

# Solution Structure of a Designed Spirocyclic Helical Ligand Binding at a Two-Base Bulge Site in DNA<sup>†</sup>

Na Zhang,<sup>‡</sup> Yiqing Lin,<sup>§</sup> Ziwei Xiao,<sup>‡</sup> Graham B. Jones,<sup>§</sup> and Irving H. Goldberg<sup>\*,‡</sup>

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, and Bioorganic and Medicinal Chemistry Laboratories, Department of Chemistry and Chemical Biology, Northeastern University, Boston, Massachusetts 02115

Received December 19, 2006; Revised Manuscript Received February 23, 2007

**ABSTRACT:** The solution structure of the complex formed between an oligodeoxynucleotide containing a two-base bulge (5'-CCATCGTCTACCTTTGGTAGGATGG) and SCA- $\alpha$ 2, a designed spirocyclic helical molecule, has been elucidated. SCA- $\alpha$ 2, a close mimic of the metabolite, NCSi-gb, of the DNA bulge-specific enediyne antibiotic neocarzinostatin, differs in possessing a more stable spirocyclic ring system and in lacking certain bulky groupings that compromise bulged DNA binding. This study provides a detailed comparison of the binding modes of the two complexes and provides new insights into the importance of shape and space, as opposed to simple nucleotide sequence, in complex formation at the bulge site. The two rigidly held aromatic rings of SCA- $\alpha$ 2 form a right-handed helical molecular wedge that specifically penetrates the bulge-binding pocket and immobilizes the two bulge residues (GT), which point toward the minor groove, rather than the major groove as in the NCSi-gb·bulged DNA complex. The ligand aromatic ring systems stack on the DNA bulge-flanking base pairs that define the long sides of the triangular prism binding pocket. Like NCSi-gb, SCA- $\alpha$ 2 possesses the natural *N*-methylfuranose moiety,  $\alpha$ -linked to the benzindanol (BI) moiety. The amino sugar anchors in the major groove of the DNA and points toward the 3'-bulge-flanking base pair. Lacking the bulky cyclocarbonate of NCSi-gb, the SCA- $\alpha$ 2·bulged DNA complex has a much less twisted and buckled 3'-bulge-flanking base pair (dG20·dC8), and the G20 residue stacks directly above the BI ring platform. Also, the absence of the methyl group and the free rotation of the methoxy group on the dihydronaphthone (NA) moiety of SCA- $\alpha$ 2 allow better stacking geometry of the NA ring above the 5'-bulge-flanking dG21·dC5 base pair. These and other considerations help to explain why NCSi-gb binds very poorly to bulged RNA and are consistent with the recent observation of good binding with SCA- $\alpha$ 2. Thus, although the two complexes resemble each other closely, they differ in important local environmental details. SCA- $\alpha$ 2 has a better hand-in-glove fit at the bulge site, making it an ideal platform for the placement of moieties that can react covalently with the DNA and for generating congeners specific for bulges in RNA.

Bulged structures in nucleic acids are of general biological significance and are potential targets for therapeutic drugs (1–4). They may arise as a consequence of recombination between DNA segments that are imperfectly homologous or of slipped mispairing during the replication of DNA and may be involved in deletion and frame shift mutagenesis. DNA bulges may also play important roles in protein recognition. HMMG-D, a high-mobility group protein, was shown to bind preferentially to DNA bulge sites (5), and the tumor suppression protein p53 also prefers binding to bulges over other base mismatches in DNA (6). Further, bulges in RNA act as intermediates in RNA splicing and as binding motifs for regulatory proteins involved in viral replication (7).

Accordingly, it would be of interest to develop small molecules that selectively interact with bulges in nucleic

acids. A way to proceed toward this goal was provided by the finding that DNA bulges are specific sites for oxidative cleavage by the enediyne antitumor antibiotic neocarzinostatin upon its activation to a radical species in a general base-catalyzed reaction (8–12). Further, a metabolite (NCSi-gb,<sup>1</sup> Figure 1) of the drug biradical species formed in this reaction binds specifically at a two-base bulge in DNA (13, 14). However, because NCSi-gb possesses a relatively labile spirocyclic ring system and other groups, such as a bulky cyclic carbonate moiety, its use as a general probe of bulged sites in nucleic acids is limited (15–17). Such considerations have led us to design and prepare small molecule analogues of NCSi-gb with more favorable binding properties (17, 18).

<sup>†</sup> This research was supported by U.S. Public Health Service Grant GM 53793 to I.H.G. from the National Institutes of Health.

\* To whom correspondence should be addressed. Telephone: (617) 432-1787. Fax: (617) 432-0471. E-mail: irving\_goldberg@hms.harvard.edu.

<sup>‡</sup> Harvard Medical School.

<sup>§</sup> Northeastern University.

<sup>1</sup> Abbreviations: 1D, one-dimensional; 2D, two-dimensional; BI, benzindanol; DQF-COSY, double-quantum-filtered correlation spectroscopy; NCS, neocarzinostatin; NCS-chrom, native form of the NCS chromophore; NCSi-gb, postactivated NCS-chrom formed under base-catalyzed conditions; NMF, 2-*N*-methylfucosamine; NMR, nuclear magnetic resonance; NA, dihydronaphthone; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect exchange spectroscopy; rmsd, root-mean-square deviation; SCA, spirocyclic alkene; THI, tetrahydroindacene; TOCSY, total correlation spectroscopy.

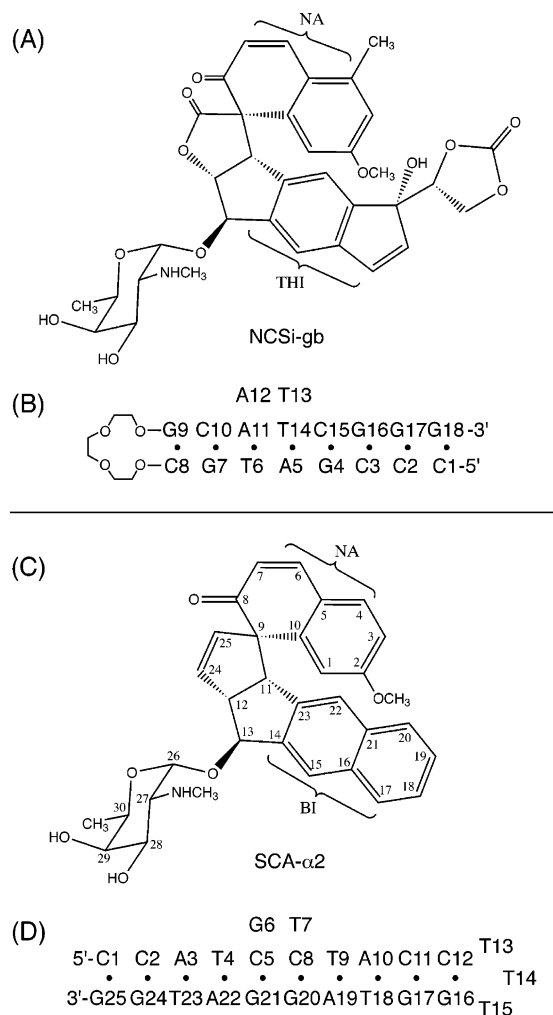


FIGURE 1: Chemical structures of (A) natural metabolite NCSi-gb, (B) the sequence of its 18-mer bulged DNA substrate with an artificial chemical-linker loop (15, 16), (C) synthetic analogue SCA- $\alpha$ 2, and (D) the sequence of its 25-mer bulged DNA substrate. The aromatic ring moieties of both ligands are marked.

These analogues share the same key structural features of NCSi-gb: a wedge-shaped aglycon moiety, consisting of two aromatic ring systems held rigidly by a stable spirocyclic ring having a right-handed helical twist of 35°, and a pendent amino sugar moiety (17, 18). Initial efforts led to the preparation of simple analogues in which an aminoglucose moiety was readily  $\beta$ -linked to the five-membered spirocyclic ring (17) or to the BI ring (18). These analogues bound specifically to two-base bulges in DNA; however, the binding affinity was more than 10-fold weaker than that of the natural product (17, 18), and there was no improvement in binding to bulged RNA over the extremely poor binding of the natural product (19, 20).

The solution structures of the complexes formed between oligodeoxynucleotides containing a two-base bulge and the natural as well as analogue ligands in which the amino sugar is  $\beta$ -linked to the spirocyclic ring (17) have been elucidated using high-resolution NMR spectroscopy and restrained molecular dynamic simulation (15, 16, 21, 22). The structures of the complexes reveal, in important details, the differences in binding modes, which provide insight into the relationship between structure and binding affinity. Despite the virtual similarity of the aglycon skeleton of these two ligands, the binding of the analogues is through the minor groove and

only one aromatic planar ring system intercalates fully into the bulge site, whereas both aromatic ring moieties of NCSi-gb insert via the major groove into the triangular bulge-binding pocket and stack with the neighboring base pairs. The less efficient stacking pattern of the analogue-bulged DNA complex likely accounts for the weaker affinity of the analogue with the bulged DNA substrate. The dissimilarity in binding modes can be attributed to the differences in the linkage positions and the stereochemical configurations of the pendent aminoglycoside moieties of the ligands, the major differences being that the analogues have a  $\beta$ -aminoglucosyl linkage to the spirocyclic ring whereas NCSi-gb has an  $\alpha$ -aminofucosyl moiety linked to the THI ring.

To improve the binding property of the ligand, we have designed and synthesized SCA- $\alpha$ 2 (Figure 1), which has the natural aminofucosyl moiety  $\alpha$ -linked to the BI ring system (counterpart of the THI system of NCSi-gb) as in NCSi-gb (23). Recently, we have successfully used SCA- $\alpha$ 2 as a probe of DNA bulged structures (23, 24), as a promoter of nucleotide expansion via DNA strand slippage (23), and as a selective inhibitor of nucleotide extension by DNA polymerase I at a two-base bulge site (24). Its binding affinity ( $K_d = 0.13 \mu\text{M}$ ) (23) for the best two-base bulged DNA substrate is close to that of NCSi-gb ( $K_d = 0.033 \mu\text{M}$ ) (17).

Due to its tight binding and the high-quality NMR spectra of the complex, we are able to elucidate the solution structure of the SCA- $\alpha$ 2•two-base bulged DNA complex and provide a detailed picture of the bulge-binding pocket. This analysis reveals that the synthetic analogue has a better hand-in-glove fit than the natural product, making the former an ideal platform for the placement of moieties that can react covalently with the DNA. Of special interest, from the solution structures, we are able to rationalize our recent finding that SCA- $\alpha$ 2 is much better than NCSi-gb in binding to RNA bulges (25).

## MATERIALS AND METHODS

**Sample Preparation.** The synthesis and purification of SCA- $\alpha$ 2 were as reported previously (23). For bulged DNA, the oligonucleotide sequence, a 25-mer DNA, containing a two-base bulge and a T-T-T loop (Figure 1D), was purchased from Integrated DNA Technologies and further purified as described previously (26). The DNA sample (~7 mL) was first dialyzed against water, then against 150 mM NaCl, 10 mM sodium phosphate buffer, and finally against 15 mM NaCl, 0.2 mM sodium phosphate buffer and lyophilized. The NMR sample solution of DNA contains 100 mM NaCl and 10 mM sodium phosphate (pH 6.8). The sample was heated at 90 °C for several minutes and then slowly annealed to room temperature. For the SCA- $\alpha$ 2•bulged DNA complex, the 1:1 drug•bulged DNA complex was formed by progressively adding microliter aliquots of an ~50 mM stock solution of SCA- $\alpha$ 2 in  $d_4$ -MeOH to the bulged DNA duplex and monitoring the titration by  $^1\text{H}$  NMR at 0 °C. Formation of the drug•bulged DNA complex was assessed by the disappearance of the resonance lines of the free bulged DNA when SCA- $\alpha$ 2 was added to the bulged DNA duplex. The complex solution was lyophilized several times and dissolved in 99.96%  $\text{D}_2\text{O}$ . The NMR sample contained a 1:1 SCA- $\alpha$ 2•bulged DNA complex at a concentration of ~4.5 mM (pH 6.8).

**NMR Methods.** Detailed NMR studies of the SCA- $\alpha$ 2•bulged DNA complex were conducted as described previously (15, 16, 21, 27). NMR spectra of the exchangeable protons were collected in aqueous buffer [100 mM NaCl and 10 mM phosphate (pH 6.8)] at 0 and 10 °C on Varian INOVA 600 MHz NMR spectrometers. For the nonexchangeable protons of the same complex, NMR spectra were collected in D<sub>2</sub>O at 0, 10, and 30 °C. NOESY, COSY, and TOCSY data sets were processed and analyzed using FELIX (Accelrys Inc.).

**Molecular Dynamics Computations.** A starting model of the two-base bulged DNA duplex with a T-T-T loop was generated manually in standard B-helical conformation using InsightII (Accelrys Inc.). The glycosidic torsion angles for all residues adopt anti conformations, which are consistent with NMR experimental observation. The resulting energy-minimized structure shows the two bulged residues in an intrahelical conformation. The conformation of SCA- $\alpha$ 2 was first built manually and then energy-minimized using InsightII without any NMR restraints. The drug SCA- $\alpha$ 2 was then placed manually more than 8 Å from the bulge packet of the bulged DNA duplex with its orientation random relative to the minor and major grooves of the DNA duplex. In total, there are 10 complexes with the drug being positioned differently outside the DNA bulge as starting models for the next structural calculation using X-PLOR (28).

The protocol outlined below has been used, as described in previous publications (16, 27). The NMR distance-restrained molecular dynamics (MD) simulations were carried out starting from the 10 initial starting structures of the SCA- $\alpha$ 2•bulged DNA complex mentioned above. X-PLOR-based restrained MD calculations for each starting structure were carried out using the simulated annealing protocol and the CHARM force field. Hydrogen bonding distance restraints were imposed to align the experimentally identified Watson–Crick pairs ( $\pm 0.10$  Å). Only the nonpaired G6 and T7 bulged residues and T13–T14–T15 loop residues were excluded on the basis of the analysis of NMR spectra. Each cycle of restrained MD simulations was initially carried out at 300 K with a force constant of 1 kcal mol<sup>-1</sup> Å<sup>-2</sup> on all experimentally obtained distance restraints. The structure was subjected to 2000 cycles of energy minimization and was slowly heated to 1000 K in 7 ps (1 ps per 100 K increase). The force constants on the experimentally obtained distance restraints were slowly scaled up to 32 (nonexchangeable protons), 16 (exchangeable protons), and 64 kcal mol<sup>-1</sup> Å<sup>-2</sup> (hydrogen bonds), over a period of 18 ps. The system was allowed to evolve for an additional 20 ps at 1000 K and next cooled gradually to 300 K over 14 ps (0.5 fs time step) with retention of the full scale of distance restraints and subsequently equilibrated for 12 ps at 300 K. The coordinates were averaged over the last 5 ps and the resulting coordinates subjected to 20 cycles (100 steps per cycle) of conjugate gradient energy minimization. A total of 10 distance-refined structures of the SCA- $\alpha$ 2•bulged DNA complex were generated, and eight of them were accepted. The acceptance criteria are based on the best structures with the least NOE and dihedral angle violations, acceptable covalent geometry, and favorable van der Waals energy.

**Distance Restraints.** The distances between nonexchangeable interprotons were obtained from the buildup curves of cross-peak intensities in NOESY spectra at five different

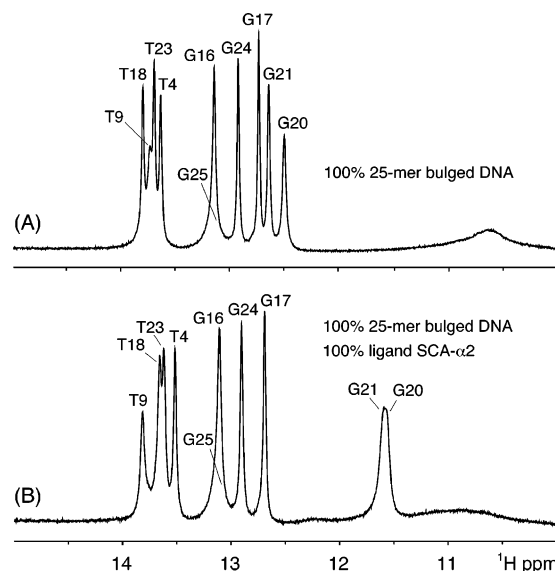


FIGURE 2: Expanded imino proton <sup>1</sup>H spectra (10.0–15.0 ppm) recorded in 100 mM NaCl, 10 mM phosphate, and H<sub>2</sub>O at pH 6.8 and 10 °C: (A) 25-mer bulged DNA alone and (B) SCA- $\alpha$ 2•bulged DNA complex formed by adding 100% SCA- $\alpha$ 2 into 100% 25-mer bulged DNA. The assignments of imino protons are given.

mixing times (50, 100, 150, 200, and 250 ms) in D<sub>2</sub>O at 10 °C and given bounds between  $\pm 20$  and  $\pm 30\%$  with distances referenced relative to the cytosine H5–H6 distance of 2.47 Å. Exchangeable proton restraints were based on NOESY data sets at two mixing times (60 and 200 ms) in H<sub>2</sub>O at 10 °C and given bounds of  $\pm 30\%$ . Cross-peaks involving exchangeable protons were classified as strong (medium to strong intensity at 60 ms), medium (weak intensity at 60 ms), and weak (observed only at a mixing time of 200 ms), and proton pairs were then restrained to distances of 3.0  $\pm$  0.9, 4.0  $\pm$  1.2, and 6.0  $\pm$  1.8 Å, respectively.

**Protein Data Bank Entry.** The structures have been deposited in the RCSB Protein Data Bank as entry 2OY2.

## RESULTS

**Exchangeable Imino Proton 1D Spectra.** The expanded one-dimensional <sup>1</sup>H NMR spectra (10–15 ppm) for the imino proton region of the bulged DNA sequence both in the absence and in the presence of SCA- $\alpha$ 2 are shown in Figure 2. The spectra were recorded in H<sub>2</sub>O buffer solution at 10 °C. The imino proton peaks of bulged DNA in both cases are well-resolved, and their assignments are labeled. In Figure 2A, in total 10 sharp imino proton peaks are characteristic of Watson–Crick base pairings and indicate the formation of duplex stems for free bulged DNA. In Figure 2B, upon addition of an equal stoichiometric amount of SCA- $\alpha$ 2, a new set of imino proton peaks is observed, which have slight to significant chemical shift changes compared to their counterparts in free bulged DNA. This high-quality spectrum demonstrates that SCA- $\alpha$ 2 associates with free bulged DNA to form a new ligand•bulged DNA complex. Ten observed sharp peaks, corresponding to the DNA imino proton components within the SCA- $\alpha$ 2•bulged DNA complex, indicate that only a single bound form exists in solution. The most noticeable chemical shift changes occur at guanine residues within the bulge-flanking dG•dC base pairs, namely, G21 and G20. Their chemical shifts are detected at 11.59 and 11.55 ppm, respectively (Figure 2B), which are shifted

~1.0 ppm upfield compared to their counterparts within the free bulged DNA form. On the other hand, there are very small chemical shift changes for those residues away from the bulge region for the two different forms before and after addition of ligand (<0.2 ppm changes). This phenomenon strongly indicates that the binding of SCA- $\alpha$ 2 to bulged DNA has virtually no effect on the DNA duplex stem, but rather, SCA- $\alpha$ 2 specifically interacts at the bulge region.

**Titration Experiments Monitored by NMR.** The titration of SCA- $\alpha$ 2 into the free bulged DNA sequence is monitored by  $^1\text{H}$  NMR at 0 °C and is shown in Figure S1 of the Supporting Information. Increasing amounts of SCA- $\alpha$ 2 are gradually added to the solution of the 25-mer bulged DNA sequence. Panels A–C of Figure S1 show the proton NMR spectra of the bulged DNA in the absence and presence of different concentrations of SCA- $\alpha$ 2. As the intensity of the peaks representing the free bulged DNA form decreases in the presence of 20 and 80% SCA- $\alpha$ 2, a new set of peaks emerges (Figure S1B,C). In the presence of 100% SCA- $\alpha$ 2, the peaks representing the free bulged DNA form totally fade away, while the intensity of peaks representing the new form of the complex reaches a maximum level (Figure S1D). The upfield shifts of the two imino protons from bulge-flanking dG•dC base pairs suggest the formation of a new form of SCA- $\alpha$ 2•bulged DNA complex. Likely, the stacking of aromatic ring moieties of SCA- $\alpha$ 2 with the two bulge-flanking base pairs of DNA results in the dramatic upfield shifts for G21 and G20 imino protons due to the ring current effect.

**NOE Analysis of the Imino Proton Region.** Expanded regions of the NOESY contour plot (mixing time of 200 ms) for the SCA- $\alpha$ 2•bulged DNA complex in  $\text{H}_2\text{O}$  buffer at 10 °C are shown in Figure 3. The imino, amino, and non-exchangeable protons have been assigned using standard methodologies described in previous publications (16, 21, 27). Imino–imino NOE connectivities are observed between all adjacent base pairs on either side of the bulge (peaks a–g in Figure 3A). This observation establishes the formation of two duplex stems, C1–C5•G21–G25 and C8–C12•G16–G20. Four thymine imino protons peaks are observed at 13.40–13.80 ppm, which are consistent with regular chemical shift values of Watson–Crick dA•dT pairs. In particular, the strong NOE cross-peaks from the deoxythymidine imino proton to its counterpart deoxyadenosine H2 proton across the dA•dT pairs (a–d in Figure 3B) prove the formation of Watson–Crick dA•dT pairs. Similarly, six Watson–Crick dG•dC pairs are identified from the observation of the characteristic NOE patterns (cross-peaks e–o in Figure 3B) from the deoxyguanosine imino proton to its counterpart deoxycytidine amino and H5 protons across the dG•dC pairs. The terminal dG25•dC1 pair is not as stable as others at 10 °C; however, we can still observe its characteristic NOEs at the lower temperature of 0 °C (data not shown). The chemical shifts of guanine imino protons in the four duplex stem dG•dC pairs maintain the regular values for Watson–Crick dG•dC pairs, which are observed at 12.70–13.10 ppm. By contrast, the chemical shifts of guanine imino protons within two bulge-flanking dG21•dC5 and dG20•dC8 pairs are shifted very much upfield to 11.59 and 11.55 ppm, respectively. Meanwhile, unlike that of other Watson–Crick base pairs away from the bulge pocket, a number of intermolecular NOEs (peaks 1–10 in Figure 3B) between SCA- $\alpha$ 2 and G21

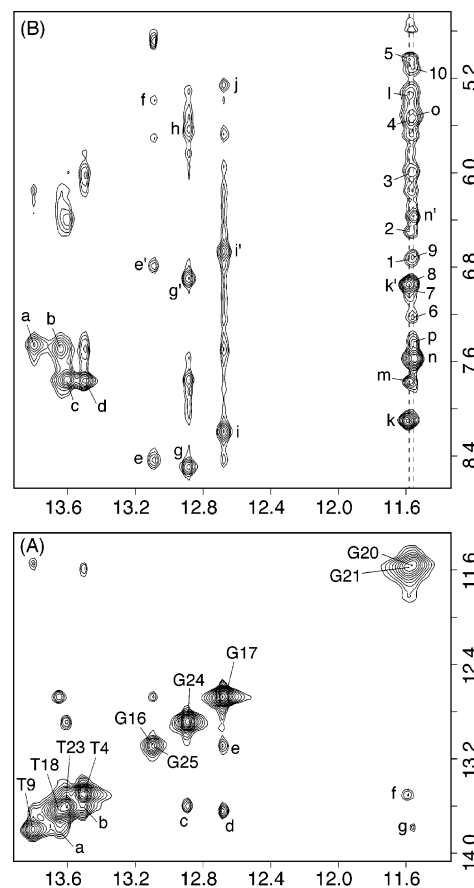


FIGURE 3: Expanded NOESY (200 ms mixing time) contour plots for the SCA- $\alpha$ 2•bulged DNA complex. Spectra were recorded in 100 mM NaCl, 10 mM phosphate, and  $\text{H}_2\text{O}$  at pH 6.8 and 10 °C. (A) NOEs between imino protons (11.3–13.9 ppm), with the cross-peaks labeled a–g, which are assigned as follows: a, T9 (NH3)–T18 (NH3); b, T23 (NH3)–T4 (NH3); c, T23 (NH3)–G24 (NH1); d, T18 (NH3)–G17 (NH1); e, G16 (NH1)–G17 (NH1); f, T4 (NH3)–G21 (NH1); and g, T9 (NH3)–G20 (NH1). The expected cross-peak between G24 (NH1) and G25 (NH1) is observed at the lower temperature. (B) NOEs between imino protons (11.3–13.9 ppm) and amino as well as nonexchangeable protons (4.8–8.6 ppm). The NOE cross-peaks between DNA protons, which characterize the formation of Watson–Crick base pairs, are labeled as a–p, while NOE cross-peaks between SCA- $\alpha$ 2 and DNA imino protons are labeled as 1–10: a, T9 (NH3)–A19 (H2); b, T18 (NH3)–A10 (H2); c, T23 (NH3)–A3 (H2); d, T4 (NH3)–A22 (H2); e and e', G16 (NH1)–C12 (NH<sub>2</sub>); f, G16 (NH1)–C12 (H5); g and g', G24 (NH1)–C2 (NH<sub>2</sub>); h, G24 (NH1)–C2 (H5); i and i', G17 (NH1)–C11 (NH<sub>2</sub>); j, G17 (NH1)–C11 (H5); k and k', G21 (NH1)–C5 (NH<sub>2</sub>); l, G21 (NH1)–C5 (H5); m, G21 (NH1)–A22 (H2); n and n', G20 (NH1)–C8 (NH<sub>2</sub>); o, G20 (NH1)–C8 (H5); and p, G20 (NH1)–A19 (H2); and 1, G21 (NH1)–SCA (H4); 2, G21 (NH1)–SCA (H3); 3, G21 (NH1)–SCA (H24); 4, G21 (NH1)–SCA (H1); 5, G21 (NH1)–SCA (H25); 6, G20 (NH1)–SCA (H17); 7, G20 (NH1)–SCA (H19); 8, G20 (NH1)–SCA (H18); 9, G20 (NH1)–SCA (H20); and 10, G20 (NH1)–SCA (H22).

and G20 imino protons are detected at this upfield-shifted region. Therefore, we conclude that SCA- $\alpha$ 2 binds specifically at the bulge pocket of DNA. Further intermolecular NOE study reveals that the G21 imino proton is close to the H1, H3, and H4 protons of the two-ring moiety, NA, of SCA- $\alpha$ 2 (peaks 1, 2, and 4 in Figure 3B); meanwhile, the G21 imino proton is also next to the H24 and H25 protons of SCA- $\alpha$ 2 (peaks 3 and 5 in Figure 3B). On the other hand, the G20 imino proton is adjacent to the H17–H20 and H22

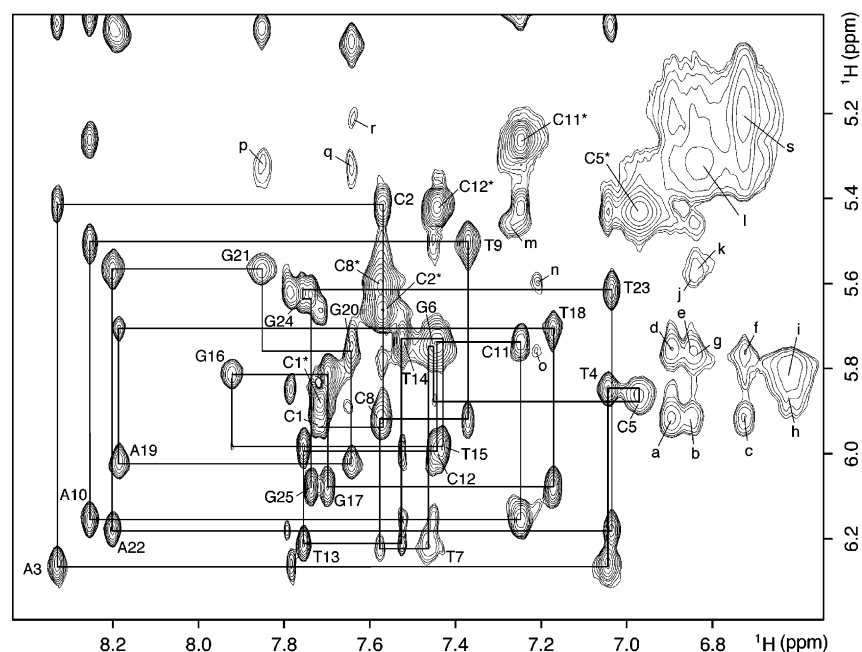


FIGURE 4: Expanded NOESY (200 ms mixing time) contour plots for the SCA- $\alpha$ 2•bulged DNA complex. Spectra were recorded in 100 mM NaCl, 10 mM phosphate, and D<sub>2</sub>O at pH 6.8 and 30 °C. The assignments and H8/H6–H1' NOE sequential walk (8.5–6.5 ppm region as base protons and 6.4–4.9 ppm region as sugar H1' protons) of the bulged DNA duplex are connected by solid lines. Asterisks denote cytosine H6–H5 NOE cross-peaks. The intermolecular NOEs between SCA- $\alpha$ 2 and DNA protons, as well as the intramolecular NOEs between SCA- $\alpha$ 2 protons, are labeled as cross-peaks a–s: a, C8 (H1')–SCA (H19); b, C8 (H1')–SCA (H18); c, C8 (H1')–SCA (H20); d, G6 (H1')–SCA (H19); e, G6 (H1')–SCA (H18); f, G20 (H1')–SCA (H20); g, G20 (H1')–SCA (H6); h, C5 (H1')–SCA (H3); i, SCA (H3)–SCA (H1); j, C8 (H5)–SCA (H18); k, G21 (H1')–SCA (H4/H6); l, SCA (H6)–SCA (H7); m, SCA (H15)–SCA (H26); n, C8 (H5)–SCA (H17); o, G6 (H1')–SCA (H17); p, G21 (H8)–SCA (H7); q, G20 (H8)–SCA (H7); r, G20 (H8)–SCA (H22); and s, SCA (H20)–SCA (H22).

protons of the three-ring moiety, BI, of SCA- $\alpha$ 2 (peaks 5–10 in Figure 3B).

#### Intermolecular NOE Analysis of the Amino Proton Region.

Unlike that of the guanine imino proton, the geometrical position of the cytosine amino proton within the Watson–Crick dG•dC pair is on the extreme edge facing toward the major groove of the DNA duplex. As expected, more binding mode information is collected on the basis of the intermolecular NOEs at 10 °C (in Figure S2 of the Supporting Information) between characteristic cytosine amino protons of the bulge-flanking dG•dC pairs and ligand protons. Cross-peaks b, e, f, and j define the closeness between the C5 amino protons and H24 and H25 protons of the ligand, consequently suggesting that the H24 and H25 protons of the ligand are adjacent to the major groove of DNA. Likely, the rigid shape of the spirocyclic junction of SCA- $\alpha$ 2 restricts the orientation of the NMF amino sugar moiety, which is a neighbor of the H24 and H25 protons, also in the major groove.

Cross-peak c shows the closeness between the C5 amino protons and the H1 proton of the NA moiety of the ligand. On the other hand, cross-peaks k and m imply the closeness between the C8 amino protons and the H15 and H17 protons on the BI moiety of the ligand. These observations are consistent with the intermolecular NOEs between guanine G21 and G20 imino protons and aromatic protons of the ligand. As a result, we conclude that the dG21•dC5 pair stacks with the two-ring moiety of the ligand, and the dG20•dC8 pair stacks with the three-ring moiety of the ligand.

**NOE Sequential Assignments and Analysis of Nonexchangeable Base and Sugar Protons.** The expanded NOESY (200 ms mixing time) contour plot establishing sequential connectivities between the base protons (7.0–8.4

ppm) and the sugar H1' and cytosine H5 protons (5.2–6.4 ppm) of the SCA- $\alpha$ 2•bulged DNA complex is given in Figure 4. The spectrum was collected in 100 mM NaCl, 10 mM phosphate, and D<sub>2</sub>O at pH 6.8 and 30 °C. The base–sugar H1' proton connectivities are traced as solid lines, from 5'-C1 to 3'-G25, along the bulge residue G6–T7 step and the T13–T14–T15 loop. The overall pattern of NOE cross-peaks and relative intensities is consistent with a right-handed double helix and a glycoside torsion angle in the anti range for the complex. The interruptions in the sequential connectivities are detected at the G20–G21 step. Also, very weak inter-residue NOE cross-peaks are observed in the C5–G6–T7–C8 bulge-containing segment. These observations imply the insertion of stacking aromatic ring moieties of the ligand with their counterpart bulge-flanking dG21•dC5 and dG20•dC8 base pairs. Due to this intercalation, a number of chemical shift changes, comparing the bound and unbound (Figure S3 of the Supporting Information) DNA forms, are observed among those residues composing the bulge-binding pocket. Especially the base H6 (7.44 ppm), the methyl proton (1.59 ppm, out of plot range here), and the H1' sugar protons (6.20 ppm) of T7 exhibit the greatest downfield shifts (comparison of Figure 4 and Figure S3 at 30 °C; Table S1 of the Supporting Information at 10 °C). This is indicative of a loss of stacking interactions and thus a strong indication of the displacement of T7 from the helical duplex.

**Intermolecular NOE Analysis of the Nonexchangeable Base and Sugar Proton Region.** Forty-six intermolecular NOEs between ligand protons and DNA nonexchangeable protons are detected. The representative NOEs within the base and sugar H1' proton region are labeled as cross-peaks a–s (Figure 4). Analysis of these NOEs provides insights

Table 1: Intermolecular NOEs Observed between SCA- $\alpha$ 2 and DNA Protons in the SCA- $\alpha$ 2•Bulged DNA Complex<sup>a</sup>

	SCA- $\alpha$ 2	bulged DNA contacts	
NA ring	H1	G21 NH(m), C5 NH <sub>2</sub> (m), C5 H5(m)	
	H3	G21 NH(m), C5 H1'(w)	
	H4	G21 NH(m), G21 H1'(m)	
	H6	G20 H1'(w), G20 H2'(m), G20 H2''(m), G21 NH(w), G21 H1'(m), G21 H2'(m), G21 H8(m)	
	H7	G20 H8(m), G20 H1'(w), G20 H2'(s), G20 H2''(m), G21 H8(m), G21 H2'(m)	
	BI ring	H13	G21 NH(w), C5 NH <sub>2</sub> (w)
		H15	G20 NH(m), C8 NH <sub>2</sub> (m,m), C8 H5(m)
H17		G6 H8(m), C8 NH <sub>2</sub> (m,m), C8 H5(m), C8 H6(m), G20 NH(m)	
H18		G6 H1'(m), G6 H2'(m), G6 H2''(m), G6 H8(m), T7 H2'(m), T7 H2''(m), C8 H1'(m), C8 H4'(w), C8 H5(w), C8 H6(m), G20 NH(m)	
H19		G6 H1'(m), G6 H8(m), C8 H1'(m), G20 NH(m)	
H20		C8 H1'(m), G20 NH(m), G20 H1'(w)	
H22		G20 NH(m), G20 H1'(w)	
spirocyclic junction		H11	G20 NH(w), G20 H8(m), C8 NH <sub>2</sub> (m)
		H12	G20 NH(w), C8 NH <sub>2</sub> (m)
		H24	C5 NH <sub>2</sub> (m,m), C5 H5(w), G21 NH(w)
	H25	C5 NH <sub>2</sub> (s,m), G21 NH(m)	
NMF	H28	C8 NH <sub>2</sub> (m,m), C8 H5(m)	
	H29	C8 NH <sub>2</sub> (m,m), C8 H5(m), G6 H2'(m), G6 H2''(m), G6 H3'(m)	
	H30	G6 H2'(m), G6 H2''(m), G6 H3'(m)	
	CH <sub>3</sub> -30	G6 H3'(m)	

<sup>a</sup> The letters s, m, and w stand for strong, medium, and weak NOEs, respectively.

into the alignment of SCA- $\alpha$ 2 inside the complex. NOEs h and k show that both H1' sugar protons, pointing to the minor groove, of the 5'-bulge-flanking dG21•dC5 pair stack with the H3 and H4 protons on one edge of the two-ring moiety of the ligand; NOE p narrows the H7 proton, on the opposite edge of the two-ring moiety of the ligand, toward the base H8 proton of dG21, which resides in the major groove. This analysis readily defines the cross-stacking pattern of the two-ring moiety onto the dG21•dC5 pair plane. Similarly, NOEs a–c and f show that both H1' sugar protons of the 3'-bulge-flanking dG20•dC8 pair stack near the minor groove with the H18–H20 protons on one edge of the three-ring moiety of the ligand. The large size of the three-ring moiety of the ligand makes it impossible to parallel-stack with the dG20•dC8 pair plane. The NMF moiety is opposite the edge of the H18–H20 protons, which define the minor groove. Therefore, the cross-stacking pattern likely leads the NMF carbohydrate moiety toward the major groove. Direct observation of the intermolecular NOEs between the NMF and the DNA (listed in Table 1) in the major groove (out of plotting range, not shown here) also supports this reasoning.

NOEs d, e, and o illustrate the closeness of the G6 residue with the H17–H19 protons. These observations place G6 toward the minor groove, but not far from one edge of the three-ring moiety. This explains the relatively smaller chemical shift changes (comparison of Figure 4 and Figure S3 at 30 °C; Table S1 at 10 °C) for bulge residue G6 between the ligand-free and bound forms, compared to those of the other bulge residue (T7).

**Sugar Conformation.** The scalar coupling constant analysis from DQF-COSY spectra indicates that most residues clearly adopt the pure 2'-endo sugar pucker of B-DNA with a large <sup>3</sup>J<sub>H1'–H2'</sub> value (7–10 Hz) and a small <sup>3</sup>J<sub>H3'–H4'</sub> value (<3 Hz or absent). Several residues exhibit more like, but not pure, patterns of C2'-endo sugar pucker, such as C5, G20 of the bulge pocket, and T13 of the loop (more flexible motion). However, there is no spectral evidence of pure C3'-endo sugar conformation. We observed a strong DQF-COSY cross-peak (with a large *J* value of 9.7 Hz) between H27 and H28 protons of SCA- $\alpha$ 2 drug sugar moiety NMF,

whereas their NOE cross-peak is weak. These data indicate a trans H<sub>ax</sub>–H<sub>ax</sub> cross-peak for H27 and H28 of drug sugar moiety NMF. We also observed medium/strong NOE cross-peaks, but weak COSY cross-peaks (with a small *J* value of <3 Hz), between other drug sugar protons (H26 and H27, H28 and H29, H29 and H30). These data indicate the H<sub>ax</sub>–H<sub>eq</sub> patterns. Therefore, we define the conformation of the drug sugar moiety NMF as being close to a chair conformation, which was also adopted for the NCSi-gb•18-mer complex (15, 16).

**Structure Calculation.** In total, 77 DNA–drug intermolecular NOEs, including both exchangeable and non-exchangeable protons, were detected and are listed in Table 1. These extensive intermolecular NOE constraints efficiently refine the bulge-binding pocket of the complex. The NMR results are incorporated as initial distance and dihedral angle restraints in the structure calculation. The final structures are selected on the basis of criteria such as low overall energies and no significant NOE (>0.2 Å) and dihedral (>5 Å) violations. Eight superimposed refinement structures of the SCA- $\alpha$ 2•bulged DNA complex are shown in Figure 5. The final refined structures converged to a pairwise rms difference of 0.93 ± 0.38 and very low constraint violation energies (Table 2). These converged structures show good bond and angle geometry and, in general, satisfy NMR-derived distances and dihedral angles.

Stereoviews, focusing on the bulge site, of a representative distance-refined structure of the SCA- $\alpha$ 2•bulged DNA complex are shown in Figure S4 of the Supporting Information.

## DISCUSSION

**Selection of Bulged DNA Sequence and Construct Comparison.** Sequence screening of various bulged DNA oligonucleotides resulted in the selection of the 25-mer shown in Figure 1D as a suitable binding substrate for SCA- $\alpha$ 2, based on its tight binding affinity and especially the quality of the ensuing NMR spectra. Figure 2 demonstrates the superior quality of the NMR spectra in both free and ligand-bound

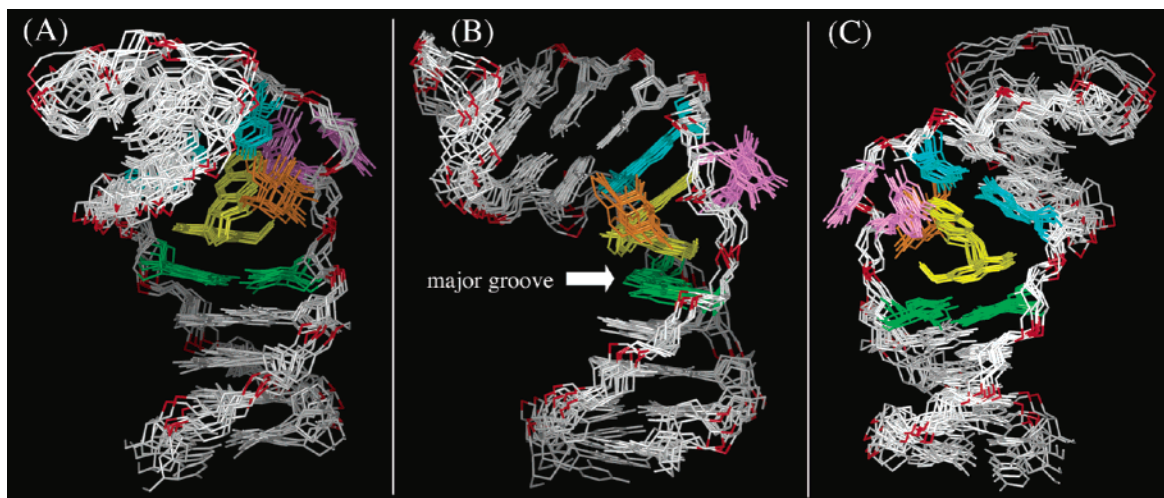


FIGURE 5: Eight superimposed refinement structures of the SCA- $\alpha$ 2•bulged DNA complex. The base residues of the T13-T14-T15 loop, the phosphate oxygen atoms of the strand backbone, and all DNA protons have been removed for clarity. SCA- $\alpha$ 2, except for its NMF moiety, is colored yellow. NMF is colored orange. The dG21•dC5 base pair is colored green. The dG20•dC8 base pair is colored cyan. Two bulge residues, G6 and T7, are colored magenta. The phosphorus atom is colored red. (A) Front view of the bulge binding site from the major groove. (B) Side view of the bulge binding site. (C) Front view of the bulge binding site from the minor groove.

Table 2: NMR Refinement Statistics of the SCA- $\alpha$ 2•Bulged 25-mer DNA Complex

NMR distance restraints	
overall no. NOE distance restraints	506
total no. of DNA distance restraints	400
no. of exchangeable distance restraints	34
no. of nonexchangeable distance restraints	326
no. of hydrogen bond restraints	40
total no. of ligand intramolecular distance restraints	29
total no. of ligand intermolecular distance restraints	77
no. of exchangeable distance restraints	31
no. of nonexchangeable distance restraints	46
structural statistics of the complex	
NOE violations	
number >0.2 Å	3.7 ± 1.4
maximum violations (Å)	0.29 ± 0.17
rmsd of violations	0.036 ± 0.002
deviation from the ideal covalent geometry	
bond lengths (Å)	0.013 ± 0.001
bond angles (deg)	3.04 ± 0.04
impropers (deg)	0.41 ± 0.05
pairwise rmsd (Å) (eight distance-refined structures)	
global structure	0.93 ± 0.38
central bulge binding site	0.77 ± 0.23

forms, with sharp and well-resolved NMR peaks. Both sequences shown in panels B and D of Figure 1 (termed 18-mer and 25-mer, respectively) possess a putative two-residue bulge with two putative short duplex stems on each side. However, the 25-mer has two bulge-flanking dG•dC base pairs instead of the two dA•dT pairs in the 18-mer (vide infra for the structural significance of this choice), in addition to a T-T-T loop, rather than an artificial chemical-linker loop. In the process of monitoring the effect of the ligands on the bulge region of the DNA sequences, we tracked the alterations of characteristic imino proton NMR peaks of bulge-flanking base pairs. In the case of the formation of the SCA- $\alpha$ 2•bulged DNA complex, the tracking of the upfield shifts of the bulge-flanking dG•dC base pairs has some advantages over that of dA•dT base pairs. In Figure 2, approximately 1.0 ppm upfield shifts allow the two imino proton peaks of bulge-flanking dG•dC base pairs in the ligand-free form (G20 at 12.50 ppm and G21 at 12.65 ppm) to move to the well-isolated region around 11.5 ppm. Alternatively, those of the dA•dT base pairs would be

expected to shift to the crowded region of the regular stem dG•dC pairs. Further, the sharpness of the two imino proton peaks of bulge-flanking dG•dC base pairs facilitates the collection of valuable intermolecular NOEs (Figure 3B) between SCA- $\alpha$ 2 and the DNA protons, which are critical in the NMR-derived structural determination.

NMR studies of the ligand-free form of the 18-mer indicate that there is a 5 bp duplex stem at the open end of the oligonucleotide with the other half of the oligonucleotide encompassing an unstructured loop (16); NCSi-gb induces the formation of a two-base bulge when it binds to it (15, 16). On the other hand, the ligand-free form of the more stable 25-mer possesses a two-base bulge with 5 bp duplex stems on each side, and SCA- $\alpha$ 2 is capable of recognizing and binding to the defined bulge region (compare Figure 4 and Figure S3). Further, NMR analysis shows that there is no effect of SCA- $\alpha$ 2 binding on the T-T-T loop segment of the 25-mer. Therefore, these data directly confirm the bulge binding specificity of SCA- $\alpha$ 2. Unlike the spectra of the NCSi-gb•18-mer complex, which also reflect the remaining 10% of the undesired free form of DNA (16), the clean spectra of the SCA- $\alpha$ 2•25-mer complex with a 1:1 ratio demonstrate only one species of complex formed in solution. This high-affinity binding property is also supported by a fluorescence quenching study, in which a  $K_d$  of 0.13  $\mu$ M was determined. Another desired feature of the SCA- $\alpha$ 2•25-mer complex is its stability within the temperature range (0–30 °C) used in this study. There was no evidence of alteration of the binding mode with a change in temperature. This feature enabled the assignment of the overlapped NMR peaks by means of temperature alteration.

We collected NMR spectra at different temperatures (from 0 to 30 °C with 5 °C increments) primarily because of partial overlaps in some spectra. The line width of the NMR signals becomes broader at low temperatures due to slower tumbling; however, at low temperatures, the binding affinity increases (supported by our fluorescence binding studies at different temperatures), the DNA base pairs themselves are more stable, and the exchange rate (imino and amino protons) with water is slower. Therefore, the quality of spectrum reflects

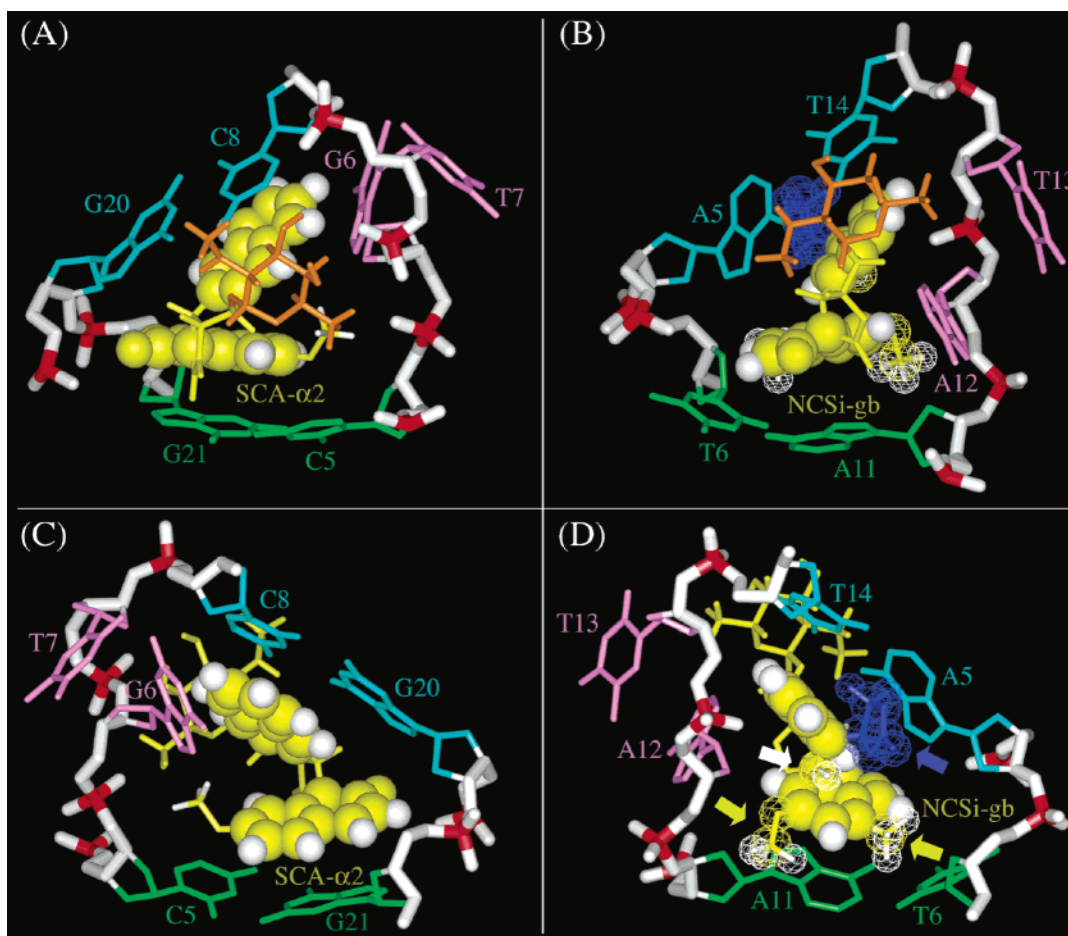


FIGURE 6: Structural comparison of the two ligand·bulged DNA complexes at the bulge binding sites. The sugar–phosphate backbone of DNA is highlighted with thicker stick drawings; the aromatic ring moieties of ligands are highlighted as yellow spheres for carbon atoms and as white spheres for protons. Ligands, except for their NMF moieties, are colored yellow. NMF moieties are colored orange. The 5′-bulge-flanking base pairs are colored green. The 3′-bulge-flanking base pairs are colored cyan. Bulge residues are colored magenta. Phosphorus atoms are colored red. The bulge binding site of the SCA- $\alpha$ 2-bulged DNA complex is shown in the left panel, with views looking into the major groove (A) and into the minor groove (C). The bulge binding site of the NCSi-gb-bulged DNA complex is shown in the right panel, with views looking into the major groove (B) and into the minor groove (D). The cyclocarbonate moiety is shown as a van der Waals line surface drawing with a pointing blue arrow; the methyl and methoxyl functional groups are shown as van der Waals line surface drawings with pointing yellow arrows, and the hydroxyl group is shown as a van der Waals line surface drawing with a pointing white arrow.

the overall effects of all of these factors. At low temperatures, the quality of spectra is good with less peak broadening but some overlaps (for example, see Figure S2 which is a NOESY spectrum in H<sub>2</sub>O at 10 °C). In this low-temperature spectrum, the intermolecular cross-peaks (6.4–7.0 ppm in the horizontal dimension, 5.0–6.0 ppm in the vertical dimension) are fairly sharp. On the other hand, the corresponding cross-peaks at 30 °C (Figure 4) are broad, basically due to the lower binding affinity at the higher temperature. Despite this, we chose to present the relatively better-resolved Figure 4 to demonstrate the full assignments of the DNA sequential walk. In addition, Figure 4 also indicates the temperature tolerance of the SCA- $\alpha$ 2·25-mer complex, making the data more physiologically relevant.

**Structural Comparison of the Two Complexes.** SCA- $\alpha$ 2 is a simplified analogue of the natural degradation product NCSi-gb. Both ligands share the same structural features of several functional moieties which contribute to their similar binding modes with bulged DNA substrates. The comparison of high-resolution NMR structures of the two different complexes at the bulge-binding sites is shown in Figure 6. In each case, the two rigidly held aromatic ring subunits of

the ligands form a molecular wedge that penetrates the bulge-binding pocket and immobilizes the two otherwise flexible bulge residues. Each ligand aromatic ring moiety stacks on the DNA bulge-flanking base pairs that define the long sides of the triangular prism binding pocket. The two aromatic ring moieties of both ligands mimic the geometry of helical bases with a right-handed twist angle of approximately 35° (measured by the long axes). The reduced enediyne systems, namely, the tetrahydroindacene (THI) moiety of NCSi-gb and the analogous benzindanol (BI) moiety of SCA- $\alpha$ 2, stack underneath the 3′-bulge-flanking base pair (colored cyan), while the dihydronaphthone (NA) moiety stacks above the 5′-bulge-flanking base pair (colored green in Figure 6). In addition, the orange-colored *N*-methylfucosamine (NMF) moiety of both ligands, with the same  $\alpha$ -glycosidic linkage to the aglycon skeleton, anchors in the major groove of DNA and points toward the 3′-bulge-flanking base pair (Figure 6A,B). Though DNA binding substrates with different bulge sequences are used, both ligands form ligand·DNA complexes with similar bulge binding modes. It appears, therefore, that the prism-shaped spirocyclic skeleton of the ligands confers the binding affinity based on shape rather than



sequence of the bulge targets. In addition, the global structures of both complexes confirm the formation of duplex stems at each side of the bulge. Due to the insertion of the ligand into the bulge-binding pocket, the whole DNA duplex stems are clearly bent (Figure 5B and ref 16).

Though basically adopting similar binding modes, the SCA- $\alpha$ 2•bulged DNA complex distinguishes itself from the NCSi-gb•bulged DNA complex in local environments. The bulky cyclocarbonate moiety [shown as a van der Waals line surface drawing (blue arrow)] on NCSi-gb sterically hinders the stacking of the THI ring platform with the 3'-bulge-flanking base pair, dA5•dT14 (Figure 6B,D). As a result, the more flexible dA5•dT14 pair of the NCSi-gb•bulged DNA complex is much more twisted and buckled than that of the dG20•dC8 pair in the SCA- $\alpha$ 2•bulged DNA complex (Figure 6A,C). The G20 residue stacks right above the BI ring platform and results in the NMR upfield shift of the G20 imino proton due to the ring current effect (Figure 6A,C); on the other hand, the T14 residue stacks poorly, off-center and tilted, on the THI ring platform (Figure 6B,D) and, therefore, exhibits an upfield shift much smaller than expected (16). Accordingly, it was proposed that the removal of the bulky cyclocarbonate moiety would likely improve the binding behavior (16, 17). Our result supports this assumption.

Another important insight derived from these data relates to the earlier observation that substitution of a dG5•dC14 base pair for the dA5•dT14 pair significantly weakened binding of NCSi-gb to the bulged DNA (10, 17). In fact, even substitution of a dT5•dA14 base pair had a similar effect. These results can be explained by the need to accommodate the cyclocarbonate moiety of NCSi-gb in the limited space. The weaker but more flexible dA•dT base pair permits distortion and space expansion that a more stable but more rigid dG•dC base pair will not. Further, the cyclocarbonate specifically interacts with the smaller pyrimidine (T14) residue, so having a larger purine (A14) residue at this position would also be expected to compromise NCSi-gb binding. Because SCA- $\alpha$ 2 lacks the cyclocarbonate, the dG•dC base pair at this site not only fails to impair binding but provides stability to formation of the bulge binding site. Thus, it is important to distinguish that the described "nucleotide sequence effects" actually derive from shape and space considerations. We have thus succeeded in our goal of designing a bulge-specific small molecule that does not have specific sequence limitations of the base pairs comprising the bulge pocket (17).

Also, the methyl and methoxyl functional groups (shown as van der Waals line surface drawings with yellow arrows) on the NA ring of NCSi-gb cause moderate in-plane disruption of the 5'-bulge-flanking dA11•dT6 base pair (Figure 6D). By contrast, the lack of a methyl group and especially the free rotation of the methoxyl group (shown as a stick drawing in Figure 6C) on the NA ring of SCA- $\alpha$ 2 allow a better stacking geometry of the NA ring right above the 5'-bulge-flanking dG21•dC5 base pair. In Figure 6C, the methoxyl group points up, away from the dG21•dC5 pair and toward the vicinity between the NA and BI rings of SCA- $\alpha$ 2. In the case of the NCSi-gb•bulged DNA complex, however, the rotation of the methoxyl group away from the dA11•dT6 base pair is restricted. Due to compression by the DNA bulge pocket, the THI ring is tilted downward and the

two discrete ring platforms of NCSi-gb are brought closer together than that of SCA- $\alpha$ 2 (Figure S5 of the Supporting Information), resulting in limited space to accommodate the methoxyl group when it rotates upward. In addition, an extra hydroxyl group (shown as a van der Waals line surface drawing with a pointing white arrow in Figure 6D) on the THI ring occupies this same region between the two ring platforms. Steric hindrance, reflected by the short distance (2.94 Å) between the two oxygen atoms of the hydroxyl and methoxyl groups, likely drives the methoxyl group to point down and toward the dA11•dT6 base pair. Importantly, compression of the same ring systems (and the ensuing intramolecular reaction between them) has also been implicated in the activation of NCS-chrom to its bulge-specific, strand-cleaving radical species (12).

These results are also relevant to an earlier proposal for why NCSi-gb binds so poorly to RNA (19, 20), since the 2'-OH group of a ribo A11 would clash with the downward-positioned methoxyl moiety on the drug. Accordingly, the ability of the methoxyl group of SCA- $\alpha$ 2 to point upward would be expected to facilitate binding to RNA bulges. Indeed, we have recently reported the formation of such complexes (25).

Both complexes have identical NMF carbohydrate moieties (colored orange), which, as a functional handle, align in the major groove of DNA (Figure 6A,B). However, the NMF moiety in the NCSi-gb•bulged DNA complex (Figure 6B) is closer to the floor center of the major groove, whereas the NMF moiety of the SCA- $\alpha$ 2•bulged DNA complex is nearer to the sugar-phosphate backbone of the bulge-containing strand (Figure 6A). A H-bond between the hydroxyl group of NMF and the backbone phosphate may potentially form, providing another driving force to facilitate the binding of SCA- $\alpha$ 2 with the bulged DNA.

When the ligand is inserted into the bulge of the DNA in the two complexes, both bulge residues become extrahelical. The striking difference, however, is that the bulge residues of the NCSi-gb•bulged DNA complex move toward the major groove while those of the SCA- $\alpha$ 2•bulged DNA complex point toward the minor groove. It appears that the positioning of the NMF moiety of the SCA- $\alpha$ 2•bulged DNA complex, which is embedded in the major groove and approaches the sugar-phosphate backbone of the bulge residue-containing segment, blocks the space in the major groove in which the extrahelical bulge residues fit. Compared to the T7 residue, the bulged G6 residue is closer to the H17-H19 protons on the BI ring of SCA- $\alpha$ 2, in accordance with the observations of corresponding intermolecular NOEs (listed in Table 1). These NOEs, between the minor groove-positioned H17-H19 protons of the ligand and the DNA G6 residue protons, define the position of the G6 residue in the minor groove.

The two aromatic ring moieties, NA and THI, of NCSi-gb are interlocked through a spirolactone pentane ring junction, while SCA- $\alpha$ 2 has a spirocyclopentene ring as the junction. This modification makes SCA- $\alpha$ 2 more stable but does not influence the binding mode. However, the two newly introduced ethylene protons, H24 and H25 in the analogue, exhibit characteristic chemical shifts at 5.5-6.5 ppm, which are readily distinguished from other ligand protons. In addition, the strong NOE between H24 and H25 can also be easily identified (as peak g in Figure S2). These

NMR features serve as a marker for this ethylene functional group. NOEs between the major groove-oriented amino proton of the bulge-flanking dC5 residue and these marker protons are observed (shown as peaks f and j in Figure S2). This ethylene group is on the extreme edge of the spiro junction; accordingly, NOEs f and j confirm the major groove binding mode of SCA- $\alpha$ 2.

## SUMMARY

SCA- $\alpha$ 2, a designed stable analogue of NCSi-gb, binds specifically and tightly at a two-base bulge in DNA via stacking of its helically oriented aromatic ring systems on the bulge-flanking base pairs that define the long sides of the triangular prism binding pocket. The structures of the complexes formed by the two ligands with DNA bulges are similar but differ importantly in local environmental details. The 2-*N*-methylfucosamine moiety, which resides in the major groove, with that of SCA- $\alpha$ 2 closer to the sugar-phosphate backbone of the bulge-containing strand, is critical in determining the proper placement of the aglycone in the bulge pocket. Whereas insertion of either ligand into the bulge causes the unpaired residues to become extrahelical, those in the SCA- $\alpha$ 2•bulge complex point to the minor rather than the major groove, which is the case with the complex formed with the natural product. The absence of the bulky cyclocarbonate moiety and other groups on SCA- $\alpha$ 2 enables it to bind in the bulge pocket with a better glove-in-hand fit than NCSi-gb and without the twisting and buckling of the base pair 3' to the bulge found with the natural product. The ability of the methoxyl moiety on the NA ring system to rotate upward away from the base pair 5' to the bulge residues, in contrast to NCSi-gb, likely plays a role in enabling SCA- $\alpha$ 2 to bind to bulges in RNA. These considerations make the designed analogue an appropriate platform for the placement of reactive moieties that react covalently (adduction or strand cleavage) at the bulge site and for generating congeners specific for bulges in both DNA and RNA. This study also provides new insights into the importance of shape and space, as opposed to nucleotide sequence, in formation of a complex at the bulge site.

## ACKNOWLEDGMENT

We appreciate the availability of NMR provided at the Harvard Medical School and especially thank Dr. Dinshaw Patel for allowing us to use the NMR facility at the Memorial Sloan-Kettering Cancer Center.

## SUPPORTING INFORMATION AVAILABLE

Proton chemical shift changes (parts per million) of the DNA bulge binding site in the SCA- $\alpha$ 2•bulged 25-mer DNA complex relative to the free bulged 25-mer DNA at 10 °C (Table S1), proton chemical shift changes (parts per million) of the bound form of SCA- $\alpha$ 2 in the SCA- $\alpha$ 2•bulged 25-mer DNA complex (in D<sub>2</sub>O at 25 °C) relative to the free form of SCA- $\alpha$ 2 (in *d*<sub>4</sub>-MeOH at 25 °C) (Table S2), exchangeable proton chemical shifts (parts per million) of DNA components within the SCA- $\alpha$ 2•bulged 25-mer DNA complex in H<sub>2</sub>O buffer at pH 6.8 and 10 °C (Table S3A), nonexchangeable proton chemical shifts (parts per million) of DNA components within the SCA- $\alpha$ 2•bulged 25-mer DNA complex in D<sub>2</sub>O buffer at pH 6.8 and 10 °C (Table

S3B), titration of various amount of SCA- $\alpha$ 2 into 25-mer bulged DNA as monitored by <sup>1</sup>H NMR shown as an expanded imino proton region (Figure S1), expanded NOESY (200 ms mixing time) contour plots of the SCA- $\alpha$ 2•bulged 25mer DNA complex, focusing on the amino and aromatic proton regions (Figure S2), expanded NOESY (200 ms mixing time) contour plots for the bulged 25-mer DNA alone at 30 °C (Figure S3), stereoviews of the bulge binding site of a representative refined structure of the SCA- $\alpha$ 2•bulged 25-mer DNA complex (Figure S4), and a comparison of ring system compression of free and bound forms of the ligands (Figure S5). This material is available free of charge via the Internet at <http://pubs.acs.org>.

## REFERENCES

- Turner, D. H. (1992) Bulges in nucleic acids, *Curr. Opin. Struct. Biol.* 2, 334–337.
- Lilley, D. M. J. (1995) Kinking of DNA and RNA by base bulges, *Proc. Natl. Acad. Sci. U.S.A.* 92, 7140–7142.
- Chastain, M., and Tinoco, I. (1991) Structural elements in RNA, in *Progress in Nucleic Acid Research and Molecular Biology* (Cohn, W. E., and Moldave, K., Eds.) Vol. 41, pp 131–177, Academic Press, New York.
- Kunkel, T. A. (1993) Slippery DNA and diseases, *Nature* 365, 207–209.
- Payet, D., Hillisch, A., Lowe, N., Diekmann, S., and Travers, A. (1999) The recognition of distorted DNA structures by HMG-D: A footprinting and molecular modelling study, *J. Mol. Biol.* 294, 79–91.
- Degtyareva, N., Subramanian, D., and Griffith, J. D. (2001) Analysis of the binding of p53 to DNAs containing mismatched and bulged bases, *J. Biol. Chem.* 276, 8778–8784.
- Dingwall, C., Ernberg, I., Gait, M. J., Green, S. M., Heaphy, S., Kam, J., Lowe, A. D., Singh, M., Skinner, M. A., and Valerio, R. (1989) HIV-1 tat protein binds trans-activation-responsive region (TAR) RNA in vitro, *Proc. Natl. Acad. Sci. U.S.A.* 86, 6925–6929.
- Xi, Z., and Goldberg, I. H. (1999) DNA-Damaging Eneidyne Compounds, in *Comprehensive Natural Products Chemistry* (Barton, D. H. R., Nakanishi, K., and Meth-Cohn, O., Eds.) Vol. 7, pp 553–592, Elsevier Science, Oxford, U.K.
- Kappen, L. S., and Goldberg, I. H. (1993) DNA conformation-induced activation of an enediyne for site-specific cleavage, *Science* 261, 1319–1321.
- 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.
- Kappen, L. S., and Goldberg, I. H. (1995) Bulge-specific cleavage in transactivation response region RNA and its DNA analogue by neocarzinostatin chromophore, *Biochemistry* 34, 5997–6002.
- Xi, Z., Mao, Q. K., and Goldberg, I. H. (1999) Mechanistic studies on the base-catalyzed transformation of neocarzinostatin chromophore: Roles of bulged DNA, *Biochemistry* 38, 4342–4354.
- Hensens, O. D., Chin, D.-H., Stassinopoulos, A., Zink, D. L., 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.
- 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, 2267–2275.
- Stassinopoulos, A., Ji, J., Gao, S., and Goldberg, I. H. (1996) Solution structure of a two-base DNA bulge complexed with an enediyne cleaving analogue, *Science* 272, 1943–1946.
- 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-post activated neocarzinostatin chromophore, *Biochemistry* 41, 5131–5143.
- 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.
18. Lin, Y., Jones, G. B., Hwang, G.-S., Kappen, L., and Goldberg, I. H. (2005) Convenient synthesis of NCS-chromophore metabolite isosteres: Binding agents for bulged DNA microenvironments, *Org. Lett.* 7, 71–74.
19. 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.
20. Kappen, L. S., Xi, Z., and Goldberg, I. H. (1997) Effect of ribonucleotide substitution on nucleic acid bulge recognition by neocarzinostatin, *Bioorg. Med. Chem.* 5, 1221–1227.
21. Hwang, G.-S., Jones, G. B., and Goldberg, I. H. (2003) Solution structure of a wedge-shaped synthetic molecule at a two-base bulge site in DNA, *Biochemistry* 42, 8472–8483.
22. Hwang, G.-S., Jones, G. B., and Goldberg, I. H. (2004) Stereochemical control of small molecule binding to bulged DNA: Comparison of structures of spirocyclic enantiomer-bulged DNA complexes, *Biochemistry* 43, 641–650.
23. Jones, G. B., Lin, Y., Xiao, Z., Kappen, L., and Goldberg, I. H. (2007) Molecular probes of DNA bulges: Functional assay and spectroscopic analysis, *Bioorg. Med. Chem.* 15, 784–790.
24. Kappen, L. S., Lin, Y., Jones, G. B., and Goldberg, I. H. (2007) Probing DNA bulges with designed helical spirocyclic molecules, *Biochemistry* 46, 561–567.
25. Xiao, Z., Zhang, N., Lin, Y., Jones, G. B., and Goldberg, I. H. (2006) Spirocyclic helical compounds as binding agents for bulged RNA, including HIV-2 TAR, *Chem. Commun.* 42, 4431–4433.
26. Wang, Y., and Patel, D. J. (1993) Solution structure of the human telomeric repeat d[AG<sub>3</sub>(T<sub>2</sub>AG<sub>3</sub>)<sub>3</sub>] G-tetraplex, *Structure* 1, 263–282.
27. Zhang, N., Lin, C., Huang, X., Kolbanovskiy, A., Hingerty, B. E., Amin, S., Broyde, S., Geacintov, N. E., and Patel, D. J. (2005) Methylation of cytosine at C5 in a CpG sequence context causes a conformational switch of a benzo[*a*]pyrene diol epoxide-*N*2-guanine adduct in DNA from a minor groove alignment to intercalation with base displacement, *J. Mol. Biol.* 346, 951–965.
28. Brunger, A. T. (1992) *X-PLOR. A System for X-ray Crystallography and NMR*, Yale University Press, New Haven, CT.

BI602599D