Induced Formation of a DNA Bulge Structure by a Molecular Wedge Ligand−Postactivated Neocarzinostatin Chromophore†

Xiaolian Gao,*,‡ Adonis Stassinopoulos,§,|| Jie Ji,‡ Youngjoo Kwon,‡ Simona Bare,‡ and Irving H. Goldberg*,§

Department of Chemistry, University of Houston, 4800 Calhoun Street, Houston, Texas 77204-5003, and Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115

Received December 3, 2001; Revised Manuscript Received February 11, 2002

ABSTRACT: Our previous structure elucidation of the complexes of DNA and postactivated neocarzinostatin chromophore (NCS-chrom) compounds revealed two distinctly different binding modes of this antitumor molecule. A thorough understanding of these results will provide the molecular basis for the binding and DNA chain cleavage properties of NCS-chrom. NCSi-gb is one of the postactivated mimics of NCS-chrom which is formed under thiol-free conditions and is able to bind to DNA. This report describes the structure refinement of the NCSi-gb-bulge-DNA complex [Stassinopoulos, A., Jie, J., Gao, X., and Goldberg, I. H. (1996) *Science 272*, 1943−1946] and the NMR characterization of the free bulge-DNA and free NCSi-gb. These results reveal that the formation of the complex involves conformational changes in both the DNA and the ligand molecule. Of mechanistic importance for the NCS-chrom−DNA interaction, the two ring systems of the drug are brought closer to each other in the complex. This conformation correlates well with the previously observed marked enhancement of the formation of a DNA bulge cleaving species in the presence of bulge-DNA sequences, due to the promotion of the intramolecular radical quenching of the activated NCS-chrom. Interestingly, the binding of NCSi-gb promotes the formation of a bulge binding pocket; this was not found in the unbound DNA. NCS-chrom is unique among the enediyne antibiotics in its ability to undergo two different mechanisms of activation to form two different DNA binding and cleaving species. The two corresponding DNA complexes are compared. One, the bulge-DNA binder NCSi-gb, involves the major groove, and the second, the duplex binder NCSi-glu which is generated by glutathione-induced activation, involves the minor groove. Since the two NCS-chrom-related ligand molecules contain some common chemical structural elements, such as the carbohydrate ring, the striking differences in their DNA recognition and chain cleavage specificity provide insights into the fundamental principles of DNA recognition and ligand design.

For more than 3 decades, efforts in understanding the antitumor agent neocarzinostatin chromophore (NCS-chrom,[1] Figure 1) have provided valuable knowledge in the unique cytotoxic molecular mechanisms and diverse chemistry of this intricate molecule (*1*, *2*). In its natural form, this small molecule exists as a protein complex, and it belongs to a family of enediyne antitumor agents. Extensive studies established that NCS-chrom causes DNA damage through hydrogen abstraction (H-abstraction) from DNA strands by a NCS-chrom diradical intermediate (*1*−*3*). The drug itself undergoes ring rearrangement to form more stable compounds (postactivated NCS-chrom, e.g., NCSi-gb, cyclospirolactone, and NCSi-glu shown in Figure 1), which have been shown to be isostructural forms of DNA binding and cleaving species (*1*, *2*). Several distinctly different DNA cleavage patterns involving single- and double-stranded sequences have been observed. In double strand cleavage of duplex DNA induced by thiol activation, the observed sequence and H-abstraction specificities suggest the formation of high-affinity precursors of the drug and DNA complexes. One of the major lesions in duplex DNA caused by treatment with thiol-activated NCS-chrom is located at trinucleotide AG<u>C</u>·GC<u>T</u> or AG<u>T</u>·AC<u>T</u> (residues of H-abstraction underlined) sites. Other drastically different, well-defined single strand lesion sites are associated with DNA sequences, which contain extrahelical residues imbedded within complementary base regions (and thus are potential bulge sites), and are associated with the general base-catalyzed intramolecular

† This research is supported by the Robert A. Welch Foundation (E-1270) and the National Foundation of Cancer Research (X.G.) and U.S. Public Health Service Grant GM53793 from the NIH (I.H.G.).

* To whom correspondence should be addressed. X.G.: tel, 713 743-2805; fax, 713 743-2709; e-mail, xgao@uh.edu. I.H.G.: tel, 617 432-1787; fax, 617 432-0471; email, irving_goldberg@hms.harvard.edu.

‡ University of Houston.

§ Harvard Medical School.

|| Current address: Cerus Corp., Concord, CA 94520.

[1] Abbreviations: 1D, one dimensional; 2D, two dimensional; CC, cyclocarbonate; COSY, correlation spectroscopy; DG, distance geometry; DQF, double quantum filter; g, gauche; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; MDS, molecular dynamics simulations; NA, dihydronaphthone; NPH, hydroxynaphthoate; NCS, neocarzinostatin; NCS-chrom, native form of NCS chromophore; NCSi-gb, postactivated NCS-chrom formed under base-catalyzed conditions; NCSi-glu, postactivated NCS-chrom formed under reductive conditions in the presence of a thiol (glutathione) agent; NMF, 2-*N*-methylfucosamine; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect and exchange spectroscopy; rMDS, restrained molecular dynamics simulations; RMSD, root-mean-square deviation; t, trans; THI, tetrahydroindacene; TOCSY, total correlation spectroscopy. Nucleic acids: A, adenosine; C, cytidine; G, guanosine; T, thymidine.

10.1021/bi012112o CCC: $22.00
Published on Web 03/27/2002

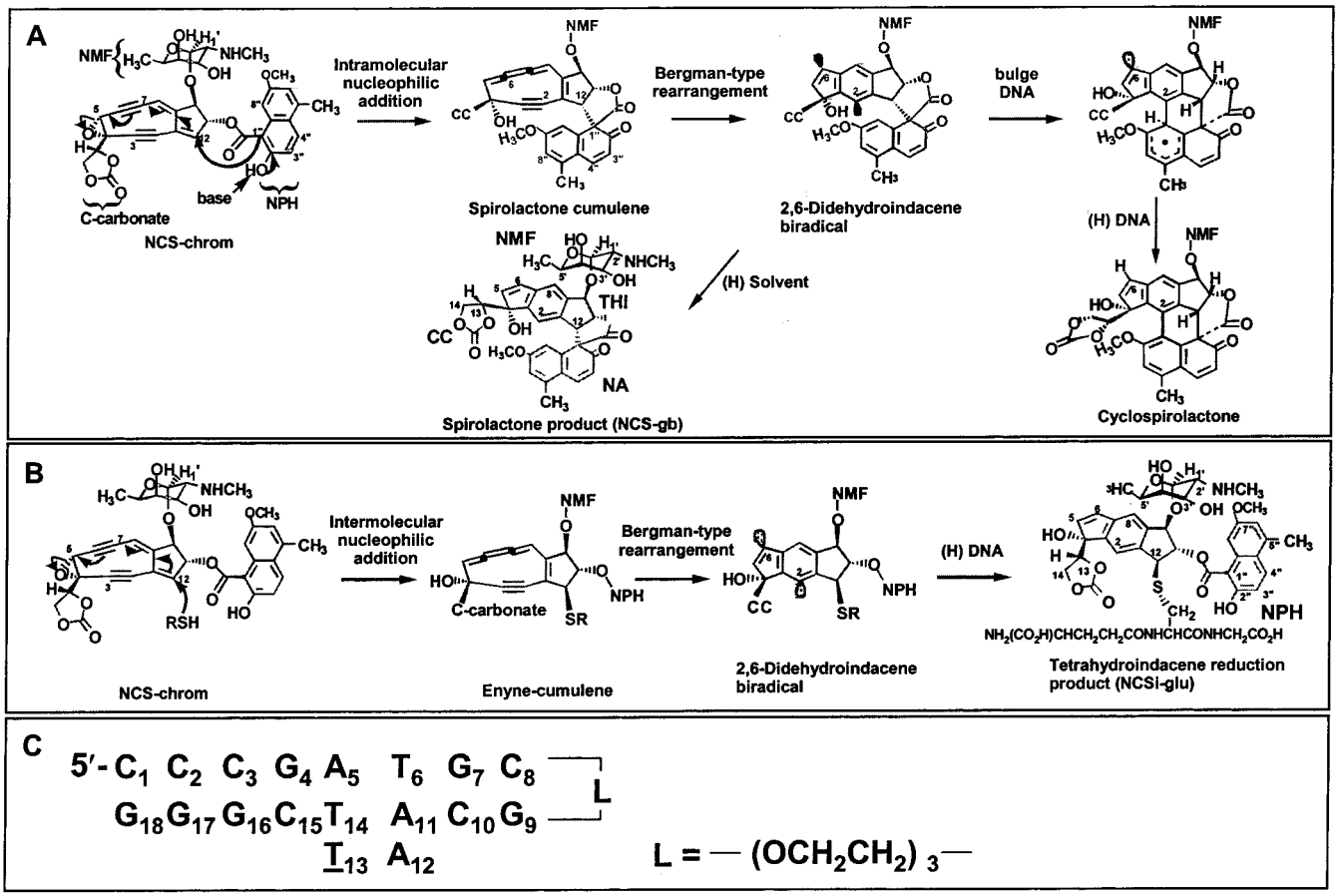

FIGURE 1: (A) Formation of NCSi-gb by postactivated NCS-chrom under basic conditions and in the absence of a thiol agent. (H) DNA and (H) solvent indicate the processes of H-abstraction by DNA or the solvent. (B) Formation of NCSi-glu in the presence of a thiol agent. (C) DNA sequence in the NCSi-gb-bulge-DNA complex.

activation of the drug (*4*). The characterization of the regio- and stereochemistry of NCS-chrom and its DNA damage products can be found in a series of studies and is reviewed in the references given (*1*, *2*).

A fundamentally important question raised by the elucidation of the distinct modes of action of NCS-chrom on DNA concerns the basis of molecular recognition in these complex systems. How can a single molecule manifest such distinctly different attack patterns of DNA? Using high-resolution solution NMR, the structures of two representative complexes of postactivated NCS-chrom (NCSi-glu and NCSi-gb, Figure 1A,B) and oligo-DNA have been elucidated (*5*−*8*). A third new structure of a related complex will be reported in a separate paper (unpublished work). These NMR studies established that the binding and cleavage of DNA by NCS-chrom could be examined in detail by using the stable postactivated forms of the drug. This was necessary since natural NCS-chrom rapidly decomposes under physiological conditions. Fortunately, certain postactivated species of the drug are isostructural forms of the binding and/or cleaving species of the activated NCS-chrom. The experimentally observed specificity in binding and cleavage are readily accounted for in the structure of each complex.

The high-resolution structure of the NCSi-glu-DNA complex (sequence shown in Figure 1C) as a model for double strand DNA cleavage by NCS-chrom has been examined in detail (*5*, *6*). This result provides an important link to a complete mechanistic description of the drug action at sequence-specific double strand lesion sites. The structure reveals a minor groove and an intercalation binding mode by NCSi-glu, which interacts with AGC on one strand and with GCTC on the other strand. The naphthoate moiety intercalates in the 5′-AG/5′-CT step, while the *N*-(methylamino)pyranose unit and the reduced enediyne ring form a binding patch in the minor groove. The formation of the complex is accompanied by a significant conformational change in the DNA duplex, which unwinds and stretches relative to the helical axis and exhibits buckled base pairs at the intercalation site. The major features of the complex are dominant van der Waals intermolecular contacts and complementary electrostatic surface match. The sequence specificity of drug binding has been attributed to a set of specific interactions, mostly associated with its hydrophobic interactions with DNA H1′ and aromatic adenine H2 and guanosine amino protons in the minor groove. The structure of the complex positions the drug H6 close to 5′ of T in GCT and drug H2 close to 1′ of C in AGC. These are the sites of radical formation upon activation of NCS-chrom by thiol, and thus the structure of the complex provides a rationale for the specific cross-strand cleavage at the staggered T and C residues. Since H-abstraction by the drug radical is a rapid reaction, the binding and cleaving species specific to duplex DNA should closely resemble its isostructural form, postactivated NCSi-glu.

A novel mode of NCS-chrom interaction with DNA in the absence of thiol activating agent is associated with single strand cleavage of unpaired sequences, which can potentially form bulge structures (*9*, *10*). When the complex formed with

postactivated drug (the isostructural form of the biradical species) under thiol-free conditions was examined, a stable bulge-DNA was found (*7*). The DNA used in this study, AAT̲T·AT (cleaved T residue underlined) (Figure 1C), contained unpaired bases and had a specific T cleavage. The most striking finding from structure elucidation of the complex was the unusual binding mode adopted by NCSi-gb. The drug molecule is a major groove binder using its stacked ring systems to fill the bulge cavity and with the binding pocket surrounded by DNA aromatic bases, forming a pyramidal binding site. The carbohydrate unit of NCSi-gb is situated in the center of the major groove floor. The H6 of NCSi-gb is in close vicinity and directed toward the 5′ of the underlined T in AAT̲T in the structure, explaining single strand cleavage of DNA at bulge residues.

The unexpected structure of the NCSi-gb-bulge-DNA complex prompts further questions. Is the bulge binding an active or passive event, i.e., does NCS-chrom induce the formation or simply enter a preformed bulge site? Given the distinct binding modes exhibited by the postactivated NCS-chrom species, how do these similar compounds discriminate minor or major groove binding? What are the mechanistic implications when different complexes are compared? To address these questions, in this work, the details of the NMR characterization of the NCSi-gb-bulge-DNA and free bulge-DNA will be described; the conformation of the free NCSi-gb and its bound form will be analyzed; structural features of the NCSi-glu-duplex DNA and NCSi-gb-bulge-DNA complexes will be contrasted. These studies also shed light on the role of the DNA bulge in generating the active DNA cleaving species.

MATERIALS AND METHODS

Sample Preparation. *(a) NCSi-gb*. NCSi-gb was prepared as described (*10*) with NCS-chrom extracted from the holoprotein (*11*). The resulting extraction solution (CH_3OH/acetic acid) was evaporated in the dark at 277 K, and the residue solution containing NCS-chrom was redissolved in 5 mM acetic acid in CH_3OH to give a solution of 2 mM, with a recovery efficiency ca. 50%. The extracted NCS-chrom was incubated in HEPES (50 mM) with less than 2% CH_3OH present for 4 h at 277 K, pH 7.0. NCSi-gb was isolated by semipreparative reverse-phase HPLC.

(b) Free DNA Sequence. 5′-d(CCCGATGC-L-GCAATTCGGG) (L = 2 × triethylene glycol) was synthesized using the standard phosphoramidite chemistry on an ABI 381A synthesizer, HPLC purified, and desalted using size exclusion and Na-exchange columns. The purity of the sample was examined using reverse-phase C_{18} HPLC and electrophoresis in a 20% denaturing polyacrylamide gel.

(c) NCSi-gb-Bulge-DNA Complex. 5′-d(CCCGATGC-L-GCAATTCGGG) was dissolved in a solution containing 0.1 M NaCl, 10 mM sodium phosphate, and 0.1 mM EDTA (unless otherwise indicated, this condition is referred as buffer solution). NCSi-gb dissolved in CH_3OH/H_2O was gradually added at 283 K, and the ^{1}H (proton) NMR spectrum was recorded to follow the formation of the complex. The formation of the complex gave rise to new NMR signals and also resulted in the disappearance of original signals. The final NMR sample contained the NCSi-gb-bulge-DNA complex (ca. 1 mM) at pH 6.3. The presence of ca. 10% of free DNA was discernible from NMR exchangeable spectra.

NMR Experiments. NMR experiments were performed on spectrometers of 500 MHz at Harvard Medical School and 600 MHz at the University of Houston. D_2O (99.96%, Cambridge Isotopes Inc.) or 10% D_2O/90% H_2O were used as solvents for observation of nonexchangeable and exchangeable ^{1}H, respectively. ^{1}H chemical shifts were referenced to the HOD resonance (4.70 ppm at 298 K, temperature correction factor −0.0109 ppm/K). ^{31}P chemical shifts were referenced relative to an external trimethyl phosphate in an aqueous solution containing 0.1 M NaCl (pH 6.5). One-dimensional (1D) ^{1}H NMR spectra were centered at the HOD resonance with spectral widths of 10 ppm for D_2O samples and 23 ppm for H_2O samples. Data points, 4096 or 8192, were acquired. The 1D ^{31}P NMR spectra were centered at −3.5 ppm with a spectral width of 4−10 ppm; 1024 or 2048 data points were acquired. These parameters were used in acquiring two-dimensional (2D) ^{1}H−^{1}H NMR spectra, but data points in t_2 were half of that for 1D spectra (for samples in D_2O, the acquisition time was 682 ms for COSY type of spectra and 341 ms for NOESY; for samples in H_2O, the acquisition time was 158 ms) and were 512 in t_1 (for samples in D_2O, t_{1max} = 171 ms; for samples in H_2O, t_{1max} = 79 ms). The 2D ^{1}H−^{31}P correlation spectra used a sweep width of 3.3 ppm in t_2 (^{1}H) and 4.6 ppm in t_1 (^{31}P), and the acquisition times were 512 and 133 ms for the ^{1}H and ^{31}P dimensions, respectively. The 2D data recorded include DQF-COSY, COSY-35, TOCSY, NOESY, and ^{1}H−^{31}P correlation spectra. Suppression of the HOD resonance was achieved by presaturation for D_2O samples or the jump−return pulse with a delay of 52 μs for H_2O samples. T_1 relaxation times of ^{1}H resonances were estimated from the null point in inversion recovery experiments using the relationship $T_1 = \tau_{null}/\ln(2)$. The estimated T_1 helped in setting repetition delays in 2D experiments and the identification of A H2 resonances. The NMR experiments used in this work are well documented in standard NMR textbooks (*12*). NMR data were processed using the UXNMR program (Bruker Instruments, Inc.) and the FELIX v2.30 and v95 programs (Accelrys Inc.). NOESY processing used a 90° phase-shifted sine bell in both t_2 and t_1 dimensions prior to Fourier transformation; TOCSY and DQF-COSY processing used a 30° phase-shifted skewed sine bell; ^{1}H−^{31}P 2D processing used a 37.5° phase-shifted skewed sine bell function. The final spectral matrix consisted of 2048 × 2048 data points in each of the ^{1}H frequency (F_2 and F_1) dimensions; ^{1}H−^{31}P contained 1024 × 1024 data points.

(a) NMR Spectra of the Free NCSi-gb. NMR spectra of the free NCSi-gb dissolved in the D_2O buffer solution, pH 6.4, were recorded at 275 and 283 K. The 2D DQF-COSY, TOCSY (78 and 120 ms mixing times), and NOESY (120 and 300 ms mixing times) were recorded. 2D data sets were acquired with a 2.5 s relaxation delay.

(b) NMR Spectra of the Free Bulge-DNA. The free duplex (∼2 mM) was dissolved in the buffer solution, pH 6.4. The 2D NOESY spectra of exchangeable protons were acquired at 270 K (0.9 s relaxation delay, 100 ms mixing time), and those of nonexchangeable protons were acquired at 278 or 283 K (2.1 s relaxation delay, 100 and 250 ms mixing times). DQF-COSY, COSY-35, TOCSY, and ^{1}H−^{31}P correlation spectra were obtained at 283 K.

(c) The NCSi-gb-Bulge-DNA Complex. The 2D exchange spectra were recorded in 90% H_2O/10% D_2O (100 and 150

ms mixing times, 270 K) and D_2O (70, 140, and 250 ms mixing times, 278 and 283 K) buffer solutions. The 2D $^1H-^{31}P$ COSY, DQF-COSY, COSY-35, and TOCSY (80 ms mixing time) were recorded in D_2O. Data acquisition and processing were performed using similar conditions as described for the free duplex (vide supra). Chemical shift assignments were based on the 250 ms mixing time NOESY at 283 K.

NMR Resonance Assignment. The ^1H and ^{31}P chemical shift assignments are provided in Table S1 (Supporting Information). (a) NCSi-gb of the free and the bound forms. NCSi-gb contains several subunits (Figure 1A), which include the reduced enediyne system, tetrahydroindacene with an attached cyclocarbonate ring (THI), the dihydronaphthone (NA), and the *N*-methylpyranose (NMF). These resonances were assigned on the basis of their coupling and NOE connectivities as described previously (*5*−*7*). The resonance assignments also benefited from several signals, such as *N*-methyl group (H2M), the NA H3″ and H4″, and the THI H5 and H6, which were easily seen in NOESY and COSY (used collectively for DQF-COSY, COSY-35, and TOCSY) spectra. (b) The bulge-DNA in the free and the bound forms. The assignments began with DQF-COSY spectra for identification of C H5 and H6 and T H5 (the 5-methyl) and H6 resonances and identification of spin systems of sugar residues. The A H2 resonances were identified from the inversion recovery experiments in which A H2 was shown as inverted peaks when most other resonances were buried within the noise level. These nonresidue-specific assignments provided the starting point for sequential assignments of DNA nonexchangeable ^1H resonances from a combination of NOESY and COSY spectral analysis and exchangeable HN (imino protons) and NH_2 (amino protons) as discussed in the literature (*13*) and in our previous publications (*5*−*7*). The final stage of the assignments was assisted using a grid search program developed in house, which required a list of cross-peaks observed from NOESY and other 2D NMR spectra and a list of chemical shifts (unnecessary to be complete, but included all available assignments). A table was generated for all possible assignments for each cross-peak. This cross-peak assignment table was analyzed to provide new chemical shift assignments and cross-peak assignments. The confirmed assignments were then used to generate the final cross-peak assignments. This process was iterated by editing the new chemical assignments into the chemical shift file.

NMR Restraints (*14*). The statistics of NMR restraints used for NCSi-gb-bulge-DNA structure calculation are summarized in Table 1. NOE cross-peak volumes of the NCSi-gb-bulge-DNA complex were measured using the FELIX program. The assigned spin pairs were converted to distance restraints using the two-spin approximation. The upper and lower distance bounds initially were at 50% of the equilibrium distances. The maximum interproton distance was generally limited to 5.0 and 5.5 Å for nonmethyl and methyl ^1H, respectively. The minimum interproton distance limit was set to 1.8 Å. Distance restraints were supplemented by the information derived from other data sets, such as the spectra recorded in H_2O for hydrogen bonds in Watson−Crick base pairing. There were 18 NOEs involving exchangeable protons, from which a set of distance restraints was estimated on the basis of peak intensities (strong, 1.8−3.8 Å; medium, 2.0−4.5 Å; weak, 2.5−4.5 Å). Twenty-two distance restraints were based on ideal values for Watson−Crick base pairs (*15*). A total of 590 NMR restraints, including more than 85 intermolecular NOE restraints, were used (Table 1). In successive iterations, distance bounds were reduced automatically using a distance analysis program developed in house, until the calculated distances matched the experimental restraints. During the calculation, some repulsion distance restraints were added to separate some proton pairs (<4 Å), which exhibited interproton contacts inconsistent with NOE data, and were responsible for serious bending and distortion of the global conformation of DNA. Relaxation matrix refinements used the same set of dihedral angle and interproton separation restraints. To stabilize the refinement calculation, the lower and upper volume bounds were reduced by one-half of the upper and lower distance bounds. The linker was not included in the calculation.

Table 1: Summary of Structure Calculation Results[a]

molecular geometry		
bond length (701) (10^{-4} Å)		4.197 ± 0.134
bond angle (1169) (deg)		1.839 ± 0.005
chirality and planetary ring (428) (improper angles, deg)		0.927 ± 0.005
RMSD of coordinates (Å)		
overall structure	0.4862 ± 0.1163[b]	0.3166 ± 0.0829[b]
binding site of DNA	0.5951 ± 0.1465	0.3875 ± 0.1040
bulge residues	0.5788 ± 0.1892	0.3773 ± 0.1289
DNA stem (4 bp from 5)	0.3107 ± 0.1045	0.2114 ± 0.0371
DNA stem (2 bp from 3)	0.1971 ± 0.1122	0.1280 ± 0.0753
NCSi-gb	0.5633 ± 0.0605	0.3698 ± 0.0284
deviation from exptl restraints[c]		
dihedral angles (39) (deg)		1.7445 ± 0.0504
R-factor ($10^{-2}R^{1/6}$)		9.841 ± 0.033
NOE (Å)		
total NOE (590)		0.3405 ± 0.0069
DNA related (462)		0.3370 ± 0.0059
NCSi-gb related (43)		0.3943 ± 0.0062
NCSi-gb-DNA (85)		0.3619 ± 0.0216

[a] Results shown in this table are mean values and standard deviations of the RMSD calculated from the final seven structures. Integers given in parentheses are the total number of entries or number of restraints. [b] The first column is the pairwise comparison, and the second column is comparison to the averaged structure. [c] Dihedral angles are used to restrain sugar pucker in DNA and in NCSi-gb and to define some DNA backbone torsion angles. $R^{1/6}$ is a statistical factor as defined in ref *23*. NOE restraints are divided into several groups: all types of interproton interactions, those involving only DNA protons, those involving only drug protons, and those involving interactions between the drug and the bulge duplex.

Sugar torsion angle restraints, such as δ (torsion of C3′−C4′) angles, were based on scalar coupling constant analyses using the COSY-35 and DQF-COSY spectra. DNA sugars were all in the C2′-endo family (*15*) with A5−C10 and G17 residues confined in $\delta = 127 \pm 10°$; C2, G4, A12, and C15 residues confined in $\delta = 120 \pm 15°$; C3 confined in $\delta = 115 \pm 10°$; and the rest of the residues loosely confined in $\delta = 110 \pm 30°$. The restraints for backbone dihedral angles (*15*) were based on spectral analysis of the $^1H-^{31}P$ correlations. The ϵ (159) (values in parentheses were measured from a B-form duplex) were derived from H3′−^{31}P couplings and restrained to $155 \pm 10°$ (weak cross-peaks), $60 \pm 10°$ (T13), and $180 \pm 10°$ (regular cross-peaks). Other backbone angle restraints used were close to canonical B-DNA values with equilibrium values at α (−39°), β (−151°), γ (31°), and ζ (−99°) and upper and lower bounds of 10−30°. Those residues associated with uncertainty in experimental data,

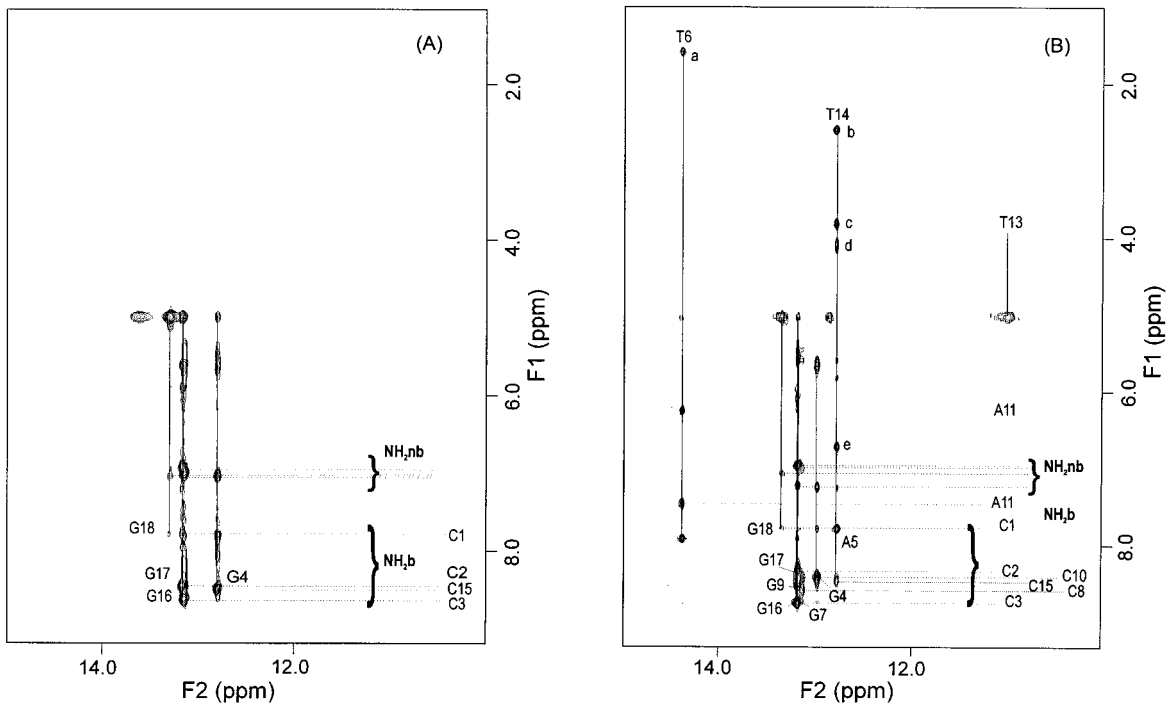

FIGURE 2: NOESY spectra recorded in H_2O and showing HN (*x*-axis) resonance correlations. (A) The free DNA sequence and (B) the NCSi-gb-bulge-DNA complex. The assigned HN resonances at the *x*-axis are labeled with residue name along the vertical lines. The horizontal lines indicate the connectivities of HN resonances of the G residues with the NH_2 resonances of the corresponding C residues and the connectivities of the HN resonances of the T residues with the H2 resonance of the corresponding A residues in the same base pair. H_2O resonance appears at 5.00 ppm. In (B), the NOE cross-peaks correlating DNA and NCSi-gb are indicated: a, T6 HN−HNM (the NA ring methyl); b, T14 HN−H13; c, T14 HN−H14a; d, T14 HN−H14b, H3′; e, T14 HN−H8.

such as T6 (α), G7 (ξ), A11 (ζ), A12 (α) and (ξ), were initially not restrained. Later, when the structures converged, the average values of these dihedral angles were computed and used as restraints in further refinement. The pyranose ring of NCSi-gb was restrained to a chair conformation and an α-glycosidic linkage. A total of 91 dihedral angle restraints were applied in the final calculation (Table 1).

Structure Calculation. Distance geometry (DG) (*16*) restrained molecular dynamics simulations (rMDS) (*17*) and relaxation matrix refinement (*18*) were used to generate a family of structures using the X-PLOR program (Accelrys Inc.). The NCSi-gb parameter file was generated from the Quanta program (Accelrys Inc.). NOE restraints in the form of square-well potential and a van der Waals nonbonded repulsion function were used in the initial DG embedding. Typical DG calculation steps involved partial metrization (using subsets of phosphorus, C3′, C5′, C1′, N9, N1, C2, C4, N3, and all atoms in NCSi-gb), embedding, regularization and refinement. DG structures (>300 structures) were further refined using rMDS which employ NMR-based distance and dihedral target functions in addition to the standard potential energy terms. A typical protocol for rMDS uses classic force fields of nonbond (cutoff 12 Å), electrostatic (dielectric constant = 1), and polar bond interactions (bond cutoff 7.5 Å and angle cutoff 80°). NOEs were normally incorporated as square-well potential functions. A biharmonic potential function was used for the dihedral angle restraints and the distance restraints related to the Watson−Crick base pairs. The shake algorithm (*19*) with a tolerance of 10^{-4} Å for bond length was invoked at all stages of the calculations. Dynamics simulations using the Verlet algorithm were carried out at 1 fs step size, and the temperature was maintained by coupling to a bath. The first stage calculation started with 1000 steps of energy minimization, followed by 4 ps of simulated annealing dynamics. During this time the temperature was increased from 100 K to between 400 and 800 K by 50 K increments, and NOE force constants were increased by a factor of 1.2 at each temperature increment to a maximum value of 100 kcal M^{-1} $Å^{-2}$. Equilibrium dynamics calculations (2 ps) were performed after the heating process, and the system was then cooled to 300 K by 50 K decrements for a total of a few picoseconds (length of time depending on the temperature differential between the initial and the final stage). After turning on the experimental dihedral angle function, additional 2 ps dynamics calculations were carried out, followed by 500 steps of energy minimization to give a new generation of calculated structures. Relaxation matrix refinement was used in the final stage to correct the effects of spin diffusion, which causes nonlinear evolution of NOE intensity for protons located in a high proton density environment. To reduce the computation time, nonbond interactions were measured to 9.0 Å. Protons were divided into three groups: H5′ and H5″, other sugar protons, and base protons. Each was assigned an order parameter ranging from 0.65 to 0.85. A 5 ns isotropic correlation time (τ) was derived from a grid research routine using X-PLOR. Volume bounds were entered in an absolute mode and equilibrium intensities and square-well functions were used. The refinement process involved initial energy minimization, followed by 1.8 ps rMDS refinement from 400 to 175 K and 0.4 ps refinement at 200 K, and final energy minimization. The satisfaction of the relaxation refinement results was judged by low $R^{1/6}$ values (Table 1), spectral comparison (Figure 2), and RMSD (root-mean-square deviations) of restrained interproton distances (Table 1).

Table 2: Structure Parameters of the DNA Duplex in the NCSi-gb-DNA Complex[a]

	base pair axis			base−base		interbase pair			
base pair	X-disp (Å)	incl (deg)	tip (deg)	buckle (deg)	propel (deg)	rise (Å)	tilt (deg)	roll (deg)	twist (deg)
C2•G17	−2.5	41.0	4.7	−28.1	−0.7				
C3•G16	−2.6	34.6	−9.0	−5.4	−2.6	3.2 (3.1)[b]	−0.1 (0.95)[b]	−19.0 (2.57)[b]	34.7 (31.1)[b]
G4•C15	−2.5	36.8	−0.3	19.2	6.4	3.5 (3.4)	−3.4 (−1.7)	−2.7 (7.5)	15.8 (20.2)
A5•T14	−2.7	28.9	−16.7	24.1	−37.7	3.5 (3.8)	7.9 (11.8)	−14.6 (16.4)	47.4 (30.1)
T6•A11	−2.4	15.1	−0.1	−40.9	−14.5	10.7 (12.8)	−52.2 (−53.9)	31.9 (43.8)	48.9 (53.6)
G7•C10	−2.4	29.2	1.8	−36.2	−17.0	2.7 (3.4)	4.9 (4.9)	4.4 (17.1)	33.1 (30.0)
B-form	−0.6	−4.6	0.0	0.0	−1.2	3.4 (3.3)	0.0 (0.0)	0.0 (−2.8)	36.0 (35.9)
A-form	−5.3	19.5	0.0	0.0	11.5	2.3 (3.2)	0.0 (0.0)	0.0 (10.8)	32.7 (31.0)

	backbone parameters (deg)								
residue	phase	sugar pucker	glyc bond	α	*β*	*γ*	*δ*	*ϵ*	*ζ*
G4	96	mixed[c]	−98	−50	−168	48	109	163	−88
A5	199	C2′-endo	−74	−89	−169	51	128	−151	−52
T6	133	mixed	−120	−78	−171	75	122	172	−88
G7	132	mixed	−114	−70	−156	56	122	172	−90
C10	129	mixed	−101	−	167	66	121	−162	−94
A11	126	mixed	−107	−122	−152	76	116	−160	−54
A12[d]	*341*	*C3′-endo*	*−145*	*−174*	*172*	*81*	*107*	*69*	*93*
T13[d]	*149*	*C2′-endo*	*−128*	*−51*	*−149*	*83*	*136*	*−31*	*−34*
T14	197	C2′-endo	−110	−113	−150	52	116	166	−51
C15	99	mixed	−105	−84	−161	64	109	160	−79
B-form	192	C2′-endo	−95	−39	−151	31	157	159	−98
A-form	14	C3′-endo	−153	−88	−149	47	83	171	−44

[a] Parameters were measured using the CURVES program from a representative structure. B- and A-form coordinates were generated using the QUANTA program. Backbone dihedral angles are as follows: α, O3′−P−O5′−C5′; *β*, P−O5′−C5′−C4′; *γ*, O5′−C5′−C4′−C3′; *δ*, C5′−C4′−C3′−O3′; *ϵ*, C4′−C3′−O3′−P; *ζ*, C3′−O3′−P−O5′. [b] Parameters preceding parentheses are based on the global helical axis, and those in parentheses are based on the local helical axis. [c] Mixed indicates an intermediate sugar conformation between C2′-endo and C3′-endo conformations. Sugar pucker range (phase angle): C3′-endo (18°), C4′-exo (54°), O4′-endo (90°), C1′-exo (126°), and C2′-endo (162°). [d] The bulge residues A12 and T13 are in italic.

The structure of the DNA duplex in the complex was characterized using the CURVES 4.1 (*20*). Structure parameters are presented in Table 2 as mean values ± RMSD derived from the seven calculated structures and as averaged values. The coordinates of the reported structures of the NCSi-gb-bulge-DNA complexes have been deposited at the Protein Data Bank. The accession number is 1KVH.

RESULTS

NMR Spectra of the Free DNA Sequence. To examine the native conformation of the DNA sequence, 5′-d(CCCGATGC-L-GCAATTCGGG) (Figure 1C), used in the NCSi-gb-bulge-DNA complex, a set of 1D and 2D NMR spectra were recorded and analyzed. The DNA sequence contains two strands; one has eight residues (the 8-mer strand); the other, 10 residues (the 10-mer strand) with an extra A and an extra T residue in the 10-mer strand (Figure 1C). The 1D and 2D NOESY spectra recorded at temperatures around 273 K show four resonances around 12.8−13.3 ppm, each of which was connected to one pair of NH_2 resonances, and these resonances were assigned to the HN of the G4, G16, G17, and G18 residues [Table S1A (Supporting Information), Figure 2A]. The NOEs were assigned revealing sequential connectivities from base H to sugar H for residues C1−A5 in the 8-mer strand and C15−G18 in the 10-mer strand for intra- and interresidue connectivities (Figure 3A). When correlated with the HN assignments, these results indicate that C1−A5 and C15−G18 strands form a duplex stem through Watson−Crick hydrogen bonding and/or base−base stacking. In the base H to sugar H spectral region, broad and/or weak NOE cross-peaks, which may be due to line width broadening, were observed for residues T6−G8 and C10−T13. These are residues on either side of the linker L. When correlated with the absence of G/T HN resonances of these residues, these results indicate random conformation in this region of the DNA sequence (Figure 4).

The most noticeable unusual chemical shifts are T6 H6 (6.85 ppm) and T6 methyl H (1.41 ppm), which are significantly shifted upfield compared to their counterparts in the T residues [Table S1A (Supporting Information)]. These values reflect that protons in the base T experience a strong ring current effect, suggesting that the T6 base is closer to other bases than in a regular duplex. T6 is a residue opposite to the extra residues in the complementary strand (Figure 1C). Since neither T6 nor T14 is base paired (Figure 4), there is a breakdown of the helical structure at the A5-T6 step in the 8-mer strand, and there is no evidence for bulge structure formation in the absence of NCSi-gb.

The ^{31}P spectrum of the free DNA sequence contains peaks of mixed line width, which were distributed over 1 ppm spectral region. A few ^{31}P resonances located within −4.2 ∼ −3.7 ppm were assigned. Among them are C1p−A5p, which belong to the residues of the duplex stem. These ^{31}P resonances are similar to what were observed for a canonical DNA duplex. The broad and more dispersed ^{31}P resonances do not show coupling cross-peaks, and they reflect conformational heterogeneity and/or mobility of part of the free DNA sequence.

NMR Spectra of the DNA Sequence in the NCSi-gb-Bulge-DNA Complex. The exchangeable H spectra of the bound DNA sequence show nine HN resonances and an expanded 2D NOESY spectrum recorded in H_2O is shown in Figure 2B. Eight HN resonances were assigned to the corresponding G•C and T•A base pairs [Table S1B (Supporting Informa-

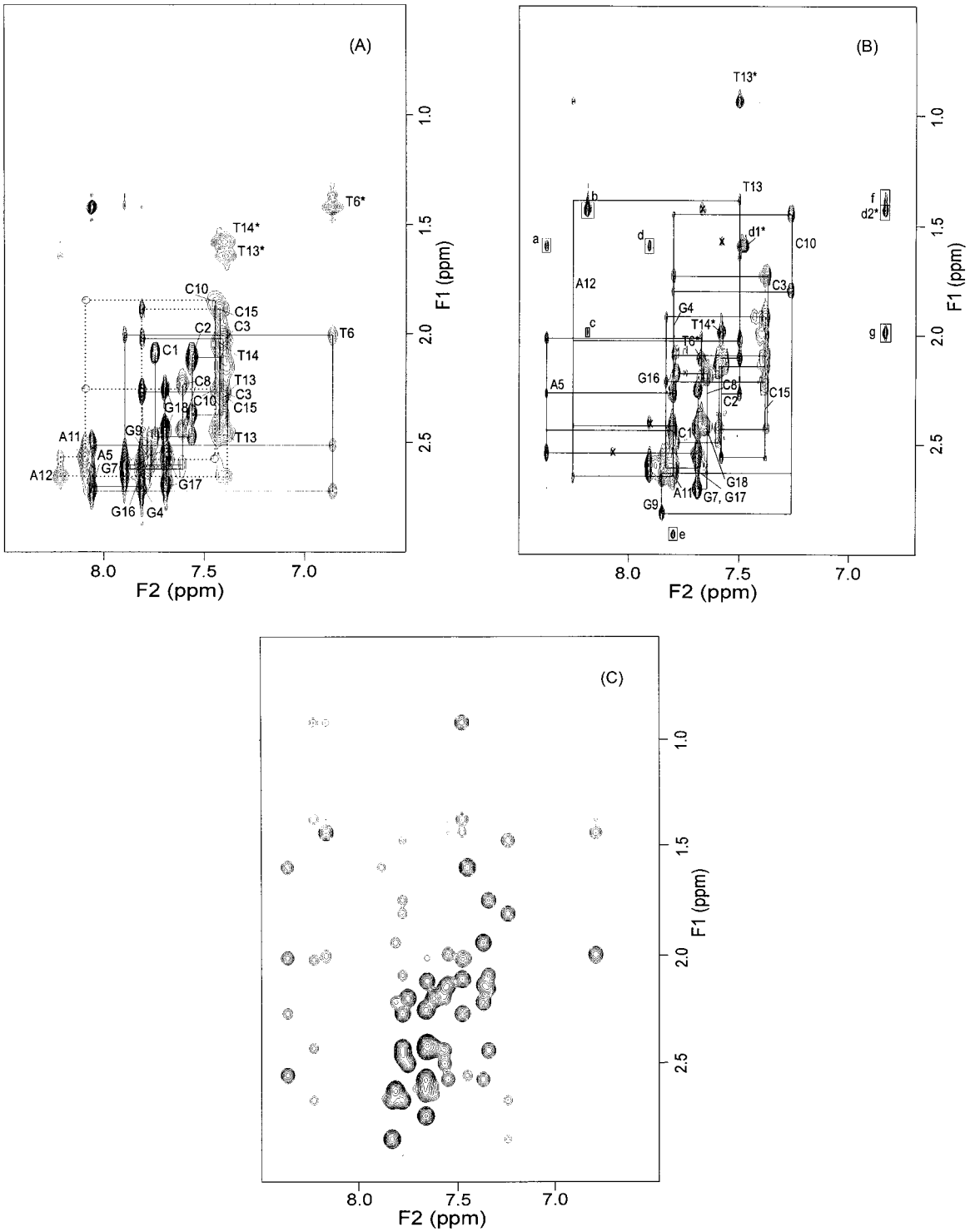

FIGURE 3: NOESY spectra recorded in D_2O and showing base proton (*x*-axis) correlations. (A) The free DNA. (B) The NCSi-gb-bulge-DNA complex. (C) The back-calculated NOESY spectrum of the complex. In (B), sequential connectivities between H2′, 2″, and base protons are traced through solid lines. Asterisks indicate NOEs from T H6 to its own H5 methyl, d1* and d2* are NCSi-gb H4″ to its own ring methyl (HNM) and H8 to the sugar ring methyl (HFM), and the crosses mark the peaks not from any resonance of the complex. The assignments of the boxed peaks are from a long-range DNA−DNA interaction (c, A12 H2−T14 5-Me) and interactions between DNA and NCSi-gb [a, A5 H8−HNM (NA ring methyl); b, A12 H2−HFM (sugar ring methyl); d, A11 H2−HNM; e, G4 H8−H2M (sugar ring *N*-methyl); f, A12 H2′−H8; g, T14 5-Me−H8]. In (C), the additional weak NOEs which are not in the experimental spectrum [see (B)] involve methyl to methyl interactions or the resonances from loop/terminal residues, reflecting heterogeneous relaxation behaviors of these residues.

tion)], revealing the formation of two duplex stems: one consists of five base pairs (C1−A5•T14−G18, 5 bp stem) and the second consists of three base pairs (T6−C8•G9−A11, 3 bp stem) (Figure 4). T13 HN in the 10-mer strand was observed in 1D spectra and showed only an exchangeable cross-peak with HOD resonance in 2D NOESY (Figure 2B). The nonexchangeable H's of the NCSi-gb-bulge-DNA complex were assigned to near completion, and the assignments are given in Table S1B (Supporting Information). In the NOESY spectra recorded in D_2O, the observed spectral

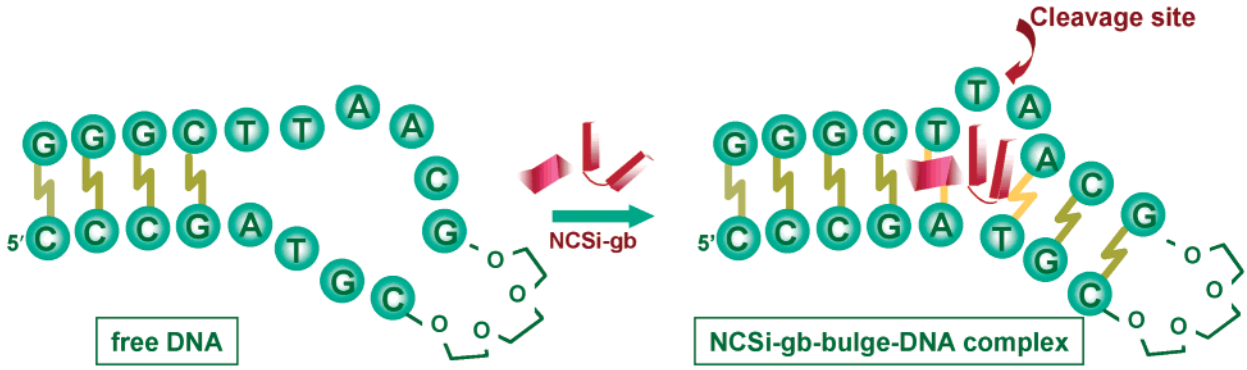

FIGURE 4: Schematic illustration of the binding-induced conformational change from free DNA to bulge-DNA upon complex formation.

pattern generally follows what is expected for a right-handed DNA duplex with intra- and interresidue NOEs connecting adjacent residues. An expanded NOESY showing base H to sugar H2′ and H2″ is displayed in Figure 3B. Although there are some overlaps, it is apparent that the NOE cross-peaks of A12 (base H to its own H1′, H2′ and H2″, and H3′) are missing or very weak as are the connectivities in the steps of A5-T6 and T13 and T14. The spectral analysis of nonexchangeable protons supports the formation of the 3 bp and 5 bp duplex stems, leaving A12 and T13 in the 10-mer strand as extrahelical residues.

COSY-35 and DQF-COSY spectra of the DNA sequence in the complex also assisted the resonance assignments and scalar coupling constant analysis. The general patterns of sugar H couplings observed are as follows: H1′−H2′ is larger than H1′−H2″, H3′−H2′ is weaker than H3′−H2″, the coupling cross-peaks related to H3′ are weaker than those related to H1′, and H3′−H4′ cross-peaks are weak or absent. These spectral patterns are typical conformations of the C2′-endo sugar family. Although there are variations among the different residues, for instance, A5, T6, C8, G9, C10, and G17 showed more consistent patterns of C2′-endo sugar pucker, there was no spectral evidence of pure C3′-endo sugar conformation found in NCSi-gb binding to DNA.

The analysis of ^{1}H−^{31}P correlation spectrum of the DNA in the complex produced ^{31}P resonance assignments [Table S1B (Supporting Information)]. The plot of the spectrum is shown in Figure S1 (Supporting Information). Except for the pG9, which is related to the hexaethylene glycol linker L, the most upfield ^{31}P resonance at 5.36 ppm is assigned to T6p, and the most downfield ^{31}P resonance at 3.60 ppm is assigned to G7p, spreading over a 1.8 ppm spectral region. The ^{31}P resonances of the helical residues are mostly concentrated in a narrow 0.5 ppm region, while those of A5p, T6p, A11p−C15p (^{31}P of A12 was not found) are shifted upfield [Table S1B (Supporting Information)]. Several spectral features related to coupling cross-peak patterns were observed. The 3′-P often shows coupling cross-peaks to H3′ of the *i* residue and H4′ and H5′/H5″ of the *i* + 1 residue, although sometimes the correlation to H4′ or H5′/H5″ may be missing. The coupling cross-peaks of T6p are extremely weak, and their patterns are distinctly different from those observed for the helical residues. The H3′−^{31}P coupling cross-peaks of G4p, T14p, and C15p are weaker compared to the rest of the H3′−^{31}P cross-peaks, but that of T13p is particularly strong. These chemical shifts and cross-peak patterns reflect variations in DNA backbone conformation and provide guidance for structure elucidation of the complex.

NMR Spectra of NCSi-gb in the Complex and as a Free Molecule. The chemical shift assignments of NCSi-gb free in an aqueous solution or in the complex are provided in Table S1D (Supporting Information). The NMR spectra of NCSi-gb in CD_3OD were previously characterized for complete chemical shift assignments (*21*). Several ^{1}H resonances of NCSi-gb in the complex can be easily identified and were used as a starting point for complete assignments. These include the H3″ and H4″ of NA and the methylene H (H14a and H14b) of the cyclocarbonate ring (Figure 1A, both showing large coupling cross-peaks) and some well-resolved ^{1}H resonances, such as H8, NA methyl (HNM), and methoxy (H7M, resonance found in complex only), and sugar *N*-methyl (H2M) and methyl (HFM), which are clearly not DNA resonances based upon their distinct patterns of NOE connectivities. NCSi-gb ^{1}H resonances exhibited most of the intramolecular connectivities as expected, permitting the assignment of the majority of chemical shifts [Table S1D (Supporting Information)]. The assignments of the free NCSi-gb (Table S1D) were guided by the literature information (*21*) and the strategy described above.

Eighteen interresidue NOEs of NCSi-gb in the complex were observed. These are mainly assigned to the interproton contacts between the NMF pyranose and the THI ring moieties and those between the THI ring and the NA ring moieties. For instance, NMF sugar H3′, H4′, and H5′ exhibited moderate NOEs to THI H8 and H6 and NMF H1′ and the *N*-methyl exhibited NOEs to THI H10, H11, and H12. In contrast, only three interresidue NOEs were observed in the free NCSi-gb, which involved NMF H5′ and THI H8, H10, and H11. These NOEs define the stacking arrangement of the three moieties (NMF, THI, and NA) of NCSi-gb and the orientation of the pyranose carbohydrate moiety. The NOE patterns of the free versus the bound forms represent different conformations of NCSi-gb.

Chemical Shift Comparison of the Free and the Bound Forms of NCSi-gb and DNA. The chemical shift differences were derived from the assignments shown in Table S1 (Supporting Information), and these results are displayed in Figure 5 with positive Δppm indicating downfield shifts upon complex formation. Figure 5A shows Δppm of the resonances of NCSi-gb according to its three subunits: NA, THI, and NMF. The resonances in the two ring systems (NA and THI) are mostly shifted upfield by more than 2 ppm (H13

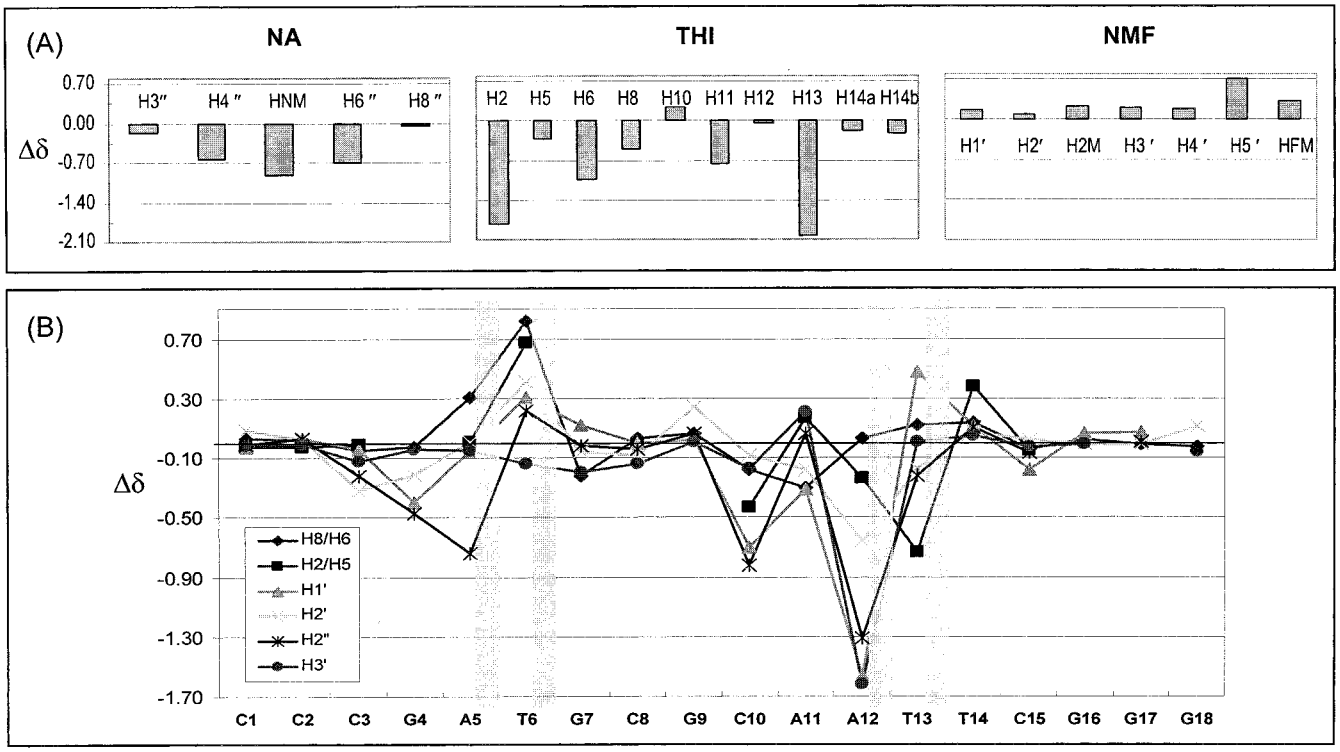

FIGURE 5: Chemical shift difference plots for (A) NCSi-gb and (B) DNA in the NCSi-gb-bulge-DNA complex. Negative values correspond to upfield shift upon complex formation. The chemical shift values can be found in Table S1 (Supporting Information).

of THI), while those in the carbohydrate unit NMF are mostly shifted downfield with smaller changes compared to the ring moiety resonances. Figure 5B shows selected Δppm of the DNA molecule according to residues and resonance assignments. The resonances of the A12 residue exhibited the largest upfield change (Δppm) upon complex formation. The resonances of the A5 and T6 of the 8-mer strand and the C10, A11, T13, and T14 of the 10-mer strand also exhibited large to modest change.

Intermolecular NOEs in the NCSi-gb-Bulge-DNA Complex. Eighty-five NOEs were assigned to the interactions between NCSi-gb and the bulge-DNA. Figure S2 (Supporting Information) summarizes the majority of intermolecular NOEs. The initial set of NOEs was by manually assigned. After the assignments of DNA and NCSi-gb ^{1}H resonance were nearly to completion, the complete assignments of the NOE cross-peaks were performed using a grid-search procedure. Essentially, all of the possible assignments to each NOE cross-peak were provided in a file, which after careful analysis allowed identification of additional intermolecular NOEs. Most interestingly, unlike most small DNA interacting ligand molecules, NCSi-gb lacks important contacts with DNA H1′ and/or H4′ in the minor groove; the few NOEs involving the minor groove H1′ and H2 of DNA and NCSi-gb are associated with the A12 and T13 residues. The majority of interproton contacts involve DNA base protons (H8, H6, T methyl, H2, NH_2, and HN). Specifically, several domains of contacts defined by intermolecular NOEs were discerned. The first domain includes those shown on the right side of the drawing in Figure S2 (Supporting Information) for the pyranose NMF ring of NCSi-gb and the bases of DNA G4, A5, A12, and T14 and the sugar of T13; the second domain contains the spiro ring linkage, the adjacent regions of THI and NA of NCSi-gb and DNA A5 base and sugar, and DNA A12 base; the third domain includes the lower side of the NCSi-gb NA ring and the bases and sugars of DNA A5 and T6 residues; the fourth domain contains the left side of the NCSi-gb NA ring, DNA A11 base and DNA A12 base and sugar; the fifth domain includes the carbocyclic ring of NCSi-gb and DNA A5 base and T14 base; the sixth domain includes the upper side of the NCSi-gb THI ring, sugars of DNA A12 and T13, and the base and sugar of DNA T14; the seventh domain includes the left side of the pyranose NMF ring of NCSi-gb, the bases of DNA A12 and T14, and DNA T13 sugar. These domain-divided intermolecular NOE analyses provide an approximate model of binding: NCSi-gb inserts into the step between A5·T14 and T6·A11 and is orientated as such that the pyranose NMF ring is pointed toward the 5 bp stem with the spiro ring linkage closer to the 8-mer strand than the 10-mer strand.

Structure of the NCSi-gb-Bulge-DNA Complex. The structure of the complex was established without using a preconceived model. The extensive NOE and scalar coupling cross-peak analyses provided a large number of distance or dihedral angle restraints (Table 1). More than 600 structures were generated using distance geometry algorithm as implemented in X-PLOR (*16*−*18*). These structures were further refined using molecular dynamics simulations, and the calculated structures were judged by their consistency with the empirical chemical structure relationships and the experimental NMR data. The fine-tuning of the structures used complete relaxation matrix back-calculation procedures (*18*). The seven best representative structures demonstrated convergent coordinates and are in agreement with NMR data (Figure 3B,C). The global structures based on pairwise or comparing to the average structure are in good agreement, but the bulge residues and the drug seem to exhibit more dynamics than the stem residues (Table 1). These structures

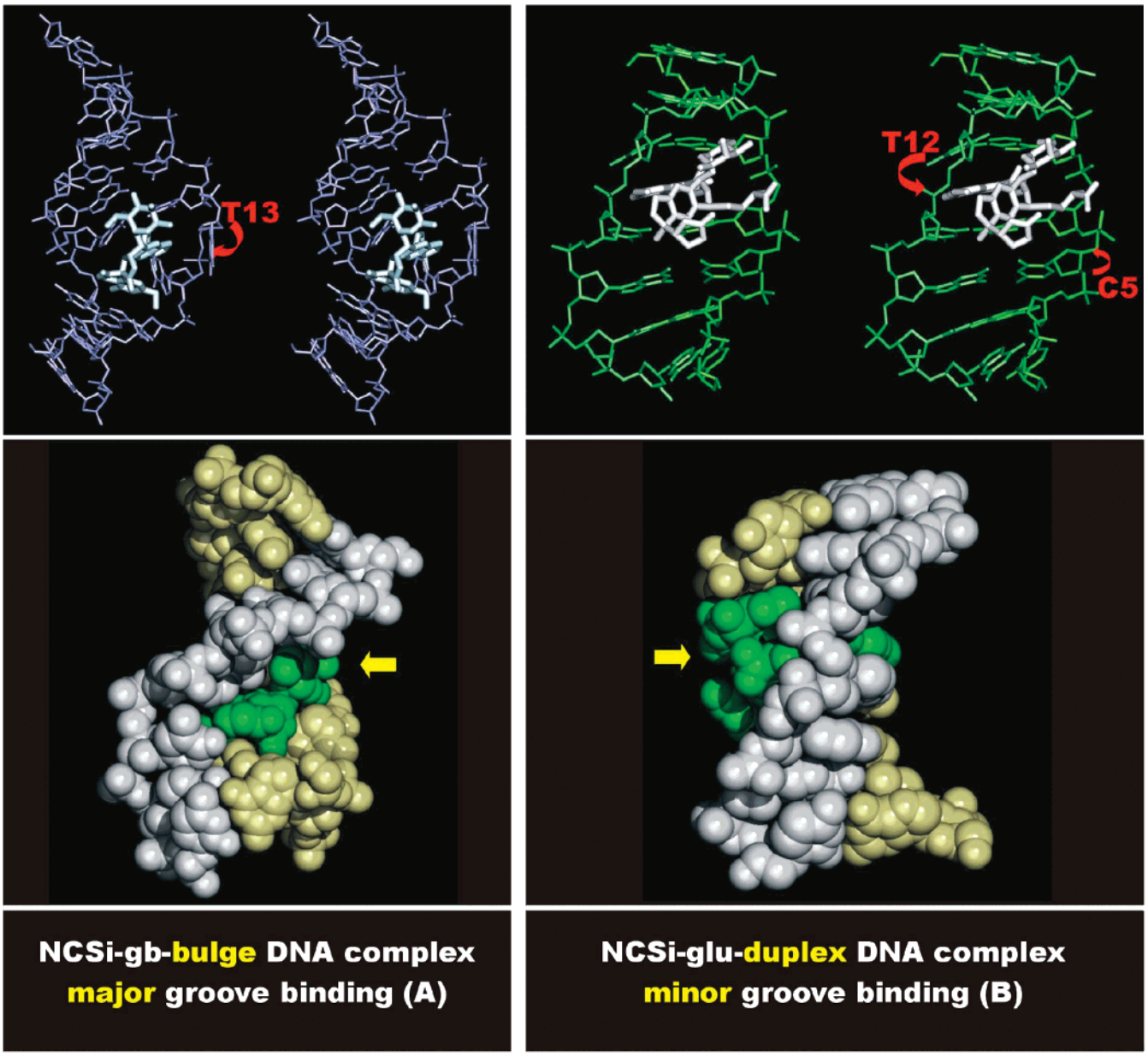

FIGURE 6: Structures of the two complexes: (A) two panels on left, the NCSi-gb-bulge-DNA complex showing major groove binding with the NA and THI rings of the drug stacking between the DNA helix; (B) two panels on right, the NCSi-glu-DNA complex showing the minor groove intercalation of the drug. In (A), the DNA cleavage site is at the bulge residue T13 (single strand cleavage); in (B), the DNA cleavage sites are at the cross-strand residues C5 and T12 (following the typical pattern of minor groove double strand cleavage). The CPK drawings (lower two panels) for the two complexes have the minor groove open to the left and the major groove to the right, contrasting the binding modes of the two complexes.

are shown in Figure 6 and are further refined from those reported previously (*7*). There is no major discrepancy between the two sets of structures, but the later set produced back-calculated NOESY which compares better with the experimental data.

Many details of the NCSi-gb-bulge-DNA complex have been described in our previous publication (*7*). To provide a basis for the comparisons of the two distinctly different complexes related to NCS-chrom, the novel features of this complex are summarized below. The binding ligand functions as a geometrically restricted molecular wedge interlocked through a spirolactone ring and with a multifunctional handle, i.e., the pyranose NMF carbohydrate ring. The binding site is a triangular pocket, rich in aromatic moieties, with two extra helical residues forming the third contacting surface with NCSi-gb. The location of the binding is solely in the major groove, aligned with the center of the groove floor. The key contact between NCSi-gb H6 and T13 H5′ (*pro-S*) suggests that the structure of the complex is approximately that of the cleavage complex, since H6 results from H-abstraction by the 6-radical in the diradical species of NCS-chrom (Figure 1A) (*7*).

DISCUSSION

Mutual Conformational Switch of the Ligand and DNA Induced by Binding. Upon complex formation, both DNA and NCSi-gb undergo profound conformational changes. These changes are clearly discernible in the differences of chemical shifts and NOE patterns of the free and bound forms of both NCSi-gb and DNA. The upfield-shifted chemical shifts of the two ring systems (NA and THI) (Figure 5A) in NCSi-gb suggest less separation of the two stacked moieties

in the bound form. Therefore, although restrained by the spirocyclic covalent bond linkages, the two rings of NCSi-gb are more compressed in the binding pocket of the complex form. In the free form, although a few interresidue NOEs were present, the data were not sufficient to derive the conformation of the free NCSi-gb. Eighteen NOEs were detected among the subunits of the bound NCSi-gb, revealing the interactions of the partially overlapped NMF carbohydrate and the THI moieties. The orientation of NMF is such that its *N*-methyl points toward the spirocyclic carbonyl. In the two forms, the orientations of NMF with respect to the THI ring are different since NMF H5′ showed NOEs to THI H6 and H8 in the complex but to H8, H10, and H11 in the free form. Overall, the observations of large chemical shift variations, new NOEs, and the changes in the NOE patterns of the bound NCSi-gb clearly signal its conformational transition in the process, as depicted in Figure 4.

The conformational transition of DNA upon formation of the complex is also illustrated in Figure 4. This is derived from comparisons of the chemical shift differences, new NOEs, and the changes in the NOE patterns for the resonances in DNA. The largest chemical shift changes (Figure 5B) are concentrated in residues A5 and T6 of the 8-mer strand and in residues C10−T13 of the 10-mer strand. Furthermore, the absence of the T6 and G7 HN resonances (Figure 2B) and NOE patterns in the free DNA as compared to those in the complex (Figure 3) suggest that the 3 bp stem is not formed in the absence of NCSi-gb. Thus, the free DNA sequence does not contain a defined bulge structure, and the 3 bp stem residues are mostly loose in solution.

The analyses of the free and the bound forms of both NCSi-gb and the DNA reveal a mutually induced fit between the loosely structured free ligand and the DNA substrate. Stabilization of the complex was achieved by the interactions, which originate from the more tightly folded NCSi-gb, newly formed Watson−Crick base pairs in the 3 bp stem, and the contacts between NCSi-gb and the DNA residues at the bulge sites. In the complex, NCSi-gb locks DNA residues A12 and T13 into a stable extrahelical conformation and initiates the formation of a new 3 bp stem, achieved by its two spirolactone-fixed rings stacked with the T6·A11 base pair in one end of the 3 bp stem (Figure 4).

The Multifaceted Features of the Interactions of NCS-chrom with DNA. There are striking differences in the binding modes of the two complexes formed by the two different activation mechanisms. The availability of their structures permits a detailed molecular view of the DNA interactions involving the chemical moieties of the two ligands, some of which are in common for NCSi-glu and NCSi-gb, both originating from NCS-chrom. Specifically, both ligands contain a naphthoate moiety (NPH in NCSi-glu and NA in NCSi-gb), a tetrahydroindacene type ring (THI), an identical 2-*N*-methylfucosamine carbohydrate unit (NMF), and an identical cyclocarbonate (CC) moiety. The major differences in the NCSi-glu-duplex-DNA and NCSi-gb-bulge-DNA complexes include (a) binding in opposite grooves, with NCSi-glu in the minor groove of the DNA duplex and NCSi-gb in the major groove of the bulge-DNA (Figure 6); (b) the carbohydrate unit in a completely different local environment, with NCSi-glu NMF contacting the sugar−phosphate backbone from the minor groove and NCSi-gb NMF lying on the major groove floor (Figure 7A,B); (c) stabilization of the bindings mostly by van der Waals contacts with sugar−phosphate backbone residues for NCSi-glu, but mostly by shape complementarity and stacking with aromatic bases for NCSi-gb (Figure 7C,D); (d) overall ligand conformation with the three chemical moieties (THI, NMF, and CC) of NCSi-glu adopting an extended fashion stretching along the DNA minor groove (Figure 6B), but they fold into an orthogonal orientation in NCSi-gb so that THI intercalates between the 3 bp and 5 bp duplex stem at the bulge junction and NMF binds in the major groove of the DNA (Figure 6A); (e) cleavage specificity with NCSi-gb associated with single strand cleavage at the 5′-position of the second bulge residue T13 and NCSi-glu associated with double strand cleavage across the duplex (Figure 6).

The distinct structures of the two complexes formed by NCS-chrom derivatives illustrate the complexity of physiochemical principles that govern intermolecular interactions; the factors affecting the binding or molecular recognition are multidimensional. Chemical structural identity of a subunit or even the overall similarity in building blocks of a molecule does not necessarily lead to the same binding specificity and mode. The structure of the complex depends on the overall energies of the system as determined by electrostatic and van der Waals interactions, the interface shape complementarity, and the intrinsic energies of the ligand and the substrate in their bound form.

Major Groove and Bulge Recognition by NCSi-gb. Since most small molecules bind to duplex DNA in the minor groove (*8*), the structure of the NCSi-gb-bulge-DNA provides valuable insights into the molecular design of bulge and/or major groove binders. The major groove has a convex surface, and thus ligand molecules making contacts with the helical bases are more likely than with the phosphate backbone. This is exactly what is displayed by the NCSi-gb-bulge-DNA complex. Other major groove binders often carry positive charges and align themselves along the groove floor as summarized in the literature (*8*). In contrast, the minor groove has sugar moieties lining either side of the concave groove surface, and thus being sandwiched in the minor groove is favored by ligand molecules. Other factors influencing the groove selection are the orientation and handedness of the ligand. In the NCSi-gb-bulge-DNA complex, the cyclospirolactone connecting THI and NA and the carbohydrate unit are oriented right-handed, and the carbohydrate unit is located at the beginning of the 5 bp stem in the major groove. If the carbohydrate unit preferred the 3 bp stem, the binding would have to occur in the minor groove in order for the right-handed ligand to fit the right-handed DNA. The NCSi-gb-bulge-DNA structure contains the drug molecule functioning as a molecular wedge at the two-base bulge site. The chromophore rings of the molecular wedge are located at the interface of the two stems of DNA. Similar to the stacking bases of DNA, they are stacked in a right-handed fashion to allow a helical transition at the bridge point of the two base bulge, although a three- or four-base bulge at the bridge point for the two discontinuous stems can also be tolerated with lower binding affinity (*22*). Since DNA bulges could be markers for frame-shift mutation and helical bending (*23*), it is of great interest to explore the underlying principle of binding revealed by the synthetic specific bulge binders. Initial results with spirocyclic compounds designed on the basis of the structure of NCSi-gb have been found

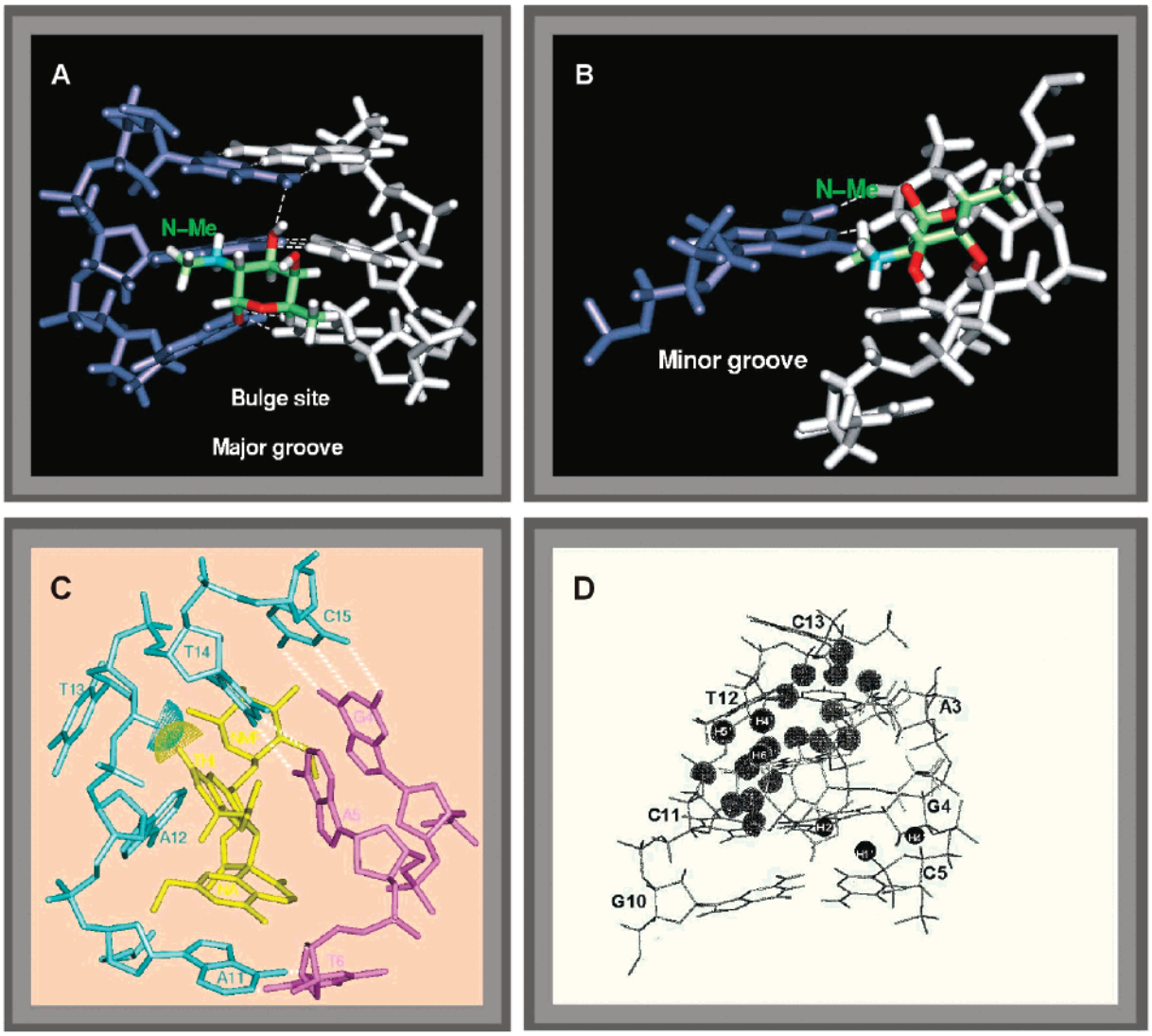

FIGURE 7: Structures of the two complexes showing features of carbohydrate recognition of DNA and details of molecular contacts in these complexes. (A) The binding of the carbohydrate moiety to the floor of the major groove of DNA in the NCSi-gb-bulge-DNA complex. (B) The binding of the carbohydrate moiety to the sugar backbone in the minor groove of DNA in the NCSi-glu-DNA complex. (C) Structure reviewed from the minor groove. The aromatic components of DNA and the formation of the triangular binding pocket in the NCSi-gb-bulge-DNA complex are shown. The major force stabilizing the NCSi-gb-bulge-DNA complex originates from stacking interactions. (D) The van der Waals contacts in the NCSi-glu-DNA complex. At the binding site, those H atoms of DNA sugar and backbone as well as of NCSi-gb, which are involved in intermolecular contacts, are shown in half van der Waals radius. The major force stabilizing the NCSi-glu-DNA complex originates from van der Waals interactions.

promising (*24*). On the basis of these results, molecular architectures of multiple stacking steps might be suitable candidates for the recognition of DNA bulges having more than two extrahelical residues. These are useful molecular templates for the development of probe molecules for the detection of unusual structures of nucleic acids.

Mechanistic Implications. NCS-chrom undergoes activation, aromatic ring cyclization, formation of the cumulene and the biradical species, and reduction to form postactivated compounds (Figure 1). In the last step of H-abstraction, DNA chain cleavage may occur due to initiation of radical reactions involving sugar and phosphate groups. In theory, stable DNA complex formation may occur with the native ligand, or it is possible that the binding of the cumulene or the biradical species has a higher affinity than that of the native form. When the structures of the complexes of NCSi-gb-bulge-DNA and the NCSi-glu-duplex-DNA are compared, it is apparent that the binding specificity must be determined by the activated species, either the cumulene (more stable) or the 2,6-biradical (less stable) species (*11*). Since the post-activated species NCSi-gb and NCSi-glu possess the major structural signatures of the various activated NCS-chrom species (Figure 1A,B), it becomes clear that the DNA chain cleavage specificity is determined by specific binding of the activated NCS-chrom species. This conclusion was supported by a DNA cleavage experiment, in which the same cleavage sites were observed upon the addition of a cumulene compound (*26*).

The finding that the two ring systems of NCSi-gb are brought closer to each other in the DNA bulge binding pocket has important mechanistic implications for the bulge-specific cleavage reaction and the generation of the bulge-specific drug reaction product (Figure 1A). Intramolecular quenching of the radical at C2 is believed to be the essential first step in generating the active DNA sugar H-abstracting species, which culminates in the final drug reaction product, the

cyclospirolactone (*4*). In fact, in the absence of bulged DNA this species is rarely formed. Further, it has been shown that the DNA functions catalytically in the drug transformation reaction and that these results cannot be explained solely by the protection afforded by the binding pocket against indiscriminate intermolecular quenching of the drug radical species (*4*). Induced conformational changes by both drug and DNA have also been recently observed by circular dichroism measurements (*26*). It appears, therefore, that by bringing the two rings together, the DNA binding pocket enables the efficient formation of a new bond between C2 of the dihydroindacene moiety and C8″ of the naphthone and the generation of the bulge-DNA-specific, monofunctional H-abstracting species.

ACKNOWLEDGMENT

The 600 MHz NMR spectrometer at the University of Houston is funded by the W. M. Keck Foundation. Acknowledgment is made to the Institute of Molecular Design and the W. M. Keck Center for Computational Biology at the University of Houston for computer resource support.

SUPPORTING INFORMATION AVAILABLE

Table S1 listing ^{1}H and ^{31}P chemical shift and $\Delta\delta$ (ppm) data and Figures S1 and S2 showing the $^{1}H-^{31}P$ spectrum of the complex and a schematic illustration of intermolecular NOEs. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Goldberg, I. H. (1991) *Acc. Chem. Res. 24*, 191−198.
2. Xi, Z., and Goldberg, I. H. (1999) in *Comprehensive Natural Products Chemistry* (Barton, D. H. R., and Nakanishi, K., Eds.) Vol. 7, pp 553−592, Elsevier Science, Oxford.
3. Myers, A. G., Proteau, P. J., and Handel, T. M. (1988) *J. Am. Chem. Soc. 110*, 7212−7214.
4. Xi, Z., Mao, Q. K., and Goldberg, I. H. (1999) *Biochemistry 38*, 4342−4354.
5. Gao, X., Stassinopoulos, A., Rice, J., and Goldberg, I. H. (1995) *Biochemistry 34*, 40−49.
6. Gao, X., Stassinopoulos, A., Gu, J., and Goldberg, I. H. (1995) *Bioorg. Med. Chem. 3*, 795−809.
7. Stassinopoulos, A., Jie, J., Gao, X., and Goldberg, I. H. (1996) *Science 272*, 1943−1946.
8. Han, X., and Gao, X. (2001) *Curr. Med. Chem. 8*, 551−579.
9. Kappen, L. S., and Goldberg, I. H. (1993) *Science 261*, 1319−1321.
10. Hensens, O. D., Chin, D.-H., Stassinopoulos, A., Zink, D. L., Kappen, L. S., and Goldberg, I. H. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 4534−4538.
11. Myers, A. G., and Proteau, P. J. (1989) *J. Am. Chem. Soc. 111*, 1146−1147.
12. Braun, S., Kalinowski, H.-O., and Berger, S. (1999) *150 and more basic NMR experiments*, 2nd ed., Wiley-VCH, New York.
13. Wuthrich, K. (1986) *NMR of proteins and nucleic acids*, John Wiley & Sons, New York.
14. Schmitz, U., and James, T. L. (1995) *Methods in Enzymology*, Vol. 261, pp 3−44, Academic Press, New York.
15. Saenger, W. (1984) *Principles of Nucleic Acid Structure*, pp 17−20 (definition of backbone torsion and glycosidic angles, sugar puckers), pp 123−124 (H-bond parameters), Springer-Verlag, New York.
16. Crippen, G. M. (1981) *Distance Geometry and Conformational Calculations*, John Wiley & Sons, New York; Kuszewski, J., Nilges, M., and Brünger, A. T. (1992) *J. Biomol. NMR 2*, 33−56.
17. Brünger, A. T., Clore, G. M., Gronenborn, A. M., and Karplus, M. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3801−3805; Nilges, M., Clore, G. M., and Gronenborn, A. M. (1988) *FEBS Lett. 239*, 129−136.
18. Nilges, M., Habazettl, J., Brünger, A. T., and Holak, T. A. (1991) *J. Mol. Biol. 219*, 499−510; James, T. L. (1994) *Curr. Opin. Struct. Biol. 4*, 275−284.
19. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) *J. Comput. Chem. 4*, 187−217.
20. Program kindly provided by R. Lavery, CNRS Institut de Biologie Physico-Chimique, France; Lavery, R., and Sklenar, H. (1988) *J. Biomol. Struct. Dyn. 6*, 63−91.
21. Hensens, O. D., Helma, G. L., Zink, D. L., Chin, D.-H., Kappen, L. S., and Goldberg, I. H. (1993) *J. Am. Chem. Soc. 115*, 11030−10031.
22. Xi, Z., and Goldberg, I. H. (2002) Enediyne antibiotic neocarzinostatin as a radical-based probe of bulged structures in nucleic acids, in *Advances in DNA Sequence Specific Agents*, Vol. IV, JAI Press (in press).
23. Rice, J. A., and Crothers, D. M. (1989) *Biochemistry 28*, 4512−4516.
24. Xi, Z., Jones, G. B., Qabaja, G., Wright, J., Johnson, F., and Goldberg, I. H. (1999) *Org. Lett. 1*, 1375−1377.
25. Myers, A. G., Cohen, S. B., and Kwon, B.-M. (1994) *J. Am. Chem. Soc. 116*, 1670−1682.
26. Yang, C. F., Jackson, P. J., Xi, Z., and Goldberg, I. H. (2002) *Bioorg. Med. Chem.* (in press).

BI012112O