Articles

New Complex of Post-Activated Neocarzinostatin Chromophore with DNA: Bulge DNA Binding from the Minor Groove†

Youngjoo Kwon,‡,§ Zhen Xi,||,§,⊥ Lizzy S. Kappen,|| Irving H. Goldberg,*,|| and Xiaolian Gao*,‡

Department of Chemistry, University of Houston, 136 Fleming Building, Houston, Texas 77004-5003, and Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115

Received October 9, 2002

ABSTRACT: Neocarzinostatin (NCS-chrom), a natural enediyne antitumor antibiotic, undergoes either thiol-dependent or thiol-independent activation, resulting in distinctly different DNA cleavage patterns. Structures of two different post-activated NCS-chrom complexes with DNA have been reported, revealing strikingly different binding modes that can be directly related to the specificity of DNA chain cleavage caused by NCS-chrom. The third structure described herein is based on recent studies demonstrating that glutathione (GSH) activated NCS-chrom efficiently cleaves DNA at specific single-base sites in sequences containing a putative single-base bulge. In this structure, the GSH post-activated NCS-chrom (NCSi-glu) binds to a decamer DNA, d(GCCAGAGAGC), from the minor groove. This binding triggers a conformational switch in DNA from a loose duplex in the free form to a single-strand, tightly folded hairpin containing a bulge adenosine embedded between a three base pair stem. The naphthoate aromatic moiety of NCSi-glu intercalates into a GG step flanked by the bulge site, and its substituent groups, the 2-*N*-methylfucosamine carbohydrate ring and the tetrahydroindacene, form a complementary minor groove binding surface, mostly interacting with the GCC strand in the duplex stem of DNA. The bulge site is stabilized by the interactions involving NCSi-glu naphthoate and GSH tripeptide. The positioning of NCSi-glu is such that only single-chain cleavage via hydrogen abstraction at the 5′-position of the third base C (which is opposite to the putative bulge base) in GCC is possible, explaining the observed single-base cleavage specificity. The reported structure of the NCSi-glu-bulge DNA complex reveals a third binding mode of the antibiotic and represents a new family of minor groove bulge DNA recognition structures. We predict analogue structures of NCSi-R (R = glu or other substituent groups) may be versatile probes for detecting the existence of various structures of nucleic acids. The NMR structure of this complex, in combination with the previously reported NCSi-gb-bulge DNA complex, offers models for specific recognition of DNA bulges of various sizes through binding to either the minor or the major groove and for single-chain cleavage of bulge DNA sequences.

Neocarzinostatin (NCS-chrom)[1], a natural enediyne antitumor antibiotic, has been shown to undergo either thiol- or nonthiol-dependent activation, resulting in distinctly different DNA chain cleavage patterns. These can be attributed to versatile chemical structures of the activated NCS-chrom species and their formation of different DNA binding specificities (Scheme 1, refs *1*−*7*). The mechanisms of NCS-chrom activation, its DNA cleavage properties, and the formation of the post-activated NCS-chrom molecules have been the subject of intensive investigations (*1*, *2*). Several DNA cleavage patterns have been identified under either activation mechanism, which include concomitant double-

† 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 Services Grant GM 53793 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; e-mail: irving_goldberg@hms.harvard.edu.

‡ University of Houston.

§ Both authors contributed equally to this work.

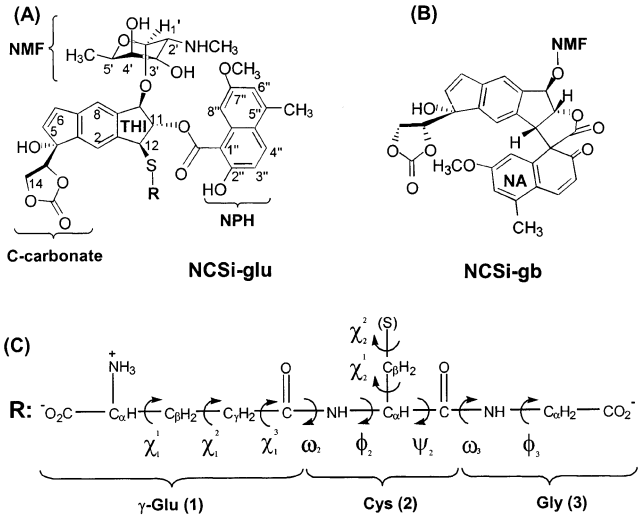
|| Harvard Medical School.

⊥ Current address: State Key Laboratory of Elemento-Organic Chemistry, Institute of Elemento-Organic Chemistry, Nankai University, Tianjin 300071, China.

[1] Abbreviations: 1-D, one-dimensional; 2-D, two-dimensional; COSY, correlation spectroscopy; Cys, cysteine; DG, distance geometry; DQF-COSY, double quantum filtered correlation spectroscopy; D_T, diffusion constant; γ-Glu, γ-glutamic acid; Gly, glycine; GSH, glutathione; HPLC, high-performance liquid chromatography; MDS, molecular dynamics simulations; NCS, neocarzinostatin; NCS-chrom, native form of NCS chromophore; NCSi-gb, post-activated NCS-chrom formed under based catalyzed conditions; NCSi-glu, post-activated NCS-chrom formed under reductive conditions in the presence of glutathione; NMF, 2-*N*-methylfucosamine; NMR, nuclear magnetic resonance; NPH, naphthoate; Nucleic acids, A: adenosine, C: cytidine, G: guanosine, T: thymidine; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect and exchange spectroscopy; rMDS, restrained molecular dynamics simulations; RMSD, root-mean-square deviation; THI, tetrahydroindacene; TOCSY, total correlation spectroscopy; torsion angles, g^+: gauche$^+$, g^-: gauche$^-$, t: trans.

10.1021/bi0206210 CCC: $25.00
Published on Web 01/15/2003

Scheme 1: Chemical Structures of (A) NCSi-glu, Formed in the Presence of Glutathione; (B) NCSi-gb, Formed under Based Catalyzed Conditions; and (C) the Substituent, Glutathione, on NCSi-glu



and single-strand cleavages in a DNA duplex (mode 1, refs *4*, *5*) and single-strand cleavage in a DNA sequence that potentially can fold to form single-stranded structures containing bulge sites (mode 2, refs *8*−*10*). In general, the cleavage sites are in A/T rich regions and allow hydrogen abstraction (H-abstraction) from DNA by an NCS-chrom diradical intermediate and chain cleavage at the position lost hydrogen. Mode 1 type of cleavages occur more frequently at 5′-G/AC**T**•AG**C**/**T**-3′ sites with the point of cleavage located at the 5′ position of the T residue and less frequently at the 1′ or 4′ position of the C/T residue in the opposite strand (residues shown in bold). Careful analysis of the DNA sequences cleaved in mode 2 revealed two subtypes of sequence patterns; one is the thiol independent mode 2a related to the sequences containing putative two-base or larger bulge sites. The specific cleavage site of mode 2a resides in the 5′ position of a T residue in the bulge. The other is the thiol-dependent mode 2b that is related to the sequences containing putative bulge sites and a specific cleavage site in the 5′ position of a residue opposite to the bulge residues. Mode 2b type of cleavage is associated with a different set of NCS-chrom and DNA contacts and does not show strong A/T preference. Therefore, the complex formation is likely to be linked to the bulge motif-directed cleavage, representing a shift in NCS-chrom DNA cleavage specificity from sequence-specific to motif-selective.

The versatility of NCS-chrom activation, formation of activated species, and the patterns of DNA chain cleavage offers a rare opportunity to gain fundamental understanding of the molecular details of DNA recognition. The first solution structure of the molecular complex of the GSH-activated NCS-chrom (NCSi-glu, Scheme 1A) and a DNA heptamer duplex (5′-CC**T**CGCG•CG**C**GAGG-3′) was reported using high-resolution NMR spectroscopy and computational methods (*4*, *5*). This NCSi-glu-duplex DNA complex (complex 1) is characterized by intercalation of NCSi-glu from the minor groove of DNA and provides a detailed model for the drug−DNA interactions that give rise to mode 1 type of DNA cleavage. The structure clearly defines the interactions of the intercalating naphthoate moiety (NPH), the minor groove binding moiety fucosamine residue (NMF), and the reduced enediyne ring (THI), all in NCS-chrom, with specific DNA residues. By discerning the close contacts between the sites of radical formation in NCSi-glu and the DNA residues, the structure explains the specific cleavages at the T and the C (above, in bold) residues. In this structure, the GSH peptide side chain attached to NCS-chrom was not defined since its ^{1}H resonances were not observed. This lack of information was attributed to a mobile peptide side chain, reflecting loose interactions between the tripeptide and the DNA in complex 1.

The second molecular complex is formed between a general-base post-activated NCS-chrom in the absence of a thiol (NCSi-gb, Scheme 1B) and a two-base bulge DNA (complex 2) (*11*, *12*). This complex relates to NCS-chrom DNA cleavage mode 2a, and its structure was a surprise (*6*, *7*). As compared to its counterpart NCSi-glu and most small DNA binding ligands (*3*), NCSi-gb adopts an unusual binding mode, which involves major groove recognition by the drug carbohydrate unit and tight fitting of the wedge-shaped drug in the triangular prism pocket formed by the two looped-out DNA bulge bases and the neighboring base pairs. In the binding cavity, the original NPH and the THI moieties now connected through a spirolactone are mimics of helical DNA bases, complementing the bent DNA structure. The positioning of the THI moiety is such that the pro-S H5′ of the target base T, which is also a bulge residue, is about 2.2 Å separated from the radical site in THI. This structure clarifies the mechanism of bulge recognition and the specific cleavage of single-stranded DNA residues by NCS-chrom. This result, among the few bulge binding complexes (*3*, *13*), establishes principles for the design of bulge-specific nucleic acid binding molecules. We have further reasoned that the effective recognition of the bulge site on the duplex DNA may have bearing on the bent DNA groove space being flexible enough to adopt the geometrically optimal conformation compatible with the wedge-shaped drug molecule rather than involving lock and key recognition (*7*, *14*). This shape-fitting principle for DNA bulge selectivity has been successfully applied to the design of synthetic bulge-specific analogues, which can form a wedge shape through covalently linked stacking aromatic moieties (*15*, *16*).

This report will describe a third, categorically distinct high-resolution structure of the complex containing the thiol post-activated NCS-chrom, NCSi-glu, with a single-base bulge bearing DNA. This complex is based on recent studies demonstrating that GSH activated NCS-chrom causes highly efficient, specific single-base cleavage in single-stranded DNA sequences containing a putative single-base bulge (*17*−*19*). The DNA sequence used in this study is d(GCCA-GAGAGC), containing a cleavage site at the 5′ position of the third C (C3) shown in the NCS-chrom DNA cleavage studies. This system is to address several interesting aspects about the observed specific DNA cleavage. It has been known that the predominant DNA cleavage sites by NCS-chrom, either in single- or double-strand mode, are T and A residues; the sequence per se is neither self-complementary (not a duplex motif) nor a bulge sequence as observed in the major groove binding bulge DNA complex 2 (*10*). Our NMR analysis now details the conformation of DNA and NCSi-glu and their intermolecular interactions and provides a structure illustrating minor groove bulge DNA recognition.

This structural analysis explains the observed single-base cleavage specificity and reveals a third binding mode of the antibiotic with noncanonical DNA. This report also describes the comparison of the conformations of the DNA in the free or the bound form. This information should be valuable for designing sequence-specific DNA binding ligands, which can find applications as antibiotic, antitumor agents and biosensor molecules.

MATERIALS AND METHODS

Sample Preparation. (a) NCSi-glu: NCS-chrom was extracted from the chromoprotein neocarzinostatin and reduced to form NCSi-glu (Scheme 1A) in the presence of GSH and purified by HPLC following published procedures (*4*, *5*). (b) Free DNA: 5′-d(GCCAGAGAGC) was synthesized using standard phosphoramidites (Glen Research) on an ABI 381A DNA synthesizer. The oligonucleotide was purified by HPLC (reverse-phase C_{18} column) and then passed through size-exclusion (Bio-Rad Econo-Pac 10 DG) and Na^+-exchange columns. The purity of the sample was examined using analytical reverse-phase C_{18} HPLC and electrophoresis on 20% denaturing polyacrylamide gel. (c) Preparation of the NCSi-glu-bulge DNA complex: d(GCCAGAGAGC) was dissolved in a solution (0.3 mL) containing 10 mM sodium phosphate and 0.1 mM EDTA, pH 6.0 (SPE10 buffer) and gradually added to a sample of NCSi-glu dissolved in d_4-MeOH at 278 K (total added volume was 80 μL). The formation of the complex was monitored by a set of one-dimensional (1-D) ^{1}H NMR spectra. The final sample contained about 0.9 mM of the complex.

NMR Experiments. NMR experiments were performed on a Bruker 600 MHz spectrometer. Spectra were recorded in either 90% H_2O/10% D_2O (Cambridge Isotopes Inc.) for observation of exchangeable protons or 99.99% D_2O for observation of nonexchangeable protons. The NMR data acquired were processed using FELIX (Accelrys Inc.) and UXNMR (Bruker Instruments Inc.). Proton chemical shifts were referenced to the HOD resonance (4.70 ppm at 298 K, temperature correction factor -0.0109 ppm/K). The ^{31}P chemical shifts were referenced to the signal of an external sample containing trimethyl phosphate in 0.1 M sodium phosphate, pH 6.3, 298 K.

(a) 1-D experiments: All 1-D ^{1}H NMR spectra recorded were centered at the HOD resonance with a spectral width of 10 ppm for D_2O samples and 23 ppm for H_2O samples. A total of 4096 or 8192 data points were acquired; ^{31}P NMR spectra were centered at -3.5 ppm with a spectral width of 4−10 ppm. A total of 1024 or 2048 data points were acquired. The inversion recovery pulse sequence (*20*) ($180_\chi - \tau - 90_\chi -$ Acq) with τ varied from 3 μs to 1.4 s, and a 15.0 s repetition delay was used to estimate T_1 (spin lattice relaxation time, $T_1 \cong \tau_{null}/\ln(2)$; τ_{null} is the time when NMR signal intensity is close to zero). The estimated values of T_1 for aromatic protons of the free DNA and the complex were about 1.6 and 1.9 s in D_2O. The inversion recovery experiments were also used to identify A H2 resonances. These resonances, which exhibit longer T_1 values than those of base and sugar protons, appeared as sharp, negative peaks, while the majority of proton signals were at null intensities.

The pulsed field gradient (PFG) experiments were performed for complex 3, the free DNA: 5′-d(GCCAGAGAGC), and a heptamer DNA duplex: 5′-d(CCTCTTG•CAAGAGG). These experiments measure the translational diffusion constant (D_T), which is a function of the size, shape, and mass of the molecules (*21*, *22*). The PFG strength was calibrated using lysozyme hen egg white (Sigma) (0.7 mM in D_2O containing 50 mM NaCl, pH 5.5), which gave 11.1 $\times 10^{-7}$ cm^2/s of D_T at 298 K (*23*). For each molecule, a set of 1-D spectra was recorded at different gradient field strengths (g, G/cm) with 10 and 14 ppm spectral widths (for DNA sample and lysozyme reference, respectively) and 16 K data points. The gradient pulse duration (δ) was 7 ms, the delay (Δ) between the two gradient pulses was 125 ms, and the PFG strengths varied from 5 to 100% (30 G/cm). In a set of spectra shown (Figure S1A, Supporting Information), the intensities (A') of several resolved ^{1}H resonances (marked with *), such as A8 H2 (8.29 ppm), G5 H8 (7.79 ppm), G9 H1′ (5.74 ppm), NCSi-glu H2M (2.88 ppm), and NCSi-glu HNM (1.90 ppm), were measured as a function of gradient field strength, and their relationship is defined by the following equation (*24*, *25*):

$$\ln(A') = -(\gamma\delta g)^2(-\delta\Delta/3)D_T + \ln(A_o)$$

where A_o is the signal intensity in the absence of gradient pulse, and γ is the gyromagnetic ratio (rad T^{-1} s^{-1}). The plots of $\ln(A')$ versus $(\gamma\delta g)^2(-\delta\Delta/3)$ follow a linear relationship, and D_T (cm^2 s^{-1}) values were derived from the slope of these lines (Figure S1B, Supporting Information).

(b) 2-D experiments: DQF-COSY, COSY-45, TOCSY, NOESY, and ^{1}H−^{31}P correlation spectra were recorded with data points in the t_2 dimension half of that used for 1-D spectra. For samples in D_2O, acquisition times in t_2 were 682 and 377 ms for COSY and NOESY, respectively, and that for samples in H_2O was 152 ms. A total of 512 increments were used in t_1 with $t_{1max} = 171$ or 79 ms for D_2O or H_2O samples. ^{1}H−^{31}P correlation spectra used a spectral width of 3.3 ppm in t_2 (^{1}H) and 4.6 ppm in t_1 (^{31}P). Acquisition time was 512 ms for the ^{1}H and 133 ms for the ^{31}P dimension. Relaxation delays were based on T_1 estimation, ranging from 1.6 to 2.5 s. Suppression of the HOD resonance was achieved by the presaturation for D_2O samples or the Jump-return pulse ($90_\chi - \tau - 90_{-\chi} -$ Acq) (*26*) for H_2O samples with a refocusing delay of 52 μs. A 90° phase-shifted sine bell function was used in the t_2 and t_1 dimensions prior to Fourier transformation of NOESY, while a 30° phase-shifted skewed sine bell function was for TOCSY and DQF-COSY; a 37.5° phase-shifted skewed sine bell function was for ^{1}H−^{31}P correlation spectra; and an unshifted sine bell function was for COSY-45. The final spectral matrix consisted of 2048 × 2048 data points in each of the ^{1}H frequency (F2 and F1) dimensions or 1024 × 1024 in ^{1}H and ^{31}P dimensions in ^{1}H−^{31}P 2-D spectra.

(c) NMR spectra of the free DNA: The free DNA (2 mM) was dissolved in the SPE10 buffer solution, and the temperatures used for data acquisition were 268 K for 2-D J-R NOESY spectra (200 ms mixing time) (Figure 1A), 273 and 280 K for NOESY spectra of nonexchangeable ^{1}H (70, 150, 250, and 400 ms mixing times) (Figure 2A) and ^{1}H−^{31}P COSY (Figure S2A, Supporting Information), and 280 K for DQF-COSY, COSY-45, and TOCSY (88 ms mixing time).

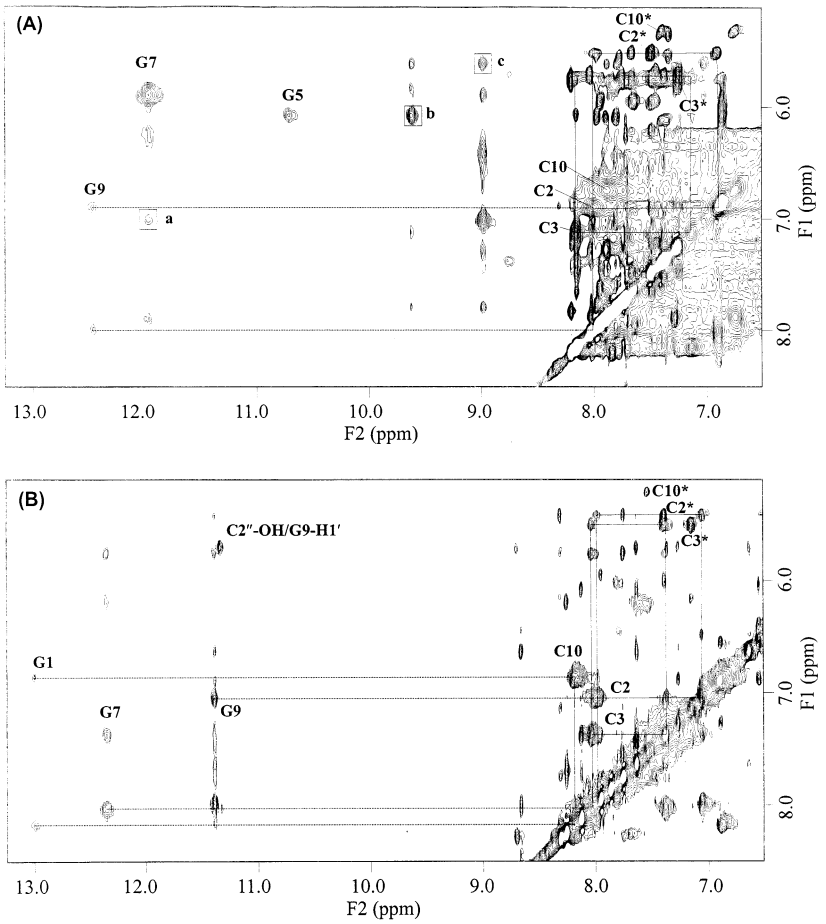


FIGURE 1: Expanded 2-D JR-NOESY spectra (200 ms mixing time) recorded in H_2O/D_2O (90:10%) at 268 K. The F2 axis covers the resonances of G HN and the hydrogen bonded and non-hydrogen bonded C NH_2 (dot line). The F1 axis covers C NH_2 (solid line) and base protons. (A) The spectrum of free DNA, showing the assignments for HN of G5, G7, and G9; amino groups of C2, C3, and C10; and H5/H6 marked with *. The cross-peaks in squares are assigned as follows: a, G7HN−A8H_2; b, G5NH_2−A6H_1'; and c, G7NH_2−A8H_1'. (B) The spectrum of complex 3 showing three sets of NOE connectivities for HN of G and amino group of C residues; amino groups of C2, C3, and C10; and H5/H6 marked with *. The intermolecular NOE between C2″ OH and G9 H1′ is indicated.

(d) The NCSi-glu-bulge DNA complex: Complex 3 (0.9 mM in SPE10 buffer solution) was used to record 2-D spectra at 268 K for 2D J-R NOESY spectra (150 and 200 ms mixing times) (Figure 1B) at 268, 278, and 283 K for NOESY of nonexchangeable ^{1}H (70, 120, 200, 250, and 400 ms) (Figure 2B); and at 268, 273, and 278 K for DQF-COSY, COSY-45, and TOCSY (88 ms mixing time) and ^{1}H−^{31}P COSY (Figure S2B, Supporting Information).

NMR Resonance Assignment. The ^{1}H and ^{31}P chemical shift assignments of complex 3 are provided in Table S1 (Supporting Information). (a) DNA in the free and the bound forms: The identification of scalar coupled spin systems, such as H5−H6 of C residues and sugar protons, from the DQF-COSY spectra provided assignments of spin systems. H2 resonances of A were identified from the inversion recovery experiments in which relaxation delay was set so that the H2 peaks remained negative when most other resonances were at the noise level. These assignments provided the starting point for sequential assignments of DNA nonexchangeable ^{1}H resonances from a combination of NOESY and COSY spectral analyses (*27*). Exchangeable amino ^{1}H (NH_2) of C residues, recognized from their connectivity to H5−H6 cross-peaks, were used to assign base-paired imino ^{1}H (HN) of G residues (Figure 1). The overlaps of H8 resonances of A6, G7, and G9 were partially resolved in NOESY spectra recorded at 278 and 283 K. H2″ resonances were specifically assigned since H1′−H2″ cross-peaks showed stronger intensities as compared to those of H1′−H2′ cross-peaks in 70 and 120 ms NOESY. (b) NCSi-glu: NCSi-glu consists of four subunits (Scheme 1A): the reduced enediyne system (i.e., tetrahydroindacene (THI) with an attached cyclocarbonate ring), the 2-*N*-methylfucosamine ring (NMF), the naphthoate ring (NPH), and the GSH tripeptide side chain attached to C12 of THI. The chemical shifts of NCSi-glu (Table S1C, Supporting Information) in complex 3 were assigned from the 250 ms mixing time NOESY spectrum for the nonexchangeable protons and from the 200 ms mixing time J-R NOESY spectrum for the exchangeable proton. Similar to what was described previously (*4*, *5*), NCSi-glu has a naphtholic proton (C2″-OH), which was easily identified at 11.38 ppm since its NOE patterns differ from those of G HN protons of DNA. Several scalar coupled cross-peaks (strong coupled: NPH H3″−H4″, THI H5−H6, and NMF H2′−H3′; not as strongly coupled: NMF H1′−H2′ and H3′−H4′) in the DQF-COSY spectra were the initiation point for complete assignments of the ^{1}H resonances. In the TOCSY spectra, the multi-bond correlations of H6−H8 in THI; H1′−H3′, H2′−H4′, and H3′−H5′ in NMF; and H6″−H8″and H6″− H5″ methyl protons (HNM) in NPH were assigned.

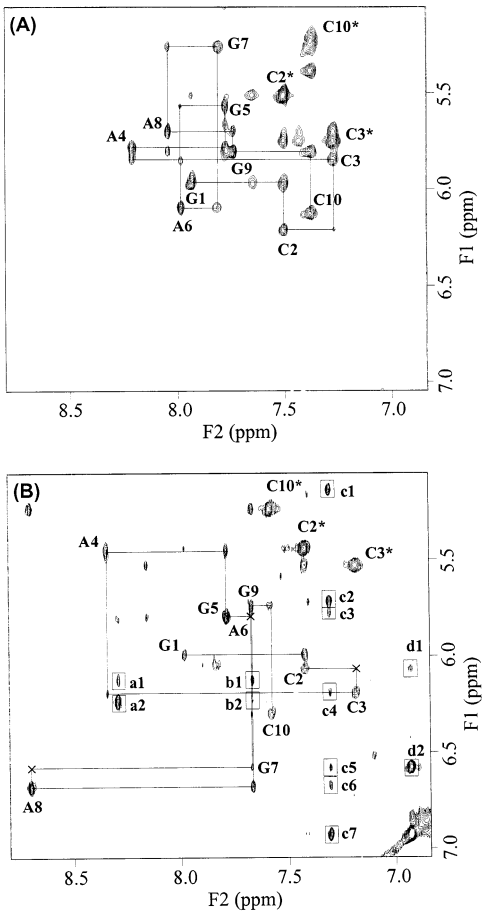


FIGURE 2: Expanded 2-D NOESY spectra recorded in D_2O of (A) the free DNA and (B) the NCSi-glu-bulge DNA complex. The sequential connectivity is shown by solid connecting lines with the cross-peaks of intramolecular NOEs labeled with residue numbers. NOEs connecting H5/H6 of C are marked with *. In B, the missed sequential connectivities of the C2−C3, G5−A6, and G7−A8 steps are marked with ×. The cross-peaks of intermolecular NOEs in complex 3 are as follows: a1, A8 H2−NPH H3″; a2, A8 H2−NPH H4″; b1, G7 H8−NPH H3″; b2, G7 H8−NPH H4″; c1, THI H8−H10; c2, THI H8−NMF H1′; c3, THI H8−H11; c4, THI H8−C3 H1′; c5, THI H8−H5; c6, THI H8−NPH H8″; c7, THI H8−H6; d1, THI H6−C2 H1′; and d2, THI H6−H5.

NMR Restraints. Distance restraints of complex 3 were derived from five D_2O NOESY spectra (70, 120, 200, 250, and 400 ms mixing times) for nonexchangeable ^1H and H_2O NOESY spectrum (200 ms mixing time) for exchangeable ^1H. The integrated volumes of the D_2O NOESY cross-peaks were obtained using the FELIX 98 program. The five sets of volumes were converted to distances using the MARDIGRAS program (Matrix Analysis of Relaxation for DIscerning GeometRy of an Aqueous Structure) (*28*, *29*). Using isotropic correlation times of 1.5, 2.5, and 3.5 ns, three sets of equilibrium distances were calculated for each volume file. This procedure generated 15 sets of distance restraints, and the average values of the 15 sets gave equilibrium distances. The upper and lower distance bounds were derived from their standard derivations but limited at 1.8−5.5 Å. The distances of exchangeable ^1H were derived from H_2O NOESY cross-peaks according to their intensities: strong (2.7 ± 0.9 Å), medium (3.4 ± 1.6 Å), or weak (4.2 ± 2.4 Å). Watson−Crick base pairing restraints further included distances connecting hydrogen-heavy atoms and donor−acceptor atoms based on canonical base-pair geometries of nucleic acids (*30*). These analyses generated 276 distance restraints—including 35 intermolecular and 18 G•C Watson−Crick hydrogen bonding distance restraints (Table 1). The quality of distance restraints was improved in the iterative calculations (e.g., after the generation of a set of structures, the average distances between each proton pair were computed and compared with the experimental distance restraints). For the distances that fell within the given restraints, the original distance bounds were reduced to narrower restraints. The distances that resided outside restraints were examined against experimental data, and their distance restraints were adjusted when appropriate. These comparisons enhanced the consistency of the restraints so gradual convergence was achieved after eight cycles of iterations.

Table 1: Statistics of NMR Data of the Nine Refined Structures of Complex 3

A. NMR Data	
NMR distance restraints	276
DNA hairpin	170
intra-residue	98
inter-residue	54
hydrogen bond	18
NCSi-glu	71
NCSi-glu-DNA (intermolecular)	35
torsion angle restraints	40
DNA hairpin	35
NCSi-glu	5
B. Structure Statistics	
NOE violations	
number (>0.4 Å)	0
rmsd of violations (Å)	0.051 ± 0.001
deviations from ideal covalent geometry	
bond length (Å)	5.70E−4 ± 5.48 E−5
bond angle (deg)	1.75 ± 2.98 E−3
impropers (deg)	1.29 ± 3.28 E−3
rmsd of coordinates (Å) among the nine refined structures (all residues and NCSi-glu)	
comparison to the averaged structure	0.36 ± 0.05
pairwise comparison	0.54 ± 0.08

Dihedral angle restraints for sugar and backbone torsion angles were obtained from combined analyses of the COSY-45, DQF-COSY, and $^1H-^{31}P$ COSY spectra. Thirty-five dihedral angle restraints related to the DNA oligonucleotide were used in the structure calculations. The DNA sugars of G1, C2, G5, A8, and C10 residues were in the C2′-endo family. For these residues, $^3J_{H1'-H2'}$ was greater than $^3J_{H1'-H2''}$ in combination with absent or very weak H2″−H3′ and H3′−H4′ cross-peaks. Therefore, these residues were restrained in δ (C5′−C4′−C3′−O3′) = 140 ± 20° (*31*). No sugar angle restraints were applied in the initial stage of the calculation to the C3, A4, A6, G7, and G9 residues since most of their COSY cross-peaks were relatively weak in intensity because of line width broadening that would be associated with local micro-motions in these resides in the mini-hairpin sequence. In the final refinement stage, δ angle restraints of 140 ± 20° were applied to these residues. The γ (O5′−C5′−C4′−C3′) angles of all residues were restrained within the gauche$^+$ (g^+) conformation (60 ± 20°) since the cross-peaks between H6/H8 and H5′/H5″ protons were weaker than those between H6/H8 and H1′ in the short mixing time NOESY spectra (70 and 120 ms) (*32*). The β (P−O−C5′−C4′) torsion angles of C2, C3, A4, G7, G9, and C10 were set within the trans (t) conformation (−151 ± 20°), based on the detection of

Table 2: Backbone Torsion Angles (deg) for D(GCCAGAGAGC)[a] in Complex 3

residue	sugar pucker[b]	α	β	γ	δ	ϵ	ζ	χ
G1	C2′-endo			61 ± 0	122 ± 2	−192 ± 4	−92 ± 2	−120 ± 6
C2	mixed[b]	−66 ± 5	−174 ± 6	59 ± 1	114 ± 2	−168 ± 3	−81 ± 3	−113 ± 3
C3	C2′-endo	−61 ± 3	−163 ± 5	66 ± 1	132 ± 3	−179 ± 3	−78 ± 2	−103 ± 5
A4	**C2′-endo**	**−50 ± 5**	**−194 ± 2**	**64 ± 1**	**138 ± 1**	**−184 ± 1**	**−131 ± 4**	**−102 ± 4**
G5	**C2′-endo**	**−49 ± 6**	**−231 ± 3**	**82 ± 3**	**126 ± 1**	**−174 ± 2**	**−184 ± 4**	**−142 ± 4**
A6	mixed	**−96 ± 4**	**−234 ± 3**	**65 ± 3**	**113 ± 2**	**−192 ± 2**	**−82 ± 4**	**−96 ± 6**
G7	mixed	−59 ± 4	−181 ± 2	61 ± 2	108 ± 4	−166 ± 3	−87 ± 3	−84 ± 2
A8	**C2′-endo**	***−61 ± 2***	***−164 ± 3***	***65 ± 1***	***120 ± 3***	***−279 ± 1***	***−82 ± 4***	***−94 ± 3***
G9	C2′-endo	−61 ± 3	−173 ± 3	65 ± 1	117 ± 1	−169 ± 2	−94 ± 2	−146 ± 2
C10	mixed	−74 ± 2	−178 ± 5	61 ± 2	112 ± 1			−111 ± 4
B-form[c]	C2′-endo	−39	−151	36	157	159	−98	−98
A-form[c]	C3′-endo	−88	−149	47	83	171	−44	−153

[a] Residues in the loop region are underlined, and the bulge residue is in italic. All torsion angles are the average of the final nine refined structures. The values for the loop and bulge residues are in bold and bold/italic, respectively. [b] Sugar pucker parameters were measured using the CURVES 5.1 program (*37*) from a representative structure. "Mixed" indicates an averaging sugar conformation between C2′-endo and C3′-endo conformations. [c] Values measured from the structures generated using the QUANTA program.

the long-range H4′−P coupling (Figure S2B, Supporting Information) combined with the g^+ conformation of the γ angles. The ϵ (C4′−C3′−O3′−P) angle is commonly found in either g^- or t; the $\epsilon(g^+)$ conformation is sterically forbidden (*33*). The $\epsilon(g^-)$ conformation would give a $^4J_{H2'-P}$ coupling cross-peak (*34*), but this was not observed in our NMR data. Therefore, the ϵ angles of all residues were restrained to a t conformation (180 ± 20°).

For NCSi-glu, five dihedral angle restraints were derived, including three ring pucker and linkage dihedral angles for NMF and two t dihedral angles for the peptide linkages. Large $J_{H2'-H3'}$ and small $J_{H1'-H2'}$ and $J_{H3'-H4'}$ of the NMF ring were observed. These data are consistent with the NMF ring in a chair conformation and an α-glycosidic linkage. NOE cross-peaks between Hα protons of the tripeptide residues were not observed, indicating the t conformation of the peptide bond (*27*).

Structure Calculation. Structure calculation used distance geometry (DG) followed by restrained molecular dynamics simulations (rMDS), which are implemented in the XPLOR program (*35*, *36*). The parameter file of NCSi-glu was essentially the same as previously used (*4*, *5*) except for the GSH tripeptide side chain, which were incorporated according to the parameters of standard amino acids (QUANTA, Accelrys Inc.). Calculation methods followed what was described previously (*7*). A calculation cycle typically began with DG calculations, which generated a large number of initial structures (>300). These structures were analyzed in terms of minimum energy, chirality of sugar protons, and the orientation of the A4, G5, and A6 loop residues and the A8 bulge residue. As discussed below, in the major structural family the bases of these residues were located in the major groove. In the second stage of structure calculation, the selected DG structures were subjected to rMDS refinement using a thermal annealing procedure and parameters as reported (*7*). The calculations were iterated, and structures were examined to ensure the correct chirality, favorable energy terms, and the consistency between the calculated structures and the experimental data. The statistics for the final nine structures of NCSi-glu-bulge DNA (complex 3) are listed in Table 1. The CURVES 5.1 program (*37*) was used to analyze the helical parameters of DNA, and these results are given in Table 2. The geometry parameters related to the torsion angles of NCSi-glu are listed in Table 3. The

Table 3: Rotatable Bond Angles and Backbone Torsion Angles (deg) of NCSi-glu in Complex 3

rotatable bond	dihedral angle	average value (deg)	standard deviation
THI−NMF linkage	(C10)(O1′)(C1′)(C2′)	−161.7	1.3
THI−NPH linkage	(C12)(C11)(O)(C═O)	−173.5	6.4
	(C11)(O)(C═O)(C1″)	−162.1	4.7
	(O)(C═O)(C1″)(C2″)	−85.1	3.2
THI−CYC linkage	(C3)(C4)(C13)(C14)	−76.0	0.8
THI−GSH linkage	(C11)(C12)(S)(CB)	−176.8	0.8
C−C single bond	(C1′)(C2′)(NHMe)(CNMe)	−72.8	6.8
	(C6″)(C7″)(O7″Me)(CO7″Me)	160.5	9.5

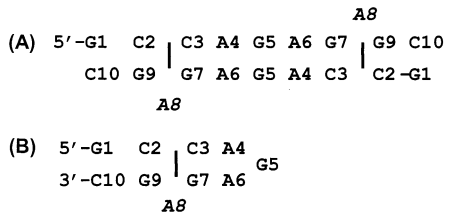
backbone and side chain torsion angle in tripeptide		average value (deg)	standard deviation
ϕ_i $(C'_{i-1})(N_i)(C^{\alpha}_i)(C'_i)$	ϕ_2	−165.3	0.5
	ϕ_3	90.3	1.5
ψ_i $(N_i)(C^{\alpha}_i)(C'_i)(N_{i+1})$	ψ_2	160.9	0.8
χ_i^1 $(N_i)(C^{\alpha}_i)(C^{\beta}_i)(X_i)$	χ_1^1	−74.3	2.7
	χ_2^1	−56.1	1.7
χ_i^2 $(C^{\alpha}_i)(C^{\beta}_i)(C^{\gamma}_i)(X_i)$	χ_1^2	176.1	2.2
	χ_2^2	164.2	1.4
χ_i^3 $(C^{\beta}_i)(C^{\gamma}_i)(C'_i)(X_{i-1})$	χ_1^3	51.0	2.3
ω_i $(C^{\alpha}_{i-1})(C'_{i-1})(N_i)(C^{\alpha}_i)$	ω_2	−142.9	1.2
	ω_3	142.7	0.8

coordinates of the structures of complex 3 have been deposited at the Protein Data Bank (accession code 1MP7).

RESULTS

NMR Spectra of the Free DNA. NMR spectral analysis of the free DNA provides a basis for comparison of the free and the drug bound conformations of DNA. ^1H and ^{31}P resonance assignments are given in Table S1A (Supporting Information). Three HN proton resonances were observed and assigned to HN of G5 (10.73 ppm), G7 (11.95 ppm), and G9 (12.46 ppm) (Figure 1A). 2-D NOEs connected G2 HN to C9 NH_2, indicating the formation of a C2•G9 base pair, but the NOEs between HN and NH_2 of C3 and G7 and between C10 and G1 were absent. An NOE observed at 11.95 and 7.91 ppm was assigned to G7 HN and A8 H2, suggesting stacking of the G7 and A8 residues. The sequential NOEs (Figure 2A) can be traced almost throughout the sequence, but those of H6/H8 to H1′ were weak at the

Scheme 2: Possible Sequence Alignments of DNA in Complex 3. (A) A Homoduplex and (B) a Self-Folded Hairpin Bulge Duplex



C2−C3 and G5−A6 steps. The DQF-COSY spectra showed stronger H1′−H2′ than H1′−H2″ cross-peaks and weak or absent H3′−H4′ cross-peaks of all residues, reflecting the free DNA predominantly in the C2′-endo conformation. The ^{31}P resonances exhibited about 1 ppm of chemical shift dispersion with pG9 at −3.95 ppm and pG5 at −4.84 ppm, respectively (Table S1A and Figure S2A, Supporting Information). pA4 was also relatively upfield shifted, and pA8 resonated at −4.12 ppm, which is within the expected canonical ^{31}P values.

Diffusion Constant Measurements. The above NMR analysis detects formation of a C•G base pair and base helical stacking, but it does not specify whether the decamer DNA sequence formed a homoduplex with an intra-helical loop (Scheme 2A) or a hairpin with a hairpin loop (Scheme 2B), because of spectral similarity of the two conformations. Previously, we showed that the translational diffusion constant (D_T) of a hairpin was 1.4 times faster than that of the corresponding homoduplex (*22*). In this experiment, D_T ($\times 10^{-7}$ cm^2/s) were measured as 6.56 ± 0.25 (free DNA), 8.44 ± 0.36 (complex 3), and 5.64 ± 0.10 (heptamer duplex) (Figure S1, Supporting Information). The D_T ratios of free DNA/heptamer duplex and complex 3/free DNA were 1.2 and 1.3, respectively. The free DNA's translational diffusion is faster than the heptamer duplex but 1.3 times slower than complex 3.

NMR Spectra of DNA in Complex 3. Analyses of the NOESY and COSY spectra of DNA in complex 3 produced nearly complete assignments of the ^{1}H and ^{31}P resonances (Table S1B, Supporting Information). Three G•C base pairs were discernible from the NOESY spectra (Figure 1B), and these were assigned to G1•C10, C2•G9, and C3•G7 base pairs. A broad resonance at 10.62 ppm was assigned to G5 HN, which became too broad to be observed at 290 K. In comparison, the three G HN resonances in the free DNA became broad at a lower temperature (280 K). The sequential connectivities of base to sugar ^{1}H were obtained, but those in the C2−C3, G5−A6, and G7−A8 steps were extremely weak or absent (marked by ×, Figure 2B). The break down in the sequential connectivities was consistent with the ratio of the diffusion constants, reflecting the formation of a mini-hairpin DNA, possibly with a bulge adenosine embedded within a three base pair stem closing the hairpin loop (Scheme 2B). DNA in complex 3 also showed poor coupling connectivities, including those correlating H3′ resonances. Since the observation of the coupling cross-peaks is sensitive to the line broadening of NMR resonances, the disappearance of these cross-peaks in complex 3 is attributed to the lack of a long stretch of helical duplex stabilizing the microscopic local motions in DNA. The $J_{H3'-P(n+1)}$ coupling cross-peaks linking C2−pC3, A6−pG7, and G7−pA8 steps were also not observed, in line with the absence of the corresponding proton coupling cross-peaks.

There were additional indications of structural alteration upon complex formation (Table S1B, Supporting Information): the H2′ chemical shifts of C3, G5, and A8 were shifted downfield as compared to those of the H2″ resonances, where the assignments of H2′ and H2″ were verified in the 70 ms NOESY spectrum. Base paired HN resonances are likely to be affected by base stacking, and the HN resonances of G7 and G9 were largely shifted to downfield by 0.40 ppm and upfield by 1.04 ppm, respectively. Changes in the chemical shifts of base ^{1}H resonances may be influenced by base stacking such as the extent of the base moiety exposed to solvent; H8 and H2 of the bulge base A8 showed largest downfield shifts (0.70 and 0.38 ppm). Chemical shift differences of sugar ^{1}H resonances upon complex formation would largely reflect the change in the relative positioning of the sugar residues to their own and 3′ adjacent base moieties. These resonances were upfield shifted in C2 and downfield shifted in C3. In A4 H1′ was upfield shifted; in G5 H1′ was downfield shifted, and H5′,5″ were upfield shifted; in A6 H1′ was upfield, and H2′,2″ were largely downfield shifted (as much as 1 ppm); in G7 H1′ and H2′,2″ were downfield shifted (as much as 1.37 ppm), and H3′ and H5′,5″ were upfield shifted. In A8 these resonances were downfield shifted (as much as 1.03 ppm) followed by noticeable changes in those of G9 and C10 residues. The ^{31}P resonances of complex 3 appeared within 1 ppm region with the resonances of pC3 and pG5 and showed significant shifted chemical shifts (upfield shifted by 0.99 ppm and downfield by 0.69 ppm) at −5.05 and −4.15 ppm, respectively (Figure S1B, Supporting Information).

NMR Spectra of NCSi-glu in Complex 3. The analysis of the intramolecular NOE and COSY connectivities provides information that defines the conformation of the drug in the complex. The assignments of the ^{1}H resonances in NCSi-glu followed a similar approach used previously (*4*) and are given in Table S1C (Supporting Information). The chemical shifts of the bound NCSi-glu showed sufficient characteristics of complex formation. The ^{1}H resonances of NPH were largely upfield shifted by more than 1.5 ppm (H4″, H8″, and 7″-methyl); THI H2 was upfield shifted by 1.03 ppm, and the resonances of the cyclocarbonate ring (H14a,b) were downfield shifted. Previously in complex 1, the ^{1}H resonances of the GSH side chain were not observed. In complex 3, these resonances are well-defined. Strong DQF-COSY cross-peaks of Hα1 and Hα2 of glycine (Gly) and Hβ1 and Hβ2 of cysteine (Cys) were unambiguously assigned. The spin systems linking Hα with Hβ resonances in the three amino acids were then identified, and the long-range scalar coupling of Hα−Hγ of the γ-glutamic acid (γ-Glu) was confirmed in the TOCSY spectra. In addition, Gly Hα and Cys Hβ resonances showed NOEs to THI H2 and H14a,b of the cyclocarbonate ring; Cys Hβ also showed NOEs to THI H11 and H12; and γ-Glu Hβ2 and Hγ1 showed NOEs to NMF H3′ and 2′-*N*-methyl (H2M). The stereospecific assignments of some methylene protons in GSH were obtained; pairs of CH_2 protons gave different NOE intensities because of their stereospecific orientations. For instance, among pairs of stereospecific protons, only Gly Hα2 showed NOE to H13, but Gly Hα1 showed strong and medium intensities of NOEs to C10 H4′ and H5″; only Cys Hβ2

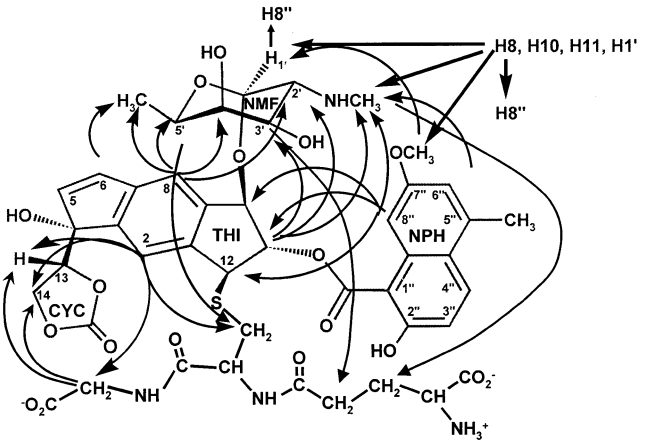


FIGURE 3: Schematic drawing showing intramolecular NOEs of NCSi-glu in complex 3.

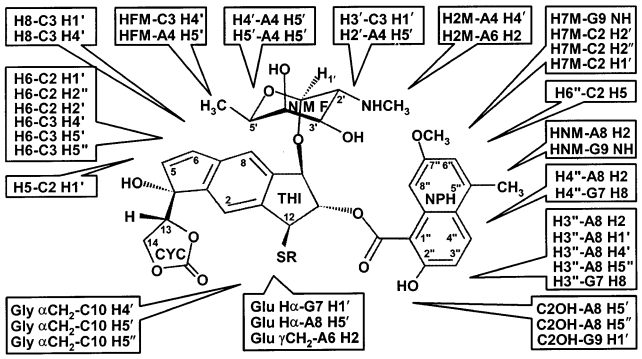


FIGURE 4: Schematic drawing showing intermolecular NOEs in complex 3.

showed NOE to NMF H5′; only γ-Glu Hγ1 and Hβ2 showed NOE to 2′-*N*-methyl protons (H2M). The observed NOE patterns allowed the assignments of pro-R and pro-S to Gly Hα1 and Hα2 and to Cys Hβ1 and Hβ2. These NOEs clearly define an extended conformation of the GSH side chain. In the orientation of the GSH side chain, the Gly moiety aligns close to the THI ring and its substituent cyclocarbonate ring, and the Glu moiety points toward the NMF ring. For the bound NCSi-glu, a complete network exists containing 71 intramolecular connectivities within the four units in NCSi-glu as displayed in Figure 3. These results show NOEs connecting H8″ of NPH with H10, H11, and H12 of THI as well as with H2M and H1′ of the carbohydrate NMF ring, revealing the overall orientation of the subunits of NCSi-glu.

Intermolecular NOEs in Complex 3. Thirty-five intermolecular interactions in complex 3 containing the hairpin bulge DNA and NCSi-glu (Scheme 1A) are displayed in Figure 4. Examination of the contacting patterns in complex 3 indicated that mainly the C2•G9, C3•G7, A4 (loop), and A8 (bulge) residues of DNA formed the site of binding with NCSi-glu. Many NOEs involved H1′, and a few involved H2 of A6 and A8 of DNA, which were diagnostic for the binding of NCSi-glu in the minor groove of the DNA, even though the minor groove, presented by the nonconsecutive three C•G base pairs, G1C2C3•G7−G9C10 (− indicates A8 bulge base), was too short to present the full features of a canonical minor groove (*30*). Collectively, the NPH ring (C2″-OH, H3″, and H4″) showed NOEs to G7 H8, bulge A8 (H2, H1′, H4′, and H5′/H5″), and G9 H1′; the NPH ring (HNM, H6″, and H7M) showed NOEs to C2 (H5, H1′, H2′, and H2″), bulge A8 H2, and G9 HN. These NOEs were consistent with the NPH ring intercalating between the C2C3•G7A8G9 helical step with the side of C2″-OH closer to the bulge A8. The site causing DNA chain cleavage was examined. The NPH ring is attached to the C11 position of the THI ring, which is a reduction product of a C2,C6-biradical species (Scheme 1A). The THI ring exhibits important NOEs connecting its H5, H6, and H8 resonances with C2 (H1′, H2′, and H2″) and with C3 (H1′, H4′, and H5/H5″). These NOEs bring the C6 radical site of THI to close vicinity of the backbone at the C2−C3 step so that THI H6 is 2.42 ± 0.05 Å away from H5′ of C3. The NMF carbohydrate moiety attached to the C10 position of THI and their NOEs indicate that NMF (H2′, H3′, H4′, H5′, HFM, and H2M) makes contacts with DNA residues C3, A4, and A6. These NOEs place the carbohydrate unit in the minor groove around C3•G7, extending toward the A4 residue at the 3′ side of the nonbulge strand. The GSH tripeptide side chain in complex 3 was found to be better accommodated as compared to complex 1 because a set of well-defined NOEs in complex 3 connected this residue with the bound DNA. These NOEs mainly involved Gly Hα with C10 (H4′/H5′ and H5″) as well as γ-Glu (Hα, Hγ) with A8 H5′, G7 H1′, and A6 H2. Therefore, the GSH chain is closer to the G7A8G9 strand and aligns against the minor groove side in an extended conformation. Anchored to the C12 position of the THI, the GSH tripeptide has its Gly residue extending to the G9 and C10 residues and its γ-Glu moiety filling the bulge cavity. While the THI and NMF moieties mainly interact with the C2C3A4 strand, the GSH moiety interacts with the opposite strand around the bulge site.

Structure Elucidation of Complex 3. The final nine structures of complex 3 were elucidated without using a preconceived model, (i.e., using DG for generation of initial structures). A total of 276 distances in combination with 40 dihedral angle restraints were applied in the calculation. These included five dihedral angles defining the conformation of NMF and the peptide linkage of GSH of NCSi-glu, 71 intramolecular inter-proton restraints of NCSi-glu, and 35 intermolecular inter-proton restraints between NCSi-glu and DNA in complex 3. An overlay drawing of these structures is shown Figure 5A. These structures agree with the experimental analysis, and the calculation statistics are given in Table 1. Although the structure calculation was straightforward (*4*, *7*), the conformations of the DNA hairpin loop and the bulge adenosine residue, which were not initially restrained, had to be resolved after the initial structures were generated. After comparing with the experimental data, a portion of the 300 structures, which contained significant NOE violations, were excluded from further calculation. Among the remaining structures, mainly three families of structures, differing in the orientation of the three loop residues and the bulge residue with respect to major (M) and minor (m) grooves, were found: family I (14 out of 300 structures) was A4(m)−G5(m)−A6(M) and A8(M), family II (39 structures) was A4(M)−G5(m)−A6(M) and A8(M), and family III (114 structures) was A4(M)−G5(M)−A6(M) and A8(M). These structures were further analyzed on the basis of experimental results. Families I and II containing a G5(m) conformation, G5 base and sugar moieties, and A6 sugar moiety would be in too close contact,

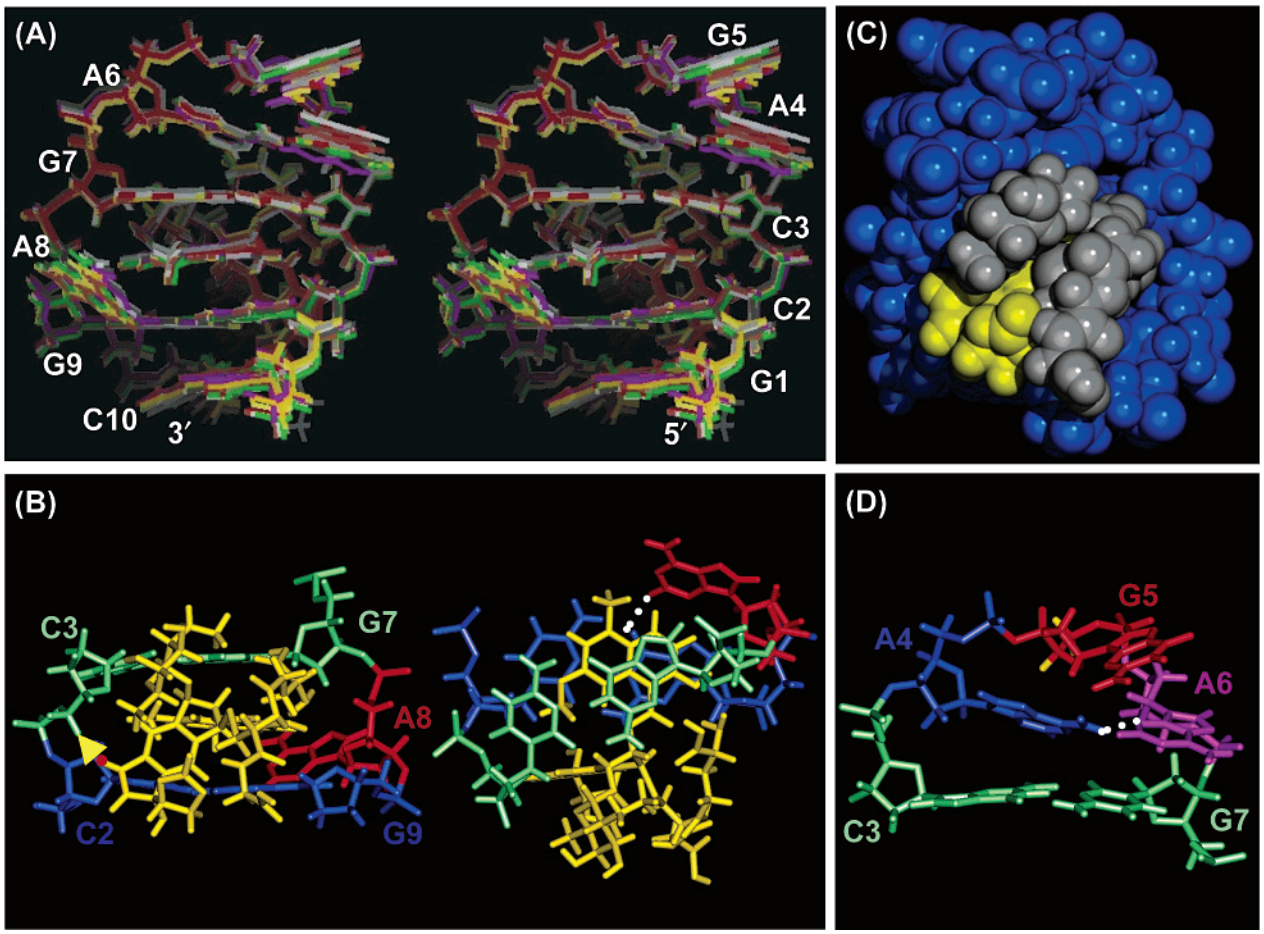


FIGURE 5: Structure of complex 3. (A) Stereoview of the superimposed final nine refined solution structures of the NCSi-glu-bulge DNA hairpin complex looking into the major groove. (B) The binding site in complex 3. The structure on the left side displays a view into the minor groove. NCSi-glu (yellow) has its naphthoate ring intercalates into the step of C2•G9 (blue) and C3•G7 (green), and the binding site contains a bulge (A8, magenta). The carbohydrate, the reduced enediyne ring, and the GSH tripeptide side chain of NCSi-glu (yellow) snugly fit in the minor groove. A yellow arrow is pointing toward the 5′ position of C3 from NCSi-glu H6 marked with a red dot. The structure on the right side is a bird view of the stacking of DNA bases and that of DNA bases with the intercalating naphthoate ring (yellow). (C) Space filling drawing of the structure of complex 3 shown in an orientation similar to that of 5B. The bulge DNA hairpin is in blue with the hairpin on top of the figure; NCSi-glu is in yellow, but its carbohydrate and GSH peptide side chain are in gray. The carbohydrate unit is close to the left side of the DNA strand; the extended GSH side chain is close to the right side of the DNA strand. (D) DNA hairpin folding in complex 3. A plausible hydrogen bond between A4 N3 and A6 NH_2 (2.60 ± 0.01 Å) is indicated by a white dot line.

but this is inconsistent with the absence of NOEs between two residues. The structures of family III represented the major conformations calculated, and they largely agreed with the NMR experimental data. These structures underwent iterative rMDS calculations to derive the final nine structures.

DISCUSSION

Global Structural Features of Complex 3. The global structure of complex 3 features the minor groove binding of NCSi-glu to a mini-hairpin bulge DNA with the NPH ring of NCSi-glu intercalating into the C2•G9/C3•G7 step (Figure 5A) and with the bulge A8 base connecting the C3•G7 base pair to form an extended stacking surface with A8 H2 and G9 O6 separated by 2.69 Å (Figure 5B). The minor groove binder, having an extended conformation, consists of THI and NMF moieties aligning in tandem mainly along the strand of C2C3A4 and the GSH tripeptide moiety aligning, on one side in parallel with THI and NMF, and on the other side mainly along the strand of G7A8G9 containing the A8 bulge site (Figure 5C). This structure is consistent with NMR characterization of the major features: (a) the loss of sequential connectivity between adjacent residues, such as C2−C3, G5−A6, and G7−A8 steps; (b) the absence of the correlation peaks of $H3'_n - P_{n+1}$ at the C2−C3, A6−G7, and G7−A8 steps and the large chemical shift perturbations at the C2−C3 and A4−G5 steps; (c) the detection of intermolecular NOEs between DNA bases (C2•G9, C3•G7, and A8) and the NPH ring; (d) intermolecular NOEs of H1′ protons of DNA with NCSi-glu; and (e) intra- and intermolecular NOEs of γ-Glu to H2M of the NMF ring and to DNA bases (A6, G7, and A8) and those of Gly to the cyclocarbonate ring (H13 and H14a,b) and to C10 (H4′ and H5′/H5″) as discussed in detail in the Results section. At the interaction site, the NPH ring aligns its long axis parallel with the long axis of the C2•G9; NPH stacks better with the purine ring of G7 than that of G9 (Figure 5B). The complex is largely stabilized by the stacking interactions between the aromatic moieties in the bulge intercalation site and the van der Waals interactions between NCSi-glu and DNA along the minor groove. There are several possible hydrogen bonds in complex, which include N3 of G7 and the carboxylic acid hydrogen of γ-Glu (1.93 Å, 175°); NH_2 of G7 and the

carbonyl oxygen of γ-Glu of the GSH tripeptide (2.14 Å, 163°); NH_2 of γ-Glu and O of NPH C2″-OH (2.10 Å, 119°); the same NH_2 of γ-Glu with the former and the carbonyl oxygen at C1″ of NPH (1.80 Å, 151°). Values of hydrogen bond distances and bond angles in the parentheses are from the average structure.

NCSi-glu Conformation in Complex 3. NCS-chrom is a molecule of multicomponents, which render the structural flexibility of the molecule as seen from the various derivatives formed upon activation (*1*, *2*) and from the various binding modes when interacting DNA (*3*). The pentacyclic moiety of THI is highly substituted (Scheme 1A), linked at C10 to the NMF carbohydrate moiety through glycosidic bonds, at C11 to the NPH aromatic moiety through an ester bond, and at C12 to the GSH tripeptide side chain through a C−S bond. Additionally, C4 of THI is linked to a pentacyclocarbonate. These linkages lack $^1H-^1H$ scalar coupling connectivities, but NOE analysis of inter-proton contacts based on both the intra- and intermolecular interactions provided restraints defining the relative orientation of these subunits. For instance, THI H10 showed weak NOEs to NMF (H3′ and H5′) but not to NMF H1′. THI H2 and H12 showed NOEs to Cys Hβ1,2 protons, while THI H11 showed NOE only to the Cys Hβ1 proton. The corresponding dihedral angles as defined for each described linkage, calculated from the nine refined structures of complex 3, are summarized in Table 3. Overall, NCSi-glu adopts an extended conformation, which is induced by binding to the DNA decamer. Large chemical shift variations (e.g., 1.54 ppm upfield shifted and 0.54 ppm downfield) between the free and the bound forms of NCSi-glu were detected (Table S1C). The H13 and H14a,b protons of the cyclocarbonate ring showed NOEs with THI H2. These NOEs were not detected in the free ligand, suggesting binding-induced changes in the relative orientation of the cyclocarbonate ring to the THI moiety. This conformational change may be important for NCSi-glu DNA recognition in the minor groove through an extended hydrophobic surface. In complex 3, the detection of 22 intra- and intermolecular NOEs related to the GSH tripeptide side chain, and scalar coupling analysis of its corresponding protons reveal the conformation of this side chain. The orientation of the tripeptide is such that the γ-Glu residue is close to the NMF carbohydrate ring, and the Gly residue is in proximity to the cyclocarbonate ring.

Hairpin Folding of DNA in Complex 3. Upon complex formation, the decamer DNA d(GCC**AGA**GAGC) folds into a mini-hairpin (hairpin loop resides in bold). The stem-loop transition is continuous at the joints of C3-**A4** and **A6**-G7 with the hairpin trinucleotide A4G5A6 in a well-defined conformation shown in Figure 5D. Within the hairpin loop residues, the bases of A4 and G5 are well-stacked, and the breakdown in helicity at the G5−A6 step to accommodate the hairpin turn is observed, pressing the sugar backbone of G5 unusually close to the base of A4 (Figure 5D). This conformation is associated with the presence of NOE sequential connectivities at the A4−G5 step and large upfield shifted chemical shifts for G5 H4′ (2.85 ppm), H5′, and H5″ (3.05 ppm, overlapping) resonances (upfield shifted from free DNA by 1.51, 0.41, and 0.54 ppm, respectively, Table S1B, Supporting Information). This continuation of base stacking from the 3′ end of the stem into the loop residue followed by turning around the point close to where the loop meets the 5′ end of the other stem appears energetically preferred in DNA hairpins (*38*). The three-purine loop bases are all pointing toward the major groove side as defined by the stem duplex. In this crowded arrangement, A4 and A6 encounter in a tail (the N3 edge) to head (the N6 edge) fashion such that N3 of A4 and the hydrogen of N6 or A6 may form a weak hydrogen bond (2.60 ± 0.01 Å). This geometry was not a result of restraints in structure calculation. The backbone perturbation was observed upon complex formation along with the hairpin loop: all residues adopt the usual $\zeta(g^-)$ $\alpha(g^-)$ β(t) dihedral conformation except for hairpin loop residues even though all ζ and α torsion angles were not constrained during calculation. For example, A4 is closer to a ζ(t) conformation; G5 is closer to a ζ(t) $\beta(g^+)$ conformation; and A6 is closer to a $\beta(g^+)$ conformation. These backbone perturbations resulted in the absence of $J_{H3'-P(n+1)}$ and $J_{H4'-P}$ peak coupling cross-peaks.

DNA Conformational Transformation upon Complex Formation. The formation of the mini-hairpin signifies a conformational transformation in DNA upon NCSi-glu binding. The free DNA under the same buffer condition does not fold back to form a hairpin as in complex 3 but exists as an antiparallel duplex, featuring terminal base pairs and an open loop in the middle of the duplex. The duplex free DNA has a translational diffusion constant 1.3 times slower than that of complex 3. In the free DNA, the A8 base is in a stacking-in orientation, showing an NOE of G7 HN to A8 H2. The conformation of the bulge consisting of an adenosine base was described in a number of DNA systems, and this type of bulge tends to have the adenosine base stacked within the helix rather than looping out (*39*).

The conformational change in DNA from a loop-containing, linear duplex to a bulge-containing, folded hairpin suggests the important role played by the hairpin loop in stabilizing the complex. Presumably, both forms can offer the bulge intercalation site, but the major difference between two conformations is whether the A4G5A6 segment is in an extended or in a folded conformation. There are only a few contacts found between NCSi-glu and the hairpin residues, mainly involving the γ-Glu residue in GSH and the backbone of A4 to G5 in DNA. These would improve the stabilization of the complex. The folding of the single-stranded DNA helps the closing of the C3•G7 base pair in a more favorable conformation for drug binding; thus, this should be important in the formation of the complex. On the basis of this assessment, we demonstrate that NCSi-glu is a unique ligand specific to single-stranded DNA structures, and its molecular architecture may serve as a lead for the design of sequence and structure specific ligands that bind to single-stranded nucleic acids.

Intercalation, Tripeptide Binding, and Efficient Single-Site Chain Cleavage. Examination of the binding site in complex 3 reveals the important role of the intercalator binding in determining the orientation of NCSi-glu. The hairpin bulge DNA in complex 3 differs from common NCS-chrom DNA chain cleavage in that the binding site does not contain a preferred T/A cleavage residue, and the site is asymmetrical with an extra A8 in one of the two strands. Thus, the intercalator NCSi-glu would have the option to adopt either orientation in the minor or major groove. NMR structures reveal that the drug molecule presents solely in one conformation, which features orientation-specific intercala-

tion of the NPH ring into the C2•G9 and C3•G7 step, partially displacing the A8 bulge into the major groove. The intercalation site in complex 3 can be characterized by a significant increase in a base rise (7.9 Å) at the C2•G9−C3•G7 step, large buckle angles: 9.1° for C2•G9 and −11.4° for C3•G7 and a large propeller angle (6.4°) within the C3•G7 base pair. The base rises of the G7−A8−G9 are similar to the base−base separation commonly seen in the intercalation binding in B-form DNA (*3*). The orientation of the intercalation thus defines the orientation of the minor groove binding components in the drug. Furthermore, the bulge site allows deeper penetration of the intercalator, NPH moiety, which in turn allows closer interactions of its substituents, especially the GSH tripeptide side chain, with DNA.

Complex 3 offers a view for peptide−DNA interactions and their allosteric effect on DNA cleavage specificity. The GSH peptide side chain was invisible in NMR spectra for the NCSi-glu-duplex DNA complex 1, where there was no bulge base present, and the side-chain binding was possibly sterically hindered; however, in the NCSi-glu-bulge hairpin DNA complex 3, the side chain was well-characterized. In complex 3, the GSH tripeptide moiety is in part embedded within the bulge site, and this binding pushes the drug toward the opposite strand, so the C6 radical is closer to the adjacent 5′-position of the C3 residue (Figure 5B). This offsets the access by the C2 radical to the second possible cleavage site by increasing the spatial separation between C2 of NCSi-glu and DNA backbone, explaining the diminishing duplex cleavage in cases of bulge binding. These results demonstrate that the binding modes found in NMR structures are directly related to the patterns of specific DNA backbone cleavage (i.e. double- and single-strand cleavage) by NCS-chrom (*2*). The tighter fit of NCSi-glu in the bulge DNA may also explain the loss of sequence selectivity at T/A residues, as the cleavage specificity can be regulated by the structure of the substrate DNA. These results clarify the observed preference of the single-base cleavage of the single-stranded decamer DNA at the 3′ side, opposite to the bugle strand (*17*).

Function of the Bulge Base. The structure of complex 3 may help to understand the results of our in vitro DNA cleavage studies using GSH activation conditions and a series of single-stranded DNA sequences. These results demonstrate that the efficiency of DNA chain cleavage depends on the type and the number of bulge residues (*17*−*19*). A single adenosine as a bulge residue causes the most specific and efficient single-base cleavage as compared to those caused by a pyrimidine bulge base or two or more bases as the bulge residues. The cleavage site is specifically located at the position opposite to the bulge residue, one base down to the 3′ end. These patterns of DNA cleavages may be attributed to the difference in the bulge binding pocket because of different bulge compositions or sizes. The A8 bulge base displays several features in complex 3 (Figure 5B): (a) it is partially displaced from the stacking-in conformation to the side alignment into the major groove, and thus, helps to release the stress in the DNA backbone caused by insertion of the drug's aromatic ring between two consecutive Watson−Crick base pairs; (b) it provides an extended stacking interface by forming a continuous surface from the C3•G9 Watson−Crick base pair so that the NPH moiety has a bias toward the G7*A8*G9 strand; and (c) it provides additional van der Waals stabilization in the minor groove and along the backbone, especially through the interactions of A8 H2 to NCSi-glu (Figure 4). Therefore, increasing the bulge size to two or more bases would reduce the stacking contacts; altering an adenosine bulge base would lead to the reduction in intermolecular stacking and/or van der Waals' attraction (either bulge base loop out or weaker ring current effect); altering bases adjacent to the bulge from purine G's to pyrimidines would suffer the same weakening in intermolecular interactions.

Two bulge DNA complexes with post-activated NCS-chrom (complex 2 with mode 2a binding and complex 3 with mode 2b binding) elucidated by NMR illustrate the structural versatility of this small ligand molecule as a model for specific recognition of DNA bulge sites. Previously in complex 2, the NMR structure revealed the formation of a prism binding cavity consisting of the two helical base pairs and the two-base bulge, where two rings of drug form a wedge that is inserted into the bulge site through the major groove of a single-stranded DNA (*6*, *7*). The bulge structures in DNA and RNA have been implicated in a number of important biological processes, such as slipped synthesis that leads to nucleotide repeat expansions in a number of neurodegenerative diseases and cancers (*40*−*45*) and frame-shift mutation. Bulges can cause kinks (*46*) and bending (*47*) in DNA helices, which may be important in the recognition of enzymes and their biological activities. Because of the importance of bulge structures, the understanding of the essentials of bulge recognition and providing effective means to achieve such recognition will be of important interest in molecular and pharmaceutical biology. It has been shown that DNA intercalators, such as ethidium, bind preferentially to DNA or RNA containing a bulge (*48*−*50*), and further footprinting studies indicate that the binding of ligands, such as *o*-phenanthroline, occurs within the bulge site (*51*). Recently, several bulge DNA sequences have been examined as the target of ligand action: nogalamycin binds to a T bulge of CTG at the terminal position, but the binding causes the bulge to slip and actually occurs in the duplex DNA (*52*); a 2-amino-1,8-naphthyridine derivative recognizes a G bulge through formation of three H-bonds and stacks in the helical DNA (*53*); the aglycone of altromycin also alkylates N7 of a G and stacks in a GG step that is opposite a single T bulge, and the resultant complex is prone to cleavage at the bulge site (*54*); macrocyclic metal ion complexes, such as those formed with ions of platinum, nickel, or cobalt, act as a chemical nuclease, exhibiting bulge specificity and induce DNA strand scission (*13*, *55*). Our NMR characterizations of the bulge DNA complexes with post-activated NCS-chrom now broaden the range of bulge sites that can be recognized by ligand molecules. Our NMR structures of complexes 2 and 3 offer distinct models for recognition of bulges of various sizes through binding at either the minor or the major grooves and for single-chain cleavage of bulge DNA sequences.

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 S1A, S1B, and S1C listing $^{1}H-^{31}P$ chemical shift; $\Delta\delta$ (ppm) data and Figure S1 and S2 showing the representative stacked plot of PFG experiment of the NCSi-glu-bulge complex; and the $^{1}H-^{31}P$ spectrum of free and bound DNA to NCSi-glu. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Goldberg, I. H. (1991) Mechanism of neocarzinostatin action: Role of DNA microstructure in determination of chemistry of bistranded oxidative damage, *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. Han, X., and Gao, X. (2001) Sequence specific recognition of ligand-DNA complexes studied by NMR, *Curr. Med. Chem. 8*, 551−581.
4. Gao, X., Stassinopoulos, A., Rice, J. S., and Goldberg, I. H. (1995) Structural basis for the sequence-specific DNA strand cleavage by the enediyne neocarzinostatin chromophore, Structure of the post-activated chromophore-DNA complex, *Biochemistry 34*, 40−49.
5. Gao, X., Stassinopoulos, A., Gu, J., and Goldberg, I. H. (1995) NMR studies of the postactivated neocarzinostatin chromophore DNA complex. Conformational changes induced in drug and DNA, *Bioorg. Med. Chem. 3*, 795−809.
6. Stassinopoulos, A., Ji, J., Gao, X., and Goldberg, I. H. (1996) Solution structure of a two-base DNA bulge complexed with an enediyne cleaving anolog, *Science 272*, 1943−1946.
7. Gao, X., Stassinopoulos, A., Ji, J., Kwon, Y., Bare, S., and Goldberg, I. H. (2002) Induced formation of a DNA bulge structure by a molecular wedge ligand-postactivated neocarzinostatin chromophore, *Biochemistry 41*, 5131−5143.
8. Kappen, L. S., and Goldberg, I. H. (1993) DNA conformation-induced activation of an enediyne for site-specific cleavage, *Science 261*, 1319−1321.
9. Kappen, L. S., and Goldberg, I. H. (1993) *Site-specific cleavage at a DNA bulge by neocarzinostatin chromophore via a novel mechanism*, *Biochemistry 32*, 13138−13145.
10. Stassinopoulos, A., and Goldberg, I. H. (1995) Probing the structure of long single-stranded DNA fragments with neocarzinostatin chromophore. Extension of the base-catalyzed bulge-specific reaction, *Biochemistry 34*, 15359−15374.
11. Hensens, O. D., Zink, D. L., Chin, D.-H., Stassinopoulos, A., Kappen, L. S., and Goldberg, I. H. (1994) Spontaneous generation of a biradical species of neocarzinostatin chromophore: Role in DNA bulge-specific cleavage, *Proc. Natl. Acad. Sci. U.S.A. 91*, 4534−4538.
12. Yang, C. F., Stassinopoulos, A., and Goldberg, I. H. (1995) Specific binding of the biradical analogue of neocarzinostatin chromophore to bulged DNA: implications for thiol-independent cleavage, *Biochemistry 34*, 2267−2271.
13. Cheng, C.-C., Kuo, Y.-N., Chuang, K.-S., Luo, C.-F., and Wang, W. J. (1999) A new Co II complex as a bulge-specific probe for DNA, *Angew. Chem., Int. Ed. 38*, 1255−1257.
14. Yang, C. F., Jackson, P. J., Xi, Z., and Goldberg, I. H. (2002) Recognition of Bulged DNA by a Neocarzinostatin Product via an Induced Fit Mechanism, *Bioorg. Med. Chem. 10*, 1329−1335.
15. Xi, Z., Jones, G. B., Qabaja, G., Wright, J., Johnson, F., and Goldberg, I. H. (1999) Synthesis and DNA binding of spirocyclic model compounds related to the neocarzinostatin chromophore, *Org. Lett. 1*, 1375−1377.
16. Xi, Z., Hwang, G.-S., Goldberg, I. H., Harris, J. L., Pennington, W. T., Fouad, F. S., Qabaja, G., Wright, J. M., and Jones, G. B. (2002) Targeting DNA Bulged Microenvironments with Synthetic Agents; Lessons from a Natural Product, *Chem. Biol. 9*, 925−931.
17. Williams, L. D., and Goldberg, I. H. (1988) Selective strand scission by intercalating drugs at DNA bulges, *Biochemistry 27*, 3004−3011.
18. Gu, F., Xi, Z., and Goldberg, I. H. (2000) DNA damage by thiol-activated neocarzinostatin chromophore at bulged sites, *Biochemistry 39*, 4881−4891.
19. Kappen, L. S., Xi, Z., and Goldberg, I. H. (2001) Probing DNA single strands for single-base bulges with neocarzinostatin chromophore, *Biochemistry 40*, 15378−15383.
20. Vold, R. L., Waugh, J. S., Klein, M. P., and Phelps, D. E. (1967) Measurement of spin relaxation in complex systems, *J. Chem. Phys. 48*, 3831−3832.
21. Altieri, S. D., Hinton, P., and Byrd, R. A. (1995) Association of biomolecular systems via pulsed field gradient NMR self-diffusion measurements, *J. Am. Chem. Soc. 117*, 7566−7567.
22. Yang, X., Sanghvi, Y. S., and Gao, X. (1997) Conformational studies of antisense DNA by PFG NMR, *J. Biomol. NMR 10*, 383−388.
23. Ilynia, E., Roongta, V., Pan, H., Woodward, C., and Mayo, K. (1997) A pulsed-field gradient NMR study of bovine pancreatic trypsin inhibitor self-association, *Biochemistry 36*, 3383−3388.
24. Stejskal, E. O., and Tanner, J. E. (1965) Spin diffusion measurements: Spin-echoes in the presence of a time-dependent field gradient, *J. Chem. Phys. 42*, 288−292.
25. Gibbs, S. J., and Jahnson, C. S., Jr. (1991) A PFG NMR experiment for accurate diffusion and flow studies in the presence of eddy currents, *J. Magn. Reson. 93*, 395−402.
26. Plateau, P., and Guéron, M. (1982) Exchangeable proton NMR without baseline distortion, using new strong-pulse sequences, *J. Am. Chem. Soc. 104*, 7301−7311.
27. Wuthrich, K. (1986) *NMR of proteins and nucleic acids*, pp 224−246 (resonance assignment strategy in nucleic acid) and pp 117−161 (resonance assignment strategy in protein), John Wiley & Sons, New York.
28. Borgias B. A., and James, T. L. (1990) MARDIGRAS—a procedure for matrix analysis of relaxation for discerning geometry of an aqueous structure, *J. Magn. Reson. 87*, 475−487.
29. Schmitz, U., and James, T. (1995) How to generate accurate solution structues of double-helical nucleic acid fragments using nuclear magnetic resonance and restrained molecular dynamics, *Methods Enzymol. 261*, 3−44.
30. Saenger, W. (1985) *Principle of nucleic acid structure*, pp 123−124 (Watson−Crick base pair geometry) and pp 25, 233−241 (minor groove features), Springer-Verlag, New York.
31. Houser, R. V., Ravikumar, M., Chary, K. V. R., Sheth, A., Govil, G., Zu-Kun, T., and Miles, H. T. (1986) Solution structure of d-GAATTCGAATTC by 2D NMR. A new approach to determination of sugar geometries in DNA segments, *FEBS Lett. 205*, 71−76.
32. Kim, S.-G., Lin, L.-J., and Reid, B. R. (1992) Determination of nucleic acid backbone conformation by 1H NMR, *Biochemistry 31*, 3564−3574.
33. Rinkel, L., and Altona, C. (1987) Conformational analysis of the deoxyribofuranose ring in DNA by means of sums of proton−proton coupling constant: A graphical method, *J. Biomol. Struct. Dyn. 4*, 621−649.
34. Mooren, M., Pulleyblank, D. E., Wijmenga, S. S., van de Ven, F. J., and Hilbers, C. W. (1994) The solution structure of the hairpin formed by d(TCTCTC-TTT-GAGAGAGA), *Biochemistry 33*, 7315−7325.
35. Kuszewski, J., Nilges, M., and Brünger, A. T. (1992) Sampling and efficiency of metric matrix distance geometry; a novel partial metrization algorithm, *J. Biomol. NMR 2*, 33−56.
36. Nilges, M., Habazettl, J., Brünger, A. T., and Holak, T. A. (1991) Relaxation matrix refinement of the solution structure of squash trypsin inhibitor, *J. Mol. Biol. 219*, 499−510.
37. Program kindly provided by R. Lavery, CNRS Institut de Biologie Physico-Chimique, France. Lavery, R., and Sklenar, H. (1988) The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids, *J. Biomol. Struct. Dyn. 6*, 63−91.
38. Varani, G. (1995) Exceptionally stable nucleic acid hairpins, *Annu. Rev. Biophys. Biomol. Struct. 24*, 379−404.
39. Hare, D., Shapiro, L., and Patel, D. J. (1986) Extrahelical adenosine stacks into right-handed DNA: solution conformation of the d(CGCAGAGCTCGCG) duplex deduced from distance geometry analysis of nuclear Overhauser effect spectra, *Biochemistry 25*, 7456−7464.
40. Ripley, L. S. (1982) Model for the participation of quasi-palindromic DNA sequences in frameshift mutation, *Proc. Natl. Acad. Sci. U.S.A. 79*, 4128−4132.

41. Streisinger, G., Okada, Y., Emrich, J., Newton, J., Tsugita, A., Terzaghi, I., and Inouye, M. (1966) Frameshift mutations and the genetic code, *Cold Spring Harbor Symp. Quantum Biol. 31*, 77−84.
42. Kunkel, T. A. (1993) Nucleotide repeats. Slippery DNA and diseases, *Nature 365*, 207−209.
43. Malkov, V. A., Biswas, I., Camerini-Otero, R. D., and Hsieh, P. (1997) Photocross-linking of the NH2-terminal region of Taq MutS protein to the major groove of a heteroduplex DNA, *J. Biol. Chem. 272*, 23811−23817.
44. Wang, Y.-H., Bortner, C. D., and Griffith, J. (1993) RecA binding to bulge- and mismatch-containing DNAs. Certain single base mismatches provide strong signals for RecA binding equal to multiple base bulges, *J. Biol. Chem. 268*, 17571−17577.
45. Ashley, C. T., Jr., and Warren, S. T. (1995) Trinucleotide repeat expansion and human disease, *Annu. Rev. Genet. 29*, 703−728.
46. Liley, D. M. J. (1995) Kinking of DNA and RNA by base bulges, *Proc. Natl. Acad. Sci. U.S.A. 92*, 334−337.
47. Rice, J. A., and Crothers, D. M. (1989) DNA bending by the bulge defect, *Biochemistry 28*, 4512−4516.
48. Nelson, J. W., and Tinoco, I., Jr. (1985) Ethidium ion binds more strongly to a DNA double helix with a bulged cytosine than to a regular double helix, *Biochemistry 24*, 6416−6421.
49. White, S. A., and Draper, D. E. (1987) Single base bulges in small RNA hairpins enhance ethidium binding and promote an allosteric transition, *Nucleic Acids Res. 15*, 4049−4064.
50. Yao, S., and Wilson, W. D. (1992) A molecular mechanics investigation of RNA complexes. I. Ethidium intercalation in an HIV-1 TAR RNA sequence with an unpaired adenosine, *J. Biomol. Struct. Dyn. 10*, 367−387.
51. Williams, L. D., Thivierge, J., and Goldberg, I. H. (1988) Specific binding of *o*-phenanthroline at a DNA structure lesion, *Nucleic Acids Res. 16*, 11607−11615.
52. Caceres-Cortes, J., and Wang, A. H.-J. (1996) Binding of the antitumor drug nogalamycin to bulged DNA structures, *Biochemistry 35*, 616−625.
53. Nakatani, K., Sando, S., and Saito, I. (2000) Recognition of single guanine bulge by 2-acylamini-1,8-naphthyridine, *J. Am. Chem. Soc.* 122, 2172−2177.
54. Nakatani, K., Okmoto, A., and Saito, I. (1999) Specific alkylation of guanine opposite to a single nucleotide bulge: A chemical probe for the bulged structure of DNA, *Angew. Chem., Int. Ed. 38*, 3378−3381.
55. Shih, H.-C., Tang, N., Burrows, C. J., and Rokita, S. E. (1998) Nickel-based probes of nucleic acid structure bind guanine N7 but do not perturb a dynamic equilibrium of extrahelical guanine residues, *J. Am. Chem. Soc. 120*, 3284−3288.

BI0206210