Antitumor Drug Nogalamycin Binds DNA in Both Grooves Simultaneously: Molecular Structure of Nogalamycin–DNA Complex†

Yen-Chywan Liaw,‡ Yi-Gui Gao,‡ Howard Robinson,‡ Gijs A. van der Marel,§ Jacques H. van Boom,§ and Andrew H.-J. Wang*,‡

Department of Physiology and Biophysics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, and Gorlaeus Laboratories, Leiden State University, Leiden 2300RA, The Netherlands

Received October 19, 1989; Revised Manuscript Received November 6, 1989

ABSTRACT: The three-dimensional molecular structures of the complexes between an interesting antitumor drug, nogalamycin, and two DNA hexamers, d[CGT(pS)ACG] and d[m^{5}CGT(pS)Am^{5}CG], were determined at high resolution by X-ray diffraction analyses. Two nogalamycins bind to the DNA double helix in a 2:1 ratio with the aglycon chromophore intercalated between the CpG steps at both ends of the helix. The nogalose and aminoglucose sugars lie in the minor and major grooves, respectively, of the distorted B-DNA double helix. The binding of nogalamycin to DNA requires that the base pairs in DNA open up transiently to allow the bulky sugars to go through. Specific hydrogen bonds are found in the complex between the drug and guanine bases. We suggest that nogalamycin may prefer GC sequences embedded in a stretch of AT sequences.

Many antitumor drugs, including anthracycline antibiotics, are compounds that bind to the DNA double helix by intercalation (Wang, 1987). Nogalamycin (Figure 1) is a unique drug, distinct from other anthracycline antibiotics such as daunomycin and adriamycin, in that it contains two sugar moieties (nogalose and aminoglucose) attached to rings A and D of the aglycon chromophore (Bhuyan & Reusser, 1970). The biological activities of these antitumor drugs are believed to be closely associated with their ability to bind to DNA in a specific manner. Recently, the elucidation of the detailed interactions between daunomycin (and adriamycin) and DNA hexamers by high-resolution X-ray diffraction analysis has provided valuable information regarding the role of various functional groups of the drug molecules (Wang et al., 1987; Frederick et al., 1990). For example, the essential amino group in daunomycin is found to lie in the minor groove of the double helix, while the O9 hydroxyl group forms hydrogen bonds to the guanine base. We are interested in obtaining similar valuable structural information for the interactions of other important anthracycline antibiotics with DNA. We have chosen nogalamycin since it is a potent antitumor drug and its complicated molecular architecture also poses an interesting question related to the exact mode of binding to the DNA molecule.

In this paper, we report the three-dimensional structure, at 1.3- and 1.7-Å resolution, respectively, of nogalamycin complexed to two DNA hexamers, d[CGT(pS)ACG] and d[m^{5}CGT(pS)Am^{5}CG], where pS is an internucleotidic phosphorothioate linkage in the *R* configuration. The molecular structure reveals that nogalamycin binds to the DNA hexamer duplex in a 2:1 ratio with the elongated aglycon chromophore intercalated between the CG base pairs of the CpG dinucleotide step at both ends of the helix. The nogalose is lying in the minor groove, and the aminoglucose is lying in the major groove of the distorted B-DNA double helix. Evidently, DNA must melt locally to allow the drug to bind. Specific hydrogen bonds are found in the complex between nogalamycin and guanine bases. A hydroxyl group in the aminoglucose forms a strong hydrogen bond to N7 of G6 in the major groove, while the N2 amino groups of the guanines above and below the aglycon ring form hydrogen bonds to the keto oxygen (O14) in the acetic methyl ester group and the glycosyl ether oxygen (O7) in the minor groove. Several van der Waals interactions also contribute to the stability of the complex.

Materials and Methods

At the onset of the crystallization experiments, no information was available regarding the prefered DNA binding sequences for nogalamycin. We therefore tested many DNA oligonucleotides (from hexamer to dodecamer), among which the two aforementioned modified DNA hexamers produced the best crystals. These two thiophospho-containing oligonucleotides were synthesized according to a procedure published earlier (Marugg et al., 1984). Nogalamycin was kindly provided by Dr. Paul Aristoff of the Upjohn Co., and it was dissolved in methanol as a stock solution for crystallization. For the nogalamycin–d[CGT(pS)ACG] complex, a typical crystallization mixture contained 0.5 mM oligonucleotide, 8 mM $MgCl_2$, 30 mM sodium cacodylate (pH 5), and 0.5 mM nogalamycin in 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. Bright orange-red crystals appeared after 1 week. The crystals were in space group $C222_1$ with unit-cell dimensions of a = 22.98, b = 47.27, and c = 64.44 Å. A 0.4 × 0.5 × 0.9 mm crystal was mounted in a sealed glass capillary with a droplet of mother liquor for data collection on a Rigaku AFC-5R rotating-anode X-ray diffractometer at 25 °C using the ω-scan mode with graphite-monochromated Cu Kα radiation. Data were collected to a resolution of 1.3 Å (2θ = 75°). A second crystal form was obtained for the nogalamycin–d[m^{5}CGT(pS)Am^{5}CG] complex under similar conditions. Large truncated hexagonal-bipyramidal crystals appeared in a few

† This work was supported by NSF (A.H.-J.W.). G.A.v.d.M. and J.H.v.B. were supported by the Netherlands Organization for the Advancement of Pure Research (ZWO). Y.-C.L. acknowledges support from the Institute of Molecular Biology, Taiwan (ROC).

* Corresponding author.
‡ University of Illinois at Urbana–Champaign.
§ Leiden State University.

0006-2960/89/0428-9913$01.50/0

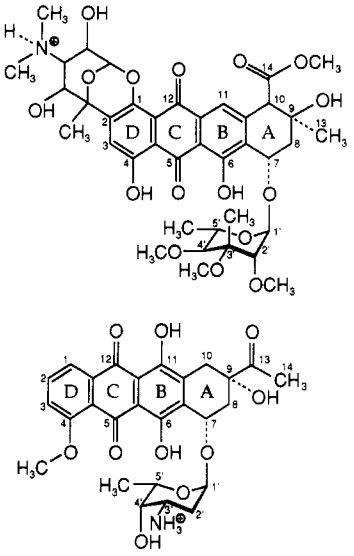

FIGURE 1: Molecular formulas of two anthracycline antibiotics, nogalamycin (top) and daunomycin (bottom). Both contain an aglycon chromophore with four fused rings (A–D). Rings B–D are unsaturated with exocyclic oxygen atoms, whereas ring A is semisaturated. Nogalamycin has two sugars attached to the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

days; they are in space group $P6_1$ (confirmed later by refinement) with unit-cell dimensions of $a = b = 26.31$ and $c = 100.25$ Å. 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. A crystal of 0.3 × 0.4 × 0.8 mm in size was used for data collection on the same diffractometer to 1.7-Å resolution. Lp, empirical absorption, and decay corrections were applied for both data sets.

The Patterson maps of the two crystal forms showed clearly the planar base stacking direction. In the $P6_1$ form, the stacking is along the a axis direction (26.3-Å repeat), while in the $C222_1$ form it is along the diagonal (52.6-Å repeat) of the ab plane. It was noted that these repeat dimensions are multiples of 8 × 3.3 = 26.4 Å, suggesting that the complex is a hexamer duplex with two nogalamycins intercalated in it. The structure of the $P6_1$ form was first solved by the molecular replacement method with the program ULTIMA (Rabinovich & Shakked, 1984) using a model derived from the molecular structure of the daunomycin–d(CGTACG) complex (Wang et al., 1987). Its structure was then refined by the Konnert–Hendrickson constrained refinement procedure (Hendrickson & Konnert, 1979) to an R factor of ~30%. At this stage we were quite confident that a correct structure was in hand. This new model was in turn used to search for the structural solution of the $C222_1$ form with the translation–rotation option in the program X-PLOR (Brunger, 1988). This procedure provided a model with a clear global maximum of correlation coefficient of 50%. This model was subsequently successfully refined by the Konnert–Hendrickson method to a final R factor of 19.0% with 5975 reflections [$F > 3.0\sigma(F)$] to 1.3-Å resolution with a root mean square (rms) difference in bond distances of 0.025 Å from the ideal values. The structure includes two nogalamycins, two DNA hexamers, two hydrated magnesium ions, and 113 water molecules in the asymmetric unit of the crystal. The structure of the $P6_1$ form was refined to a current R factor of 21.5% with 2869 reflections [$F > 3.0\sigma(F)$] to 1.7-Å resolution. The detailed structural analyses of these two crystal forms will be reported more fully elsewhere, and at that time, the final atomic coordinates will be deposited in the Brookhaven Protein Data Bank. As the structure derived from the $C222_1$ form is more accurate due to its higher resolution, it will be used for discussion in this paper.

Results and Discussion

Molecular Structure. The overall structure of the 2:1 complex of nogalamycin–d[CGT(pS)ACG] is shown in Figure 2, which reveals a number of interesting features. First, the DNA molecule is a distorted B-DNA double helix with the two nogalamycins intercalated in it. The complex has a noncrystallographic molecular 2-fold symmetry. The rms deviation between the two halves (one drug plus one DNA hexamer) of the complex is 0.306 Å. The elongated aglycon chromophore skewers the DNA molecule 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 drug molecule has a conformation very similar to that of the free drug structure (Arora, 1983). The two sugars are on the same side of the flat aglycon ring, and they both point toward the AT region in the middle of the helix. 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. In ring A, the C9 atom has the largest displacement from the mean plane (0.53 Å). The consequence of this pucker in ring A is that the O9 hydroxyl is in an equatorial position pointing away from the DNA and it does not interact with DNA, in contrast to that seen in daunomycin (Wang et al., 1987). Interestingly, the acetic methyl ester on C10 is in an axial position almost perpendicular to the plane of the aglycon, and it is in a good position to receive a hydrogen bond from DNA as will be discussed later. Therefore, the juxtaposition of the various groups in ring A is quite different from that of daunomycin, leading to different binding interactions for the two drugs.

Figure 2 (bottom) depicts the van der Waals models of the complex in which the close fit between the drug and DNA is evident. In particular, the hydrophobic nogalose completely occupies the minor groove, making several van der Waals contacts to the bases and the backbones of DNA. This close fit expands the minor groove slightly as judged by the phosphorus to phosphorus distances (~15 Å) across the groove, compared to that in the uncomplexed B-DNA helix (11.5–13.0 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 that the drug occupies (Kopka et al., 1985; Coll et al., 1987; Wang & Teng, 1989). However, the binding of the bulky nogalamycin to DNA requires other significant rearrangements of DNA backbone torsion angles as shown in Table I. The largest deviations of these angles from the normal, uncomplexed B-DNA are associated with the C-G residues involved in the intercalation as expected. These include ϵ of C5 and C11 (−104.7° and −102.2°) with the concomitant changes of ζ to −178.1° and −170.2°, respectively. The puckers of all the sugars remain in the C2′-endo type except for the C1 residue, which is C3′-endo type. 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).

The detailed interactions between nogalamycin and DNA are shown more clearly in the skeletal view in Figure 3. The

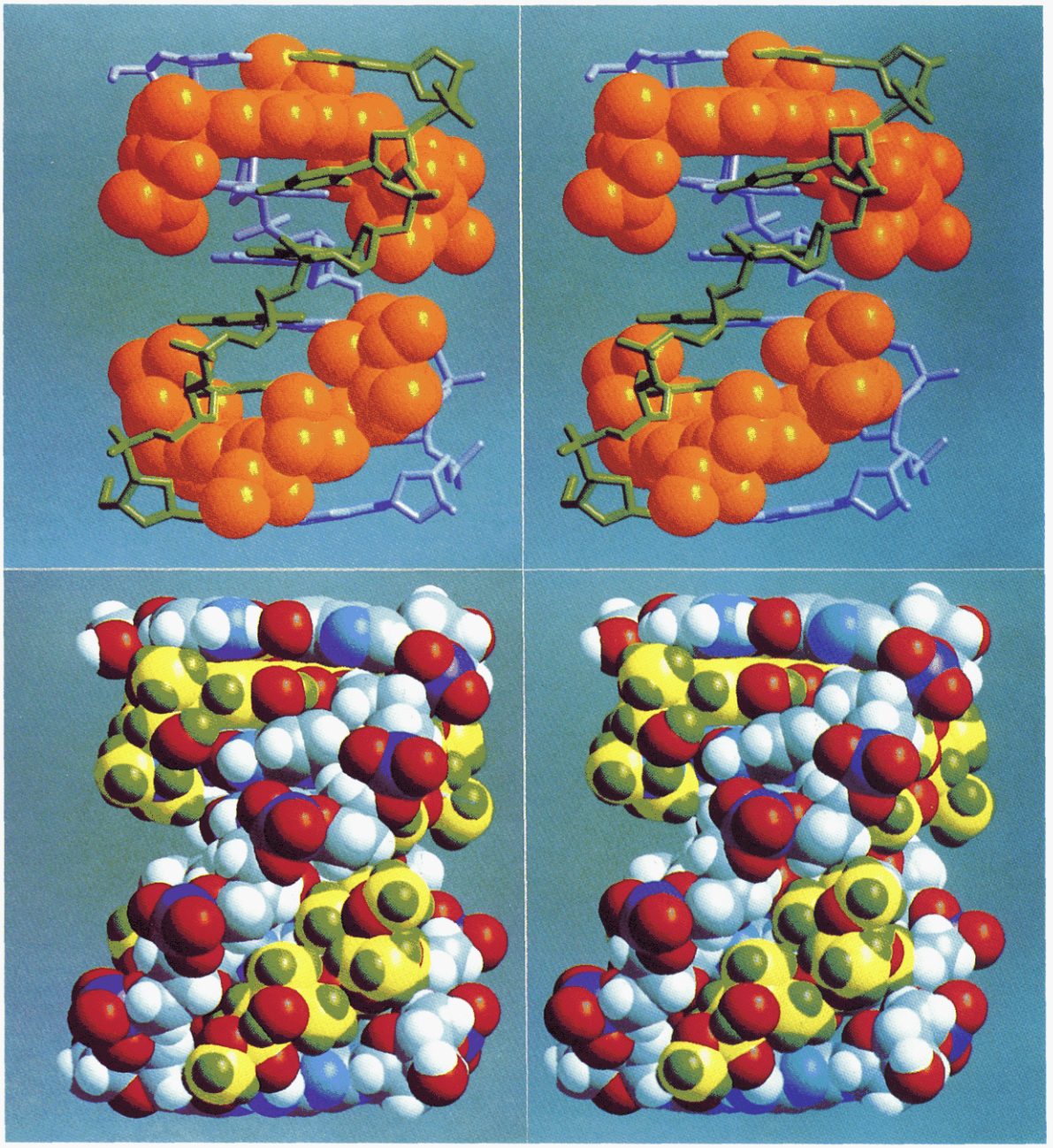

FIGURE 2: (Top) Stereoscopic drawing of the nogalamycin–d[CGT(pS)ACG] complex, which has a noncrystallographic molecular 2-fold symmetry. Two nogalamycins (in orange) intercalate between the CpG steps at both ends of the distorted right-handed B-DNA hexamer duplex (wire model) with the aglycon chromophore penetrating through the helix between the CG 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. Due to these bulky sugar groups, the binding of nogalamycin likely requires a transient opening of a few bases to allow the drug to span both grooves. (Bottom) van der Waals diagram of the complex. Nogalamycin (in yellow with green hydrogens) hugs tightly to the double helix with a combination of stacking, hydrogen-bonding, and van der Waals interactions. These figures were generated by molecular graphic program MARS (Molecular Architecture Rendering System) on an IRIS computer (Silicon Graphics Inc.) developed by us at UIUC.

inner GC base pairs (G2-C11 and C5-G8) in the intercalated CpG steps are highly buckled (−25.4° and 25.9°) as they are in close contact with the nogalose and aminoglucose sugars and as they adjust to accommodate the bulky drug molecules. The terminal two G-C base pairs, which are less buckled (12.8° for C1-G12 and -17.5° 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 diagonal of the *ab* plane in the unit cell. 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. There are three direct hydrogen bonds between nogalamycin and DNA, anchoring the drug in the helix with a specific orientation. In the major groove, the O2G hydroxyl of the aminoglucose is 2.84 Å from N7 of guanine G2. In the minor groove, the keto oxygen O14 from the acetic methyl ester at C10 receives 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) above the aglycon intercalator. 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). The DNA adjusts its conformation to align these hydrogen bonds and optimize other hydrophobic interactions, resulting in a slightly asymmetrical backbone con-

Table I: Torsion Angles (deg) and Some Helical Parameters[a] of the Nogalamycin–d[CGT(pS)ACG] Complex

	α	β	γ	δ	ε	ζ	χ	pucker	ω	κ
C1			69.5	88.6	−147.8	−83.7	−157.7	C4′-exo	39.2	12.8
C7			16.8	139.5	−134.0	−96.6	−147.1	C3′-exo		
G2	−53.8	−166.0	55.8	136.1	−174.0	−107.6	−101.3	C2′-endo	21.8	−25.4
G8	−72.0	−170.3	68.7	130.5	−167.5	−106.4	−102.1	C2′-endo		
T3	−79.4	167.8	66.8	118.8	−172.3	−131.4	−118.9	C1′-exo	39.0	1.2
T9	−67.7	167.5	55.1	113.2	179.7	−123.2	−123.9	C1′-exo		
A4	−42.1	159.5	51.8	118.5	−152.1	−88.5	−130.4	C1′-exo	23.0	2.7
A10	−53.7	157.1	54.7	103.8	−157.2	−99.6	−137.4	O4′-endo		
C5	−74.6	177.4	49.6	140.0	−104.7	−178.1	−101.4	C2′-endo	35.7	25.9
C11	127.9	−177.9	−157.4	158.6	−102.2	−170.2	−104.3	C2′-endo		
G6	−62.0	164.9	34.4	126.7			−113.6	C2′-endo		−17.5
G12	−75.6	165.7	57.0	136.3			−117.3	C2′-endo		
B-DNA[b]	−63	171	54	123	−169	−108	−117			

[a] Torsion angles along the backbone of the oligonucleotide are defined as P—O5′(α)—C5′(β)—C4′(γ)—C3′(δ)—O3′(ε)—P(ζ) and χ is the glycosyl angle. The nucleotides are numbered from C1 to G6 in one strand and from 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 with the program CURVE (Lavery & Sklevar, 1989). [b] Average values from DNA dodecamer d(CGCGAATTCGCG) (Drew & Dickerson, 1981).

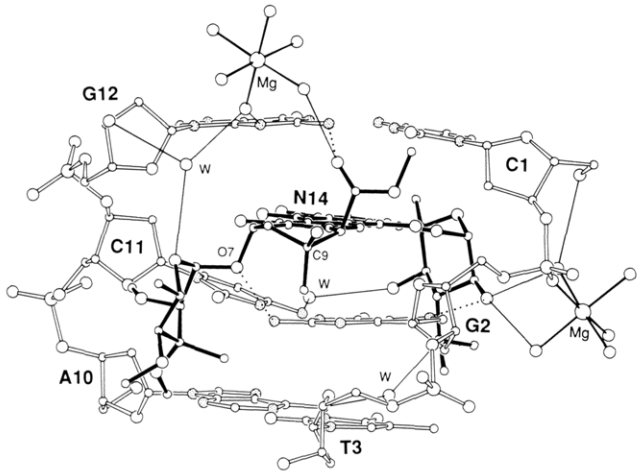

FIGURE 3: Skeletal diagram of the detailed surroundings of the intercalated nogalamycin. Three base pairs of the hexamer helix are shown. The other half of the complex is similar. Hydrogen bonds between the nogalamycin and DNA are shown as dotted lines. In the major groove, the hydroxyl O4G of the aminoglucose forms a strong hydrogen bond (2.84 Å) to the N7 of guanine G2. The other hydroxyl O4G is close to N4 of cytosine C11 (3.58 Å) but not directly hydrogen bonded to it. In the minor groove, two hydrogen bonds are found. The keto oxygen O14 of the acetic methyl ester receives a hydrogen bond (3.02 Å) from the N2 amino group of 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 large buckle in the C11-G2 base pair (−25.4°). A number of bridging water molecules and two hydrated magnesium ions that bridge between two complexes in the crystal lattice are also shown.

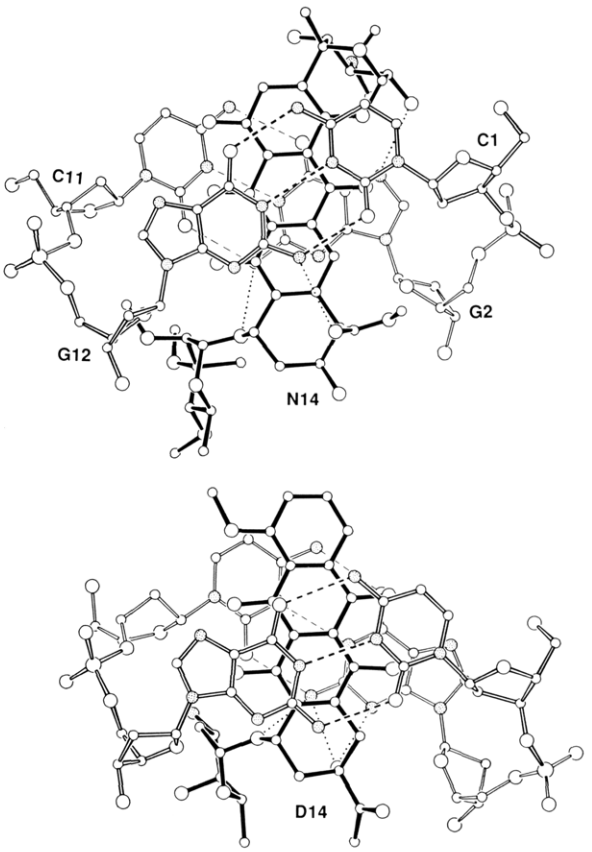

FIGURE 4: Comparative view of the nogalamycin (top) and daunomycin (bottom) molecules and the two adjacent G-C base pairs from a direction perpendicular to the base plane. Both drugs intercalate such that the long dimension of the aglycon chromophore is nearly perpendicular to the C1′–C1′ vectors of the base pairs. However, nogalamycin seems to be "pulled" down toward the minor groove with N2 of G12 over ring B of the aglycon. In the daunomycin complex, N2 of G12 is over ring A of the aglycon. The difference in the orientation is due primarily to the different hydrogen-bonding interactions between the drugs and DNA.

formation for the two DNA strands straddling the aglycon. It is interesting to note that three groups, including acetic methyl ester, ring A, and nogalose, are connected in a linear manner that follows the contour of the right-handed minor groove covering three base pairs. A number of first-shell bridging water molecules and hydrated magnesium ions are clearly visible in the electron density map. Some of them are shown in Figure 3.

Comparison with Daunomycin. It is interesting to compare the binding interactions of two related, but distinct, anthracycline drugs (nogalamycin vs daunomycin) to DNA. Both structures, nogalamycin + d[CGT(pS)ACG] and daunomycin + d(CGTACG) (Wang et al., 1987), were determined at high resolution (1.3 and 1.2 Å, respectively), which affords a reliable comparison. Figure 4 shows the views of the two complexes from a direction perpendicular to the plane of the aglycon ring. At first glance, they seem to be almost identical, as the long dimension of both aglycons lies across the base pairs and reaches both grooves. However, due to the different juxtaposition of the hydrogen bonds between the drug and DNA, significant differences are observed in the complexes. Spe-

cifically, if we focus on the location of the aglycon ring relative to the base pairs (Figure 4, top), it is obvious that the nogalamycin is "pulled" toward the minor groove by about 2.0 Å, in comparison to that of the daunomycin–d(CGTACG) complex (Figure 4, bottom). Ring D of nogalamycin is now stacked underneath the N4 amino group of the C1 residue. The average helical twist angle of the two CpG steps across the intercalator in the nogalamycin complex is 37°, indicative of a very small overwinding. The rest of the steps have a combined unwinding angle of −11°. 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 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.

Biological Implications. Nogalamycin is an active agent against tumor cells both in vitro and in vivo (Bhuyan & Reusser, 1970; Wiley, 1979). Like other anthracyline antibiotics, its cellular target is presumed to be nuclear DNA through the binding mechanism of intercalation (Crooke & Reich, 1980). This raises an interesting question regarding the mechanism of intercalation, since nogalamycin has two bulky sugar residues at both ends of the aglycon ring which are expected to present great hindrance for the drug to penetrate between base pairs. In addition, there have been conflicting results regarding the base sequence specificity of nogalamycin (Kersten et al., 1966; Bhuyan & Smith, 1965). Our results suggest a reasonable explanation for this confusion. The structure of the complex described above clearly shows that the two sugars of the drug reside on opposite sides of the G-C base pairs with specific hydrogen bonds between the drug and DNA, both in the major groove and in the minor groove. If nogalamycin is intercalated between A-T base pairs, the hydrogen-bonding interactions in the minor groove will be completely lost due to the lack of the NH_2 group in adenine. Therefore, G-C sequences are preferred for the base pairs immediately surrounding the drug. However, in order for the drug to slide through between base pairs, it is imperative that the helix be open transiently with sufficient room for the bulky sugars. This requirement favors the A-T sequences, since they open up more readily. This binding process is expected to be slow, as has recently been shown (Fox et al., 1985). Our preliminary NMR study of the nogalamycin–d(CGTACG) complex (0.6:1 drug/duplex ratio) reveals that the complex and DNA are in slow equilibrium on the NMR time scale (unpublished result). We have noticed that the same drug–DNA complex at a 2:1 ratio exhibited an NMR spectrum consistent with nogalamycin binding at the CpG sites. These studies also suggest that the binding interactions of nogalamycin to DNA seen in the crystal structure are not affected by the chemical modification of the phosphate group in DNA.

From this analysis, we suggest that nogalamycin prefers isolated G-C sequences (e.g., CpG) embedded in a stretch of A-T sequences. (It may be noted that DNA sequences in the promoter regions often possess such features, i.e., AT-rich sequences sprinkled with GC sequences.) This is a rather unusual sequence preference for a DNA binding compound. Compounds that are capable of binding to both grooves of the helix have been synthesized (Atwell et al., 1984; Zimmerman et al., 1989), and they may have a similar sequence preference. In comparison, echinomycin and triostin (two antitumor quinoxaline bis-intercalators) bind to tetranucleotide sequences (A/T)CG(A/T) (Wang et al., 1984).

This unique sequence specificity may be tested by binding nogalamycin to different oligonucleotides with defined sequences. For example, the NMR study of the 1:1 complex of nogalamycin and d(TATATACGTATATA) should show that the chemical shifts associated with CG protons have the largest changes upon binding of the drug.

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). 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., 7-con-O-methylnogarol, do not have nogalose at the C7 position, and they have significant antitumor activity in vivo (Wiley, 1979). It would be interesting to determine 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 DNA octamer d(ACGTACGT), and the structural solution of this complex will provide additional structural information regarding base pairs further away from the intercalation site. Our ultimate goal is to be able to correlate the functions of various substituents on different anthracycline antibiotics to their biological activities. Structural analysis like this is the first step.

ACKNOWLEDGMENTS

We thank M.-k. Teng for his assistance.

REFERENCES

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, F. (1970) *Cancer Res. 30*, 984–989.

Brunger, A. (1988) X-PLOR software, The Howard Hughes Medical Institute and Yale University, New Haven, CT.

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

Crooke, S. T., & Reich, S. D., Eds. (1980) *Anthracyclines*, Academic Press, New York.

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., Brassett, C., & Waring, M. J. (1985) *Biochim. Biophy. 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* (in press).

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.

Kopka, M. L., Yoon, C., Goodsell, D., Pjura, P., & Dickerson, R. E. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 1376–1380.

Lavery, R., & Sklenar, H. (1989) *J. Biomol. Struct. Dyn. 6*, 655-667.
Marugg, J. E., van den Bergh, C., Tromp, M., van der Marel, G. A., van Zoest, W. J., & van Boom, J. H. (1984) *Nucleic Acids Res. 12*, 9095-9110.
Rabinovich, D., & Shakked, Z. (1984) *Acta Crystallogr. A40*, 195-200.
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.
Wang, A. H.-J. (1987) in *Nucleic Acids and Molecular Biology* (Eckstein, F., & Lilley, D. M. Eds.) Vol. 1, pp 32-54, Springer, Berlin.
Wang, A. H.-J., & Teng, M.-k. (1989) in *Crystallographic and Modeling Methods in Molecular Design*, Springer-Verlag, New York (in press).
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, J. C. (1974) *J. Mol. Biol. 89*, 783-801.
Wiley, P. F. (1979) *J. Nat. Prod. 42*, 569-582.
Zimmerman, S. C., Lamberson, C. R., Cory, M., & Fairley, T. A. (1989) *J. Am. Chem. Soc. 111*, 6805-6809.

Articles

Electrical Potentials in Trypsin Isozymes†

Kizhake Soman,‡,§ An-Suei Yang,‡ Barry Honig,*,‡ and Robert Fletterick‖

Department of Biochemistry and Molecular Biophysics, Columbia University, New York, New York 10027, and Department of Biochemistry and Biophysics, University of California, San Francisco, California 94143

Received May 25, 1989; Revised Manuscript Received August 9, 1989

ABSTRACT: Cow and rat trypsin differ in net charge by 12.5 units yet have the same enzymatic mechanism. The role of electrical potentials in the catalytic mechanism of these trypsin isozymes is investigated by using the finite difference Poisson-Boltzmann method. The calculations reveal that the active sites are effectively shielded from surface charge, thus making it possible for the two enzymes to have essentially identical potentials in their catalytically important regions. The potentials in both active sites are dominated by local interactions arising both from partial charges and from the negative charge on Asp-102. The latter is found to stabilize the transition state by about 4 kcal/mol, a value that is consistent with the extent of reduced catalytic activity in the variant Asn-102 trypsin, in which the negative charge is absent. The calculations predict that Asp-102 is ionized and that His-57 is neutral in the resting state of the enzyme. In contrast to their negligible effect on catalytic activity, the cumulative effect of surface charges is found to raise the pK of the N-terminal α-amino group of Ile-16 in the rat enzyme by about 1.5 units relative to that of cow trypsin. This charged amino acid forms an ionic bond with Asp-194, which stabilizes the active conformation of the enzyme. An increase in pK of Ile-16 thus provides a possible explanation for the retention of activity of rat trypsin at high pH. The results of this study could not have been obtained from an electrostatic model based on Coulombic potentials.

Electrostatic fields around the surface of proteins have been shown to play an important role in molecular recognition and binding. For example, the positive potential of DNA polymerase aids in the assembly of the protein-DNA complex (Warwicker et al., 1985). The electric field of Cu, Zn superoxide dismutase is responsible for enhanced diffusion rates of its superoxide substrate to the active site (Koppenol, 1981; Klapper et al., 1986; Sharp et al., 1987) while the electric field of cytochrome *c* plays a similar role in the interaction of this protein with cytochrome *c* peroxidase (Koppenol & Margoliash, 1982; Northrup et al., 1988). These fields, which extend out into solution, result from the asymmetric distribution of charged side chains on the protein surface. They do not appear to be due to a single amino acid but, rather, arise from the cumulative effects of a number of residues. It is of interest to consider whether surface charges might also play a role in influencing function in the interior of proteins, particularly in active sites.

In considering this question in "protein design", it is important to distinguish the effects of specific charged groups that are part of the active site from the aggregate effects of surface charge some distance from the active site. That charged residues near the active site might be functionally important has been suspected for some time. For example, it was suggested over 30 years ago that charged amino acids in the retinal binding site play a role in determining the colors of visual pigments (Kropf & Hubbard, 1958). Quantum mechanical calculations and experiments on artificial pigments have identified the location of these groups with respect to the retinal chromophore (Honig et al., 1976; Honig et al., 1979).

† This work was supported by NIH Grants GM-30518 (B.H.) and DK-39304 (R.F.).
‡ Columbia University.
§ Present address: Department of Molecular Biology, Research Institute of Scripps Clinic, La Jolla, CA 92037.
‖ University of California.

0006-2960/89/0428-9918$01.50/0