Binding of the Antitumor Drug Nogalamycin to Bulged DNA Structures†

Janet Caceres-Cortes and Andrew H.-J. Wang*

Biophysics Division, and Department of Cell and Structural Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801

Received August 7, 1995; Revised Manuscript Received November 13, 1995⊗

ABSTRACT: Defects in DNA, e.g., unpaired/bulged nucleotides, are repaired by specific repair enzymes. Understanding the dynamics and structure of DNA defects is important. Two DNA heptamers, CT_b-GTACG and $CGTACT_bG$, each containing a bulged T nucleotide embedded in the CpG step, have been studied by NMR. Both duplexes are significantly destabilized, and the bulged T remains intrahelical. Binding of the anthracycline antitumor antibiotic nogalamycin (Ng) to these two heptamers stabilizes the duplex structure. The solution structures of the 2:1 complexes of Ng-d(CT_bGTACG) and Ng-d($CGTACT_bG$) have been determined by the NOE-restrained refinement procedure. In both structures the elongated aglycon of Ng is intercalated between base pairs, and the nogalose and aminoglucose lie in the minor and major grooves, respectively. The bulged T behaves differently upon the binding of Ng. In Ng-CT_bGTACG wobble G_6:T_b base pairs are formed, leaving two dangling 5′-C_1 nucleotides; whereas in Ng-$CGTACT_bG$ weak C_1:T_b base pairs are formed, leaving two dangling 3′-G_6 nucleotides. Thus Ng induces the bulged T and the opposing base in the duplex to stack on the aglycon and causes the base next to T_b to unpair, mimicking a "frame-shift". Such structural rearrangement of a bulged DNA site due to the binding of an intercalator drug may perturb the recognition of DNA defects by repair enzymes or may cause mutation during replication.

Nucleic acids are capable of forming complex three-dimensional structures, e.g., Holliday junctions (Lilley & Clegg, 1993), triplexes (Radhakrishnan & Patel, 1994), tRNA (Kim et al., 1974a,b), pseudoknots (Pleij et al., 1986; Puglisi et al., 1990; Turek et al., 1992; Taylor et al., 1994), and ribozymes (Pley et al., 1994; Scott et al., 1995). These higher-ordered nucleic acid structures are often associated with important biological functions. Some structural motifs, e.g., bulges and hairpins, are ubiquitous components of the complex folded assembly, and they have been a topic of great interest (Turner, 1992). The behavior of bulges in DNA duplexes depends on many factors, including the sequence context, temperature, the length of the bulge region, etc. (Turner, 1992; Kalnik et al., 1989; Morden et al., 1982, 1990; Woodson et al., 1987, 1988; Rice & Crothers, 1989).

The property of bulged DNA structures may be influenced by binding of ligands. Of particular interest is whether many of the DNA-binding anticancer drugs affect the structure and dynamics of those higher-ordered structures. Several earlier studies showed that some intercalators (e.g., ethidium) bind to bulged structures more tightly than to normal B-DNA duplexes (Nelson & Tinoco, 1985; White & Draper, 1987; Williams & Goldberg, 1988). Recently, other more complicated drugs have been studied. For example, neocarzinostatin has been shown to cleave at a bulge site effectively (Kappan & Goldberg, 1993). Duocarmycin and SN-07 bound and were cross-linked to an unique site in a triple helix (Lin & Patel, 1992; Ye et al., 1993). The binding of actinomycin D to the stem region of a tetraloop hairpin has been studied by NMR (Brown et al., 1994). McConnaughie et al. (1994) have designed small ligands that bind to the HIV TAR and RRE RNA sequences. These research activities clearly showed that the interactions between DNA-binding anticancer drugs and unusual structures are being recognized as important issues for understanding the mode of action of those drugs.

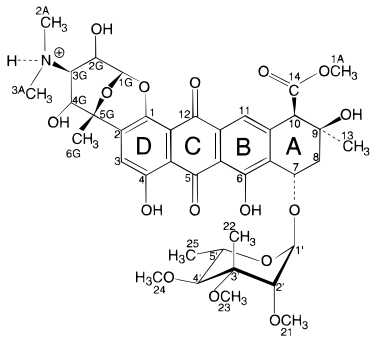

FIGURE 1: Molecular formula of the Ng drug with the numbering system used in the paper. The molecule contains an aglycon chromophore with four fused rings (A−D). Two sugars are attached to the aglycon with nogalose at C^7 and a positively charged α-D-3,6-dideoxy-3-dimethylaminoglucose (abbreviated aminoglucose in text) at C^1/C^2 positions.

Nogalamycin (Ng)[1] (Figure 1) is a member of the family of anthracycline antibiotics which includes several important antitumor drugs such as daunorubicin, doxorubicin, and

† This work was supported by a grant from the American Cancer Society (DHP-114) to A.H.-J.W. Support to J.C.-C. by an NIH Radiation, Biophysics, and Bioengineering in Oncology Traineeship (PHS 5 T32 CA09067) is acknowledged.

* To whom correspondence should be addressed.

⊗ Abstract published in *Advance ACS Abstracts,* December 15, 1995.

[1] Abbreviations: bp, base pair(s); T_b, bulged T nucleotide; oligo, oligonucleotides; Ng, nogalamycin; Ng-CGTACG, nogalamycin−d($CGTACG)_2$ complex; Ng-CT_bGTACG, nogalamycin−d($CT_bGTACG)_2$ complex; Ng-$CGTACT_bG$, nogalamycin−d($CGTACT_bG)_2$ complex; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; COSY, correlated spectroscopy; ppm, parts per million; T_1IR, T_1 relaxation inversion recovery; RMSD, root mean square deviation; SPEDREF, spectral-driven refinement; 1D and 2D, one and two dimensional.

0006-2960/96/0435-616$12.00/0

aclacinomycin A (Crooke & Reich, 1980; Lown, 1988; Denny, 1989; Wang, 1992). Ng is unique in that it contains two sugar moieties (nogalose and aminoglucose) attached to rings A and D, respectively, of the aglycon chromophore. How an intercalator with two bulky groups threads through the space between base pairs has generated considerable interest. The issue has been addressed in several ways. The DNA binding affinity and sequence specificity of Ng has been studied by DNase I footprinting experiments (Fox & Alam, 1992), theoretical studies (Brown & Neidle, 1988; Trinquier et al., 1988), NMR (Searle et al., 1988; Zhang & Patel, 1990; Searle & Bicknell, 1992; Robinson et al., 1994), and X-ray crystallographic analyses (Liaw et al., 1989; Gao et al., 1990; Williams et al., 1990; Smith et al., 1995). These studies showed that Ng molecules are intercalated between the 5′-NpG (or 5′-CpN) steps in a B-DNA double helix. The elongated aglycon chromophore (rings A−D) penetrates the DNA double helix such that it is almost perpendicular to the $C^{1'}-C^{1'}$ vectors of the two base pairs above and below the intercalator. The drug spans the two grooves of the helix with the nogalose and the aminoglucose occupying the minor and major grooves, respectively.

Here we address the issue of how intercalators bind to a bulged site in DNA duplexes using Ng as a probe. The solution structure of two DNA heptamers CT_bGTACG and $CGTACT_bG$, each containing a bulged-T sequence embedded in a CpG step of different sequence context, has been analyzed by NMR. The binding of Ng to such bulged DNA duplexes has been analyzed by determining the structures of the 2:1 Ng-CT_bGTACG and Ng-$CGTACT_bG$ complexes using NOE-restrained refinement procedures. Special attention has been paid to the positional effect of the bulged site.

MATERIALS AND METHODS

Sample Preparation. The oligonucleotide heptamers were synthesized on a DNA synthesizer in the Genetic Facility at UIUC. Nogalamycin was a gift from Dr. Paul Aristoff of the Upjohn Co. (Kalamazoo, MI). Solutions of Ng-CT_b-GTACG and Ng-$CGTACT_bG$ for NMR studies were prepared by dissolving the ammonium salt of the DNA heptamers plus appropriate amounts of Ng stock solution (in methanol) in 500 μL of phosphate buffer solution (25 mM sodium phosphate, pH 7.0, 0.15 M NaCl in 99.8% D_2O). The final complex (duplex) concentrations were 3 and 4 mM, respectively. Each solution was lyophilized three times with 99.8% D_2O and then dried in an NMR tube with a stream of nitrogen gas, and finally 500 μL of 99.96% D_2O was added to produce the samples. The free DNA samples of CT_bGTACG (1.5 mM) and $CGTACT_bG$ (2 mM) were similarly prepared.

NMR Data Collection. Both 1D and 2D NMR spectra were recorded on a Varian VXR500 spectrometer operating at 500 MHz for ^1H. T_1 relaxation experiments (T_1IR) were carried out with the standard $180°-\tau-90°$ inversion−recovery sequence. Phase-sensitive 2D NOESY experiments for nonexchangeable protons were performed at 2 °C for the free DNA samples, at 10 °C for Ng-CT_bGTACG, and at 20 °C for Ng-$CGTACT_bG$. The NOESY spectra were recorded using the TPPI−States technique (States et al., 1982) for phase cycling. A total of 512 t_1 increments each with 2048 t_2 complex points were collected over a sweep width of 5000 Hz and averaged for 32 transients per FID. The recycle delay was 4.41 s, and the mixing time was 100 ms. The 2D data sets were processed with the program FELIX (Hare Research, Woodinville, WA) using the Silicon Graphics workstations. Apodization of the data in the t_2 and t_1 dimensions consisted of 6 Hz exponential multiplication with a sine-bell squared function extending over the last quarter of the data to reduce truncation effects. The data set was zero-filled in t_1 to 2048 points prior to transformation. In addition to 2D NOESY, standard TOCSY spectra were collected to aid in the assignment process. Resonance assignments were made using standard methodology (Hare et al., 1983).

1D and 2D spectra of the exchangeable protons were collected in 90% H_2O/10% D_2O at 2 °C, unless otherwise noted. The experiments were performed using the 1:1 pulse sequence (Hore, 1983) as the read sequence to give on-resonance suppression of the solvent peak. Radiation damping was avoided using the method described by Sklenar et al. (1987). The offset was set to one quarter the sweep width to get maximum excitation. The 2D spectra were recorded as 512 t_1 increments of 2048 complex points in t_2 over a sweep width of 12 000 Hz and averaged over 16 or 32 transients per FID with a mixing time of 80 ms. The chemical shifts (in ppm) are referenced to the HDO peak which is calibrated to 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) at different temperatures.

Structural Refinement. Starting models of the free DNA and Ng-bound DNA complexes were built using QUANTA (version 4.0, MSI, Massachusetts, MA). Structural refinement of these molecules were carried out by the procedure SPEDREF (Robinson & Wang, 1992). The experimental 2D NOESY data sets were analyzed using MYLOR (Robinson & Wang, 1992) to define the line shapes and chemical shifts for each spin in each frequency domain. These line shapes were then used to determine the volumes of the NOESY cross peaks. The inversion−recovery experiment determined the T_1 relaxation time for every spin. The correlation time τ_c was determined using the SPEDREF procedure and found to be 7.6, 3.7, 3.0, and 4.4 ns for Ng-CT_bGTACG, Ng-$CGTACT_bG$, CT_bGTACG, and $CGTACT_bG$, respectively. The starting models were refined within the program X-PLOR (Brünger, 1992) in conjunction with the SPEDREF procedure. During the first 40 refinement cycles, the molecules were given a random set of velocities equivalent to 300 K. NOE-restrained molecular dynamics were then run for 0.1 ps, the temperature coupling bath lowered by 25 K, and the process repeated until the molecules had been cooled to 40 K. Each molecular dynamics cycle was carried out for a total of 10.4 ps. At this point, molecular dyad symmetry was slowly imposed while running an additional 100 cycles of NOE-restrained conjugate gradient minimization. During the last 20 cycles, the molecules were refined with only NOE-restrained conjugate gradient minimization. The simulated NOE relaxation rates and NOE intensities for the refined models were calculated by the program MORASS (Post et al., 1990). Simulated 2D NOESY spectra were produced by the program CSL (in SPEDREF package) using the NOE intensities from the simulation and the line shapes and chemical shifts previously obtained from MYLOR. The NMR R factor is defined as R factor $= \sum|N_o - N_c|/\sum N_o$, where N_o and N_c are the experimental and calculated NOE integrals, respectively.

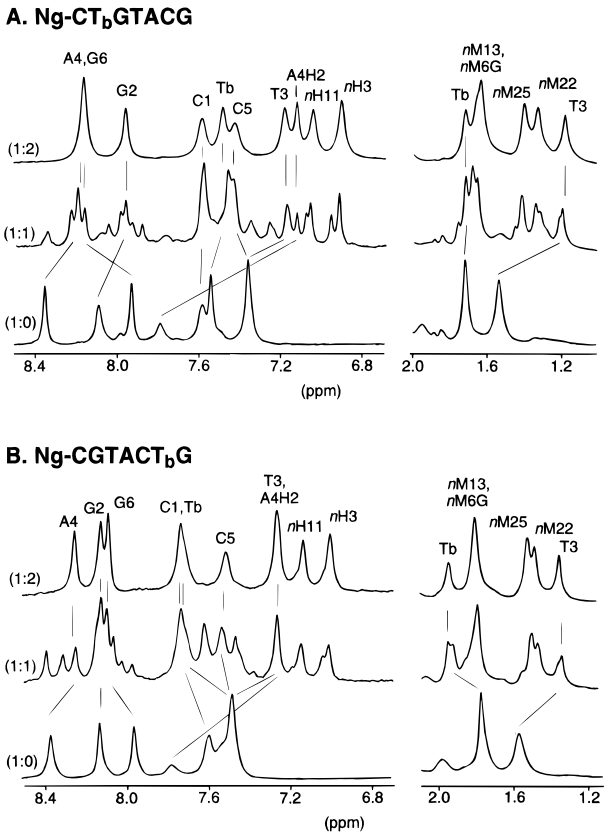

FIGURE 2: Titration of Ng to two bulged-T DNA heptamers, d(CT_b-GTACG)$_2$ and d($CGTACT_bG$)$_2$ monitored by proton 1D NMR spectroscopy. Methyl and aromatic regions of the free DNA (1:0), one Ng molecule per DNA duplex (1:1) and two Ng molecules per duplex (1:2) are shown. The chemical shift of some resonances (e.g., A_4H^2 and T_3Me) changed dramatically in going from the free DNA to the Ng-DNA form. The spectra were recorded in D_2O at 10 °C.

RESULTS

NMR Analysis. Binding of Ng to two bulged-T heptamers has been monitored by 1D NMR spectroscopy, Figure 2. Both DNAs formed well-defined 2:1 complexes with Ng. When Ng is mixed with the DNA duplex in a 1:1 ratio, the spectrum of each complex shows resonances from the free DNA, 1:1 Ng-DNA complex, and 2:1 Ng-DNA complex. This indicates that the binding equilibrium is slow on the NMR time scale since resonances from all three species coexist. The Ng-CT_bGTACG complex has sharper resonances (average 15 Hz) than those of the Ng-$CGTACT_bG$ complex (average 20 Hz).

The complete experimental 2D NOESY spectra of all four systems (CT_bGTACG, $CGTACT_bG$, Ng-CT_bGTACG, Ng-$CGTACT_bG$) are shown in Figure S1. All resonances, except those of the $H^{5'}/H^{5''}$ protons, have been assigned stereospecifically using a standard sequential assignment procedure similar to that used in the analysis of Ng-CGTACG (Robinson et al., 1990). The 2D TOCSY data (not shown) are consistent with the 2D NOESY assignment. Figure 3 shows the aromatic to $H^{1'}/H^5$ fingerprint region of the four molecules. The complete chemical shifts are listed in Table 1 for CT_bGTACG and Ng-CT_bGTACG and in Table 2 for $CGTACT_bG$ and Ng-$CGTACT_bG$, respectively. The chemical shifts of most of the resonances of the two Ng-DNA complexes vary only slightly in raising the temperature from 5 to 50 °C (data not shown), suggesting that they are stable.

CT_bGTACG and $CGTACT_bG$. The 2D NOESY spectra clearly indicated that the resonances from $CGTACT_bG$ are significantly broader than the corresponding resonances from CT_bGTACG, suggesting that the former molecule is a poorly structured duplex. This is corroborated by their NMR spectra recorded in H_2O (Figure 4) and the T_1IR measurement. Comparison of exchangeable proton NMR spectra reveals that the imino proton resonances from CT_bGTACG are significantly stronger than those from $CGTACT_bG$. Thus the influence of a bulged-T on the structure depends on the sequence context. The T_b nucleotide inserted in the CpG step on the 5′-end of CGTACG has a less destabilizing effect than that inserted in the 3′-end.

The intensity pattern of the cross peaks in Figure 3A,B suggests that T_b in both heptamers remains intrahelical. For example, the $C_1H^{1'}-T_bH^6$ and $G_2H^8-T_bH^{1'}$ cross peaks in

Table 1: ^{1}H NMR Chemical Shifts in the Nogalamycin-d(C-T_b-G-T-A-C-G)$_2$ Complex

		Proton Chemical Shifts of Free and Bound d(CT_bGTACG)$_2$[a]									
		H8/H6	H2/H5/M5	H1′	H2′	H2″	H3′	H4′	H5′/H5″	NH	NH2
C1	f-DNA	7.58	5.73	6.01	2.01	2.41	4.55	4.02	3.64/3.60		6.73, 7.07
	b-DNA	7.57	5.65	5.99	2.14	2.51	4.64	4.12	3.74/3.68		
T_b	f-DNA	7.56	1.71	5.92	2.18	2.32	4.81	4.18	4.05/3.96		
	b-DNA	7.48	1.72	6.04	2.70	2.64	5.01	4.25	4.19/4.12		
G2	f-DNA	8.11		6.09	2.82	2.92	4.99	4.46	4.19/4.06	12.99	
	b-DNA	7.95		5.75	2.56	2.74	4.98	4.69	4.32/4.21	12.53	8.45, 4.09
T3	f-DNA	7.38	1.53	5.67	2.16	2.45	4.91	4.06	4.19/4.00	13.59	
	b-DNA	7.17	1.18	5.80	2.56	2.15	4.87	4.10	4.18/4.10	13.45	
A4	f-DNA	8.37	7.79	6.31	2.74	2.92	5.08	4.48	4.18/4.03		8.09, 6.40
	b-DNA	8.16	7.11	6.24	2.42	2.84	5.05	4.29	4.20/4.17		7.50, 5.80
C5	f-DNA	7.37	5.40	5.88	1.96	2.34	4.78	4.26	4.18/4.07		8.35, 6.86
	b-DNA	7.41	5.27	5.92	2.37	2.16	4.99	4.17	4.19/4.16		7.99, 6.15
G6	f-DNA	7.95		6.13	2.72	2.45	4.67	4.19	4.14/4.09		
	b-DNA	8.16		6.29	2.54	2.48	4.73	4.32	4.18/4.15		

Proton Chemical Shifts of Bound Nogalamycin[a]											
H3	6.89	H8E	2.74	H1′	5.20	H4′	3.18	H1G	5.66	M3A	2.87
H11	7.03	M13	1.66	H2′	3.13	M24	3.58	H2G	4.30	H4G	3.92
H10	4.33	M1A	3.80	M21	3.61	H5′	3.84	OH2G	8.79	OH4G	8.95
H7	4.85	OH4	11.56	M22	1.32	M25	1.40	H3G	4.85	M6G	1.63
H8X	2.01	OH6	11.42	M23	3.25			M2A	3.10		

[a] Buffer was 0.1 NaCl/25 mM phosphate aqueous solution. Free DNA proton data at 2 °C. Bound DNA and drug proton data at 10 °C.

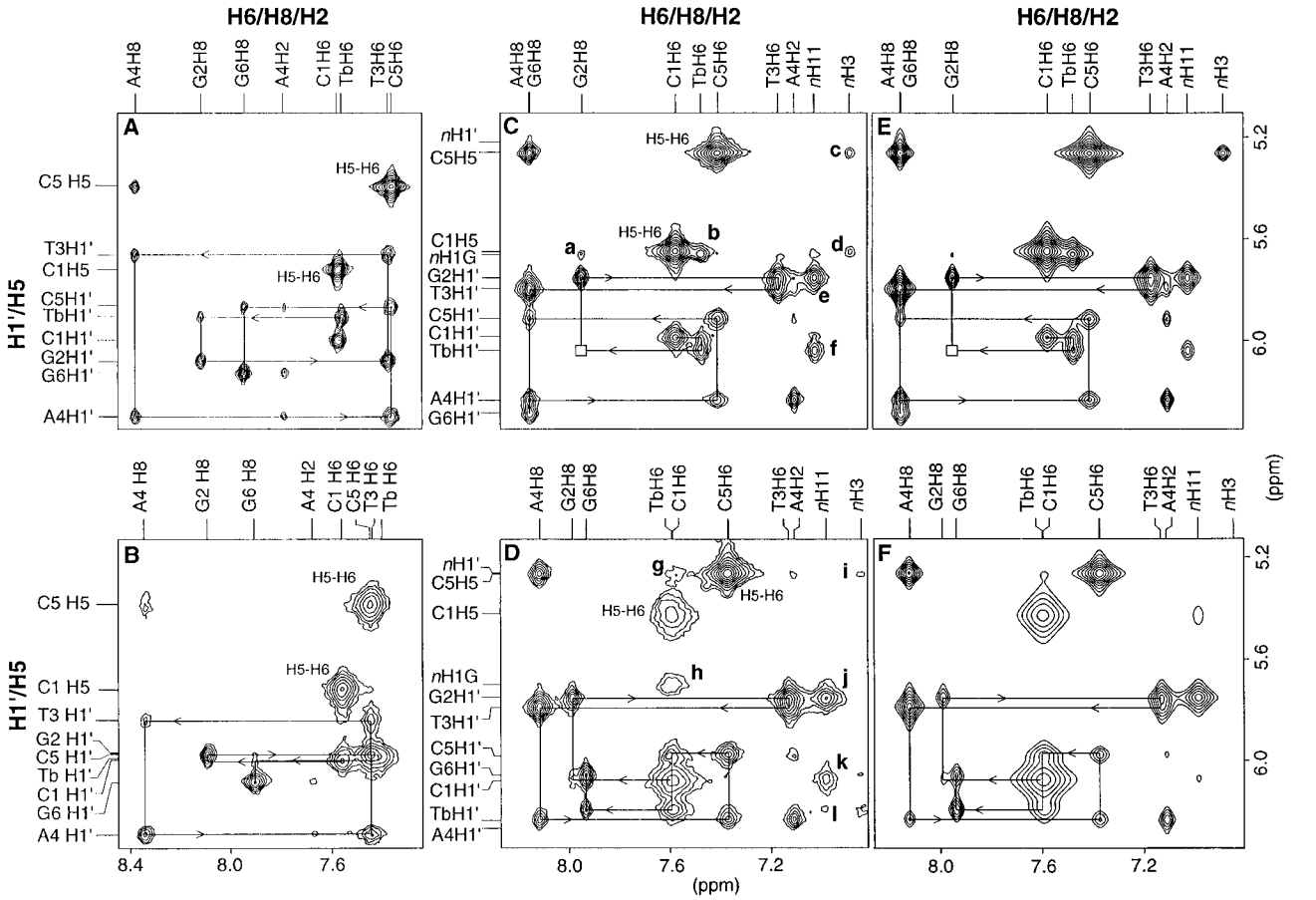

FIGURE 3: Experimental and simulated 2D-NOESY spectra showing the crosspeaks between $H^{1'}/H^5$ and aromatic protons. (A) CT_bGTACG. (B) $CGTACT_bG$. (C) 2:1 Ng-CT_bGTACG complex. (D) 2:1 Ng-$CGTACT_bG$ complex. (E) Simulated spectra of 2:1 Ng-CT_bGTACG complex (*R* factor 21.2%). (F) Simulated spectra of 2:1 Ng-$CGTACT_bG$ complex (*R* factor 26.0%). The sequential assignment pathway is illustrated. The cytidine H^5-H^6 cross peaks are labeled. Cross peaks between Ng and DNA are labeled a−f for Ng-CT_bGTACG and g−l for Ng-$CGTACT_bG$. The unusually strong $C_1H^{1'}-C_1H^6$ cross peak in D suggests a *syn*-glycosyl conformation for the C_1 nucleotide in Ng-$CGTACT_bG$.

Table 2: ^{1}H NMR Chemical Shifts in the Nogalamycin-d(C-G-T-A-C-T_b-G)$_2$ Complex

		Proton Chemical Shifts of Free and Bound d($CGTACT_bG$)$_2$[a]									
		H8/H6	H2/H5/M5	H1′	H2′	H2″	H3′	H4′	H5′/H5″	NH	NH2
C1	f-DNA	7.57	5.73	6.00	1.87	2.35	4.62	4.11	3.73/3.70		
	b-DNA	7.59	5.42	6.08	2.72	2.72	4.92	4.30	3.89/3.81		
G2	f-DNA	8.09		5.98	2.80	2.84	4.98	4.43	4.11/4.09	12.85	
	b-DNA	7.99		5.76	2.60	2.73	5.03	4.69	4.19/4.12	12.49	8.32, 4.08
T3	f-DNA	7.46	1.54	5.83	2.25	2.56	4.92	4.24	4.18/4.06	13.69	
	b-DNA	7.13	1.20	5.79	2.55	2.13	4.86	4.08	4.20/4.15	13.48	
A4	f-DNA	8.33	7.68	6.29	2.69	2.88	5.05	4.43	4.20/4.07		8.08, 6.42
	b-DNA	8.12	7.11	6.23	2.39	2.80	5.03	4.26	4.15/4.07		8.03, 6.19
C5	f-DNA	7.46	5.39	5.99	2.20	2.36	4.83	4.24	4.27/4.15		8.33, 6.99
	b-DNA	7.37	5.26	5.98	2.33	2.09	4.94	4.16	4.25/4.12		7.58, 5.70
T_b	f-DNA	7.41	1.75	6.00	1.95	2.27	4.73	4.11	4.05/4.02		
	b-DNA	7.59	1.78	6.19	2.33	1.96	4.85	4.33	4.14/4.08		
G6	f-DNA	7.91		6.08	2.68	2.44	4.65	4.11	4.06/4.03		
	b-DNA	7.93		6.06	2.62	2.39	4.67	4.16	4.09/4.07		

Proton Chemical Shifts of Bound Nogalamycin[a]											
H3	6.84	H8E	2.71	H1′	5.26	H4′	3.15	H1G	5.70	M3A	2.89
H11	6.99	M13	1.64	H2′	3.15	M24	3.56	H2G	4.30	H4G	3.93
H10	4.30	M1A	3.80	M21	3.64	H5′	3.81	OH2G	8.82	OH4G	8.94
H7	4.91	OH4	11.65	M22	1.32	M25	1.37	H3G	4.89	M6G	1.61
H8X	1.99	OH6	11.44	M23	3.25			M2A	3.11		

[a] Buffer was 0.1 M NaCl/25 mM phosphate aqueous solution. Exchangeable proton data and free DNA nonexchangeable proton data at 2 °C. Bound DNA and drug nonexchangeable proton data at 20 °C.

CT_bGTACG are clearly detectable. Additional cross peaks from T_b to both C_1 and G_2 are also observed (Figure S2). For CT_bGTACG and $CGTACT_bG$, 843 and 933 NOE cross peak integrals, respectively, were measured for use in the SPEDREF refinement. Some of the relevant NOEs are shown schematically in Figure S3, and their refined structures

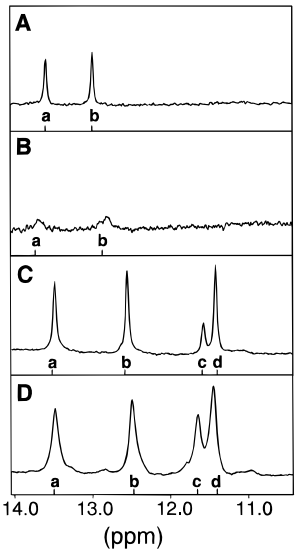

FIGURE 4: 1D NMR spectra of the imino/phenolic proton region. Peaks a, b, c, and d are from T_3N^3H, G_2N^1H, nO^4H, and nO^6H, respectively. (A) CT_bGTACG. (B) $CGTACT_bG$. (C) 2:1 Ng-CT_b-GTACG complex. (D) 2:1 Ng-$CGTACT_bG$ complex.

shown in Figure S4. Even though in both structures T_b is wedged between the base pairs of a CpG step, the structural consequences are different. In CT_bGTACG the terminal G_6:C_1 base pair is partially disrupted so that G_6 is in a position to base pair with T_b. Similarly, C_1 is in a position to base pair with T_b in $CGTACT_bG$. Since a wobble G:T base pair is more stable than a C:T base pair, the latter duplex is destabilized to a greater extent.

Ng-CT_bGTACG. The 2D NOESY data of both 2:1 complexes suggest that their conformations are significantly perturbed, but to a different degree on the two DNA strands. Many cross peaks between the Ng drug resonances and those from T_b, G_2, C_5, and G_6, but not C_1, nucleotides from DNA are found in the Ng-CT_bGTACG spectra, (e.g., peaks a−f in Figure 3C). Some of those cross peaks have their counterparts from the 2:1 Ng-CGTACG complex (Robinson et al., 1994). It was concluded that the intercalation sites are at the T_bpG_2:C_5pG_6 steps. The strong cross peaks between nH^{11} (of Ng) and $T_bH^{1'}$ (peak f) and $G_2H^{1'}$ (peak e) (Figure 3C) indicate that the aglycon ring is oriented such that the edge with the nH^{11} proton is facing the T_bpG_2 backbone, as in the crystal structure (Liaw et al., 1989; Gao et al., 1990) and solution structure (Robinson et al., 1994) of the 2:1 Ng-CGTACG complex. NOE cross peak intensities from other regions are also consistent with the interpretation (Figure S2), and many relevant NOE cross peaks are summarized in Figure 5A. A starting model of the 2:1 Ng-CT_bGTACG complex was prepared by modifying the 2:1 Ng-CGTACG structure (Robinson et al., 1994) to include two T_b:G_6 wobble base pairs, with the terminal C_1 base stacked beyond the T_b:G_6 base pair. The NMR *R* factor for the refined Ng-CT_bGTACG model is 21.2%. The agreement between the experimental and calculated NOE data can partly be seen in Figure 3C,E.

Ng-$CGTACT_bG$. Binding of Ng to $CGTACT_bG$ appears to stabilize the DNA duplex significantly, as revealed from the comparison of the 1D imino proton resonances (Figure 4). In free DNA, the imino protons of T_3 (13.69 ppm) and G_2 (12.85 ppm) are barely detectable (Figure 4B). In contrast, the corresponding resonances are readily detected at 13.48 and 12.49 ppm, respectively, in the 2:1 Ng-$CGTACT_bG$ complex (Figure 4D). In addition, the phenolic proton nO^4H and nO^6H resonances appeared strongly at 11.65 and 11.44 ppm, respectively. It should be noted, however, that the resonances of the Ng-$CGTACT_bG$ complex (Figure 4D) are still somewhat broader than those of the Ng-CT_bGTACG complex (Figure 4C).

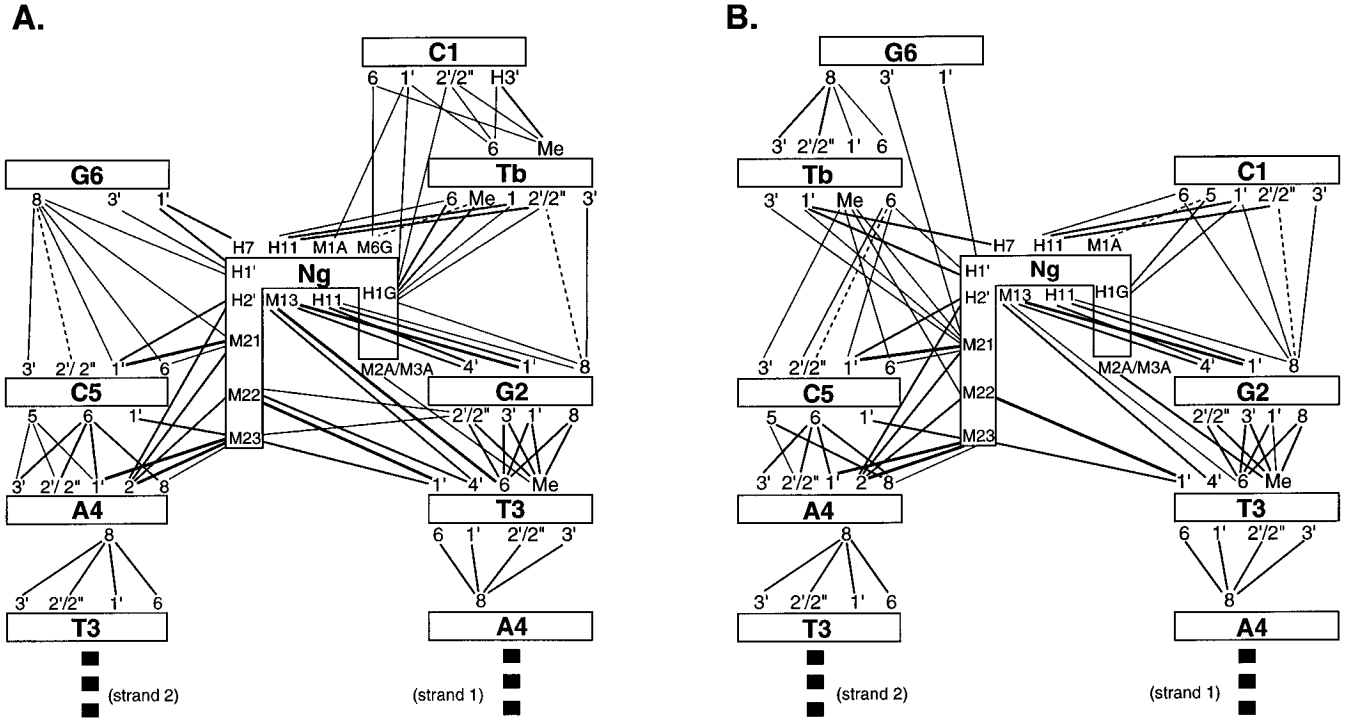

FIGURE 5: Schematic diagram showing important internucleotide and Ng-DNA cross peaks observed in the experimental 2D NOESY spectra of Ng-CT_bGTACG (A) and Ng-$CGTACT_bG$ (B). Straight lines with various thickness denote corresponding NOE strength. Dashed lines indicate overlapped cross peaks.

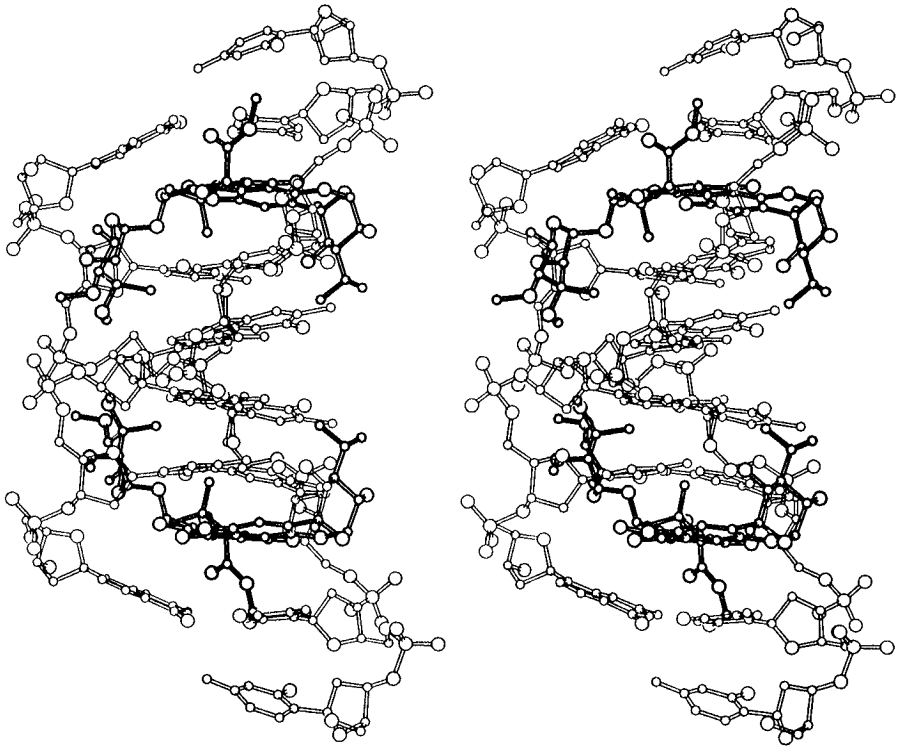

FIGURE 6: Stereodrawing of the Ng-CT$_b$GTACG complex. Nogalamycin (darker bonds) intercalates between the T$_b$pG steps, positioning the nogalose sugar in the minor groove and the aminoglucose sugar in the major groove. A G:T wobble base pair is formed causing the cytosine on the 5′-end to dangle.

There are unusual features in the 2D NOESY spectrum of the 2:1 complex (Figure 3D). Note that the resonances associated with the C_1 residue are especially broad. In addition, the $H^{1'}-H^6$ cross peak of C_1 is stronger than its H^5-H^6 cross peak, suggesting that C_1 nucleotide is in the *syn* conformation to a large extent. There are many cross peaks from Ng to T_b, C_5, C_1, and G_2, residues from DNA, but only a few to G_6, some of which are shown as peaks g−k in Figure 3D. Similar strong cross peaks between *n*H^{11} (of Ng) and $C_1H^{1'}$ (peak k) and $G_2H^{1'}$ (peak j) (Figure 3D) indicate that Ng is intercalated at the C_1pG_2 and the C_5pT_b steps with its aglycon ring oriented in such a way that the edge with the *n*H^{11} proton is facing the C_1pG_2 backbone. NOE cross peak intensities from other regions are also consistent with the interpretation and some relevant NOE cross peaks are summarized in Figure 5B. A starting model of the 2:1 Ng-CGTACT$_b$G complex was similarly prepared to include two possible C_1:T_b base pairs and subjected to NOE-restrained refinement. The refined Ng-CGTACT$_b$G model that has C_1 glycosyl bond in a *syn* conformation (Figure S5) has an NMR *R* factor of 26.0%. Portions of the simulated spectra of this model appear in Figure 3E and Figure S6. Note that a few observed NOE cross peaks (e.g., peaks h and k) in Figure 3D do not appear in the simulated spectra of this model. They are attributed to the minor DNA conformer with the terminal C_1 in the *anti* glycosyl conformation (Figure S5 and S6).

Structures of Bulged DNA Bound with Ng. Binding of Ng to bulged-T DNA stabilizes both duplexes. The imino proton spectra (Figure 4) indicate that the duplex is significantly more destabilized in Ng-CGTACT$_b$G than in Ng-CT$_b$GTACG. Note that the Watson−Crick base paired imino protons in the Ng-DNA complexes are substantially longer-lived than those in free DNA.

The structure of the refined 2:1 Ng-CT$_b$GTACG complex is shown in Figure 6. Two Ng drugs are intercalated in the two symmetry-related T_bG_2:C_5pG_6 sites. The large buckle between the C_5 and G_6 bases in Ng-CT$_b$GTACG, as that seen in the Ng-CGTACG, remains a conspicuous feature. The relatively strong $C_5H^{1'}-G_6H^8$ NOE cross peak (Figure 3C) suggests that the sugar of C_5 is close to the G_6 base, despite an intercalated drug. Furthermore, the G_6:T_b base pair, due to its wobble configuration, allows the bases to have good stacking interactions with the Ng aglycon ring (Figure 7A). The C^6-O^6 bond is stacked below the G_{12} (i.e., G_6) base, whereas the C^1, C^{12} region of ring D is stacked below the T_b base. The backbone of G_{12} moves closer to the nogalose, making the minor groove narrower. The nogalose together with the C^{13} methyl group enclose the N^2 amino group of G_2 and shield it from the solvent region. This may explain the unusual observation of the G_2N^2 amino protons (Figure 8). In the major groove, there are possible hydrogen bonds between the O^{2G} hydroxyl and N^7 of G_2 (2.84 Å) and between the O^{4G} hydroxyl and N^4 of C_{11} (3.11 Å).

Ring A of Ng has the α conformation in which the carbomethoxy group on the C^{10} position adopts an axial position, almost perpendicular to the plane of the aglycon, and its keto oxygen O^{14} atom receives a hydrogen bond from the NH_2 of G_{12}. Due to the wobble G:T configuration, the Ng's methyl group (M^{1A}) moves slightly away from the sugar of C_1 residue. The aglycon ring of Ng in both structures is slightly bent so that the nogalose and the aminoglucose are brought closer toward each other (Liaw et al., 1989; Gao et al., 1990). The smooth bending of Ng chromophore was detected in the crystal structure and other NMR studies (Zhang & Patel, 1990). However, such subtle distortion derived from NMR refinement may be influenced by the applied energy functions and thus should be taken cautiously.

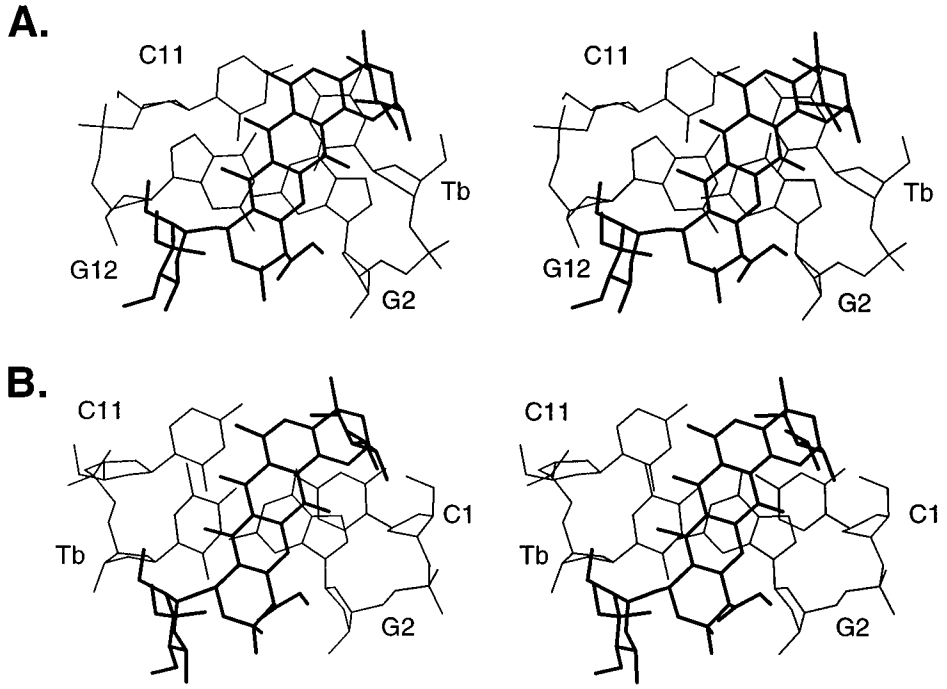

FIGURE 7: (A) Stacking diagrams (top view) surrounding the aglycon of Ng. (A) Ng-CT_bGTACG complex. (B) Ng-$CGTACT_bG$ complex. For clarity, only the two base pairs directly adjacent to the Ng molecules are shown and the dangling 5′-C or 3′-G residues are omitted. The Ng molecules appear as thick lines, and the nucleotides (thin lines) are labeled.

The structure of the 2:1 Ng-$CGTACT_bG$ complex is not as well-refined as that of the Ng-CT_bGTACG complex due to several factors. As mentioned before, the terminal C_1 nucleotide is in equilibrium between two conformations, predominantly in the *syn* conformation and less frequently in the *anti* conformation. This has resulted in two sets of NOE cross peaks (Figure S6); thus any single model would not account for the observed NOEs. Figure 6B shows the stacking of Ng with its neighboring base pairs. The (*syn*)-C_1 and T_b bases, due to their small size, need to come closer together in order to form a reasonable base pair. However, the process may be hindered by the steric van der Waals clashes between the Ng and DNA deoxyriboses. Consequently the (*syn*)C_1:T_b base pair would be less stable than the wobble T_b:G_6 bp in Ng-CT_bGTACG. Nevertheless, despite the somewhat unfavorable C:T bp, the intercalation of Ng in $CGTACT_bG$ helped stabilize the DNA structure without forcing the bulged T_b to become extrahelical.

The quality of the refined structures was evaluated using the observed NOE cross peaks of the exchangeable proton NOE spectra (Figure 8). Since the NOE cross peaks involving the exchangeable protons were not included in the refinement, they served as experimental controls in checking the validity of the refined models. Figure 8 shows expanded regions of the exchangeable proton 2D NOESY of Ng-CT_b-GTACG. Several inter-drug−DNA cross peaks are selected for analysis. The cross peaks between nH^3 and C_5N^4H (peaks n and p) are moderately strong, consistent with the distances ($C_5N^4Ha-nH^3 = 2.61$ Å, and $C_5N^4Hb-nH^3$=2.63 Å) derived from the refined structure (Figure 6).

The strong appearance of the cross peak between the geminal protons of the G_2 amino group (peak u) is interesting, since it is normally not seen due to its broadness resulting from the slow rotation of the C^2-N^2 bond in G. Here, the $G-N^2$ amino group is hydrogen bonded to Ng and is shielded from solvent by Ng, thereby making the $G-N^2$ protons readily detectable. A number of strong NOE cross peaks between G_2N^2 protons and Ng are detected, including G_2N^2-Ha−nH2′ (2.81 Å) and $G_2N^2Ha-nM^{22}$ (3.48 Å) (data not shown). Consequently, it is concluded that the refined structure can satisfactorily explain all experimental cross peaks in both the nonexchangeable and exchangeable 2D NOESY spectra.

DISCUSSION

The structural basis of the sequence specificity of Ng binding to DNA has been significantly clarified by several recent studies, including NMR studies of the 2:1 complexes of Ng-$d(CGTACG)_2$ (Robinson et al., 1994), Ng-$d(GCATGC)_2$ (Searle et al., 1988), Ng-$d(AGCATGCT)_2$ (Zhang & Patel, 1990), and 1:1 complexes of Ng-$d(GACGTC)_2$ (Searle & Bicknell, 1992) and Ng-d(GCGT)·d(ACGC) (van Houte et al., 1993), and the crystal structures of Ng-d[m^5-$CGT(pS)m^5CG]_2$ (Gao et al., 1990; Williams et al., 1990), Ng-$d[CGT(pS)ACG]_2$ (Liaw et al. 1989), and Ng-$d(TGATCA)_2$ (Smith et al., 1995). It has been suggested that Ng has a DNA sequence preference for 5′-NpG or 5′-CpN steps. More specifically, the aglycon prefers to intercalate at the 5′-side of a guanine (between NpG) or at the 3′-side of a cytosine (between CpN) with the sugars facing toward the G-C base pair (Gao et al., 1990). This is likely due to the hydrogen bonds between the drug and DNA both in the major groove [i.e., the two hydroxyl groups (O^{2G} and O^{4G}) and N^7 and O^6 of G] and minor grooves (i.e., the carbomethoxy of Ng and N^2 of G).

Here we ask what is the consequence when the preferred Ng binding sequence, CpG, is interrupted by a bulged-T nucleotide. Previous work has shown that the structural consequence of a bulged nucleotide depends on the type of nucleotide (e.g., purines G and A or pyrimidines T and C). Bulged G or A tend to remain intrahelical and cause a kink in the helix (Kalnik et al., 1989; Woodson et al., 1988; Patel et al., 1982), whereas the behavior of a bulged C or T is more variable; either intrahelical or extrahelical depending

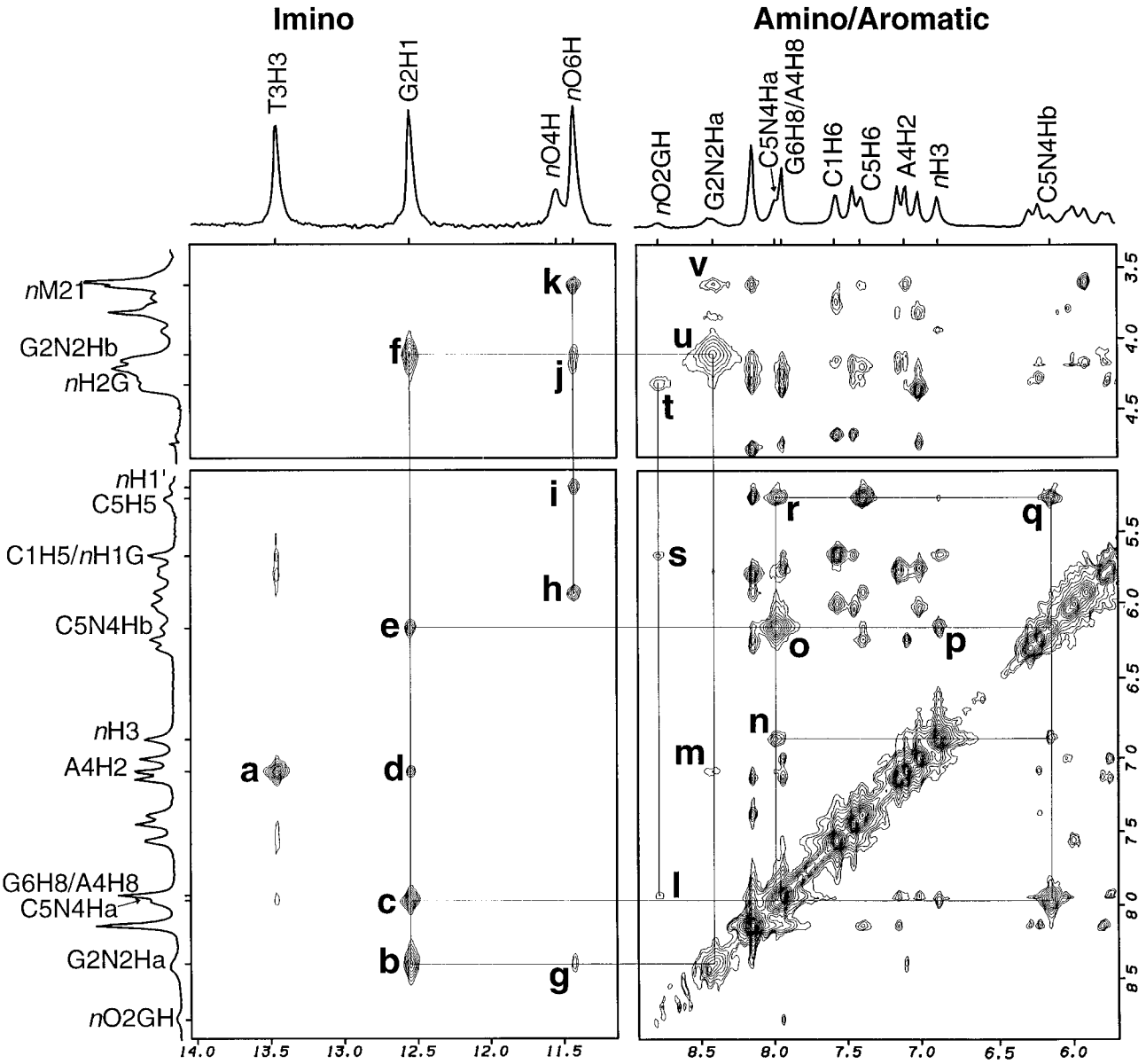

FIGURE 8: Portions of the exchangeable proton 2D NOESY spectra of the 2:1 Ng-CT_bGTACG complex. Selected cross peaks are marked. Note that the cross peak from the geminal proton of G_2-N^2 amino group appears prominently (peak u) whereas the cross peak from the geminal proton of A_3-N^6 amino group is not evident.

on sequence context (Kalnik et al., 1990; Turner, 1992) and temperature (Kalnik et al., 1989, 1990). The results from our analysis of two free DNA heptamers, CT_bGTACG and $CGTACT_bG$, suggested that the bulged-T destabilizes the duplexes significantly, with the latter DNA being more perturbed. In both molecules the bulged T is intrahelical, resulting in a severe distortion of the terminal C_1:G_6 base pairs.

We focused here on what happens to the bulged-T DNA duplex upon binding of Ng. We were interested in determining whether in the two bulged-T heptamers, the bulged-T would be forced out of the duplex (making it extrahelical) such that Ng becomes intercalated between two C:G base pairs. Our results showed, on the contrary, that DNA prefers to maintain an uninterrupted backbone at the intercalator site surrounding the bound Ng with the aglycon stacked over the central C:G base pair and the sugars (nogalose and aminoglucose) wrapped around the base pair. The other face of the aglycon is stacked with modified base pairs. In Ng-CT_bGTACG, the 3′-terminal G_6/G_{12} is "pulled" down to form a wobble G:T_b base pair, leaving the C_1 residue dangling. In Ng-$CGTACT_bG$, the 5′-terminal C_1 is "pulled" down to form a C:T_b base pair, leaving the G_6 residue dangling. It should be noted that a pyrimidine−pyrimidine C:T base pair is significantly less stable than a normal Watson−Crick (A-T or G-C) base pair, or even a wobble G-T base pair. Furthermore, the size of a C-T base pair is insufficiently large to span across the aglycon. Both factors affect the stability of the Ng-$CGTACT_bG$ complex. However, it should be emphasized that despite the imperfection of the mismatched base pairs, Ng still substantially stabilizes the bulged-T DNA structure, as evident from the strong imino proton resonances (Figure 4). The binding of Ng to different DNA sequences is schematically shown in Figure 9.

The present results may be used to understand the possible effect of Ng to an isolated bulged-T site in DNA. Normally, an intrahelical bulged-T site in free DNA causes DNA to bend and the duplex is destabilized (Rice & Crothers, 1989). Binding of an intercalator to DNA bulges may induce the bulged base to form a base pair over the intercalator ring,

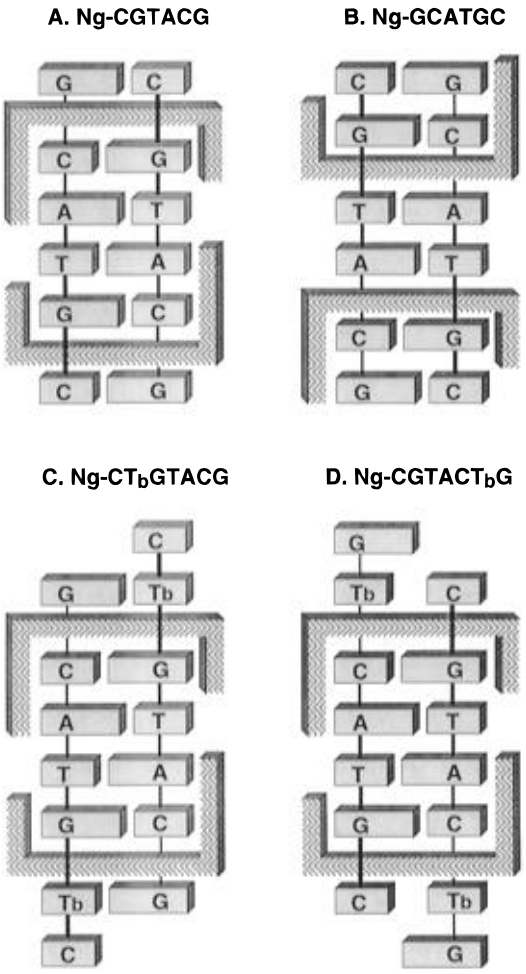

FIGURE 9: Schematic representations showing the binding of Ng to four different DNA duplexes. (A) Ng-CGTACG (Robinson et al., 1990) contains two Ng molecules intercalated between CpG steps with the sugars directed toward the center of the duplex. (B) Ng-GCATGC (Searle et al. 1988) contains two Ng molecules intercalated between CpA steps with the Ng sugars pointing toward the ends of the duplex. In both (C) Ng-CT_bGTACG and (D).CG-$TACT_bG$ Ng intercalated between NpG sites. A G:T wobble base pair is formed in (C) and a weaker T:C base pair is formed in (D) causing 5′-C and 3′-G dangling nucleotides, respectively.

thereby shifting the bulged distortion away from the original location. Thus a long-range effect of conformational change, propagated away from the initial bulged site due to the binding of intercalators, should be taken into account regarding the biological function of intercalator antitumor drugs.

Some intercalators are suspected to be mutagens. It is interesting to note, based on the results of this work, that at the replication fork or the site of transcription, an intercalator may facilitate a mismatched base pair to form. Such process is likely to introduce a mutation in the resulting DNA or RNA daughter strand.

A number of new issues remain to be addressed. For example, what will be the consequence when a larger bulged site (i.e., longer than one base) is bound with an intercalator? Is the effect dependent on the type of intercalators (e.g., actinomycin D) or the type of the bulged bases? We have recently discovered that actinomycin D significantly stabilizes the duplex structure with the $(CAG)_n$ and $(CTG)_n$ triplet sequences (C. Lian, H. Robinson, and A. H.-J. Wang, manuscript submitted). Additional studies will enhance our understanding on the biological activities of many DNA-binding antitumor drugs.

ACKNOWLEDGMENT

We thank Dr. Howard Robinson for his technical assistance.

SUPPORTING INFORMATION AVAILABLE

Six figures showing the complete experimental 2D NOESY spectra of the DNA and Ng-bound DNA molecules, expanded regions of the 2D NOESY spectra of the nonexchangeable protons, schematic diagram showing cross peak intensities for CT_bGTACG and $CGTACT_bG$, refined structures of CT_bGTACG, $CGTACT_bG$ and Ng-$CGTACT_bG$ with C_1 in both a *syn* and an *anti* glycosyl conformation, and three tables listing the torsion angles of the nogalamycin bound molecules (11 pages). Ordering information is given on any current masthead page.

REFERENCES

Brown, D. R., Kurz, M., Hsu, V. L., & Kearns, D. R. (1994) *Biochemistry 33*, 651−664.

Brown, J. R., & Neidle, S. (1988) in *Anthracycline and Anthracenedione-Based Anticancer Agents* (Lown, J. W., Ed.) Elsevier, New York.

Brünger, A. T. (1993) X-PLOR, version 3.2, The Howard Hughes Medical Institute and Yale University, New Haven, CT.

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

Denny, W. A. (1989) *Anti-Cancer Drug Des. 4*, 241−263.

Fox, K. R., & Alam, Z. (1992) *Eur. J. Biochem. 209*, 31−36.

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

Gold, L., Polisky, B., Uhlenbeck, O., & Yanus, M. (1995) *Annu. Rev. Biochem. 64*, 763−797.

Hare, D. R., Wemmer, D. E., Chou, S. H., Drobny, G., & Reid, B. R. (1983) *J. Mol. Biol. 171*, 319−336.

Hore, P. J. (1983) *J. Magn. Reson. 54*, 539−542.

Kalnik, M. W., Norman, D. G., Zagorski, M. G., Swann, P. F., & Patel, D. J. (1989a) *Biochemistry 28*, 294−303.

Kalnik, M. W., Norman, D. G., Swann, P. F., & Patel, D. J. (1989b) *J. Biol. Chem. 264*, 3702−3712.

Kalnik, M. W., Norman, D. G., Li, B. F., Swann, P. F., & Patel, D. J. (1990) *J. Biol. Chem. 265*, 636−647.

Kappen, L. S., & Goldberg, I. H. (1993) *Biochemistry 32*, 13138−13145.

Kim, S.-H., Suddath, F. L., Quigley, G. J., McPherson, A., Sussman, J. L., Wang, A. H.-J., Seeman, N. C., & Rich, A. (1974a) *Science 185*, 435−440.

Kim, S.-H., Sussman, F. L., Quigley, G. J., McPherson, A., Wang, A. H. J., Seeman, N. C., & Rich, A. (1974b) *Proc. Natl. Acad. Sci. U.S.A. 71*, 4970−4974.

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.

Lilley, D. M. J., & Clegg, R. M. (1993) *Q. Rev. Biophys. 26*, 131−175.

Lin, C. H., & Patel, D. J. (1992) *J. Am. Chem. Soc. 114*, 10658−10660.

Lown, J. W., Ed. (1988) *Anthracycline and Anthracenedione-Based Anticancer Agents*, Elsevier, New York.

McConnaughie, A. W., Spychala, J., Zhao, M., Boykin, D., & Wilson, D. W. (1994) *J. Med. Chem. 37*, 1063−1069.

Morden, K. M., Chu, Y. G., Martin, F. H., & Tinoco, I., Jr. (1982) *Biochemistry 22*, 5557−5563.

Morden, K. M., Gunn, B. M., & Maskos, K. (1990) *Biochemistry 29*, 8835−8845.

Nelson, J. W., & Tinoco, I., Jr. (1985) *Biochemistry 24*, 6416−6421.

Patel, D. J., Kozlowski, S. A., Marky, L. A., Rice, J. A., Broka, C., Itakura, K., & Breslauer, K. J. (1982) *Biochemistry 21*, 445−451.
Pleij, C. W. A., van Belkum, A., Rietveld, K., & Bosch, L. (1986) Pseudoknots in RNA: A Novel Folding Principle, in *Structure and Dynamics of RNA* (van Knippenberg, P. H., Hilbers, C. W., Eds.) Plenum Press, New York.
Pley, H. W., Flaherty, K. M., & Mckay, D. B. (1994) *Nature 372*, 68−74.
Post, C. B., Meadows, R. P., & Gorenstein, D. G. (1990) *J. Am. Chem. Soc. 112*, 6796−6803.
Puglisi, J. D., Wyatt, J. R., & Tinoco, I., Jr. (1990) *J. Mol. Biol. 214*, 437−453.
Radhakrishnan, I., & Patel, D. (1994) *Biochemistry 33*, 11405−11416.
Rice, J. A., & Crothers, D. M. (1989) *Biochemistry 28*, 4512−4516.
Robinson, H., & Wang, A. H.-J. (1992) *Biochemistry 31*, 3524−3533.
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.
Robinson, H., Yang, D., & Wang, A. H.-J. (1994) *Gene 149*, 179−188.
Scott, W. S., Finch, J. T., & Klug, A. (1995) *Cell 81*, 991−1002.
Searle, M. S., & Bicknell, W. (1992) *Eur. J. Biochem. 205*, 45−58.
Searle, M. S., Hall, J. G., Denny, W. A., & Wakelin, L. P. G. (1988) *Biochemistry 27*, 4340−4349.
Smith, C. K., Davis, G. J., Dudson, E. J., & Moore, M. H. (1995) *Biochemistry 34*, 415−425.
Sklenar, V., Brooks, B. R., Zon, G., & Bax, A. (1987) *J. Magn. Reson. 75*, 352−357.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286−292.
Taylor, E. W., Ramanathan, C. S., Jalluri, R. K., & Nadimpalli, R. G. (1994) *J. Med. Chem. 37*, 2637−2654.
Trinquier, G., Chen, K. X., & Gresh, N. (1988) *Biopolymers 27*, 1491−1517.
Tuerk, C., MacDougal, S., & Gold, L. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 6988−6992.
Turner, D. (1992) *Curr. Opin. Struct. Biol. 2*, 334−337.
van Houte, L. P. A., van Garderen, C. J., & Patel, D. J. (1993) *Biochemistry 32*, 1667−1674.
Wang, A. H.-J. (1992) *Curr. Opin. Struct. Biol. 2*, 361−368.
White, S. A., & Draper, D. E. (1987) *Nucleic Acids Res. 15*, 4049−4064.
Williams, L. D., & Goldberg, I. H. (1988) *Biochemistry 27*, 3004−3011.
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.
Woodson, S. A., & Crothers, D. M. (1987) *Biochemistry 26*, 904−912.
Woodson, S. A., & Crothers, D. M. (1988) *Biochemistry 27*, 3130−3141.
Ye, X. M., Kimura, K., & Patel, D. J. (1993) *J. Am. Chem. Soc. 115*, 9325−9326.
Zhang, X., & Patel, D. J. (1990) *Biochemistry 29*, 9451−9466.

BI9518398