DNA−Nogalamycin Interactions: The Crystal Structure of d(TGATCA) Complexed with Nogalamycin[†,‡]

Camilla K. Smith, Gideon J. Davies, Eleanor J. Dodson, and Madeleine H. Moore*

York Protein Structure Research Group, Department of Chemistry, University of York, Heslington, York YO1 5DD, England

Received August 9, 1994; Revised Manuscript Received October 25, 1994[⊗]

ABSTRACT: The structure of the self-complementary deoxyoligonucleotide $d^{5'}$(TGATCA) complexed with nogalamycin, an antitumor anthracycline, has been solved to 1.8 Å resolution using X-ray crystallographic methods. The technique of single isomorphous replacement, utilizing the anomalous signal of bromine in derivative data collected at three different wavelengths, Cu Kα, Mo Kα, and 0.91 Å synchrotron radiation, was used. The complex crystallized in space group $P4_12_12$ with unit cell dimensions a = 37.2 Å and c = 70.1 Å. The final structure including 116 water molecules has an overall R factor of 19.5% for the 4767 reflections with $F \geq 1\sigma F$ in the resolution range 10.0−1.8 Å. One nogalamycin molecule intercalates between each of the $d^{5'}$(TpG) steps at both ends of a distorted B DNA double helix. This structure provides the first three-dimensional picture of nogalamycin bound to the triplet sequence $d^{5'}$(TGA), one of its favorable natural binding sites. The drug exhibits a strict requirement for binding to the 3′ side of a pyrimidine and the 5′ side of a purine. Nogalamycin has bulky sugar groups at either end of a planar aglycon chromophore; therefore, in order for intercalation to occur, the DNA must either transiently open or flex along the helix axis to allow insertion of the chromophore between the base pairs. Conformational change in nogalamycin is observed in the drug−DNA complex with respect to free nogalamycin. Nogalamycin binding to DNA induces severe deformation to the intercalation site base pairs. In comparison to previously reported anthracycline−DNA structures significant differences in base-pair geometry, drug hydrogen−bonding patterns, and the extent of hydration are observed. The position of the drug in this complex is stabilized by a number of nonbonded forces including van der Waals interactions and extensive direct and solvent-mediated hydrogen bonds to the DNA duplex.

Among DNA intercalating compounds, anthracycline antibiotics are the most effective and widely used chemotherapeutic agents, yet their mechanism of action is only partially understood. Their major effect is believed to be inhibition of transcription and replication. This may be achieved in two ways. First, binding to promoter regions in DNA inhibits the binding of polymerases to initiate transcription. Second, inhibition of the DNA unwinding enzyme topoisomerase II, either by forming a stable DNA−drug−enzyme complex or by preventing the binding of the enzyme to DNA (Fisher & Aristoff, 1988). The biochemical effects of anthracyclines are highly dependent upon their chemical structure.

Nogalamycin is a naturally occurring anthracycline antibiotic produced by *Streptomyces nogalator* var. *nogalator*. In addition to its antibacterial effects, it has been found to have antitumor activity in mice (Bhuyan & Dietz, 1965; Bhuyan & Reusser, 1970; Li *et al.*, 1979). Nogalamycin is a DNA intercalator which selectively inhibits DNA-directed RNA synthesis *in vivo* and *in vitro* (Ennis, 1981).

As a member of the anthracycline family of molecules, nogalamycin possesses a planar aglycon chromophore (rings A−D in Figure 1). In contrast to certain members of this group, including daunomycin and adriamycin, the sugar substituent at position 7 on the nonaromatic ring A of the chromophore is the more hydrophobic noncharged nogalose in nogalamycin. The structure of nogalamycin is unique in that a bicyclic aminoglucose moiety [α-D-3,6−dideoxy-3-(dimethylamino)glucose] is substituted as a fused ring system at C1 and C2 of ring D. Other anthracyclines lack a bulky substituent on ring D which poses a mechanistic problem for this drug. How does a molecule, essentially the shape

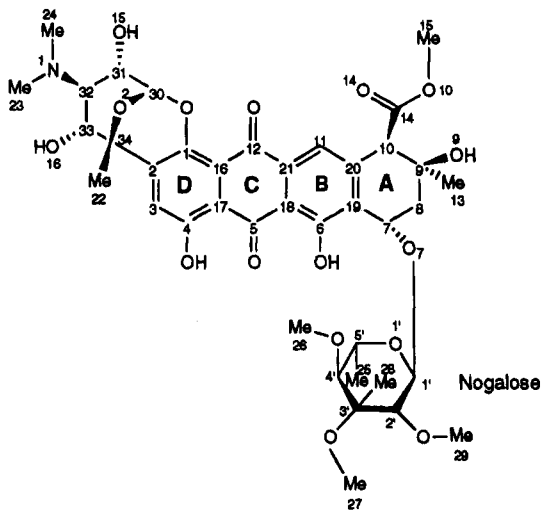


FIGURE 1: Molecular formula and nomenclature of nogalamycin $C_{39}H_{49}NO_{16}$. The aglycon chromophore comprises four fused rings, A−D. The nogalose sugar is attached to the only unsaturated ring, A, and the aminoglucose sugar is fused to ring D.

[†] This work was funded by the United Kingdom Science and Engineering Research Council and the Yorkshire Cancer Research Campaign.

[‡] Coordinates have been deposited in the Brookhaven Protein Data Bank (file name 182d.pdb).

* Author to whom correspondence should be addressed.

[⊗] Abstract published in *Advance ACS Abstracts,* December 1, 1994.

0006-2960/95/0434-415$09.00/0

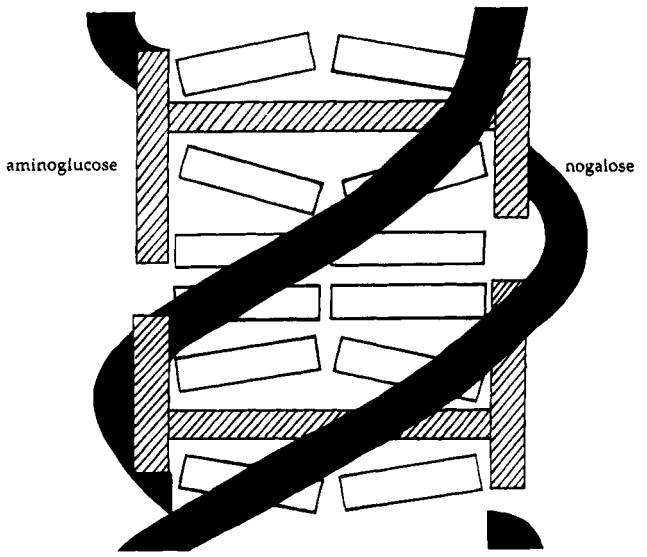


FIGURE 2: Schematic mode of binding of two nogalamycin molecules in a hexamer DNA. This illustrates the mechanistic problem of intercalating this bulky dumbell-shaped molecule between base pairs of DNA and the effect it has on the DNA structure in the complex formed. The intercalation site base pairs cup to accommodate the drug as represented by the buckling of the base pairs in this diagram.

of a "dumbbell" with a minimum bell-width of 5.5 Å, intercalate between DNA base pairs (3.4 Å apart in normal B DNA) with one sugar in the minor groove and one in the major groove, eventually spanning three base pairs (Figure 2)? Two suggestions have been made. The first is that transient melting of the DNA helix may be required to allow the drug to nestle into place. Re-formation of the helix then occurs around the drug (Collier *et al.*, 1984). The second binding hypothesis involves elongation of the DNA along its helix axis with associated unstacking and buckling of the base pairs without disruption of base-pairing hydrogen bonds (Williams *et al.*, 1990). This would require the bulky sugar substituents of the drug to flip from axial to equatorial conformations to facilitate intercalation. Major conformational changes of both the drug and the DNA would be required for the latter theory.

Single-crystal X-ray diffraction studies have shown that, in anthracycline–DNA complex structures, the long axis of the intercalating chromophore is perpendicular to the long axis of the DNA base pairs (Quigley *et al.*, 1980). For nogalamycin, there is no steric preference for sugar accommodation in either groove (Collier *et al.*, 1984). In all reported structures of oligonucleotide–nogalamycin complexes, including d$^{5'}$(TGATCA)–nogalamycin, the aminoglucose lies in the major groove and the nogalose sugar lies in the minor groove (Searle *et al.*, 1988; Liaw *et al.*, 1989; Zhang & Patel, 1990; Robinson *et al.*, 1990; Williams *et al.*, 1990; Gao *et al.*, 1990; Egli *et al.*, 1991). A combination of nonbonded forces including van der Waals interactions as well as direct and solvent-mediated hydrogen bonds between nogalamycin and DNA stabilizes the DNA–drug complex.

Experiments on natural long DNA in solution have demonstrated that nogalamycin has DNA sequence binding preferences. Footprinting studies have shown that the drug exhibits a sequence preference for d$^{5'}$(TpG) pyrimidine–purine steps within a DNA duplex (Fox & Waring, 1986). The X-ray crystal structures of d$^{5'}$(^{m5}CGT(pS)A^{m5}CG), where cytosine is methylated at the C5 position, and d$^{5'}$(CGT(pS)ACG) complexed with nogalamycin have been determined (Liaw *et al.*, 1989; Williams *et al.*, 1990; Gao *et al.*, 1990; Egli *et al.*, 1991). In these structures the drug intercalates between the d$^{5'}$(^{m5}CpG) and d$^{5'}$(CpG) steps, respectively. NMR studies on nogalamycin complexed with d$^{5'}$(AGCATGCT), d$^{5'}$(GCATGC), and d$^{5'}$(CGTACG) revealed the drug intercalated at both the d$^{5'}$(CpA), d$^{5'}$(TpG), and d$^{5'}$(CpG) pyrimidine–purine steps (Zhang & Patel, 1990; Searle *et al.*, 1988; Robinson *et al.*, 1990). In summary, previous work has shown that intercalation of nogalamycin is favored at the 5′ side of a purine and at the 3′ side of a pyrimidine. To add to these findings and to initiate examination on the binding of nogalamycin to different triplet DNA sequences, we have carried out the structure determination of nogalamycin complexed with the deoxyoligonucleotide hexamer d$^{5'}$(TGATCA). The intention was that the structure of the nogalamycin–d$^{5'}$(TGATCA) complex, subsequently referred to as TGA–NOG,[1] would reveal in atomic detail the interactions the drug makes with one of its preferred natural substrate binding sites.

Cytotoxicity is related to the binding affinity of the drug to its target DNA in the cell, nuclear DNA. Of all the anthracyclines, nogalamycin has the slowest association and dissociation rates yet determined, and this drug exhibits higher levels of cytotoxicity than other members of the anthracycline family (Fox *et al.*, 1985). It also shows high levels of cardiotoxicity, rendering nogalamycin of little use clinically. The long-term aim is to try to ascertain the importance of different functional groups on anthracyclines and the biological implications of their interactions with DNA, in order to design derivatives with increased cytotoxicity and decreased cardiotoxicity. Structure determinations of oligonucleotide–drug complexes are steps toward this goal.

MATERIALS AND METHODS

The hexamer self-complementary deoxyoligonucleotide d$^{5'}$(TGATCA) was synthesized, on a 10 μM scale leaving the final trityl protecting group bound, using an Applied Biosystems 392 DNA synthesizer, which employs phosphoramidite chemistry (Mcbride & Caruthers, 1983). The crude product was purified by a combination of reverse-phase and strong anion-exchange high-pressure liquid chromatography. The pure oligonucleotide solution was desalted using a Sephadex column and lyophilized. Approximately 8–10 mg of pure lyophilized product could be obtained as a result of this procedure. The bromine derivative d$^{5'}$(TGA^{Br}UCA) was synthesized by replacing the fourth thymine residue in d$^{5'}$(TGATCA) with 5-bromouracil.

Nogalamycin was purchased from Sigma Chemicals Ltd. This proved to be insoluble in water; therefore, a 3.8 mM stock solution was made in 50% methanol–water. Three milligrams of oligonucleotide was added to 870 μL of stock nogalamycin solution. The mixture was then lyophilized and reconstituted in 470 μL of water and 400 μL of 30 mM sodium cacodylate buffer (pH 6.5).

[1] Abbreviations: TGA–NOG, nogalamycin–d$^{5'}$(TGATCA) complex; TGABr–NOG, nogalamycin–d$^{5'}$(TGA^{Br}UCA) complex; ^{m5}CG-TpS–NOG, nogalamycin–d$^{5'}$(^{m5}CGT(pS)A^{m5}CG)$_2$ complex; TGA–4′EPIADR, 4-epiadriamycin–d$^{5'}$(TGATCA) complex.

Table 1: Crystal and Diffraction Data for DNA−Nogalamycin Complexes

	TGA−NOG	TGABr−NOG	TGABr−NOG	TGABr−NOG
crystal data	native data	derivative 1	derivative 2	derivative 3
crystal size, mm	0.2 × 0.2 × 0.4	0.15 × 0.2 × 0.4	0.15 × 0.15 × 0.6	0.15 × 0.15 × 0.6
growth temp, °C	4	4	4	4
solvent content, %	48	48	48	48
pH	6.5	6.5	6.5	6.5
initial crystallization concentrations	0.6 mM oligonucleotide, 2.3 mM nogalamycin, 10 mM sodium cacodylate, pH 6.5, 8 mM magnesium chloride, 10−20% 2-methyl-2,4-pentanediol, 0.6 mM spermine	as for native crystals	as for native crystals	as for native crystals
unit cell dimensions, Å	$a = 37.20$, $b = 37.20$, $c = 70.10$	$a = 37.23$, $b = 37.23$, $c = 70.20$	$a = 37.22$, $b = 37.22$, $c = 70.54$	$a = 37.20$, $b = 37.20$, $c = 70.10$
crystal system	tetragonal	tetragonal	tetragonal	tetragonal
space group	$P4_12_12$	$P4_12_12$	$P4_12_12$	$P4_12_12$
vol/AU, Å^3	12125	12163	12163	12125
vol/base pair, Å^3	1515	1520	1520	1515
diffraction data				
X-ray source	EMBL beamline X31, Hamburg	Rigaku RU200HB Cu rotating anode	Rigaku RU200HB Mo rotating anode	EMBL beamline X31, Hamburg
wavelength, Å	0.910	Cu Kα, 1.5418	Mo Kα, 0.7107	0.910
apparatus	Hendrix−Lentfer image plate	Rigaku R-axis IIC image plate	Rigaku R-axis IIC image plate	Hendrix−Lentfer image plate
temp, °C	4	4	4	4
resolution, Å	1.80	2.00	2.00	2.56
no. of unique reflections	4953	3668	3681	1796
no. of observed reflections	4788	3663	3664	1768
completeness, %	81	100	99	98
multiplicity	4.4	4.4	3.4	3.3
$R_{merge} = \sum\|I - \langle I\rangle\|/\sum\|I\|$	0.057	0.035	0.067	0.081
$R_{deriv} = \sum\|F_{deriv} - F_{nat}\|/\sum\|F_{deriv}\|$		0.222	0.231	0.258

Initial concentrations of the constituents in the crystallization solution were as described in Table 1. Crystallization was carried out in sitting drops containing the crystallization mixture which was equilibrated with 100% MPD at 4 °C, allowing crystals to be grown by vapor diffusion. Bright orange, well-diffracting crystals grew within a few weeks. The crystals used for data collection were mounted and sealed in 0.7-mm glass capillaries (Figure 3). The crystals are in space group $P4_12_12$ with unit cell dimensions $a = 37.2$ Å and $c = 70.1$ Å.

All crystal and data collection information is listed in Table 1. One single crystal of TGA−NOG (Figure 3) was used to collect a native data set to 1.8 Å resolution with an image plate on the EMBL beamline X31 at the Hamburg outstation. Crystals seemed relatively stable to data collection. Three derivative data sets were collected at three different wavelengths, using three different crystals of nogalamycin complexed with d$^{5'}$(TGA^{Br}UCA), referred to as TGABr−NOG crystals. Two of these data sets, one with 0.7107 Å (Mo Kα) radiation and the other with 1.5418 Å (Cu Kα) radiation, were collected on an R-axis IIC image plate using rotating anode sources. The third data set was collected at EMBL beamline X31 on an image plate using synchrotron radiation at a wavelength of 0.91 Å, which lies immediately below the K absorption edge of bromine; this served to maximize the f'' component of the anomalous signal caused by bromine in the complex. All data were collected at 4 °C by positioning the capillary in a stream of cold air. All data processing was carried out using DENZO (Z. Otwinowski, Dallas), and data reduction was performed by using programs included in the CCP4 suite.

The two bromine atom positions were located in difference Patterson syntheses. These positions were refined, and an initial set of phases was calculated by using MLPHARE in the CCP4 program suite. All heavy atom refinement statistics are listed in Table 2. Solvent flattening and phase improvement were carried out using SQUASH (Cowtan &

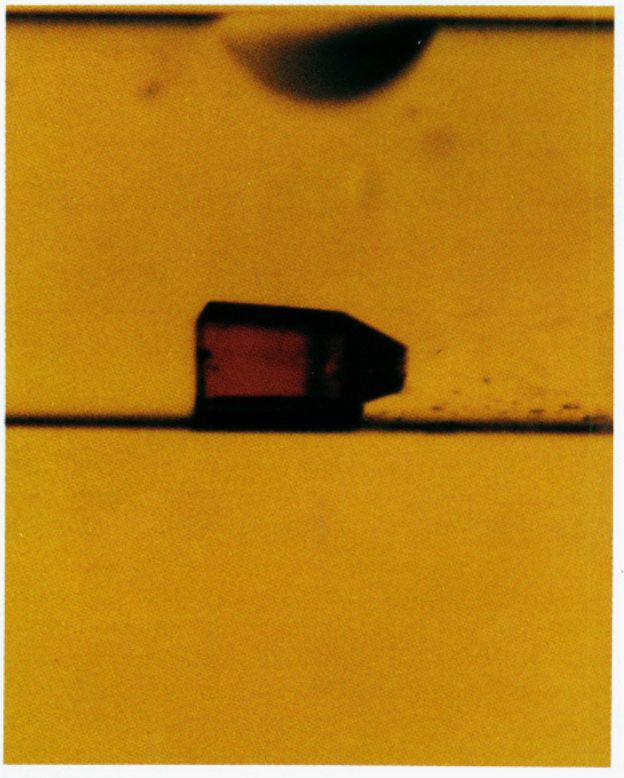


FIGURE 3: Crystal of d$^{5'}$(TGATCA)−nogalamycin which was mounted in a glass capillary and used for data collection at EMBL in Hamburg (more details are given in Table 1).

Table 2: Phasing in MLPHARE

derivative	TGA−NOG	TGABr−NOG	TGABr−NOG	TGABr−NOG
data	native data	derivative 1	derivative 2	derivative 3
wavelength, Å	0.91	Cu Kα, 1.5418	Mo Kα, 0.7107	0.91 (Br K edge)
resolution, Å	1.8	6.3−1.8	6.3−1.8	6.3−2.3
fractional coordinates [Br(1), Br(2)]				
x		0.139, 0.605	0.139, 0.605	0.139, 0.604
y		0.839, 0.200	0.839, 0.202	0.844, 0.205
z		0.028, 0.070	0.027, 0.070	0.025, 0.071
real occupancy		0.582, 0.605	0.584, 0.603	0.504, 0.522
anomalous occupancy		0.424, 0.495	1.230, 1.232	1.536, 1.768
isomorphous phasing				
no. of reflections used for phasing (acentric, centric)		2665, 902	2662, 901	1247, 551
phasing power[a] (acentric, centric)		4.0, 3.2	3.7, 2.9	1.9, 1.4
R_{Cullis}[b] (acentric, centric)		0.36, 0.33	0.39, 0.36	0.62, 0.61
anomalous phasing				
no. of reflections		1563	2633	1230
anomalous difference (observed, calculated)		4.9, 1.0	9.4, 3.0	12.3, 5.9
R_{Cullis}[b] (anomalous)		1.0	0.9	0.9
overall phasing				
no. of reflections (acentric, centric)	2665, 903			
figure of merit[c] (acentric, centric)	0.607, 0.812			

[a] Phasing power = $\sum|F_{hc}|/\sum[||F_{po}| \exp(i\theta) + F_{hc}| - |F_{pho}|]$, i.e., heavy atom structure factor amplitude/residual lack of closure. $|F_{hc}|$ is the calculated diffraction amplitude of the heavy atom; $|F_{po}|$ and $|F_{pho}|$ are the observed amplitudes for the native DNA and heavy atom derivatives, respectively; θ is the calculated phase. Summations are carried out over all observations. [b] $R_{Cullis} = \sum||F_{pho} \pm F_{po}| - F_{hc}|/\sum|F_{pho} - F_{po}|$. [c] Overall figure of merit = $\int P(\theta) \exp(i\theta)\, d\theta / \int P(\theta)\, d\theta$.

Main, 1993). The initial SIRAS map clearly indicated the position of the DNA and nogalamycins in the complex. For the DNA strands, model coordinates of the d5′(TGATCA)−epiadriamycin complex (Langlois d'Estaintot *et al.*, 1992) were taken from the Brookhaven Protein Data Bank (Bernstein *et al.*, 1977). To build the drug into the map, coordinates from the small molecule structure of nogalamycin were extracted from the Cambridge Structural Database (Arora, 1983). By combination of these two sets of coordinates, an initial model was built into the map using FRODO (Jones, 1978).

Rigid body refinement was carried out with these initial coordinates using the program AMORE (Navaza, 1994). Data between 10 and 3.5 Å were initially used; then the outer limit of this was gradually extended to 1.8 Å. The initial model which had been built into the SIRAS map was successively split into smaller and smaller rigid bodies. First, the drug molecules were treated as separate rigid bodies, and the duplex DNA was treated as one. The DNA coordinates were then split into two double-stranded triplets, d(TGA) and d(TCA), followed by three doublets, d(TG), d(AT), and d(CA), and then into individual base pairs. This refinement resulted in an overall reduction in *R* factor from 36% to 33% for all data in the 10−1.8 Å resolution range.

Idealized models for each of the four nucleotide residues and the drug were generated using the Cambridge Structural Database (Cambridge Crystallographic Data Centre, Cambridge, U.K.) and Quanta (Molecular Simulations, Burlington, MA). These models were used to prepare restraints for Hendrickson−Konnert restrained least squares refinement as implemented in PROLSQ (CCP4, 1994). A comparison of PROLSQ refinement, starting with the initial model coordinates and coordinates after rigid body refinement, was carried out. Preliminary AMORE rigid body refinement did not appear to be essential. The initial *R* factor in PROLSQ without previous rigid body refinement was 36.7% whereas with rigid body refinement the *R* factor was 35.4%, not significantly less. During the first 10 cycles of PROLSQ with and without prior rigid body refinement, the *R* factor quickly dropped to the same value in both refinements. The overall and free *R* factors were monitored throughout refinement. The free *R* factor was calculated by flagging a random 10% of reflections for omission from refinement and use in free *R* calculation only. Positional refinement of the model yielded an overall *R* value of 30.3% (free *R* = 38.9%) which dropped to 30.1% (free *R* = 37.4%) with *B* value refinement included. Addition of solvent produced further reductions in both overall and free *R* values. Throughout refinement the free *R* factor remained 10−12% higher than the overall *R* factor. Water molecules were modeled into the density using FRODO. It was initially possible to locate 105 waters with confidence. The final *B* values for the waters ranged from 20 to 92 Å^2, even though some of the waters had high *B* values, probably due to partial occupancy. However, the density for these waters in $2F_o - F_c$ maps was quite clear, so these waters were retained; no attempt to refine occupancy was made. The final PROLSQ refinement resulted in an overall *R* factor of 19.5% including 468 non-hydrogen atoms, 116 of these were waters, and all 4767 reflections between 10 and 1.8 Å. The final refinement parameters and quality of the final structure are given in Table 3.

RESULTS AND DISCUSSION

Global Conformation and Crystal Packing. The asymmetric unit contains a hexamer DNA duplex, two nogalamycins, and 116 water molecules (Figure 4). No metal ions could be located in the density. The two DNA strands form a distorted right-handed B DNA-like duplex with Watson−Crick base pairs. A nogalamycin is intercalated at each d5′(TpG) step. Nucleosides are labeled in the 5′ to 3′ direction from T1 to A6 for strand 1 and T7 to A12 for strand 2. Nogalamycins are labeled NOG1 at the d5′(T7pG8) intercalation site, equivalent to the d5′(C5pA6) step, and NOG2 at d5′(T1pG2), also corresponding to the d5′(C11pA12) step, respectively. Atomic nomenclature used for nogalamycin

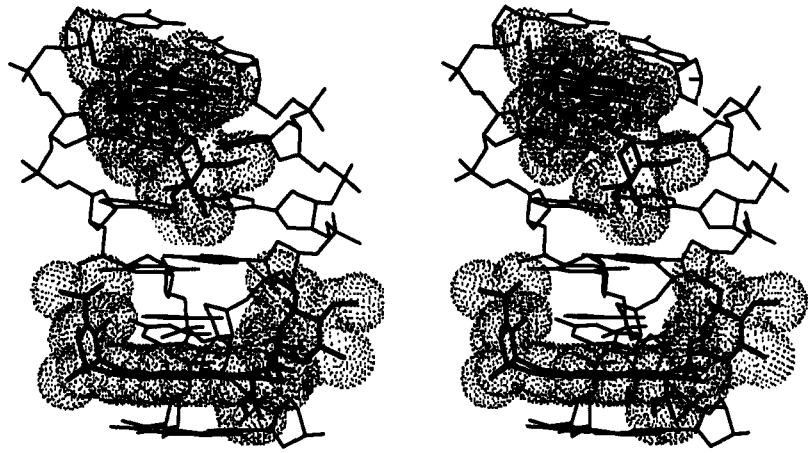


FIGURE 4: Stereoview stick drawing of the d$^{5'}$(TGATCA) duplex complexed with two molecules of nogalamycin. A dot surface represents the van der Waals surface of each nogalamycin.

Table 3: Final Refinement Parameters of TGA−NOG in PROLSQ

no. of atoms (non-hydrogen)	468
no. of water molecules	116
range of spacings, Å	10−18
no. of reflections	4767 ($F \geq 1\sigma F$)
overall R factor = $\Sigma\|F_o - F_c\|/\Sigma\|F_o\|$	0.195
rms deviations from ideal geometry of final model distances	
bonds (Å)	0.020 (0.020)[a]
angles (deg)	0.076 (0.050)
chiral volumes (Å^3)	0.037 (0.060)
planes (Å)	0.025 (0.020)
range of B values (Å^2)	
bases	15.0−21.0
sugars	24.0−30.0
phosphates	26.0−42.0
nogalamycin	12.0−35.0
water	20.0−90.0

[a] Target values are in parentheses.

is consistent with previous work (Liaw *et al.*, 1989; Williams *et al.*, 1990) and is given in Figure 1.

The TGA−NOG complexes pack end to end with a distance between complexes of 3.6 Å, which is similar to the normal rise between base pairs in B DNA (3.4 Å). Two overall helical axes are generated by these complexes in this packing arrangement. Figure 5a illustrates the perpendicular orientation of the two global helical axes which lie almost directly along both diagonals of the *xy* plane separated by a translation of one helical diameter along the *z* axis. A native Patterson map shows the vectors generated by bases stacking along the two **a**•**b** diagonals (Figure 5b). The self-rotation function revealed the position of the noncrystallographic 2-fold axis to be 50° to the **a**•**b** diagonal (Figure 5c). There are no direct hydrogen bonds between complexes in the same continuous helix or to complexes in adjacent stacks. Solvent-mediated interactions link not only helices in the same stack but also the helices in adjacent columns.

DNA Conformation. The sugar−phosphate backbone and glycosidic torsion angles as well as furanose conformations of the DNA in the TGA−NOG structure are listed and defined in Table 4. The same parameters of d$^{5'}$(^{m5}CGT(pS)-A^{m5}CG) complexed with nogalamycin (Gao *et al.*, 1990), ^{m5}CGTpS−NOG, and the same sequence, d$^{5'}$(TGATCA), used for this structure bound to the anthracycline antibiotic 4′-epiadriamycin (Langlois d'Estaintot *et al.*, 1992), TGA−4′EPIADR, as well as average values for B DNA (Dickerson & Drew, 1981; Leonard & Hunter, 1993) are listed in Table 4. The average torsion angles for most of the TGA−NOG structures fall into the range observed for B DNA. Modifications are confined to the locality of the drugs at the ends of the double helix. Overall, these torsion angles resemble more closely those of d$^{5'}$(TGATCA) complexed with 4′-epiadriamycin than those of d(^{m5}CGT(pS)A^{m5}CG) bound to nogalamycin. Likewise, the furanose conformations in TGA−NOG are B DNA-like and are more close to those found in TGA−4′EPIADR than in ^{m5}CGTpS−NOG.

Table 5 lists the geometrical properties of base pairs and base-pair steps in the same three structures. These figures illustrate that d$^{5'}$(TGATCA) is more distorted from a B DNA-like helix (Kennard & Hunter, 1989) with nogalamycin bound than with 4′-epiadriamycin. The helical rise at the intercalation site base pairs, steps 1 and 5, is 5.8 and 5.6 Å which is similar to what has been observed in other nogalamycin−DNA complexes and slightly larger than in complexes of the simpler anthracyclines (Nunn *et al.*, 1991; Langlois d'Estaintot *et al.*, 1992; Leonard *et al.*, 1992, 1993). In the nogalamycin complexes both steps 2 and 3 have a larger than normal rise for B DNA independent of DNA sequence. As previously observed in all anthracycline−DNA complexes, most of the DNA unwinding is confined to one intercalation site base-pair step per drug; steps 2 and 4 are each unwound by approximately 7°. There is severe buckling of the intercalation site base pairs, creating larger cup values for the TGA−NOG structure than observed in other nogalamycin−DNA complexes. This may partially be due to two repulsive forces felt. First, the close proximity of the O5 carbonyl substituent of ring C in nogalamycin and the N3 of the intercalation site cytosine is unfavorable, and second, the O10 of the acetic methyl ester group on ring A and the O2 carbonyl of the intercalation site thymine are situated close enough to have a repulsive effect. The markedly different values for roll and slide at the central ApT step of the TGA−NOG complex serve to optimize intrastrand stacking of six-membered rings.

In spite of some significant differences in base-pair geometries, the overall rms difference between the sugar−phosphate backbones of the TGA−NOG and ^{m5}CGTpS−NOG complexes is 1.2 Å. The greatest differences are at the terminal base pairs.

The snug fit of the hydrophobic nogalose sugar in the minor groove serves to maximize van der Waals interactions

and widen the groove to 14−15 Å from 11.5−13 Å as found in normal B DNA.

Nogalamycin Conformation and DNA Interactions. The overall intercalation geometry in the d$^{5'}$(TGATCA)−nogalamycin complex is similar to that observed in both d$^{5'}$(^{m5}C-GT(pS)A^{m5}CG)− and d$^{5'}$(CGT(pS)ACG)−nogalamycin complexes. Nogalamycin intercalates with its nogalose sugar in the minor groove and its aminoglucose sugar in the major groove, and both sugars point toward the central AT base pairs. In spite of the overall similarity of intercalation configuration in all these complexes, the interactions anchoring the drug in each DNA groove differ significantly. The orientation and conformation of the aminoglucose sugar moiety in the major groove are more rigid since it is involved in at least one strong direct hydrogen bond to the DNA. In contrast, the position of the nogalose sugar in the minor groove appears to be completely determined by steric and van der Waals forces.

The sugar groups of both nogalamycin residues in TGA−NOG and ^{m5}CGTpS−NOG point toward the center of the complex. It is possible that this orientation of the drug in these complexes is a consequence of oligonucleotide sequence and perhaps also length. If the two sugars of each intercalated nogalamycin pointed away from each other toward the ends of the complex, they would overhang the ends of the DNA and perhaps interfere with the crystal packing. In the NMR structures of d$^{5'}$(AGCATGCT) and d$^{5'}$(GCATGC) complexed to nogalamycin, one drug is intercalated between each of the two pyrimidine−purine steps in both sequences. In contrast to the X-ray crystallographic structures, in these NMR structures the sugars of each intercalated nogalamycin point in opposite directions outward toward the ends of the DNA. However, this is probably due to steric effects which have arisen because of the two adjacent intercalation sites in the DNA sequences chosen for the NMR work. Possibly in longer DNA sequences, with intercalation sites situated further apart, both 180° related, head-on nogalamycin intercalations may be observed.

In the TGA−NOG complex each drug, NOG1 and NOG2, has different hydrogen-bonding patterns (Figure 6). Interactions are observed for at least one of the drugs with the intercalation site G•C base pair, the backbone of the intercalation site thymine residue, and the thymine base in the neighboring A•T base pair. More differences in both direct and solvent-mediated interactions occur in the major groove than in the minor groove.

The only good direct drug−DNA hydrogen bond exists in the major groove. The O15 hydroxyl group of the aminoglucose in NOG1 is 2.71 Å from the N7 of G8, and NOG2 makes a similar strong 2.54 Å hydrogen bond to the N7 of G2. The same interaction was found in both previous nogalamycin−DNA complex structures. NOG1 also interacts directly with the O5′ of T7, making a weak 3.29 Å hydrogen bond to the O1 atom in the aminoglucose moiety. In the previous structures a similar interaction was observed with the O15 substituent.

In the crystal structure of d$^{5'}$(^{m5}CGT(pS)A^{m5}CG) complexed with nogalamycin (Egli *et al.*, 1991), it was proposed that the NH_2 group of guanine is important for complex stability due to its participation in the formation of a minor-groove hydrogen bond to the O14 of nogalamycin. In TGA−NOG no such hydrogen bond is possible since an A•T base pair forms the terminal intercalation site base pair.

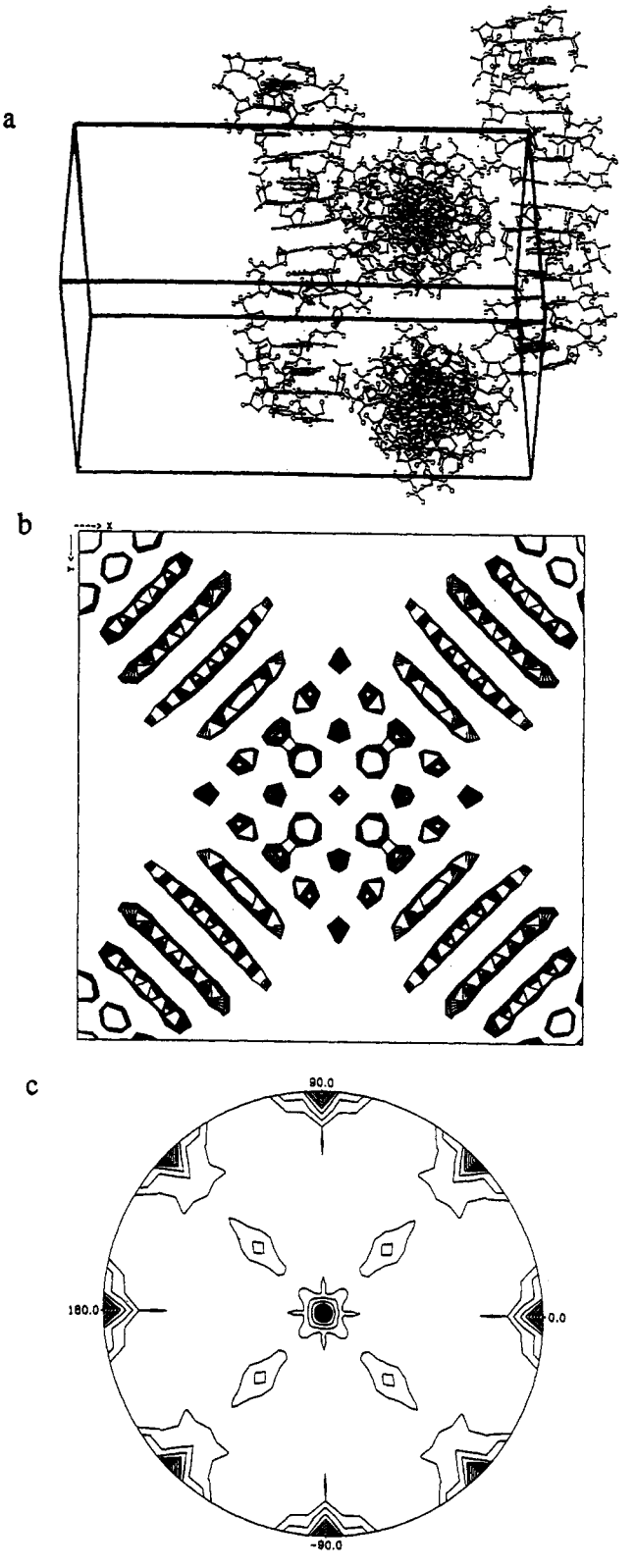
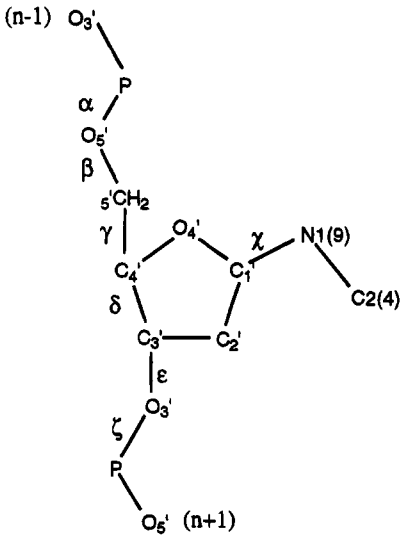


FIGURE 5: (a) Packing arrangement of d$^{5'}$(TGATCA)−nogalamycin complexes in the crystal lattice. A view almost along the unit cell **a•b** diagonal of eight symmetry-related complexes, drawn in ball and stick, with the unit cell superimposed. Stacking of complexes produces continuous helices which are almost at right angles to each other. (b) The $z = 0$ section of the native Patterson map illustrates the predominant 3.5 Å spaced vectors. This indicated the helices packed approximately at right angles to each other and along the **a•b** diagonals of the unit cell. (c) Stereogram of the $\kappa = 180°$ section of the self-rotation function giving the position of the noncrystallographic 2-fold axis at 50° to the **a•b** diagonal and 40° to the crystallographic 4-fold axis.

Table 4: Sugar−Phosphate Torsion Angles (deg)[a] and Furanose Conformations for the d(TGATCA)−Nogalamycin (TGA−NOG), d(^{m5}CGTpSA^{m5}CG)−Nogalamycin (^{m5}CGTpS−NOG), and d(TGATCA)−4′-Epiadriamycin (TGA−4′EPIADR) Complexes[b]



residue	α	β	γ	δ	ϵ	ζ	χ	P	pucker
TGA−NOG (this work)									
T1			67	100	212	286	213	93	O4′ endo
G2	287	195	55	142	227	185	259	144	C2′ endo
A3	315	135	43	125	194	255	260	138	C2′ endo
T4	290	160	60	91	194	283	232	87	O4′ endo
C5	285	176	49	144	261	176	269	152	C2′ endo
A6	293	157	43	145			255	161	C2′ endo
T7			346	146	232	304	190	251	C4′ endo
G8	298	169	51	118	193	227	256	114	C1′ exo
A9	295	157	63	125	188	255	253	134	C1′ exo
T10	301	169	53	103	188	270	230	96	O4′ endo
C11	297	166	50	137	257	172	271	150	C2′ endo
A12	308	146	42	146			261	183	C3′ exo
^{m5}CGTpS−NOG (Gao *et al.*, 1991)									
^{m5}C1			46	79	212	239	195	51	C4′ exo
G2	341	154	48	144	270	144	269	140	C1′ exo/C2′ endo
T3	285	132	71	104	190	265	218	118	C1′ exo
A4	282	166	64	87	191	289	232	128	C1′ exo
^{m5}C5	284	178	55	116	275	185	262	101	O4′ endo
G6	285	195	17	133			239	241	C4′ endo
^{m5}C7			298	139	267	298	171	−88	O4′ exo
G8	302	164	57	130	185	222	257	137	C1′ exo
T9	314	141	56	109	197	269	234	119	C1′ exo
A10	328	190	1	135	195	275	248	163	C2′ endo
^{m5}C11	145	170	177	124	270	188	229	188	C3′ exo
G12	316	128	71	118			234	144	C1′ exo
TGA−4′EPIADR (Langlois d'Estaintot *et al.*, 1992)									
T1			58	128	236	272	208	136	C1′ exo
G2	294	174	48	147	229	174	288	152	C2′ endo
A3	294	137	49	100	181	273	238	121	C1′ exo
T4	296	186	60	122	192	283	242	139	C1′ exo
C5	284	176	43	147	263	167	290	164	C2′ endo
A6	288	173	52	143			274	164	C2′ endo
av B DNA[c]	295	167	50	126	194	234	249	141	C2′ endo

[a] Definitions of torsion angles: $(n-1)$ refers to the preceding nucleotide; $(n+1)$ refers to the next nucleotide. The χ angle is defined by N1 and C2 in pyrimidines and by N9 and C4 in purines. P = pseudorotation phase angle. [b] All parameters were calculated using the NEWHEL94 program distributed by R. E. Dickerson. [c] Average values calculated using d(CGCGAATTCGCG) (Dickerson & Drew, 1981) and d(CGTAGATC-TACG) (Leonard & Hunter, 1993). For the γ average the first residue was not included as the flexibility of the terminal 5′-hydroxyl produces large variations for this torsion angle.

All solvent-mediated drug−DNA hydrogen bonds were observed in the major groove only. O16 hydroxyl on the aminoglucose of NOG2 is linked *via* one water molecule to N4 of C11. Interestingly, this was found to be a direct interaction in the previous complex structures. The O4 hydroxyl substituent in ring D of the NOG1 chromophore interacts *via* two water molecules with N4 of C5. One solvent-mediated hydrogen bond, between N1 on aminoglucose of NOG1 and O4 of neighboring base T4, would not be possible with a G·C neighboring base pair.

A least squares fit of the two nogalamycin molecules in the TGA−NOG complex gives an rms difference of 0.27 Å in average atomic positions. The smallest differences are in the chromophore region, with an overall rms displacement

Table 5: Geometrical Properties of Base-Pair Steps and Base Pairs for TGA−NOG, ^{m5}CGT−NOG, and TGA−4′EPIADR Structures[a]

base pair	step	roll (deg)	slide (Å)	twist (deg)	rise (Å)	tilt (deg)	propellor twist (deg)	buckle (deg)	cup (deg)
TGA−NOG (this work)									
T1•A12							−0.8	−12.2	
	1	6.5	−1.5	34.6	5.8	5.5			36.9
G2•C11							5.6	24.7	
	2	2.2	−1.3	29.2	4.1	−0.0			−20.4
A3•T10							−3.0	4.3	
	3	−7.5	3.0	32.9	3.9	0.3			−10.6
T4•A9							−3.8	−6.2	
	4	4.1	−1.3	28.8	4.2	−0.6			−18.7
C5•G8							2.8	−24.9	
	5	3.7	−1.3	37.0	5.6	−7.4			33.1
A6•T7							1.1	8.2	
^{m5}CGT−NOG (Gao *et al.*, 1991)									
^{m5}C1•G12							4.1	−5.7	
	2	1.0	−0.1	35.0	5.8	2.5			20.8
G2•^{m5}C11							4.8	15.2	
	2	1.2	−0.1	27.9	−1.0				−11.6
T3•A10							−2.7	3.5	
	3	1.6	−1.1	32.8	3.8	−0.7			−7.2
A4•T9							−4.3	−3.7	
	4	−2.6	−0.2	26.1	3.9	1.4			−15.4
^{m5}C5•G8							3.4	−19.1	
	5	2.3	0.4	40.3	5.8	−4.1			26.1
G6•^{m5}C7							5.6	7.1	
TGA−4′EPIADR (Langlois d'Estaintot *et al.*, 1992)									
T1•A12							−0.4	−8.9	
	1	−1.7	−2.0	37.3	5.3	−1.2			24.3
G2•C11							0.3	15.4	
	2	−0.9	0.6	29.3	3.5	−3.1			−6.4
A3•T10							−2.2	9.0	
	3	3.5	3.1	33.2	3.7	0.0			−18.0
T4•A9							−2.2	−9.0	
	4	−0.9	0.6	29.3	3.5	3.1			−6.4
C5•G8							0.3	−15.4	
	5	−1.7	−2.0	37.3	5.3	1.2			24.3
A6•T7							−0.4	8.9	

[a] All parameters were calculated using the NEWHEL94 program distributed by R. E. Dickerson.

of 0.08 Å and a maximum of 0.20 Å. Larger differences are observed in the nogalose and aminoglucose sugars. The nogalose sugars differ the most, but retain the same conformation, with a maximum displacement of 0.78 Å between the methyl groups and an overall rms displacement of 0.29 Å for all atoms. The aminoglucoses have the same conformation, with an overall rms displacement of 0.25 Å and a maximum displacement of 0.38 Å. The conformation of nogalamycin in this TGA−NOG complex differs significantly from that found in the small molecule crystal structure (Arora, 1983). The two sugar moieties of nogalamycin in the complex are brought in closer together to make the molecule more compact than in the free drug (Figure 7).

Solvent Interactions. Solvent plays a very important part in stabilizing the TGA−NOG complex. There are many interactions between the drug and solvent molecules forming a network of hydrogen bonds. Certain arrangements of water molecules are seen within the structure. Trigonal arrangements are particularly common, but a noticeable feature is the pentagonal arrangement of hydrogen bonds to water molecules located on the major-groove side of the A6•T7 base pair. A similar solvent structure has been observed in other anthracycline−DNA crystal structures (Leonard *et al.*, 1993). In this particular structure most of the positions in the major groove where solvent molecules may interact with the DNA in this fashion are occupied by the aminoglucose moiety of the drug; therefore, this pentagonal arrangement is not seen at any other base pair.

SUMMARY

The work here presents the first X-ray structure of nogalamycin complexed to one of its natural DNA binding sites. Previously reported nogalamycin−DNA structures contain DNA with modifications to either the backbone, the bases, or both. This structure corroborates earlier work and demonstrates that the drug intercalates with similar overall binding geometry at 5′pyrimidine−purine base pairs irrespective of minor modifications to the DNA. The N7 atom of a purine in the major groove appears to be a major factor in stabilizing the binding of nogalamycin to DNA. In spite of the overall similarity of intercalation geometry between our nogalamycin complex and previously reported X-ray structures, significant differences in base-pair geometry, hydrogen-bonding patterns of the drugs, and solvent structure are observed. Although, as a consequence of nucleotide sequence, one direct minor-groove hydrogen bond has been lost in the TGA−NOG complex, there are more waters located in the structure. Therefore, the hydrogen-bonding potential of nogalamycin is more fully satisfied. One drug makes an extra sequence-dependent major-groove direct hydrogen bond, and many more solvent-mediated hydrogen bonds are made by both drugs to the DNA. The importance

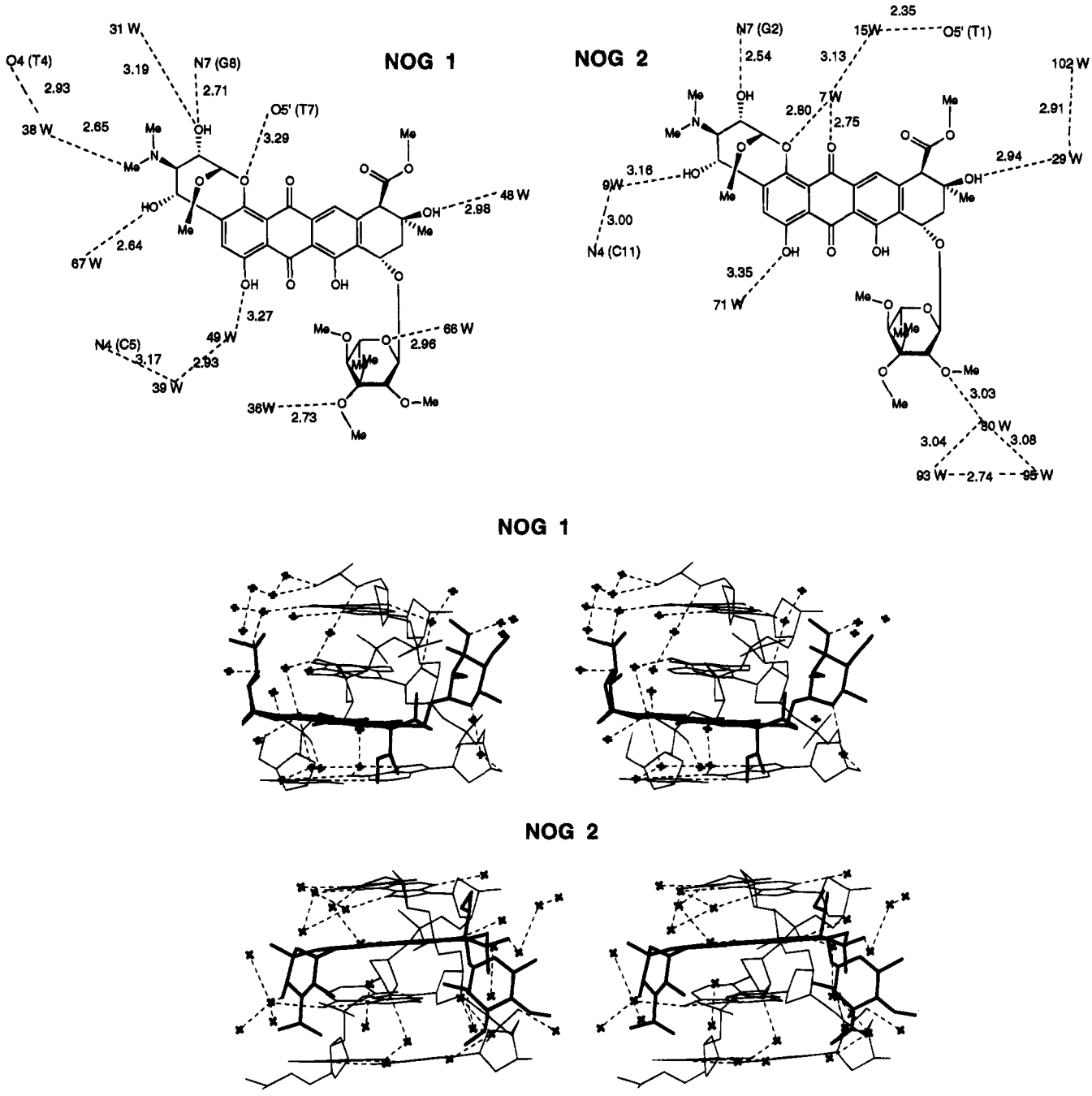


FIGURE 6: (a) Two-dimensional line drawings of the nogalamycins and all hydrogen bonds made to them are shown by dashed lines. Hydrogen-bonding patterns of both nogalamycin molecules, NOG1 and NOG2, in the $d^5'(TGATCA)$–nogalamycin complex illustrate direct and solvent-mediated interactions between the drug and DNA. The distances of the hydrogen bonds are also given to illustrate the strength of these interactions. (b) Stereoviews of the corresponding three-dimensional line drawings illustrating all hydrogen-bonding interactions made with each drug.

of water in this respect was first reported in DNA–protein interactions (Otwinowski *et al.*, 1988). 5'Pyrimidine–purine bases may be preferred sites in the immediate vicinity of the chromophore, but when it comes to the orientation of the sugar groups, the third base which is protected by the drug may also be important. The neighboring A·T base pair provides flexibility and the required functional group to participate in a solvent-mediated drug–DNA interaction.

In comparison with the less substituted anthracycline antibiotics, daunomycin and adriamycin, which are currently used in the treatment of leukemias and solid tumors, the increased cytotoxicity of nogalamycin is related to its increased residence time of intercalation in DNA. This is due to its slower dissociation from nuclear DNA (Fok & Waring, 1972; Fox *et al.*, 1985). This is not surprising since our structure clearly demonstrates the combination of extensive hydration, direct and solvent-mediated drug–DNA hydrogen bonds, and van der Waals interactions which serve to stabilize the complex. Extensive functionalization in both the A and D rings of nogalamycin requires more conformational change for either intercalation or dissociation of the drug to occur. The binding of nogalamycin to natural sequences of DNA therefore requires combinations of the flexibility of A·T base pairs interspersed with stability-providing G·C pairs.

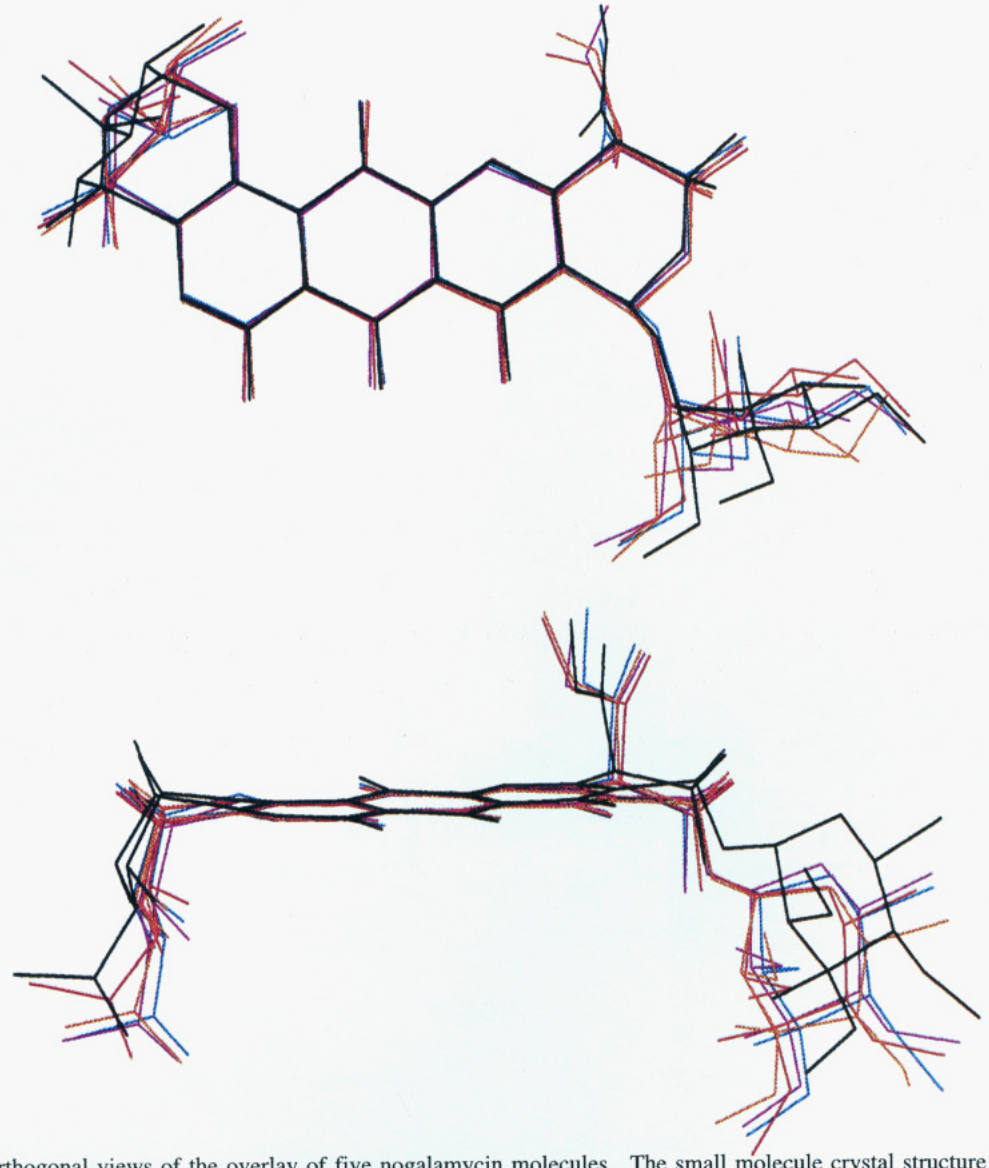


FIGURE 7: Two orthogonal views of the overlay of five nogalamycin molecules. The small molecule crystal structure is shown in dark green (Arora, 1983). The two nogalamycin molecules from the $d^{5'}(^{m5}CGTpSA^{m5}CG)$−nogalamycin complex are shown in red and orange (Williams *et al.*, 1990). NOG1 and NOG2 from $d^{5'}(TGATCA)$−nogalamycin are drawn in blue and purple, respectively.

ACKNOWLEDGMENT

The authors acknowledge the help of J. Brannigan, Z. Dauter, A. Dettmar, W. N. Hunter, L. van Meervelt, P. C. E. Moody, Z. Otwinowski, G. Wilson, K. S. Wilson, and the SERC- and MRC-supported Protein Structure Group at York.

REFERENCES

Arora, S. K. (1983) *J. Am. Chem. Soc.* **105**, 1328−1332.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977) *J. Mol. Biol. 112*, 535−542.

Bhuyan, B. K., & Dietz, A. (1965) *Antimicrob. Agents Chemother.*, 836−844.

Bhuyan, B. K., & Reusser, F. (1970) *Cancer Res. 30*, 984−989.

CCP4 (1994) *Acta Crystallogr. D50*, 760−763.

Collier, D. A., Neidle, S., & Brown, J. R. (1984) *Biochem. Pharmacol. 33*, 2877−2880.

Cowtan, K. D., & Main, P. (1993) *Acta Crystallogr. D49*, 148−157.

Dickerson, R. E., & Drew, H. R. (1981) *J. Mol. Biol. 149*, 761−786.

Egli, M., Williams, L. D., Frederick, C. A., & Rich, A. (1991) *Biochemistry 30*, 1364−1372.

Ennis, H. (1981) *Antimicrob. Agents Chemother. 19*, 657−665.

Fisher, J. F., & Aristoff, P. A. (1988) *Prog. Drug Res. 32*, 411−498.

Fok, J., & Waring, M. J. (1972) *Mol. Pharmacol. 8*, 65−74.

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.

Gao, Y.-G., Liaw, Y.-C., Robinson, H., & Wang, A. H.-J. (1990) *Biochemistry 29*, 10307−10316.

Jones, T. A. (1978) *J. Appl. Crystallogr. 11*, 268−272.

Kennard, O., & Hunter, W. N. (1989) *Q. Rev. Biophys. 22*, 327−379.

Langlois d'Estaintot, B., Gallois, B., Brown, T., & Hunter, W. N. (1992) *Nucleic Acids Res. 20*, 3561−3566.

Leonard, G. A., & Hunter, W. N. (1993) *J. Mol. Biol. 234*, 198−208.

Leonard, G. A., Brown, T., & Hunter, W. N. (1992) *Eur. J. Biochem. 204*, 69−74.

Leonard, G. A., Hambley, T. W., Macauley-Hecht, K., Brown, T., & Hunter, W. N. (1993) *Acta Crystallogr. D49*, 458−467.

Li, L. H., Kuentzel, S. L., Murch, L. L., Pschigoda, L. M., & Krueger, W. C. (1979) *Cancer Res. 39*, 4816−4822.

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.

Mcbridge, L. J., & Caruthers, M. H. (1983) *Tetrahedron Lett. 24*, 245−248.

Navaza, J. (1994) *Acta Crystallogr. 50*, 157−163.

Nunn, C. M., van Meervelt, L., Zhang, S., Moore, M. H., & Kennard, O. (1991) *J. Mol. Biol. 222*, 167−177.

Otwinowski, Z., Schevitz, R. W., Zhang, R.-G., Lawson, C. L., Joachimiak, A., Marmorstein, R. Q., Luisi, B. F., & Sigler, P. B. (1988) *Nature 335*, 321−329.

Quigley, G. J., Wang, A. H.-J., Ughetto, G., van der Marel, G., van Boom, J. H., & Rich, A. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 7204−7208.

Robinson, H., Liaw, Y.-C., van der Marel, G. A., van Boom, J. H., & Wang, A. H.-J. (1990) *Nucleic Acids 18*, 4851−4857.

Searle, M. S., Hall, J. G., Denny, W. A., & Wakelin, L. P. G. (1988) *Biochemistry 27*, 4340−4349.

Williams, L. D., Egli, M., Gao, Q., Bash, P., van der Marel, G. A., van Boom, J. H., Rich, A., & Frederick, C. A. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 2225−2229.

Zhang, X., & Patel, D. J. (1990) *Biochemistry 29*, 9451−9566.

BI941835N