Binding of Daunorubicin to β-D-Glucosylated DNA Found in Protozoa *Tryponosoma brucei* Studied by X-ray Crystallography

Yi-gui Gao,† Howard Robinson,† Eric R. Wijsman,‡ Gijs A. van der Marel,‡ Jacques H. van Boom,‡ and Andrew H.-J. Wang*,†

Department of Cell & Structural Biology University of Illinois at Urbana-Champaign Urbana, Illinois 61801 Gorlaeus Laboratories Leiden Institute of Chemistry Leiden 2300RA, The Netherlands

Received October 31, 1996

Modified bases are not uncommon in DNA, and they play important biological roles. *C*5-Methylcytosine (m^5C) or *N*6-methyladenine (m^6A) are used in bacteria for the restriction-modification system.[1] In eucaryotic systems m^5C may be involved in transcriptional repression.[2] Interestingly, unusual bases including 5-hydroxyuracil,[3a] uracil,[3b] α-putrescinylthymine,[3c] 2-aminoadenine,[3d] *N*4-methylcytosine,[3e,f] *N*7-methylguanine,[3g] *N*6-(carbamoylmethyl)adenine,[3h] and finally hexosylated bases[3i] are found in the DNA of many bacteriophages. Glucosylated 5-(hydroxymethyl)cytosines and 5-hydroxycytosines are found in *Escherichia coli* phage T3, T4, and T6, and in *Rhizobium* phage RL38J1,[4] respectively. Another interesting case involves glucosylated-T at the C5 position of the bases in protozoa *Trypanosoma brucei*.[5] Those modified bases may be involved in the regulation of gene expression. In the long telomeric repeat $(GGGTTA)_n$ of *T. brucei* about 13% of T is replaced by *b*-glucosyl(hydroxymethyl)uracil (denoted g^5T); this is compared to 0.8% modification in total DNA.[6] How these DNA molecules with bulky substituents occupying the major groove interact with relevant proteins remains an intriguing question to be answered.

As a first step to address these questions, we probe the accessibility of the glucosylated DNA duplex using the anticancer intercalator drug daunorubicin (denoted DNR). The modified DNA chosen was that of the g^5T-containing DNA from *T. brucei*. It is of interest to note that ethidium bromide and chloroquine, two commonly used intercalators, have been shown to have a cytotoxic effect toward trypanosomes.[7] If the cytotoxic effect would be related to certain specific interactions between the intercalator and the glucosylated region of the trypanosome DNA, it would be of interest to investigate other clinically available drugs as potential anti-trypanosome drugs.

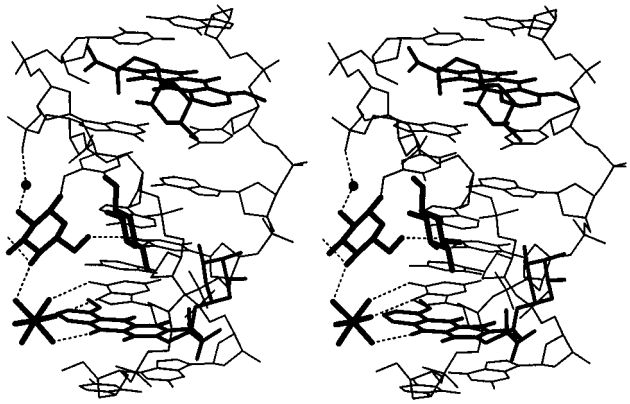

Figure 1. Stereoscopic drawings of the 2:1 DNR−CG[g^5T]ACG complex. The glucoses and the hydrated Mg^{2+} ion are shaded in thick lines, and the bridging water is shown as a dark circle. Relevant hydrogen bonds are indicated by dashed lines. The O3′ of the ordered glucose is hydrogen bonded to the O2P of C5 from the neighboring complex.

We have undertaken a structural study in which the intercalator anticancer drug daunorubicin was added to the glucosylated DNA hexamer CG[g^5T]ACG.[8] The 3D structure of the 2:1 complex (Figure 1), determined at 1.5 Å resolution by X-ray diffraction analyses, is very similar to the DNR−CGTACG complex[9] with a rmsd of 0.51 Å between the two structures from different crystal lattices (Figure 2). In the complex, two DNR molecules are intercalated in two symmetry-equivalent CpG steps with the drug aglycon chromophores lying in the

* To whom correspondence should be addressed.
† University of Illinois at Urbana-Champaign.
‡ Leiden Institute of Chemistry.

(1) Wilson, G. G. *Trends Genet.* **1988**, *4*, 314−318.
(2) Bird, A. P. *Cell* **1992**, *70*, 5−8.
(3) (a) Hemphill, H. E.; Whiteley, H. R. *Bacteriol. Rev.* **1975**, 257−315. (b) Takahashi, I.; Marmur, J. *Nature* **1963**, *197*, 794−795. (c) Kropinski, A. M. B.; Bose, R. J.; Warren, R. A. J. *Biochemistry* **1973**, *12*, 151−157. (d) Kirnos, M. D.; Khudyakov, I. Y.; Alexandrushkira, N. I.; Vanyushin, B. F. *Nature* **1977**, *270*, 369−370. (e) Janulaitis, A.; Klimasauskas, S.; Petrusyte, M.; Butkus, V. *FEBS* **1983**, *161*, 131−134. (f) Ehrlich, M.; Gama-Sosa, M. A.; Carreira, L. H.; Ljungdahl, L. G.; Kuo, K. C.; Gehrke, C. W. *Nucleic Acids Res.* **1975**, *13*, 1399−1412. (g) Nikolskaya, I. I.; Lopatina, N. G.; Debov, S. S. *Biochim. Biophys. Acta* **1976**, *435*, 206−210. (h) Swinton, D.; Hattman, S.; Crain, P. F.; Cheng, S.-S.; Smith, D. L.; McCloskey, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 7400−7404. (i) Wyatt, G. R.; Cohen, S. S. *Biochem. J.* **1953**, *55*, 774−782.
(4) Hsu, F. F.; Crain, P. F.; McCloskey, J. A.; Swinton, D. L.; Hattman, S. *Adv. Mass Spectrom.* **1989**, *11B*, 1450−1451.
(5) Gommers-Ampt, J. H.; Van Leeuwen, F.; de Beer, A. L. J.; Vliegenthart, J. F. G.; Dizdaroglu, M.; Kowalak, J. A.; Crain, P. F.; Borst, P. *Cell* **1993**, *75*, 1129−1136.
(6) van Leeuwen, F.; Wijsman, E. R.; Kuyl-Yeheskiely, E.; van der Marel, G. A.; van Boom, J. H.; Borst, P. *Nucleic Acids Res.* **1996**, *24*, 2476−2482.
(7) Merck Index, 12th ed.; Merck & Co., Inc.: Rahway, NJ, pp 334 and 809.

(8) Oligonucleotide d(CG[g^5T]ACG) was synthesized according to the procedure of Wijsman et al.[11] and purified by Sepharose G50 column chromatography. Crystals of the 2:1 complexes of DNR and the glucosylated DNA hexamer were obtained from a mixture containing 1.2 mM DNA hexamer (single strand), 4 mM $MgCl_2$, 35 mM sodium cacodylate (pH 6.0), 2.5 mM spermine, 1.2 mM drug, 0.02% dioxane, and 5% (v/v) 2-methyl-2,4-pentanediol (2-MPD). The solution was equilibrated with 30 mL of 50% 2-MPD reservoir at room temperature (∼20 °C) by vapor diffusion according to the procedure of Wang and Gao.[12] Orange-red crystals in the form of irregular clusters appeared after several weeks. Crystal data: space group *P*1, $a = 18.62$ Å, $b = 20.10$ Å, $c = 26.55$ Å, $\alpha = 69.09°$, $\beta = 90.04°$, $\gamma = 108.03°$, *R*-factor 17.3%, 4678 reflections at $2\sigma(F_o)$ (Table 1S in the Supporting Information). A crystal with the size of $0.2 \times 0.2 \times 0.3$ mm was selected and mounted in a sealed glass capillary with a droplet of mother liquor for data collection to 1.5 Å resolution. A Rigaku R-Axis IIc image plate area detector system mounted on a RU200 rotating anode X-ray generator at 20 °C with CuKα radiation was used. Data were processed with the Molecular Structure Corporation programs to obtain the structure factor amplitudes F_o(hkl). The crystal structure was solved by the molecular replacement method using ULTIMA.[13] The atomic coordinates from the structure of the DNR−CGTACG complex[9] were used as the starting search model. The Patterson map showed that the base stacking direction is nearly along the *a*-axis, and such information facilitates the placement of the complex in the unit cell. A clear solution was obtained from the rotation search. No translation search was needed for the *P*1 space group. The crystal structure was refined by the simulated annealing procedure using X-PLOR.[14] A series of refinement stages located the two glucoses. One glucose is well-defined as shown in the $(F_o - F_c)$ omit electron density map (Figure 1S), but the other one that is associated with the T9 base is disordered. The latter glucose was placed in the electron density at half-occupancy. The possibility that one of the glucoses was deglycosylated from DNA due to the influence of the crystallization condition, lattice forces, binding of daunorubicin, or a combination of the above factors was considered. We washed and dissolved several crystals and subjected the solution to electrospray mass spectrometry. Only the peak of mass 1971 from the intact CG[g^5T]ACG could be detected. The results suggested that both glucoses are intact in the crystal lattice, but one on the T9 is disordered. Careful examination of the $2F_o - F_c$ map permitted the identification of three hydrated Mg^{2+} ions. Inclusion of the DNA, drug, magnesium ions, and 89 water molecules in the refinement produced a final *R*-factor of 17.3% with root mean square deviations (rmsd) in bond distances of 0.017 Å from the ideal values. The final atomic coordinates of the structure have been deposited in the Nucleic Acids Database (ID No. DDFB70). For the discussion of the structures, nucleotides are numbered from C1 to G6 in one strand; and C7 to G12 in the other strand, the two daunorubicin molecules are numbered D13 and D14.

(9) Wang, A. H.-J.; Ughetto, G.; Quigley, G. J.; Rich, A. *Biochemistry* **1987**, *26*, 1152−1163.

S0002-7863(96)03793-6 CCC: $14.00

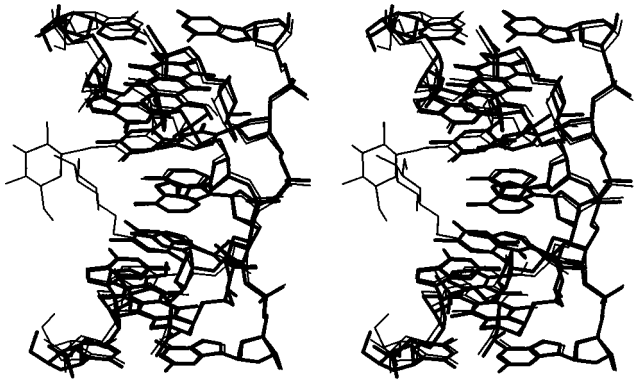

Figure 2. Stereoscopic skeletal drawing of the least square fits between the DNR−CGg^{5}TACG (from crystal space group *P*1) and DNR−CGTACG (from crystal space group $P4_12_12$) complexes. The rmsd between them is 0.51 Å, indicating very similar structures despite different DNA and different space groups.

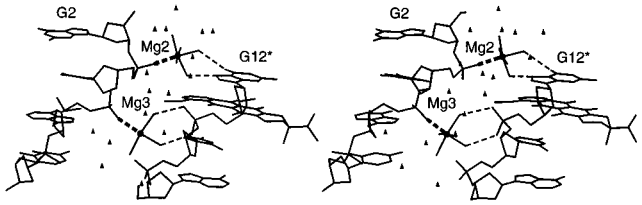

Figure 3. Stereoscopic skeletal drawing of the interactions between the two pentahydrated Mg^{2+} ions and the drug−DNA complex. The DNA strand of G12* is from the symmetry-related molecule at (x, $y-1$, z).

helix with the ring D protruding into the major groove and the daunosamine in the minor groove. The glucose moiety did not interfere with the binding of DNR to DNA. The detailed torsion angles and helix parameters of the complex are listed in Tables 2S and 3S in the Supporting Information.

Both glucoses adopt the chair conformation, and they have different glycosidic C5−M5−O1G−C1G torsion conformations with respect to the thymine base. The well-ordered glucose of g^{5}T3 is anchored firmly to the DNA surface using hydrogen bonds both directly and via a bridging water and a hydrated Mg^{2+} ion (Figure 1). The other disordered glucose of g^{5}T9 occupies two positions without specific hydrogen bonds to the DNA surface. Both glucoses extend from the hydroxymethylene tail of g^{5}T toward the 3′ direction, covering the major groove of the next two base pairs (A4/C5 and A10/C11, respectively).

There is no direct interaction between DNR and the glucose moieties. However, inspection of the structure suggests that binding of an intercalator at the (g^{5}T)pA or (g^{5}C)pG site will have a profound effect on the structure of the DNA. In the case of DNR binding to glucosylated DNA the protrusion of its ring D into the major groove may cause it to clash and interfere with the movement of the glucoses. Moreover, it is not clear what will be the effect on the juxtaposition of glucose and DNA caused by the helix unwinding associated with the intercalation.[15] (A helix unwinding would bring the glucose closer to the next 3′ base.) It would be of interest to investigate the binding of intercalators to fully-glucosylated DNA.

The complex was crystallized in a new P1 lattice (Figure 2S in the Supporting Information). In addition to the hexahydrated Mg^{2+} ion already mentioned (Figure 1), two pentahydrated Mg^{2+} ions were found to bind directly to non-ester phosphate oxygens (Figure 3). Interestingly one (Mg3) is directly bound to the *pro-S* of A4O1P, while the other (Mg2) to the *pro-R* of T3O2P oxygen atoms, respectively. These two hydrated Mg^{2+} ions are near each other and surrounded by negatively-charged phosphates in a manner that might be relevant in the cleaving/joining of phosphodiester bond by enzymes (polymerases or ribozyme).[10]

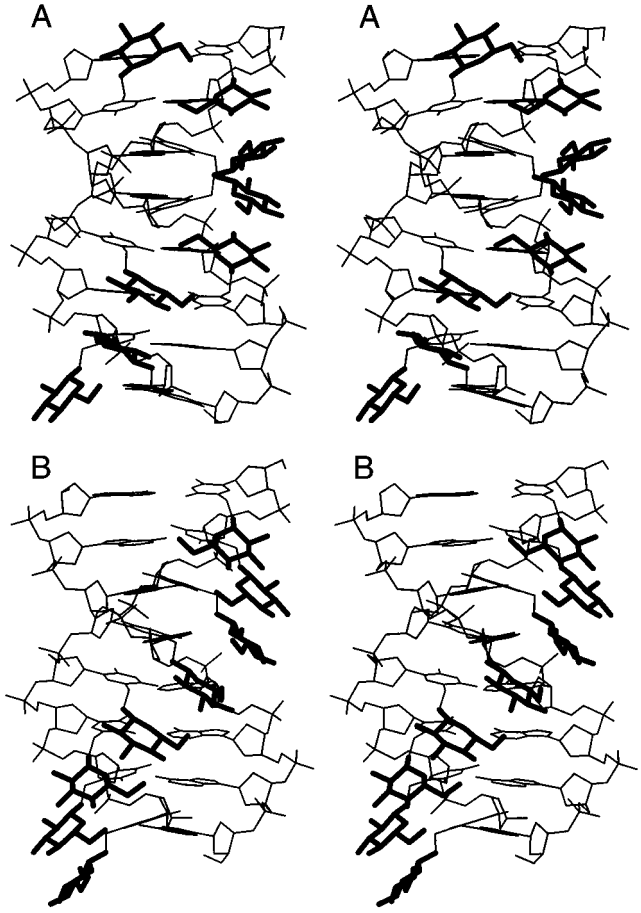

Figure 4. Stereoscopic skeletal drawing of the models of the fully-glucosylated d(g^{5}T-A)$_n$ and d(g^{5}T)$_n$.d(A)$_n$ duplexes. Both glucose conformations that were seen in the crystal structure could be used as the starting conformation to generate the fully-glucosylated DNA. The glucoses can be accommodated in the major groove of B-DNA without any steric hindrance.

In conclusion, our X-ray crystallographic structural studies reveal for the first time that glucoses attached at the C5 of pyrimidine sites adopt certain multiple, but preferred, conformations, covering about 3 base pairs toward the 3′ direction. Binding of an intercalator close to, but probably not at, the glucosylated site is possible. Further studies of other types of glucosylated (e.g., α- vs β-linked glucose, fully-glucosylated, other hexosylated, or oligosaccharide-attached) DNA duplexes will shed light on the possible biological roles of those highly unusual DNA modifications.[16]

Acknowledgment. This work was supported by grants from the American Cancer Society (DHP-114A) and NIH (GM-41612) to A.H.-J.W.

Supporting Information Available: Two figures showing the omit electron density map and the crystal packing diagram and three tables listing the crystal data, torsion angles, and helix parameters (6 pages). See any current masthead page for ordering and Internet access instructions.

JA963793R

(10) (a) Joyce, C. M.; Steitz, T. A. *J. Bacteriol.* **1995**, *177*, 6321−6329. (b) Scott, W. G.; Klug, A. *Trends Biochem. Sci.* **1996**, *21*, 220−224.

(11) Wijsman, E. R.; van den Berg, O.; Kuyl-Yeheskiely, E.; van der Marel, G. A.; van Boom, J. H. *Recl. Trav. Chim. Pays-Bas* **1994**, *113*, 337−338.

(12) Wang, A. H.-J.; Gao, Y.-G. *Methods* **1990**, *1*, 91−99.

(13) Rabinovich, D.; Shakked, Z. *Acta Crystallogr.* **1984**, *A40*, 195−200.

(14) Brünger, A. *X-PLOR*, version 3.1; The Howard Hughes Medical Institute and Yale University: New Haven, CT, 1993.

(15) Wang, A. H.-J. *Curr. Opin. Struct. Biol.* **1992**, *2*, 361−368.

(16) We have constructed models of the fully-glucosylated d(g^{5}T-A)$_n$ and d(g^{5}T)$_n$.d(A)$_n$ duplexes and showed that the glucoses can be accommodated in the major groove of B-DNA without any steric hindrance (Figure 4).