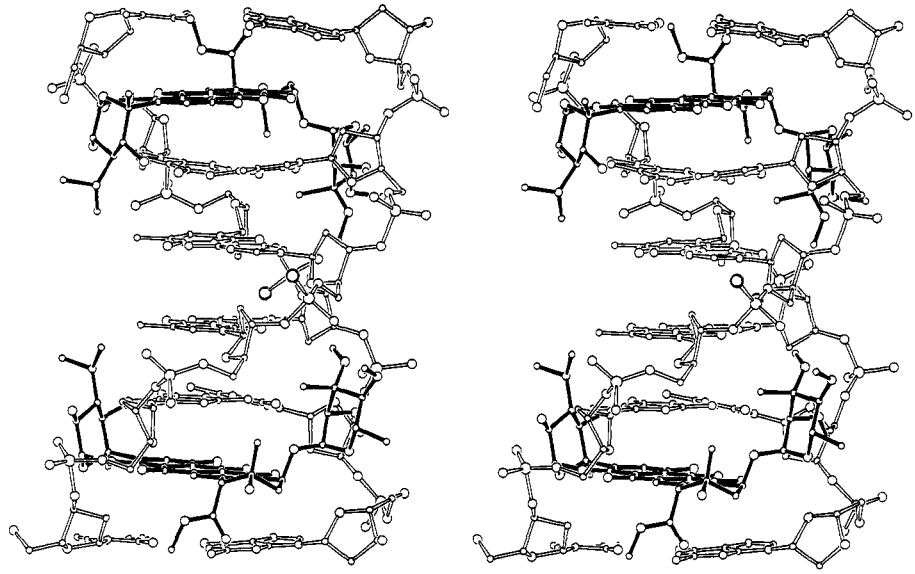


FIGURE 2: Stereoscopic drawing of the nogalamycin–d[m^{5}CGT(pS)am^{5}CG] complex viewed perpendicular to the noncrystallographic molecular 2-fold symmetry. Two nogalamycins (in filled bonds) intercalate between the CpG steps at both ends of the distorted right-handed B-DNA hexamer duplex (open bonds), with the aglycon chromophore penetrating through the helix between the C·G base pairs. The nogalose lies in the minor groove (right side), while the aminoglucose is located in the major groove (left side). These sugars nearly fill up both grooves of the hexamer duplex completely, displacing many first shell water molecules. Each nogalamycin covers a little over three base pairs.

molecules in the crystal lattice. During the refinement, the model was periodically compared to the structure of the $C222_1$ crystal form, which was refined at significantly higher (1.3-Å) resolution, to detect any apparent deviations. The structure of the $P6_1$ form was refined to a final R factor of 20.8% at 1.7-Å resolution with the root mean square (rms) differences in bond distances of 0.032 Å from the ideal values. The structure of the nogalamycin derivatives (U-58872) complexed to the same hexamer was similarly refined to an R factor of 0.196 at 1.8-Å resolution. The refinement parameters are listed in Table 1S of the supplementary material. The final atomic coordinates of these two $P6_1$ forms have been deposited in the Brookhaven Protein Data Bank. The structure of the a related $P6_122$ crystal form has been independently determined and refined to a final R factor of 0.204 at 2.0-Å resolution by using 809 reflections (Williams et al., 1990a).

The structure of the $C222_1$ form has been described briefly recently (Liaw et al., 1989). Here we compare the structure of the two hexagonal crystal forms containing nogalamycin and U-58872 (formamido derivative of nogalamycin; see Figure 1) and address the effect of chemical modification of the drug molecule on its DNA binding interactions.

Results

Molecular Structure. The overall structure of the 2:1 complex of nogalamycin–d[m^{5}CGT(pS)Am^{5}CG] is shown in Figure 2, which reveals a number of important features. It can be seen that the two nogalamycin molecules are intercalated between the CpG steps at both ends of a distorted B-DNA double helix. The complex maintains an approximate noncrystallographic molecular 2-fold symmetry. The rms deviation between the two halves (one drug plus one DNA hexamer) of the complex is 0.778 Å. The elongated aglycon chromophore (rings A–D) penetrates the DNA double helix such that it is almost perpendicular to the C1′–C1′ vectors of the two GC base pairs above and below the intercalator. The drug spans the two grooves of the helix with the nogalose in the minor groove and the aminoglucose in the major groove. The aminoglucose is in an orientation so that the flat surface of the six-membered ring is facing toward the GC base pair of the major groove. In contrast, the nogalose nestles in the minor groove with the sugar six-membered ring going sideways approaching the bottom of the minor groove. The two sugars on the same side of the flat aglycon chromophore, wrapping around the second (and the fourth) GC base pair, and they both point toward the AT region in the middle of the helix.

It is interesting to note that although the drug molecules in the complex have an overall conformation similar to that of the free drug (Arora, 1983) due to the rigidity of the ring systems, they differ in a subtle way. For example, the torsion angles around the glycosyl ether linkage (C8–C7–O7–C1′ and C7–O7–C1′–C2′) are 101° and 163°, respectively, in comparison to 105° and 165° in the free drug. All the methoxy groups in nogalose (in chair conformation) are also oriented with respect to one another in almost exactly the same manner as found in the free drug. However, when the two molecules are overlaid on each other by a least-squares fitting of rings B–D and ring A (Figure 3), the aglycon chromophore of the bound drug in the complex seems to bend down, making the aminoglucose slightly closer to the nogalose. Detailed analysis of this bending suggests that it is caused by a combination of changes of torsion angles around C1–C2 (by ~15°) and C1–O1 (by ~10°) bonds and smaller but concerted twisting at several other places along the aglycon ring. This reason for this bending of the bound drug will be discussed (vide infra).

Figure 4 shows the interactions between nogalamycin and DNA viewed from the minor groove (Figure 4A) and from the major groove (Figure 4B). In Figure 4A, the C9 atom in ring A can be seen to deviate from the mean plane with the largest displacement (0.53 Å). This pucker in ring A aligns the O9 hydroxyl in the equatorial position pointing away from the DNA without interacting with DNA directly, in contrast to that seen in the daunomycin–d(CGTACG) complex, which has the O9 hydroxyl group forming two hydrogen bonds with the N2 amino group and the N3 atom of the guanine base in the minor groove (Wang et al., 1987). The C13 methyl group attached to the C9 position is in the axial position with the C9–C13 bond almost parallel to the C7–O7 bond. In this orientation, this methyl group approaches another methyl group on C3′ of the nogalose with a distance of 4.2 Å. Those methyl groups, on both ring A and nogalose, constitute a large

Table II: Torsion Angles (deg) and Some Helical Parameters[a] of the Nogalamycin–d[m^{5}CGT(pS)Am^{5}CG] Complex

	α	β	γ	δ	ϵ	ζ	χ	pucker	ω	κ
C1			45.5	78.5	−147.8	−121.3	−165.4	C4′-exo	35.0	−5.8
C7			−62.3	138.8	−92.7	−61.5	171.0	O1′-exo		
G2	−19.4	154.4	48.0	144.3	−90.0	144.0	−90.6	C1′-exo	27.9	14.9
G8	−57.6	164.4	57.0	130.2	−174.6	−138.2	−102.8	C1′-exo		
T3	−74.7	131.6	71.4	103.6	−170.0	−94.9	−141.7	C1′-exo	32.8	3.5
T9	−46.4	141.2	56.4	108.9	−163.1	−91.0	−126.6	C1′-exo		
A4	−78.5	165.7	63.5	87.3	−168.7	−79.2	−128.4	C1′-exo	26.1	−3.7
A10	−31.6	−170.5	1.0	135.3	−165.0	−85.1	−111.8	C2′-endo		
C5	−76.5	178.4	54.7	116.4	−85.0	−174.8	−98.4	O1′-endo	40.3	−18.4
C11	144.6	169.7	176.6	124.1	−90.2	−171.6	−131.0	C3′-exo		
G6	−75.2	−165.1	17.0	132.9			−121.3	C4′-endo		6.7
G12	−44.5	128.4	70.7	118.3			−125.9	C1′-exo		
B-DNA[b]	−63	171	54	123	−169	−108	−117			

[a] Torsion angles along the backbone of the oligonucleotide are defined as $P\underline{^{\alpha}}O5'\underline{^{\beta}}C5'\underline{^{\gamma}}C4'\underline{^{\delta}}C3'\underline{^{\epsilon}}O3'\underline{^{\zeta}}P$, and χ is the glycosyl angle. The nucleotides are numbered from C1 to G6 in one strand and C7 to G12 in the other strand. The nomenclature of the helical parameters follows that of Dickerson et al. (1989). ω is the helical twist angle, and κ is the base pair buckle. The helical parameters are calculated by using the program NEWHELIX (Dickerson, 1990). [b] Average values from DNA dodecamer d(CGCGAATTCGCG) (Drew & Dickerson, 1981).

hydrophobic patch in the minor groove, which is expected to present hindrances for solvent water molecules to interact with DNA. Another substituent group in ring A, the carbomethoxy on C10, is in the axial position almost perpendicular to the plane of the aglycon, and its carbonyl oxygen O14 atom is in a good position to receive a hydrogen bond from the NH_2 of G12 (3.02 Å; a similar bond of 3.06 Å for the second nogalamycin from G6) as evident in this figure. In addition, the O7 of the glycosyl linkage of nogalose receives another weaker hydrogen bond from the NH_2 of G2 (3.38 Å; 3.28 Å for the second nogalamycin from G8). Therefore, the juxtaposition of the various groups in ring A is quite different from that of daunomycin, leading to different DNA binding interactions for the two drugs.

In the major groove, there are two direct hydrogen bonds between drug and DNA (Figure 4B). First, the O2G hydroxyl of the aminoglucose is 2.84 Å from N7 of guanine G2. A somewhat weaker bond is found between O4G and N4 of cytosine m^{5}C11 (3.10 and 3.21 Å for the symmetry mate). The disposition of this pair of hydrogen bonds appears to have excellent complementarity with a single G·C base pair in the major groove. The dimethylamino group in the aminoglucose is bridged through a water molecule to the N6 of adenine A10 one base pair below. It is worthy to note that, in the U-58872–DNA complex, this bridging water is replaced by an oxygen atom of the formamido NCHO group (Figure 4C) and a direct hydrogen bond is formed between N6 of A10 and the NCHO oxygen.

Figure 5 shows several van der Waals views of the complex. Figure 5A is a stereoscopic van der Waals drawing viewed from the same orientation as in Figure 2. Here the close fit between the drug and DNA is evident. Especially the hydrophobic nogalose appears to completely occupy the minor groove. The hydrophobic methyl groups make van der Waals contacts to the bases and the backbones of DNA in Figure 5B. For example, the methoxy group on the C3′ of the nogalose approaches the H2 atom of adenine A10, with the distance between the methyl and H2 being 2.8 Å. Interestingly, the minor groove width is only slightly affected by the occupation of the nogalose in the minor groove. For example, the phosphorus to phosphorus distances across the groove (e.g., 15.3 Å for P3–P10, 14.9 Å for P4–P11) are not significantly different from the corresponding values found in the uncomplexed B-DNA double helix (15.0 Å) (Arnott et al., 1982). In contrast, complexes of the DNA and minor groove binders like netropsin have much narrower DNA minor grooves (~10 Å) in the AT region where the drug occupies (Kopka et al., 1985; Coll et al., 1987; Wang & Teng, 1990). The aminoglucoses from two nogalamycins cover much of the major groove surface as shown in Figure 5C. These van der Waals drawing clearly illustrate the tight fit between nogalamycin and the DNA hexamer and suggest the complex should be very stable.

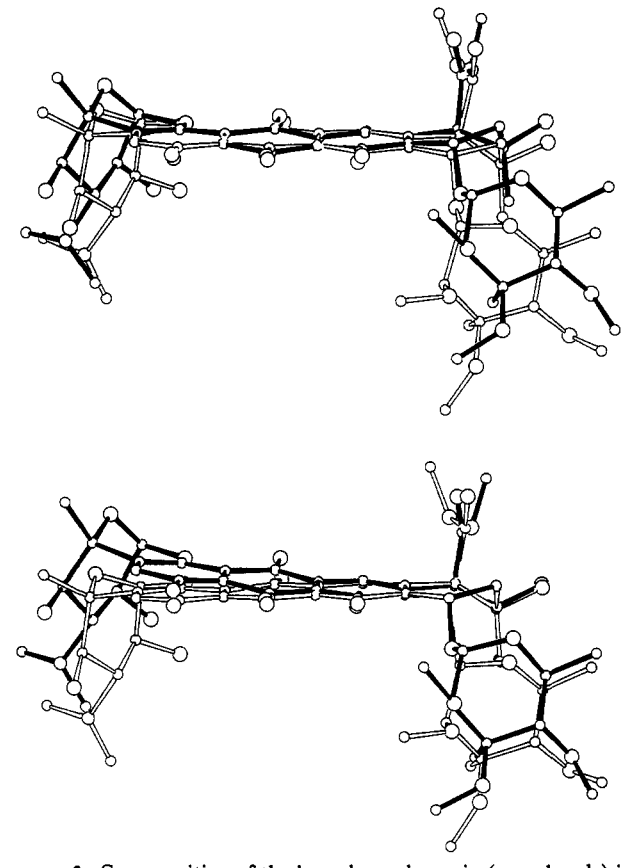


FIGURE 3: Superposition of the bound nogalamycin (open bonds) in the nogalamycin–d[m^{5}CGT(pS)Am^{5}CG] complex and the free nogalamycin (filled bonds) (Arora, 1983). (Top) Least-squares fitting of rings B–D. (Bottom) Least-squares fitting of ring A only. The bound nogalamycin bends slightly along the long aglycon chromophore, bringing the nogalose and aminoglucose closer.

DNA Conformation. It is interesting to ask what kinds of significant rearrangements of the DNA backbone torsion

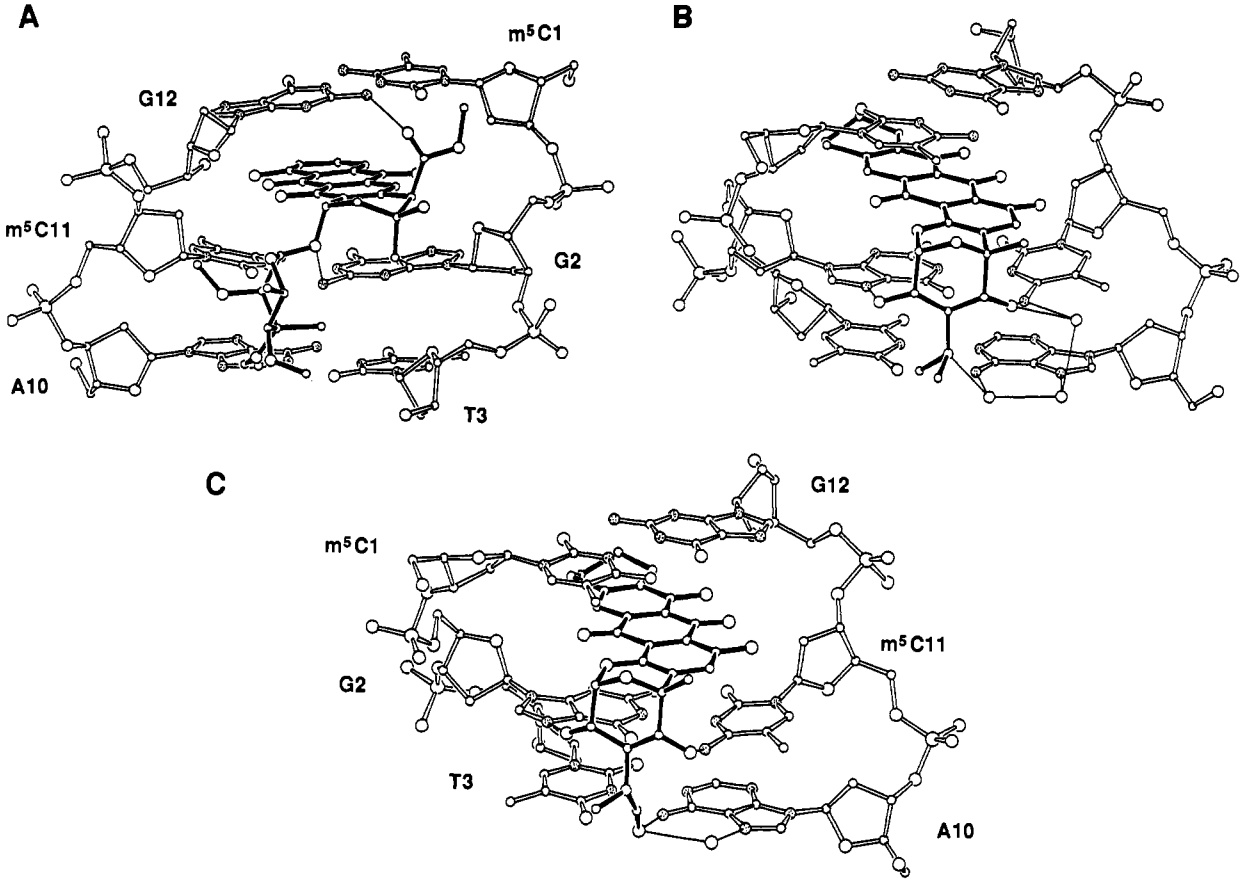


FIGURE 4: Skeletal diagram of the detailed surroundings of the intercalated nogalamycin. Three base pairs of the hexamer helix are shown. The other half of the complex is similar. Hydrogen bonds between the nogalamycin and DNA are shown as thin lines. (A) In the minor groove, two hydrogen bonds are found. The carbonyl oxygen O14 of the carbomethoxy receives a hydrogen bond (3.02 Å) from the N2 amino group of the G12 residue. In addition, N2 of G2 donates a weaker hydrogen bond (3.38 Å) to the O7 atom of nogalose. The tight fit of drug to DNA results a significant buckle in the C11–G2 base pair (14.9°). (B) In the major groove, the hydroxyl O2G of the aminoglucose forms a strong hydrogen bond (2.84 Å) to the N7 of guanine G2, and the other hydroxyl O4G is hydrogen bonded to N4 of cytosine m^5C11 (3.10 Å). Three bridging water molecules are also shown. (C) The major groove view of the U-58872 complex. A direct hydrogen bond is formed between formamido NCHO of U-58872 and the N6 of adenine A10.

angles have occurred around the drug molecule due to the binding of bulky nogalamycin to DNA. The torsion angles of the distorted right-handed double helical DNA are listed in Table II and Table 2S (supplementary material) for the nogalamycin and U-58872 complexes, respectively. When compared to the corresponding values of the orthorhombic form, they are more variable. Nevertheless, some common trends are notable. The largest deviations of these angles from the normal, uncomplexed B-DNA are associated with the C-G residues involved in the intercalation sites, as expected. These include the ϵ of C5 and C11 (−85.0° and −90.2°) with the concomitant changes of ζ to −174.8° and −171.6°, respectively. The puckers of all the sugars remain in the C2′-endo type, except for the C1, C5, and C7 residues. No obvious pattern of the so-called C3′-endo-(5′,3′)-C2′-endo mixed sugar pucker (Tsai et al., 1977; Wang et al., 1978) is seen. This is quite similar to that observed in the daunomycin–d(CGTACG) complex (Wang et al., 1987).

Other distortions involve the inner G·C base pairs (G2·C11 and C5·G8) in the intercalated CpG steps which are significantly buckled (14.9° and −18.4°), as they are in close contact with the nogalose and aminoglucose sugars and they adjust to accommodate the bulky drug molecules. These buckles in the two complexes of the $P6_1$ form are slightly less than the corresponding values of the $C222_1$ form (Liaw et al., 1989). This may be due to the different crystal lattice forces and/or the lower resolution of the structure in the $P6_1$ form. The terminal two G·C base pairs, which are less buckled (−5.8° for C1·G12 and 6.7° for C7-G6), are located away from the drug sugars, and they are involved in the end-to-end crystal packing of the complexes along the unit cell *a* and *b* axes. The two A·T base pairs in the central T(pS)A step are quite normal in spite of the thiophospho linkage and the bound drug in the complex. Table 3S (supplementary material) lists the complete helical parameters of the DNA double helices for both complexes.

Comparison with Other Anthracycline Drugs. Figure 6 shows the view of the complex from a direction perpendicular to the plane of the aglycon ring. The long dimension of both aglycons lies across the base pairs, reaching both grooves. As pointed out before, the location of the aglycon ring relative to the base pairs in the nogalamycin complex is "pulled" toward the minor groove by about 2.0 Å, in comparison to that of the daunomycin–d(CGTACG) complex (Liaw et al., 1989). Ring D of nogalamycin is stacked underneath the N4 amino group of C1 residue.

An apparent feature from the stacking diagram is the symmetry between the two DNA strands. The long direction of the aglycon nearly coincides with the local 2-fold axis (between the two base pairs) that relates the two strands. In fact, the average helical twist angle of the two CpG steps across the intercalator in the nogalamycin complex is 38°, very close to that of the B-DNA. The rest of the steps have a combined unwinding angle of −12°. Therefore, the overall unwinding angle of the DNA helix due to the intercalation of nogalamycin is estimated to be about −10°, significantly lower than the

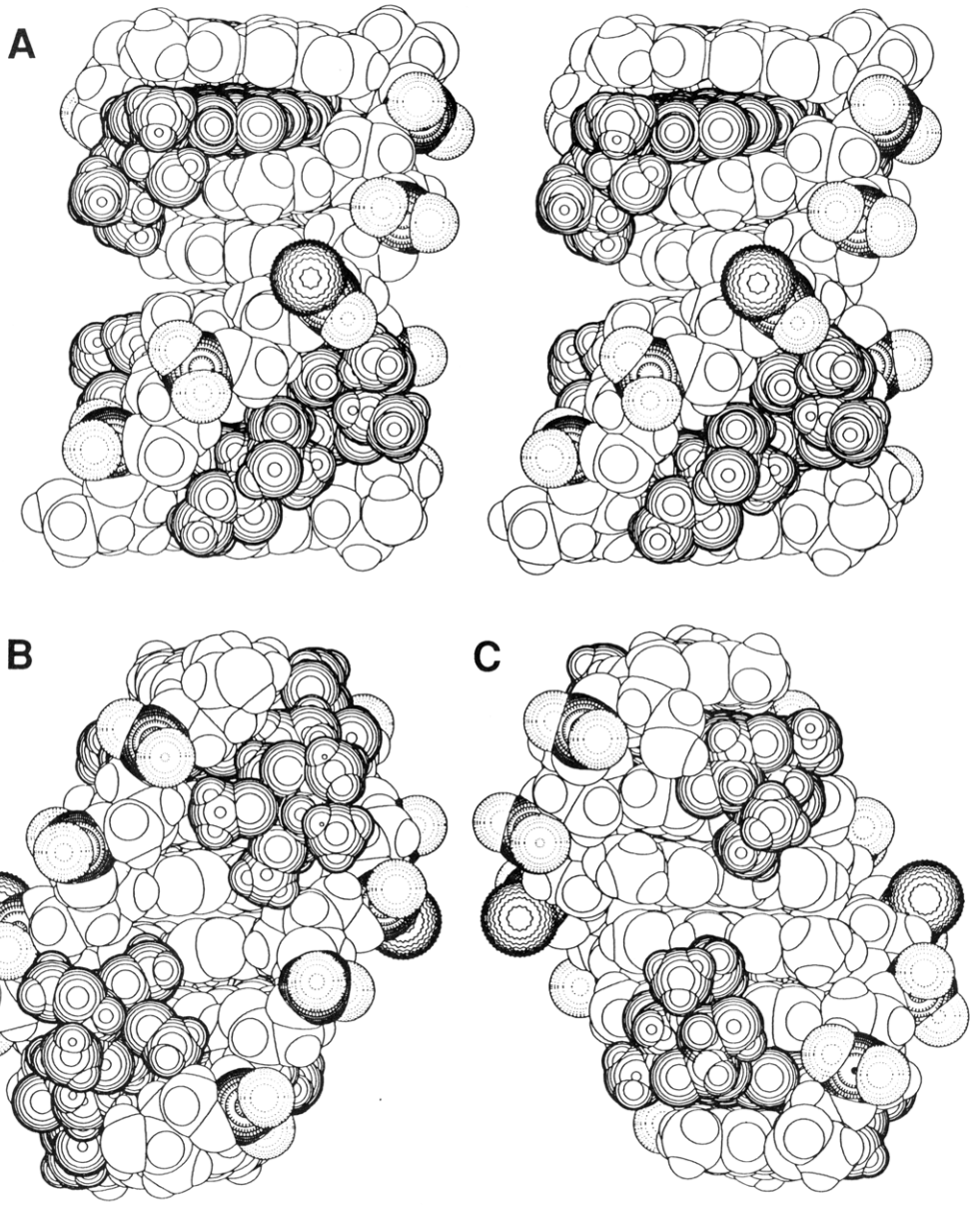


FIGURE 5: (A) Stereoscopic van der Waals diagram of the complex as viewed from the same direction as in Figure 2. Nogalamycin (in concentric solid circles) binds tightly to the double helix by a combination of the stacking, hydrogen-bonding, and van der Waals interactions. The sulfur atom in the thiophospho T(pS)A linkage is shown with jagged concentric circles. (B) View into the minor groove in which the nogalose, ring A, and carbomethoxy group are combined to cover three base pairs. (C) View into the major groove.

value observed for the simple intercalator ethidium (−26°) (Wang, 1974). This low DNA unwinding angle associated with nogalamycin is not unlike that seen in the case of daunomycin (Wang et al., 1987), suggesting that it may be a common feature of anthracycline drugs.

Comparison of Three Crystal Forms. The present methylated DNA–drug complexes crystallized in the hexagonal $P6_1$ space group, but their diffraction pattern showed strong pseudo 6_122 symmetry. The complexes possess only approximate, but not exact, 2-fold symmetry as evident from the torsion angles listed in Table II. Notably, the same nogalamycin–d-[m^{5}CGT(pS)Am^{5}CG] complex also crystallized in the $P6_122$ space group from different crystallization conditions (Williams et al., 1990a). In comparison, the unmethylated DNA–drug complex crystallized in the orthorhombic $C222_1$ space group. Despite different crystal lattice environments, most of the conformational features and interactions in the complex are remarkably similar in those three crystal forms.

The drug–DNA complexes are packed in the crystal lattices with the end-over-end base pair stacking interactions. In the hexagonal $P6_1$ form, this stacking interaction is along the *a* axis shown in Figure 7A, thus making a long column of molecules in that direction. This column is packed against another 6_1 symmetry-related column in the *c*-axis direction (Figure 7B). Two complexes approach each other such that the aminoglucose of one complex is in contact with the thiophosphate group of the other complex. Strikingly similar packing interactions are also found in the orthorhombic form of the nogalamycin–d[CGT(pS)ACG] complex (Liaw et al., 1989). This may explain why only complexes containing thiophosphate linkages can be crystallized since the presence of normal phosphate linkages may introduce additional, undesirable nonspecific hydrogen-bonding interactions with different functional groups, preventing the formation of the lattice.

Sequence Specificity of Nogalamycin. The three different nogalamycin–DNA complexes provide an opportunity to have a detailed view of the interactions between the drug and DNA. It is apparent from Figures 4 and 5 that several factors contribute to the stability of the complex. The most conspicuous

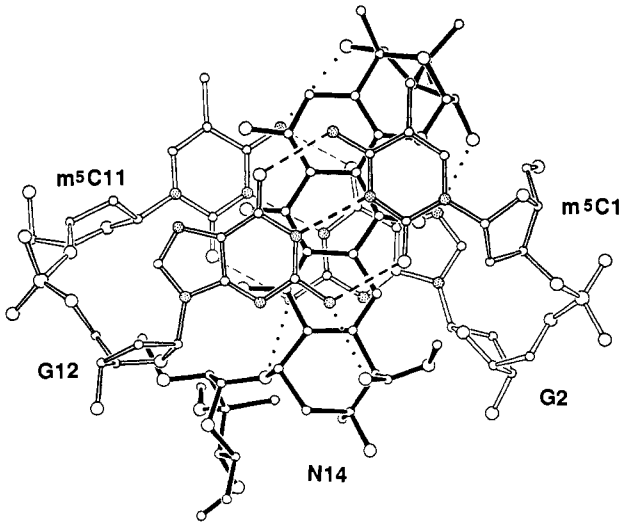


FIGURE 6: A view of the intercalated nogalamycin molecule and the two adjacent G·C base pairs from a direction perpendicular to the base plane. The long dimension of the aglycon chromophore is nearly perpendicular to the C1′–C1′ vectors of the base pairs.

one is the stacking between the aglycon chromophore and the G·C base pairs. However, it is the hydrogen-bonding interactions that determine the sequence specificity as mentioned earlier. In Figure 8 we compare the interactions between nogalamycin and a G·C base pair seen in the crystal structure with those of the other three possible base pairs (C·G, A·T, and T·A) using model building. Figure 8A shows that there are three good hydrogen bonds (two in major groove and one in minor groove) between nogalamycin and a single G·C base pair with a favorable bonding geometry. O6 of the guanine base is in contact with a hydrogen atom (attached to a carbon) from the aminoglucose. The nogalose gets close to the deoxycytidine residue (base and sugar), with the O7 oxygen receiving a hydrogen bond from the NH_2 of the guanine mate. When the G·C is reversed to C·G (Figure 8B), the aglycon chromophore is required to change its position in order to form the same hydrogen bonds. The new position moves the hydrophobic nogalose away from dG residue, leaving a gap between them, and puts the nogalose further into the solvent region. Furthermore, the carbomethoxy group on the C10 position clashes with the base pair (not shown in the figure) below the aglycon ring. For both A·T and T·A base pairs (Figure 8, panels C and D), an even less favorable situation exists. Here the N6 amino group from adenine makes a bad contact with a CH hydrogen atom from the aminoglucose. In addition, there is a space between the nogalose and the A·T or T·A base pair due to the missing N2 amino group of adenine in the minor groove.

The importance of the two hydroxyl groups (O2G and O4G) on the aminoglucose for the recognition of a G·C base pair should be stressed here. In the major groove of the DNA double helix, nogalamycin O2G donates a hydrogen bond to the N7 of guanine, while O4G receives a hydrogen bond from the N4 amino group of cytosine. This exact matching of hydrogen bond donor/acceptor between the drug and DNA is critical, as suggested from a newly purified antitumor antibiotic viriplanin A which belongs to the nogalamycin family (Kind et al., 1989). In viriplanin A, O4G has a trisaccharide attached, but O2G remains as a free hydroxyl group. Therefore, its O2G and O4G pair retains the identical donor/acceptor capacity to recognize a G·C base pair. A reversed modification, i.e., trisaccharide attached to O2G, is likely to perturb significantly the binding of the drug to DNA.

On the basis of the above reasonings, we suggest that nogalamycins has a DNA sequence preference for NpG or CpN steps. More specifically, *the nogalamcyin aglycon chromophore prefers to intercalate at the 5′-side of a guanine (i.e., between NpG) or at the 3′-side of a cytosine (between CpN)*

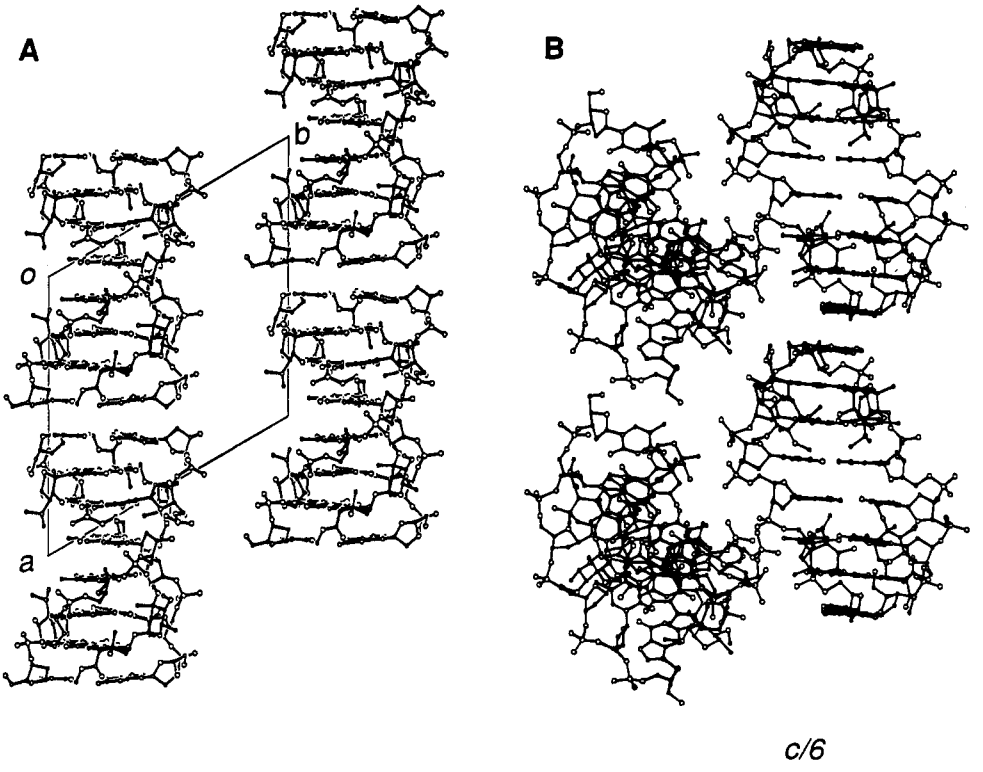


FIGURE 7: Packing diagram of the $P6_1$ crystal form of the nogalamycin–d[m^{5}CGT(pS)Am^{5}CG] complexes (four are shown). (A) The complexes are stacked end over end along the *a* and *b* axes, forming a sheet. (B) Successive sheets in the *c* direction are packed against one another, bringing the thiophospho group close to the adjacent complexes.

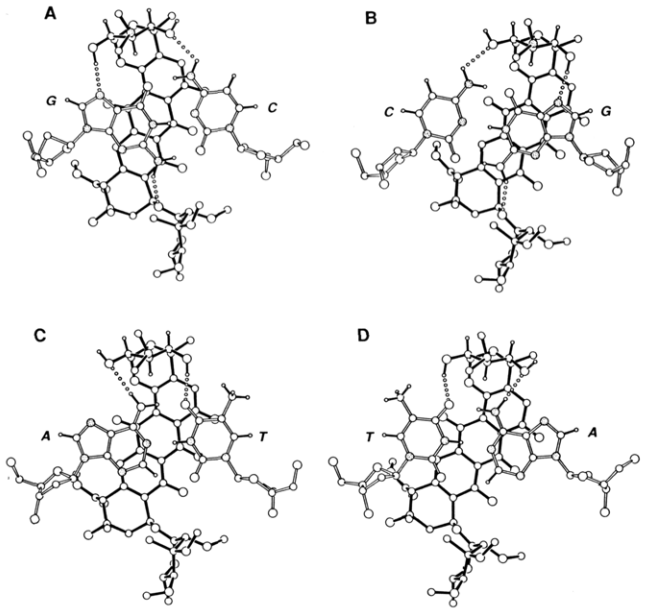


FIGURE 8: Modeling of the binding of nogalamycin to different base pairs. (A) Binding to a G·C base pair with the aglycon on the 5′-side of guanine. This is observed in the crystal structure. (B) Binding to a C·G base pair with the aglycon on the 3′-side of guanine. The aglycin is rotated clockwise by about 20° in order to re-form the hydrogen bonds. This new position creates a gap between nogalose and the backbone on the guanine strand. In addition, the carbomethoxy group clashes with the base pair below the aglycon ring. (C and D) Binding to A·T and T·A base pairs, respectively. The two hydrogen bonds in the major groove are probably not likely because there is a bad contact between the N6 amino group of adenine to a CH hydrogen atom in the aminoglucose. There is a hole between nogalose and the A·T base pair due to the missing N2 amino group. Consequently, only the binding mode shown in (A) is favored.

with the sugars facing toward the G·C base pair.

Discussion

Intercalator antitumor drugs constitute an important class of compounds for cancer chemotherapy (Denny, 1989). Although it is generally believed that the ability of these compounds to insert their planar chromophores between DNA base pairs plays an essential role in their biological activities, it is not the only requirement for those compounds to be effective as useful drugs, since many intercalators are not active anticancer agents. Other components in the molecules besides the intercalator chromophore are critical in determining whether they are active antitumor compounds or not. Presumably these components contribute in making the compounds have different DNA binding affinity or DNA sequence specificity. Additionally, they may alter the ways in which proteins (e.g., polymerases or topoisomerases) interact with the drug–DNA complexes (Wang, 1986). By solving the structure of the drug–DNA complexes, one can start to gain insights on the roles of various functional components in these compounds.

X-ray diffraction analysis is the most definitive way to visualize the molecular interactions between ligand and receptor such as those found in anticancer drug–DNA complexes. Up to the present time, only a few structures of these types have been determined by X-ray diffraction. These include the DNA minor groove binding antitumor drugs (e.g., netropsin, distamycin, Hoechst 33258) complexed to a series of related DNA dodecamers (Kopka et al., 1985; Coll et al., 1987; Wang & Teng, 1990). The only intercalator antitumor drug–DNA structures elucidated so far are those of simpler anthracyclines (daunomycin and adriamycin) complexed to DNA hexamers d(CGTACG) and d(CGATCG) (Wang et al., 1987; Moore et al., 1989; Frederick et al., 1990; Williams et al., 1990b).

The present two crystal structures involving methylated cytosine DNA hexamer expand the knowledge of the interactions between nogalamycin (and derivatives) and DNA. The comparison of these two structures with the nogalamycin–d[CGT(pS)ACG] structure (Liaw et al., 1989) shows that they are very similar with only minor differences. For example, there are two direct hydrogen bonds between the aminoglucose and the G2·C11 base pair in the methylated complex, but only one direct hydrogen bond (to N7 of G2) in the unmethylated hexamer. In the minor groove, the interactions are identical. In the U-58872 complex, an additional direct hydrogen bond is found between the NCHO group of the aminoglucose and the NH_2 of A10 adenine in the major groove. Although this provides extra bonding interaction, the U-58872 complex may not be more stable due to the loss of the positive charge of the dimethylamino group. The buckle in the G2·C11 base pair is smaller in the methylated hexamer complex (av 17°) than that in the unmethylated hexamer (av 25°). This difference is probably not due to the methylation of cytosine, since the same methylated hexamer complex in a different $P6_122$ lattice has a large (25°) buckle in the corresponding base pair (Williams et al., 1990a).

We have also shown that the thiophospho modification in the TpA steps does not affect the binding interactions of the drug. A two-dimensional NMR (500 MHz) study of the 2:1 complex between nogalamycin and d(CGTACG) showed that the distance information derived from the crystal structure is fully consistent with the NOE data, suggesting the crystal structure is likely to be preserved in solution (Robinson et al., 1990). This result also agrees with the recent observation in the daunomycin–DNA complexes in which the overall binding mode of the interaction between 11-deoxydaunomycin with d[CGT(pS)ACG] (Williams et al., 1990b) is very similar to that of the daunomycin–d(CGTACG) complex (Wang et la., 1987).

The structure determination of the more complicated nogalamycin–DNA complexes allows the interesting question regarding the mechanism of intercalation of nogalamycin to be addressed. The structure clearly shows that the drug binds to DNA with its bulky sugar residues at both ends of the aglycon ring lying in the major and minor grooves simultaneously. How does the drug overcome the substantial hindrance in order to thread through the base pairs? What kind of binding sequences are preferred for nogalamycin? There have been conflicting results to those questions. Some suggested G+C sequences are preferred, while others suggested A+T sequences (Bhuyan & Reusser, 1970; Bhuyan & Smith, 1965; Kersten et al., 1966). More recently, footprinting experiments have shown that nogalamycin appeared to bind go GpT or TpG sequences more frequently (Fox & Waring, 1986). However, Searle et al. (1988) have reinterpreted the same footprinting data and suggested a 5′-GCA sequence preference for nogalamycin. As discussed before, the crystal structure showed specific hydrogen bonds between nogalamycin and the G·C base pair, suggesting the drug prefers to have the aglycon chromophore intercalating between NpG or CpN steps with its sugars pointing in the direction of the G·C base pair. However, the drug actually covers three base pairs, as shown clearly in Figures 4 and 5. The influence of the sequence of the third base pair on the binding of nogalamycin is not clear, though we notice that the C3′ methoxy group of nogalose is in close contact with the H2 atom of A10 adenine

base. This may suggest that triplets like 5′-ACN or 5′-NGT are slightly favored. Finally, this reasoning immediately implies that nogalamycin can bind to the symmetrical CpG sequence with two possible orientations.

In order for the drug to intercalate the DNA, the double helix needs to open transiently with sufficient room for the bulky sugars to slide through between base pairs without hindrance. This requirement should favor the A-T sequences, since they open up more readily. This binding process is expected to be slow, as has recently been shown (Fox & Waring, 1986). Our recent NMR study of the solution containing nogalamycin and unmodified d(CGTACG) duplex in a 1:1 ratio reveals that there are two forms of complexes (both 2:1 and 1:1) as well as the free DNA in slow equilibrium on the NMR time scale (Robinson et al., 1990). From this analysis, we suggest that nogalamycin prefers G-C sequences (e.g., CpG) embedded in a stretch of A-T sequences. It is interesting to note that DNA sequences in the promoter regions often possess such features, i.e., AT-rich sequence sprinkled with GC sequences.

Our model seems to provide a reasonable explanation for the footprinting data (Fox & Waring, 1986). For example, the most apparent protected regions of the *tyr*T DNA to DNase I cleavage are located in the center of the following sequences: 5′-ACGCAACC, 5′-AACGTAAC, 5′-ACAGCG, and 5′-ATGCG. Here the triplet sequence 5′-ACN occurs five times and 5′-NGT once. In fact, out of the ~110-nucleotide sequence (nucleotides 25–135 where the footprinting data are reliable) of the *tyr*T DNA, only four other 5′-ACN or 5′-NGT sequences are found. In addition, these nogalamycin-protected sequences are flanked by AT-rich sequences, further supporting our proposal above. This type of sequence preference for a DNA binding compound is unique and may be common to compounds that are capable of binding to both grooves of the helix (Atwell et al., 1984; Zimmerman et al., 1989; Wakelin, 1986). In comparison, echinomycin and triostin A (two antitumor quinoxaline bis-intercalators) bind to tetranucleotide sequences (A/T)CG(A/T) (Wang et al., 1984).

A significant observation in the present structure is that in the complex both nogalamycin (guest) and DNA (host) molecules have changed their respective conformations in order to form a tight complex. The change in the nogalamycin is particularly surprising since the free drug seems to possess little flexibility. Instead, as shown in Figure 3, the bound nogalamycin bends slightly along the long direction of the aglycon chromophore so that the juxtaposition of various substituents at two ends of the molecules is significantly different from that of the free drug (Arora, 1983). Presumably, this bending occurs so that (1) the pair of hydroxyl groups of the aminoglucose can form hydrogen bonds to the G·C base pair in the major groove, (2) the O7 and the carbonyl group of the carbomethoxy at the C10 position of the ring A can receive a hydrogen bond in the minor groove, and finally (3) maximum van der Waals contacts can occur in both grooves. Using semiempirical and ab initio calculations, we estimate the energy cost for this bending in anthracyclines is not very great (manuscript in preparation). This bending in conjugated ring systems has also recently been found in an anthraquinone-containing antibiotic dynemicin, suggesting the facile nature of this bending property (Konishi et al., 1990). This points out that when doing model building studies of drug–receptor binding interactions, the possible conformational flexibilities in both drug and receptor have to be taken into consideration carefully. This concept of mutual conformational adaptation in drug–DNA complexes has been discussed more fully elsewhere (Wang et al., 1990).

Many antibiotics bind exclusively in the minor groove. We have suggested that this may be due to the natural selection process for the microbes to develop secondary metabolite agents that attack the minor groove where few proteins bind specifically (Ughetto et al., 1985; Wang, 1987). Interestingly, nogalamycin binds to DNA with its positively charged aminoglucose in the major groove, a rare example for natural product antibiotics. Some derivatives of nogalamycin, e.g., menogaril, do not have nogalose at the C7 position, and they have significant antitumor activity in vivo (Wiley, 1979; Adams et al., 1989). We are in the process of determining the structure of this derivative complexed to DNA to see in which groove the aminoglucose resides. We have also crystallized other derivatives of nogalamycin with a DNA octamer d(ACGTACGT), which will provide additional structural information regarding base pairs further away from the intercalation site.

In summary, the recent structural determination of several anthracycline antibiotics complexed to DNA (Wang, 1987; Moore et al., 1989; Liaw et al., 1989; Frederick et al., 1990; Williams et al., 1990a,b) allows some generalizations on the binding of those drugs to DNA to be made. First, the elongated aglycon chromophore intercalates between base pairs with its long direction almost perpendicular to the C1′–C1′ vector of the base pairs above and below the intercalator ring. Second, specific hydrogen bonds between DNA and drug are formed using hydroxyls (as donor and acceptor) and ether or carbonyl group(s) (as acceptor only) of the drug molecule and the respective acceptor atoms (N7 and N3 of purines and O2 of pyrimidines) and donors (N2 of guanine, N4 of cytosine, and N6 to adenine) of the DNA. The DNA backbone oxygen atoms, including the negatively charged phosphate oxygens, are less frequently involved in the direct hydrogen-bonding interactions. Third, water molecules or hydrated metal ions can further stabilize the complex by bridging functional groups. Fourth, van der Waals interactions between hydrophobic groups (e.g., methyls) and DNA in the grooves are important. We can now use these general rules to predict and model the binding of different natural (e.g., aclacinomycin, steffimycin) and synthetic (e.g., mitoxanthrones) compounds to DNA. More structural analyses like this would enable us to contemplate designing new compounds with different DNA binding properties.

Acknowledgments

We thank Professor J. H. van Boom and Dr. G. A. van der Marel of Leiden State University for their continuous encouragement and support. We also thank Dr. P. Aristoff of Upjohn Co. for providing nogalamycin and its derivatives.

Supplementary Material Available

Three tables: Table 1S, listing the crystallographic and stereochemical refinement parameters; Table 2S, listing the torsion angles and some helical parameters of the U-58872–d[m^{5}CGT(pS)Am^{5}CG] complex; and Table 3S, listing the complete helical parameters of both DNA–nogalamycin complexes (3 pages). Ordering information is given on any current masthead page.

References

Adams, W. J., McGovern, J. P., Dalm, E. A., Brewer, J. E., & Hosley, J. D. (1989) *Cancer Res. 49*, 6328–6336.

Arnott, S., Chandrasekaran, R., Hall, I. H., Puigjaner, L. C., Walker, J. K., & Wang, M. (1982) *Cold Spring Harbor Symp. Quant. Biol. 47*, 53–65.

Arora, S. K. (1983) *J. Am. Chem. Soc. 105*, 1328–1332.
Atwell, G. J., Cain, B. F., Baguley, B. C., Finlay, G. J., & Denny, W. A. (1984) *J. Med. Chem. 27*, 1481–1485.
Bhuyan, B. K., & Smith, C. G. (1965) *Proc. Natl. Acad. Sci. U.S.A. 54*, 566–572.
Bhuyan, B. K., & Reusser, R. (1970) *Cancer Res. 30*, 984–989.
Coll, M., Frederick, C. A., Wang, A. H.-J., & Rich, A. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8385–8389.
Collier, D. A., Neidle, S., & Brown, J. R. (1984) *Biochem. Pharmacol. 33*, 2877–2880.
Crooke, S. T., & Reich, S. D., Eds. (1980) *Anthracyclines*, Academic Press, New York.
Denny, W. A. (1989) *Anti-Cancer Drug Des. 4*, 241–263.
Dickerson, R. E., Bansal, M., Calladine, C. R., Diekman, S., Hunter, W. N., Kennard, O., von Kitzing, E., Lavery, R., Nelson, H. C. M., Olson, W. K., Saenger, W., Shakked, Z., Sklenar, H., Soumpasis, D. M., Tung, C.-S., Wang, A. H.-J., & Zhurkin, V. B. (1989) *Nucleic Acids Res. 15*, 247–265.
Drew, H. R., & Dickerson, R. E. (1981) *J. Mol. Biol. 152*, 723–736.
Fox, K. R., & Waring, M. J. (1986) *Biochemistry 25*, 4349–4356.
Fox, K. R., Brassett, C., & Waring, M. J. (1985) *Biochim. Biophys. Acta 840*, 383–392.
Frederick, C. A., Williams, L. D., Ughetto, G., van der Marel, G. A., van Boom, J. H., Rich, A., & Wang, A. H.-J. (1990) *Biochemistry 29*, 2538–2549.
Hendrickson, W. A., & Konnert, J. (1979) in *Biomolecular Structure, Conformation, Function and Evolution* (Srinvasan, R., Ed.) pp 43–57, Pergamon, Oxford.
Kersten, W., Kersten, H., & Szybalski, W. (1966) *Biochemistry 5*, 236–244.
Kind, R., Hutter, K., Zeeck, A., Schmidt-Base, K., & Egert, E. (1989) *J. Antibiot. 42*, 7–13.
Konishi, M., Ohkuma, H., Tsuno, T., Oki, T., VanDuyne, G. D., & Clardy, J. (1990) *J. Am. Chem. Soc. 112*, 3715–3716.
Kopka, M. L., Yoon, C., Goodsell, D., Pjura, P., & Dickerson, R. E. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 1376–1380.
Liaw, Y.-C., Gao, Y.-G., Robinson, H., van der Marel, G. A., van Boom, J. H., & Wang, A. H.-J. (1989) *Biochemistry 28*, 9913–9918.
Marugg, J. E., de Vroom, E., Dreef, C. E., Tromp, M., van der Marel, G. A., & van Boom, J. H. (1986) *Nucleic Acids Res. 14*, 2171–2185.
Moore, M. H., Hunter, W. N., Langlois d'Estaintot, B., & Kennard, O. (1989) *J. Mol. Biol. 206*, 693–705.
Rabinovich, D., & Shakked, Z. (1984) *Acta Crystallogr. A40*, 195–200.
Robinson, H., Liaw, Y.-C., van der Marel, G. A., van Boom, J. H., & Wang, A. H.-J. (1990) *Nucleic Acids Res. 18*, 4851–4858.
Tsai, C.-C., Jain, S. C., & Sobell, H. M. (1977) *J. Mol. Biol. 114*, 301–315.
Ughetto, G., Wang, A. H.-J.; Quigley, G. J., & Rich, A. (1985) *Nucleic Acids Res. 13*, 2305–2323.
Wakelin, L. P. G. (1986) *Med. Res. Rev. 6*, 275–340.
Wang, A. H.-J. (1987) in *Nucleic Acids and Molecular Biology* (Eckstein, F., & Lilley, D. M.) Vol. 1, pp 32–54, Springer, Berlin.
Wang, A. H.-J., & Teng, M. (1990) in *Crystallographic and Modeling Methods in Molecular Design* (Bugg, C. E., & Ealick, S. E., Eds.) pp 123–150, Springer-Verlag, New York.
Wang, A. H.-J., Nathan, J., van der Marel, G. A., van Boom, J. H. & Rich, A. (1978) *Nature 276*, 471–474.
Wang, A. H.-J., Ughetto, G., Quigley, G. J., Hakoshima, T., van der Marel, G. A., van Boom, J. H., & Rich, A. (1984) *Science 225*, 1115–1121.
Wang, A. H.-J.; Ughetto, G., Quigley, G. J., & Rich, A. (1987) *Biochemistry 26*, 1152–1163.
Wang, A. H.-J., Liaw, Y.-C., Robinson, H., & Gao, Y.-G. (1990) in *23rd Jerusalem Symposium in Quantum Chemistry and Biochemistry* (Pullman, B., & Jortner, J., Eds.). Kluwer Academic Publishers, Dordrecht, Holland (in press).
Wang, J. C. (1974) *J. Mol. Biol. 89*, 783–801.
Wiley, P. F. (1979) *J. Nat. Prod. 42*, 569–582.
Wiley, P. F., Elrod, D. W., Houser, D. J., & Richard, F. A. (1982) *J. Med. Chemother. 25*, 560–567.
Williams, L. D., Egli, M., Gao, Q., Bash, P., van der Marel, G. A., van Boom, J. H., Rich, A., & Frederick, C. A. (1990a) *Proc. Natl. Acad. Sci. U.S.A. 87*, 2225–2229.
Williams, L. D., Egli, M., Ughetto, G., van der Marel, G. A., van Boom, J. H., Rich, A., Wang, A. H.-J., & Frederick, C. A. (1990b) *J. Mol. Biol.* (in press).
Zimmerman, S. C., Lamberson, C. R., Cory, M., & Fairley, T. A. (1989) *J. Am. Chem. Soc. 111*, 6805–6809.