Binding-Induced Activation of DNA Alkylation by Duocarmycin SA: Insights from the Structure of an Indole Derivative−DNA Adduct

Jason R. Schnell,[†] Randal R. Ketchem,[†] Dale L. Boger,[‡] and Walter J. Chazin*[,†]

Contribution from the Departments of Molecular Biology and Chemistry and the Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037

Received October 7, 1998

Abstract: The mechanism for catalysis of DNA alkylation by the potent antitumor antibiotic duocarmycin SA (DSA) has been probed by determining the structure of a DNA adduct of the indole analogue (DSA-indole, DSI) lacking three methoxy functional groups. The three-dimensional structure of DSI covalently bound to A_{19} in d-(G_1AC<u>TAATT</u>GAC_{11})·d-(G_{12}TC<u>AATTA</u>GTC_{22}) was determined by ^{1}H NMR spectroscopy using a total of 935 experimental distance and dihedral angle constraints. The representative ensemble of 20 conformers has no distance restraint violations greater than 0.03 Å, no torsional restraint violations greater than 0.7°, and a pairwise rmsd over all atoms in the binding site of 0.48 Å. Comparison of the structures of the DSA and DSI adducts reveals a structural basis for the critical role of one of the trimethoxy-indole functional groups in alkylation reactivity. A deeper penetration into the DNA minor groove in the vicinty of the indole subunit is observed for the DSI versus the DSA adduct, along with some variations in the width and depth of the minor groove throughout the binding site. The most significant difference between the DSI and DSA addducts is the 8° smaller twist of the two ligand subunits in DSI, which correlates with its ∼20-fold slower rate of DNA alkylation. This comparison of the structures of the DSI and DSA adducts to the same DNA duplex provides the most direct evidence to date in support of the proposal that the binding of the ligand in the DNA minor groove and consequent twisting of the two ligand subunts, disrupting vinylogous amide stabilization and thereby activating the conjugated cyclopropane electrophile, plays a central role in controlling DNA alkylation reactivity.

Introduction

Duocarmycin SA (DSA) is a potent antitumor antibiotic whose activity is associated with sequence-selective alkylation of DNA. The mechanism of alkylation by DSA and other members of the CC-1065 class of DNA alkylating agents has been extensively studied, but the details of binding site recognition and alkylation catalysis are not yet fully understood. Extensive studies have shown that these and related molecules bind in the minor groove of duplex DNA in AT-rich regions, and react via nucleophilic attack on the least substituted carbon of the spirocyclopropyl ring of the ligand by the N3 of an adenine residue. What is less clear is the origin of the sequence selectivity and especially the source of catalysis of the alkylation reaction. Recently, important insights have been made in recognizing that binding of the ligand to DNA results in changes in the ligand conformation and an accompanying increase in its alkylation reactivity.[1] Herein, we describe structural studies designed to provide further insights into the details of this binding-induced conformational change and its correlation with the consequent DNA alkylation reactivity, efficiency, and selectivity.

The proposal that DNA binding induces activation of the ligand was introduced on the basis of several key observations about the site selectivity, alkylation rate, and overall alkylation efficiencies of numerous CC-1065 and duocarmycin analogues. The most unique feature of these molecules is the extraordinary stability of the alkylation subunit cyclopropane. Under aqueous solvolytic conditions, the conjugated cyclopropane of DSA has a half-life of 177 h at pH 3.[2] This unusual stability is attributed to an intrinsic feature of the structure, the vinylogous amide at N12 (Figure 1). Acid-promoted solvolysis studies revealed a large Hammett coefficient ($\rho = -3.0$) for the N12 substituent,[3] indicating that even small changes in the relative extent of the vinylogous amide conjugation are accompanied by large perturbations in the reactivity of the agent. X-ray crystal structures of DSA and several analogues revealed a flat coplanar arrangement of the two subunits with the three torsions making up the linker between them (χ_1, χ_2, χ_3) all approximately 0°.[4] This conformation of the ligand maximizes the conjugation throughout the molecule. However, binding in the minor groove of duplex DNA would not be possible in this geometric arrangement of the subunits, and one or more of the linking torsional angles must be twisted when the ligand binds to DNA. In the binding-induced activation hypothesis, the value of χ_1 increases significantly upon binding of the ligand to the DNA. The change in χ_1 decreases the cross-conjugated vinylogous amide stabilization of the conjugated cyclopropane ring, thereby leading to the observed increase in its alkylation reactivity, and an increase in the rate of DNA alkylation. X-ray crystal structures of a series of natural and analogue alkylation subunits have exhibited a good correlation between the degree of vinylogous amide

[†] Department of Molecular Biology.

[‡] Department of Chemistry and the Skaggs Institute for Chemical Biology.

(1) Boger, D. L.; Johnson, D. S. *Angew. Chem., Int. Ed. Eng.* **1996**, *35*, 1438−1474.

(2) Boger, D. L.; Machiya, K.; Hertzog, D. L.; Kitos, P. A.; Holmes, D. *J. Am. Chem. Soc.* **1993**, *115*, 9025−9036.

(3) Boger, D. L.; Yun, W. *J. Am. Chem. Soc.* **1994**, *116*, 5523−5524.

(4) Boger, D. L.; Goldberg, J. A.; McKie, J. A. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 1955−1960.

10.1021/ja983556j CCC: $18.00
Published on Web 06/06/1999

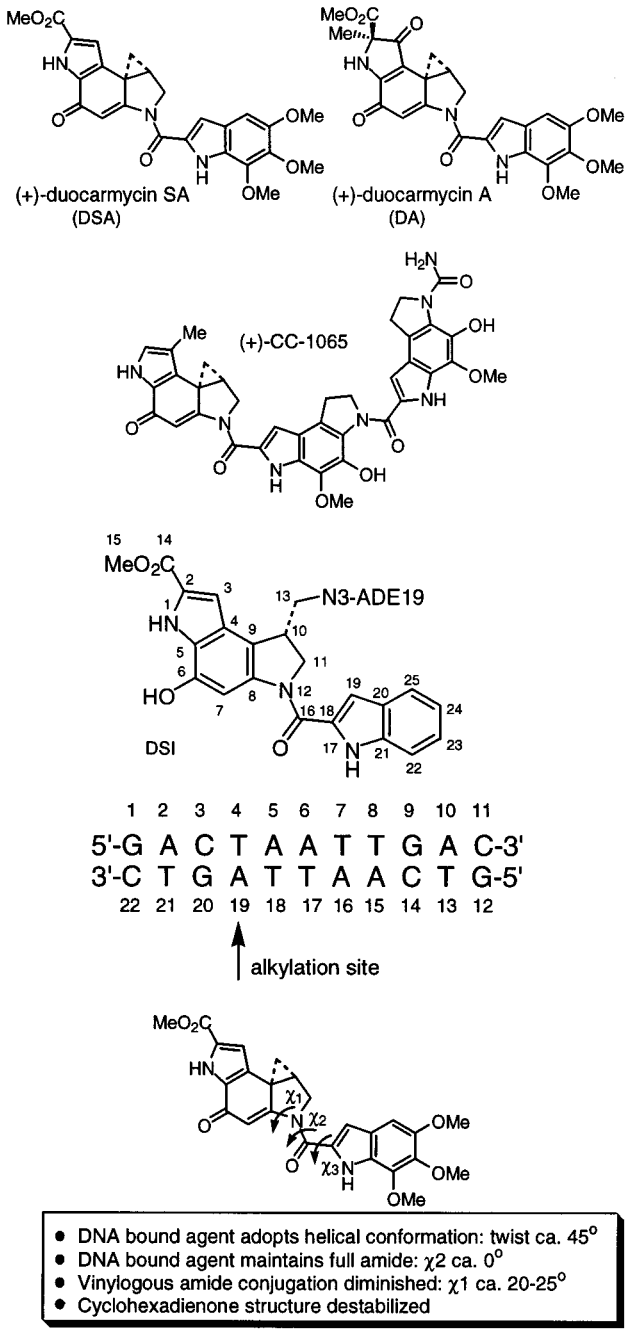

Figure 1. Molecular structures of (+)-duocarmycin SA, (+)-duocarmycin A, and (+)-CC1065, schematic diagram of the covalent (+)-duocarmycin SA-indole adduct, complete sequence of the "AATTA" 11 mer DNA duplex, and molecular structure of (+)-duocarmycin SA-indole indicating torsion angles χ_1, χ_2, and χ_3.

conjugation, the χ_1 dihedral angle, and the relative reactivities of the agents. Sufficient reactivity changes to account for DNA alkylation catalysis were observed with even partial disruption of the vinylogous amide. In addition, it has been proposed that the role of the binding subunit methoxyls of DSA is to extend the rigid length of the agent, thereby forcing a larger twist in order to accommodate DNA groove binding and resulting in an increased reactivity when bound. An important ramification of these observations is that this source of activation is independent of pH, consistent with the lack of acid catalysis of the DNA alkylation reaction.

We have enlisted the structural analysis of the covalent DNA complexes formed upon alkylation by DSA and selected analogues to probe the DNA binding-induced activation hypothesis and ultimately define the alkylation mechanism. Here, we report the solution structure of (+)-duocarmycin SA-indole (DSI) bound covalently to the 11 base pair "AATTA" DNA duplex (Figure 1), and compare it to the previously determined structure of (+)-duocarmycin SA (DSA) bound covalently to the same duplex.[5] DSA and DSI differ chemically in that the trimethoxyindole-binding subunit of DSA has been replaced with a simple indole in DSI. As a result of the minor simplification and as predicted, it binds DNA with a diminished perturbation in the χ_1 dihedral angle (8°). This smaller perturbation leads to less disruption of the ligand vinylogous amide, resulting in less activation for DNA alkylation consistent with its observed 20-fold slower rate of DNA alkylation.[6]

Experimental Section

Sample Preparation. The DNA oligomers were purchased from Biosource International (Menlo Park, CA) and purified by anion exchange chromotography on a Biocad Sprint (PerCeptives Biosystems, Cambridge, MA) using HQ POROS resin. To compare the overall alkylation efficiency and reactivity of (+)-DSI to the parent compound (+)-DSA, we carried out small-scale survey reactions on a panel of DNA duplexes containing different AT-rich binding sites, as described previously.[5] While site selectivity and overall efficiencies were very similar, the rate of reaction was slowed considerably: the time to half completion of the reaction with d-(GACTAATTGAC)·d-(GTCAATTAGTC) was 39 min for (+)-DSA and 973 min for (+)-DSI. Thus, DSI was 25 times slower at alkylating the deoxynucleotide, in excellent agreement with the prior studies.

To prepare the NMR sample, we reacted 1 μmol of the AATTA duplex with 2 μmol of (+)-DSI at 37 °C in a buffer containing 10 mM KH_2PO_4 and 10 mM KCl at pH 3, as described for the corresponding (+)-DSA adduct.[5] The only difference in preparing the DSI adduct was that, rather than isolating the alkylated DNA strand by HPLC and titrating back the complementary DNA strand to form the duplex, once the DSI reaction was deemed complete (~2 days), the solution was centrifuged and the supernatant transferred directly to an NMR tube without further purification.

NMR Experiments and Structural Constraints. NMR experiments were performed on Bruker AMX-600 spectrometers. The temperature was set to 300 K, except as indicated otherwise. The scalar correlated experiments recorded were two-quantum (τ_{mix}= 30 ms), P−COSY (35° observe pulse), TOCSY (τ_{mix}= 60, 100 ms), and natural abundance ^{13}C-^{1}H correlation spectra. These were complemented by two NOESY spectra recorded in D_2O with τ_{mix}= 50, 150 ms and a third NOESY spectrum recorded in H_2O with τ_{mix}= 200 ms. A detailed description of the acquisition parameters can be found elsewhere.[5] All data were processed on a Sun Sparcstation LX with Felix97 software (Molecular Simulations Inc., San Diego, CA).

To obtain distance constraints, we measured NOE intensities in the 150 ms D_2O NOESY spectrum and supplemented them with additional data from 50 ms D_2O and 200 ms H_2O NOESY spectra. In the initial rounds of structure calculations, the NOE intensities were classified into strong, medium, and weak categories and converted into upper bound distance constraints based on empirical calibrations. Once the first set of reasonable structures was obtained, the upper bounds were readjusted on

(5) Eis, P. S.; Smith, J. A.; Rydzewski, J. M.; Case, D. A.; Boger, D. L.; Chazin, W. J. *J. Mol. Biol.* **1997**, *272*, 237−252.

(6) Boger, D. L.; Hertzog, D. L.; Bollinger, B.; Johnson, D. S.; Cai, H.; Goldberg, J.; Turnbull, P. *J. Am. Chem. Soc.* **1997**, *119*, 4977−4986.

the basis of hybrid relaxation matrix calculations using the MARDIGRAS program.[7,8] The DNA from the structure of the (+)-DSA adduct was used for the first round of MARDIGRAS calculations; then each subsequent round used the best structure from the most recent structural ensemble. Due to the experimental uncertainties in the measurement of cross-peak intensities, the MARDIGRAS target distances were increased by a variable amount, in proportion to the magnitude of the NOE intensity. The distance constraint upper bounds were adjusted as follows: <2.0, +0.35 Å; 2−3, +0.7 Å; 3−4, +1.1 Å; 4−5, +1.5 Å; >5.0, +1.9 Å. All distance constraints assigned from NOEs were given a lower bound of 1.8 Å.

Watson−Crick hydrogen bonding constraints were included on the basis of observation of resonance frequencies in the appropriate regions of the ^{1}H NMR spectrum and characteristic NOEs between labile protons in the H_2O NOESY spectrum. Hydrogen bonds were identified for each guanosine and thymidine imino group and all cytosine amino groups. N···N or N···O distances,[9] as well as H···O or H···N distances, were constrained. The lower and upper bound distances were as follows: G−O6 to C−N4, 2.81−3.01 Å; G−N1 to C−N3, 2.85−3.05 Å; G−N1H to C−N3, 1.80−2.20 Å; G−O6 to C−N4H1, 1.76−2.16 Å; A−N1 to T−N3, 2.72−2.92 Å; A−N1 to T−N3H, 1.67−2.07 Å. Two additional H-bond constraints were included for the *A19*:T4 base pair (A−N6 to T−O4, 2.85−3.05 Å; and A−N6H1 to T−O4, 1.80−2.20 Å) because characteristic NOEs to the *A19*-6NH_2 resonances were observed in the NOESY spectrum acquired in H_2O.

Dihedral angle constraints were applied to the pseudorotational phase angle (*P*) and the sugar phosphate backbone angles β, γ, χ, and ϵ. Backbone angle constraint ranges were determined according to the method of Kim et al.[10] on the basis of scalar couplings observed in the PE-COSY and characteristic cross-peaks in the 150 ms NOESY. This analysis provided angular constraints with the following bounds: β, 105°−255°; γ, 20°−80°; χ, (−180°)−(−15°); ϵ, 130°−350°. *P* angles were determined from the $J_{1'2'}$ and $J_{1'2''}$ scalar couplings measured in the PE-COSY spectrum and a semiquantitative analysis of $J_{2'3'}$, $J_{2''3'}$, and $J_{3'4'}$. The constraints for *P* angles were set to the target values determined from the appropriate Karplus curve, ±30°. In a few instances, the values of the *P* angle for a given residue were found to tightly cluster at one of the extremes of the constraint boundaries in all structures. In these instances, the constraint range was expanded by an additional 10°−20°. In each of these cases, the angle in subsequent structure calculations then exhibited a preferred value comfortably within the constraint range.

The full list of all distance and dihedral angle constraints is available from the Brookhaven Data Bank under accession code 1dsi.

Structure Calculations. The computational protocol consisted of four primary steps. First, 40 starting DNA structures (average pairwise rmsd over all atoms = 3.04 Å) with varying degrees of x-displacement, incline, rise, and twist were constructed using the program NAB.[11] Holding base pair geometry fixed, we optimized the backbone geometry via 1000 steps of steepest descent energy minimization in the AMBER 4.1 force field[12] (average all atom pairwise rmsd = 2.99 Å). Next, the structures were refined by 20 ps of high temperature (1000 K) rMD using only the DNA intramolecular distance and torsion angle constraints. The resulting structures had an average all atom pairwise rmsd of 2.65 Å. Starting DSI conformations were generated by taking snapshots of a 50 ps high-temperature rMD simulation in vacuo. Randomly selected pairs of DNA and DSI conformers were then positioned using NAB, such that the DSI molecule was >10 Å away from the DNA minor groove. The long axis of each molecule was adjusted to be parallel, with the relative orientation set on the basis of the ligand−DNA NOEs. The ligand was docked during a 15 ps rMD simulation with the intra-DNA and intra-DSI restraints imposed for the duration of the simulation and the DNA-DSI intermolecular restraints ramped on linearly from 0 to 15 ps. These docked structures (pairwise rmsd of 1.18 Å) were then refined by two rounds of 20 ps of rMD simulated annealing (1 ps heating to 1000 K, 4 ps of high temperature rMD, followed by a gradual cooling to 0 K). All but one of the 40 starting structures was successfully refined. The resulting family of 39 structures had a pairwise rmsd over all atoms of 0.88 Å. All molecular dynamics (MD) and restrained molecular dynamics (rMD) calculations were carried out in vacuo with reduced phosphate charges and a distance-dependent dielectric using the Sander module of the AMBER simulation package.[12]

The 39 final structures were ordered by increasing residual constraint violation energy, and then a calculation of the number of structures required to accurately represent all of the conformational space consistent with the experimental data was made using FINDFAM.[13] This calculation indicated that 15 structures would be sufficient to represent the structure of the complex. However, the top 20 structures were included in the final ensemble to be consistent with the (+)-DSA complex. The 20 structures are essentially identical in terms of agreement with the experimental data and energies as calculated from the AMBER 4.1 force field. The all atom pairwise rmsd of these structures is 0.82 Å. The coordinates for this ensemble have been deposited with the Brookhaven Databank under accession code 1dsi.

The molecular structures were examined using Insight II (Molecular Simulations Inc., San Diego, CA). The helical parameters and groove widths and depths were calculated using CURVES.[14]

Results and Discussion

^{1}H NMR Characterization of the Adduct. The general strategies for ^{1}H NMR analysis of DNA duplexes and ligand−DNA complexes are well-established. The specific protocols used for the (+)-DSI adduct followed that described for the (+)-DSA adduct.[5] The characterization of the d-(GACTAATTGAC)·d-(GTCAATTAGTC) [AATTA] duplex in the absence of the ligand is also described in that report. The assignment of the ^{1}H NMR resonances of the DNA (Table S1) followed the usual approach of spin system identification followed by sequence-specific assignment on the basis of NOEs.[5,15] Significant chemical shift differences between the DSI-bound and the free DNA were centered around base pairs C3·G20 and T8·A15 (Figure 2), confirming the location of the ligand-binding

(7) Borgias, B. A.; James, T. L. *Methods Enzymol.* **1989**, *176*, 169−183.
(8) Borgias, B. A.; James, T. L. *J. Magn. Reson.* **1990**, *87*, 475−487.
(9) Schmitz, U.; Sethson, I.; Egan, W. M.; James, T. L. *J. Mol. Biol.* **1992**, *227*, 510−531.
(10) Kim, S.-G.; Lin, L.-J.; Reid, B. R. *Biochemistry* **1992**, *31*, 3564−3574.
(11) Macke, T., Ph.D. Thesis, The Scripps Research Institute, 1996.
(12) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E., III; Ferguson, D. M.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. AMBER 4.1. University of California, San Francisco, CA, 1994.
(13) Smith, J. A.; Paloma, L. G.; Case, D. A.; Chazin, W. J. *Magn. Reson. Chem.* **1996**, *34S*, 147−155.
(14) Stofer, E.; Lavery, R. *Biopolymers* **1994**, *34*, 337−346.
(15) Chazin, W. J.; Wuthrich, K.; Hyberts, S.; Rance, M.; Denny, W. A.; Leupin, W. *J. Mol. Biol.* **1986**, *190*, 439−453.

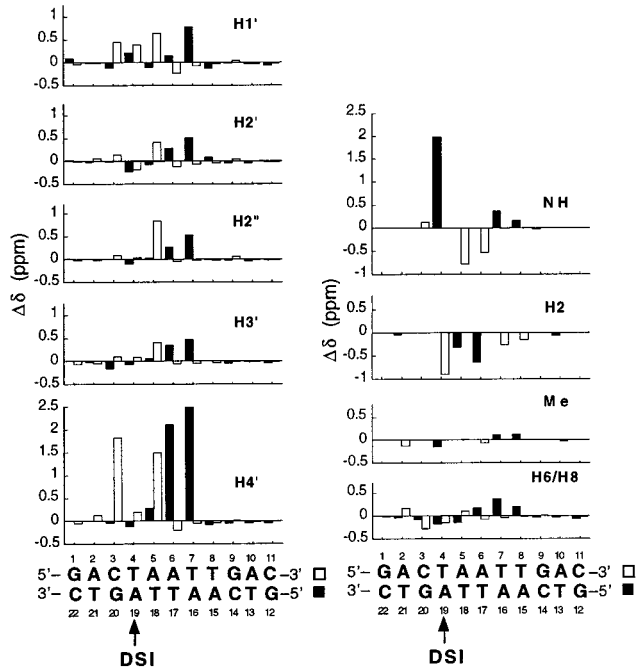

Figure 2. Chemical shift difference ($\Delta\delta = \delta_{free} - \delta_{adduct}$) plots for selected deoxyribose sugar and base protons of the AATTA duplex. Filled bars are for residues 1−11 and open bars for residues 12−22.

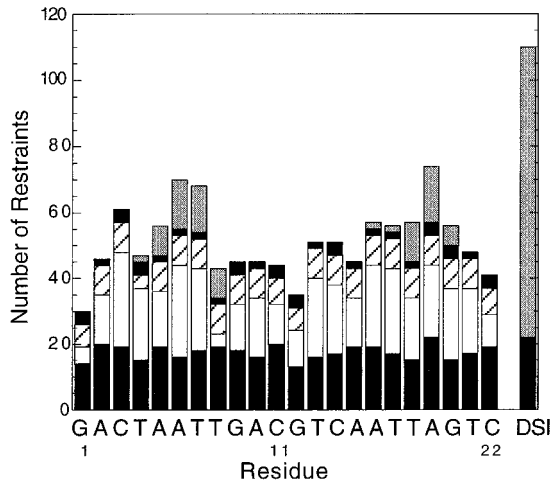

Figure 3. Distribution of experimental constraints for the DSI adduct: intraresidue NOEs (filled), inter-residue NOEs (open), torsions (hatched), hydrogen bonds (dark shading), and intermolecular ligand−DNA NOEs (light shading).

site as identical to that of the (+)-DSA adduct. The ^{1}H chemical shifts of the DSI and DSA adducts were similar except in the region of the indole ring (Figure S1).

Three-Dimensional Structure of the Complex. The excellent ^{1}H NMR chemical shift dispersion and the availability of virtually complete resonance assignments for the complex enabled a large number of constraints to be assigned for the structure calculations. The distribution of these structural constraints is shown in Figure 3, grouped into intraresidue, inter-residue, and intermolecular categories. A total of 716 distance constraints were assigned from the NOESY spectra, 105 angle constraints on the conformations of the sugar rings from the PE-COSY and 2Q spectra, and 82 angle constraints on the DNA backbone using the approach of Kim et al.[10] In addition, 32 hydrogen bond constraints were assigned for the guanosine and thymidine imino and cytosine amino protons, on the basis of their ^{1}H NMR resonances appearing in the spectral regions characteristic of Watson−Crick base pairing. In all, there were 935 constraints, corresponding to an average of 41/residue (57/residue in the binding site), counting the ligand as a single residue. This compares favorably to the 1047 constraints identified for the (+)-DSA adduct, particularly when noting that a majority of the "extra" constraints for the (+)-DSA adduct were of the intraresidue category. The computational protocol used to determine the structure of the (+)-DSI adduct followed that used for the (+)-DSA adduct,[5,11] so that comparisons between the two structures would not be biased by the method used to generate the structure. As noted previously, there is an unusually high density of experimental constraints for the DSI (and DSA) adduct, which provides well-defined three-dimensional structures that are essential to our goal of elucidating fundamental aspects of the alkylation chemistry of these agents.

The ensemble of 20 structures representing the (+)-DSI adduct is shown in Figure 4. The most well-defined regions in the structure are the DSI and its binding site in the DNA minor groove (base pairs G3·C20 to T8·A15). The average pairwise rmsd over all atoms is 0.82 Å and drops to 0.48 Å (0.33 Å from the mean) for the DSI and DNA atoms in the binding site. Structural statistics for the adduct are provided in Table 1. These parameters show excellent agreement with the input constraints (low violation energy <0.03 kcal/mol and only small residual constraint violations <0.03 Å and 0.7°), high precision (low rmsds), and chemically reasonable molecular structures (large negative total molecular energies and Lennard-Jones energies). Figures S2 and S3 show the variation in the backbone and pseudorotation angles and the helicoidal parameters for the DSI adduct, which fall predominantly in the range of B-DNA geometry. The DNA is fully base paired and falls well within the B-DNA structural family, except for a limited set of residues at and adjacent to the site of alkylation. The most dramatic variations in DNA conformation are observed in the vicinity of the site of alkylation and the opposite end of DNA binding site, in accord with the observations made for the DSA adduct.

Unique Features of the DNA in the (+)-DSI Adduct. The conformations of both the ligand and the DNA in the DSI adduct are very similar overall to those of the DSA adduct with pairwise rmsds of the mean structures of 0.91 Å (0.65 Å in the binding site). There are, however, important differences in the details, so the two adducts have been compared in a number of additional ways to accurately determine the structural effects of removing the three methoxyl groups on the indole ring. The most striking characteristic of the alkylated DNA is the relative dislocation of the first four base pairs (G1·C22 to T4·A19) from the rest of the duplex. This was seen previously in the DSA adduct, as well as in the structure of (+)-CC-1065 bound covalently to the duplex d(GGCGG*AGTTA**GG)·d(CC*T-AACT*CCGCC).[16,17] The shift in the DNA is reflected in an underwinding of the duplex in all three structures, and is presumably an intrinsic feature of duplex DNA alkylated by this class of agent. However, in the DSI adduct, there is a 45° lower value for the T18 pseudorotation angle and slightly larger values for tilt (4°) and roll (10°) at A5·T18 that leads to a straightening out of the helix relative to the DSA adduct, which remains somewhat kinked.

Another significant difference between the DSI and DSA adducts is the width of DNA minor groove (Figure 5). The narrowing of the minor groove in the AT-rich central tract in

(16) Lin, C. H.; Hill, G. C.; Hurley, L. H. *Chem. Res. Toxicol.* **1992**, *5*, 167−182; Sun, D.; Lin, C. H.; Hurley, L. H. *Biochemistry* **1993**, *32*, 4487−4495.

(17) Hurley, L. H.; Warpehoski, M. A. *Chem. Res. Toxicol.* **1988**, *1*, 315−333.

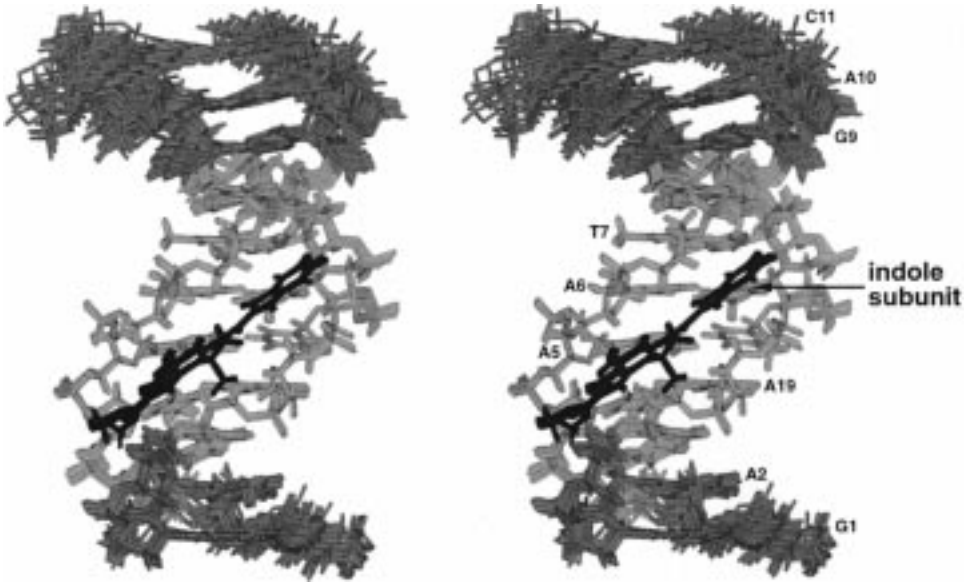

Figure 4. Stereoview of the conformational ensemble representing the three-dimensional structure of the covalent complex of DSI and the AATTA duplex. The 20 structures were superimposed by fitting to all atoms in the binding site (the DSI molecule and base pairs 3G·C20-T8·A15).

Table 1. Structural Statistics for the Final Structural Ensemble of the DSI Complex

Average AMBER Energies (kcal/mol)	
E_{total}	-834.3 ± 0.6
$E_{distance\ violation}$	0.004 ± 0.004
$E_{torsion\ violation}$	0.002 ± 0.001
$E_{Lennard-Jones}$	-411.5 ± 0.9
Average Number of Constraint Violations for the Ensemble[a]	
NOE violations ≥ 0.03 Å	0
PPA violations, $\|\theta\| = 0-0.7°$	2.75
backbone ($\chi,\beta,\gamma,\delta,\epsilon$) violations, $\|\theta\| > 0°$	0.05
Progression of the NMR Refinement (rmsd, Å)[b]	
40 starting DNA structures	3.04
rMD, 40 DNA structures only	2.65
rMD, DSA docking to 40 DNA structures	1.18
2 rMDs, 39 DSA-DNA complex structures	0.88
20 lowest constraint energy structures	0.82 (0.57)
DSA-binding site[c]	0.48 (0.33)

[a] Reported as the average number of violations in each structure for the given range. [b] Mean pairwise rms difference between the structures. The rmsd from the mean structure is given in parentheses. [c] The binding site region includes residues C3-T8, A15-G20, and the entire DSI molecule.

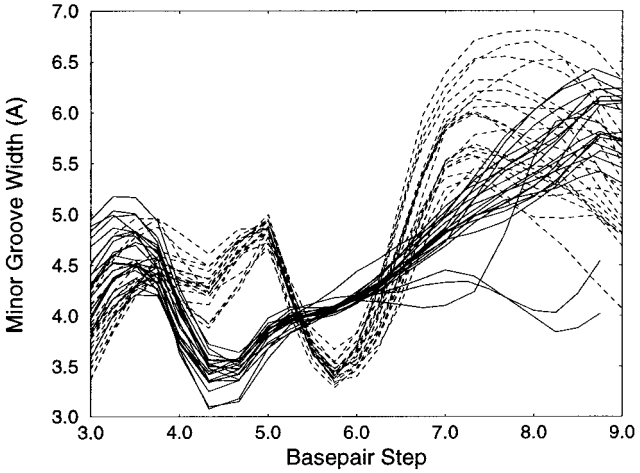

Figure 5. Variations in the DNA minor groove width in the covalent complexes of DSI (dashed lines) and DSA (solid lines) bound to the AATTA duplex. The plots were generated using the program Curves.[14]

the DSI adduct is centered at base pairs A6·T17 and T7·A16. In contrast, the width of the minor groove in the DSA adduct is at its minimum at the site of alkylation (A19·T4) and increases through to the trimethoxy indole-binding site (G9·C14). These differences in the minor groove are reflected in certain helicoidal parameters. For example, the small magnitude of the propeller twist at T4·A19 in the DSA adduct (−5°) is even more unusual in the DSI adduct, increasing to a very atypical value of +5°. The differences in the DNA structure of the DSA and the DSI adducts demonstrate the intrinsic plasticity of the DNA, a property that permits both the recognition and binding of the large variety of ligands required for DNA processing, as well as the effectiveness of chemotherapeutic agents targeted to DNA.

Unique Features of the Ligand in the (+)-DSI Adduct. The effect of removing the three methoxy groups of the TMI subunit of DSA appears not to be simply localized to the immediate region of the methoxy groups themselves, but rather extends to the vicinity of the covalent bond formed upon alkylation. In the DSI adduct, the ligand is seen to adopt a fit into the DNA minor groove, in which as snug as that in the DSA adduct. However, the indole ring has clearly penetrated more deeply into the groove (Figure 6). The absence of the more bulky methoxy groups of DSA enables greater accessibility to the minor groove. The deeper penetration of the indole ring into the minor groove is presumably driven by the increased stabilization available from more extensive favorable van der Waals interactions.

The most striking differences between the DSI and DSA adducts is an 8° *decrease* in the relative twist of the two subunits in DSI (45° for DSA, 37° for DSI), which is best viewed by looking down the respective minor grooves of two adducts, as shown in Figure 7. The intersubunit twist angle is extremely well-defined (±2.0° or better) in the two adducts because of the very large number of experimental data that position the ligands in the DNA minor groove, so the difference in twist angle between DSI and DSA adducts is accurately determined. Interestingly, the smaller twist angle in the DSI adduct is observed despite the fact that its deeper groove penetration accentuates the twisting requirement; that is, the twist angle would be even smaller if it were not for the depth of the groove

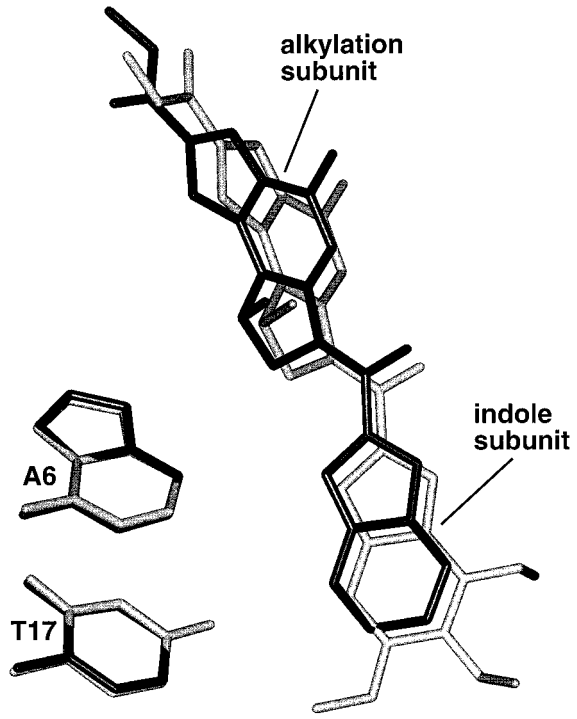

Figure 6. View into the plane of base pair A6·T17 showing the relative penetration of DSI (black) and DSA (grey) into the helix minor groove. Structures were superimposed over all heavy atoms of the bases shown.

penetration. This is integrally associated with the shorter effective length of DSI relative to DSA due to the absence of the C24 methoxyl group.

The twisting of the ligand is distributed into the three torsions (χ_1, χ_2, χ_3) that constitute the linker between the two subunits (Figure 1). In the DSA adduct, the χ_1 angle was most perturbed from the ground-state value of 0°, consistent with expectations based on the energy barriers to rotation in simple model compounds. The lower value of the intersubunit twist angle in the DSI adduct is almost completely due to a reduction in the χ_1 angle (14.2° ± 0.6° for DSI, 22.4° ± 0.7° for DSA). This observation is central to our interpretation of the origin of the differences in the relative rates of DNA alkylation of these two agents (vide infra).

Implications for the Hypothesis of Acid Catalysis Mediated by the DNA Backbone. One potential source of catalysis of the alkylation reaction advanced in the work of Hurley and Warpehoski[17,18] involves autocatalytic C6 carbonyl protonation of the ligand by a DNA backbone phosphate. Our structures show that the positionings of the proximal backbone phosphates in both the DSA and DSI adducts are much too far from the phenol C6 carbonyl center to be engaged in a direct hydrogen bond. The distance from the nearest backbone phosphate oxygen of the alkylated strand (T21) is 4.7 ± 0.2 and 5.0 ± 0.3 Å, and to the nearest backbone phosphate oxygen of the opposing strand (A7) is 4.4 ± 0.2 and 5.5 ± 0.3 Å, for the DSA and DSI adducts, respectively. Of course, the assumption that the structure of the bound product reveals the conformations and positioning of the reactants at the time of catalysis is implicit in this analysis.

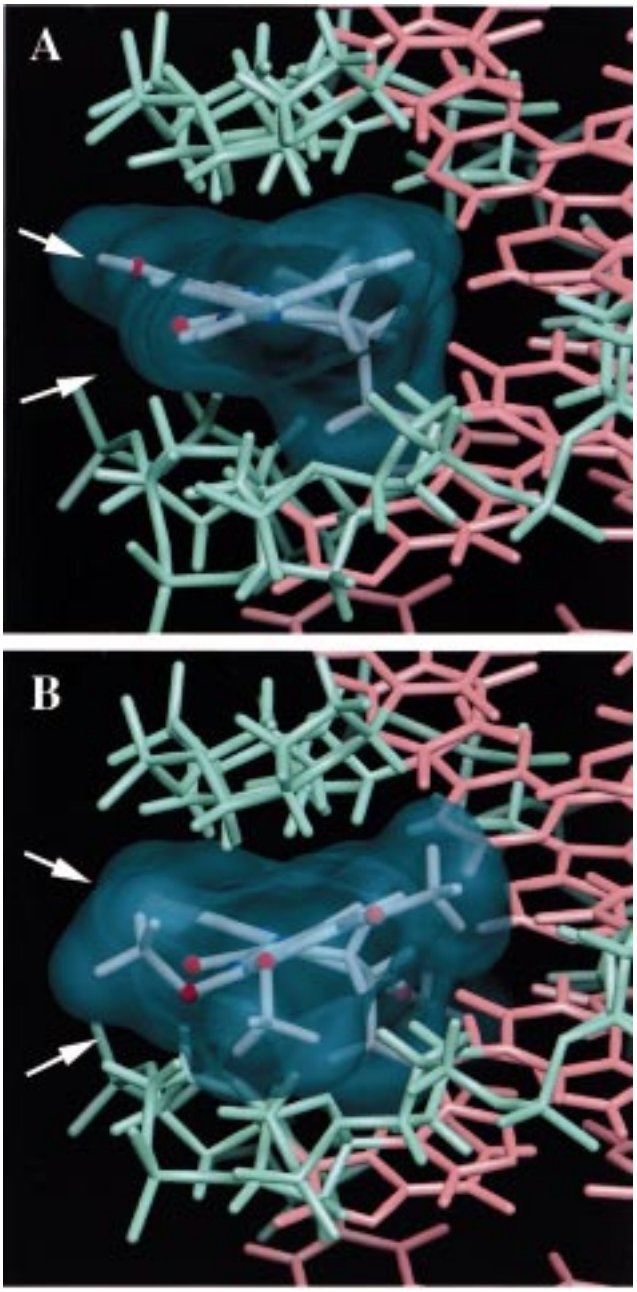

Figure 7. Difference in the intersubunit twist of the adducts of DSI (A) and DSA (B) bound to the AATTA 11mer duplex. Side view looking down the minor groove with a full surface representation of the ligand. The arrows point in the direction of the planes of the two subunits. The average values of the intersubunit twist angles over the 20 structures in the ensembles are 37° ± 1° and 45° ± 2° for the DSI and DSA adducts, respectively. The representative structure selected is that closest to the geometric mean. The DNA backbone is cyan, the DNA bases are salmon, and the DSI and DSA ligands are white with oxygen atoms red and nitrogen atoms blue.

Evidence for an ordered water molecule in the structure of a (+)-CC1065 DNA adduct has led to a related hypothesis, which invokes water relayed general acid catalysis of the alkylation reaction.[19] Although the structural data reported here can neither confirm nor refute this hypothesis, we note a significant body of evidence has accumulated, which shows that the DNA alkylation is not an acid-catalyzed reaction and would not require such an activation event.[6,20−21]

Implications for the Binding-Induced Activation Hypothesis. The binding of DSA in the DNA minor groove cannot be

(18) Warpehoski, M. A.; Harper, D. E. *J. Am. Chem. Soc.* **1995**, *117*, 2951−2952.
(19) Lin, C. H.; Beale, J. M.; Hurley, L. H. *Biochemistry* **1991**, *30*, 3597−3602.
(20) Boger, D. L.; Garbaccio, R. M. *Bioorg. Med. Chem.* **1997**, *5*, 263−276.
(21) Boger, D. L.; Bollinger, B.; Hertzog, D. L.; Johnson, D. S.; Cai, H.; Mesini, P.; Garbaccio, R. M.; Jin, Q.; Kitos, P. A.; Goldberg, J. *J. Am. Chem. Soc.* **1997**, *119*, 4987−4998.

accommodated without some perturbation of the structure of either the DNA or the ligand because the length of the ligand exceeds the length of the minor groove that is accessible across one face of the DNA. The two subunits of DSA, which are coplanar in the X-ray crystal structure of the ligand alone, are twisted when bound to the DNA substrate (Figure 7) via changes in the χ_1, χ_2, and χ_3 angles (Figure 1). The shift in χ_1 from the coplanar arrangement reduces the resonance stabilization by the vinylogous amide at N12, thereby increasing the electrophilicity of the alkylating cyclopropyl ring, that is, activating the ligand for the alkylation reaction. Thus, as the cross-conjugated vinylogous amide is diminished as χ_1 increases, the cyclopropane π conjugation decreases and reactivity increases. To the extent that established structures of the bound product might reveal insights into the bound conformation of the reacting substrate or the reaction transition state itself, our studies directly probe this binding-induced activation hypothesis.

Previous studies have shown that the C24 indole methoxy group is essential for attaining the full alkylation activity of DSA, while the adjacent C22 and C23 methoxy groups have almost no effect on alkylation.[6] Elimination of the C24 methoxyl results in a 5−20-fold drop in alkylation efficiency and a similar 20-fold reduction in the rate of alkylation of the AATTA site, but there is no change in alkylation properties upon removal of both the C22 and C23 methoxyls. Examination of the DSA adduct revealed that only the C24 methoxyl is deeply imbedded in the minor groove, effectively extending the length of the ligand on the side facing the minor groove. The extent of binding-induced twist observed in the ligand is dependent on this effective length. Thus, removing the C24 methoxy group reduces the "effective" rigid length of the ligand, and less binding-induced twist is expected. In contrast, the removal of the C22 and C23 methoxy groups, which have little contact with the DNA and face out into solvent, does not affect the binding-induced twist of the ligand. In the structure of the DSI adduct, the removal of the methoxy groups does result in a smaller twist angle than that for the more reactive DSA. Assuming that the degree of twist can be extrapolated to the bound conformation of the starting ligand or the transition state itself, this suggests that DSI is less highly activated than DSA upon binding to DNA. This is fully consistent with the experimental observation that DSI has a slower rate of DNA alkylation and lower overall efficiency than DSA, while retaining the same site selectivity.

Consistent with the observation that the agents preferentially alkylate AT-rich sequences, the binding-induced twist is largest within the narrower, deeper AT-rich versus the wider, shallower GC-rich minor groove. Consequently, the dependence of activation of the alkylation reaction on the shape of the minor groove at the binding site and its enforced conformational change on the bound ligand is a process we have come to refer to as shape-dependent catalysis. The binding-induced (in situ) activation of DSA/CC-1065 agents demonstrates an exquisite level of fine-tuning and suggests a mechanism that we believe can be incorporated into a range of molecular designs.

Conclusions

The concept of in situ activation is a potentially powerful means of delivering a therapeutic agent to a target site in a form that has reduced reactivity until it binds to its substrate. There is a growing body of evidence indicating that duocarmycin SA, CC-1065, and related agents fall into this category; these molecules exhibit an exceptional increase in alkylation reactivity upon binding to their DNA substrate. The change in ligand conformation runs counter to the common belief that DNA is a flexible molecule which usually alters *its* conformation to accommodate the binding of small ligands or macromolecules such as DNA-binding proteins.

The results from our study also run counter to the proposal of Hurley and co-workers that DNA alkylation is controlled by DNA bonding and not by noncovalent DNA binding.[22] They maintain that the alkylating subunit alone contains sufficient structural information to encode for sequence selectivity. According to their hypothesis, alkylation site selectivity arises from sequence-dependent activation of the ligand via protonation of the C6 carbonyl mediated by a DNA backbone phosphate and/or via unique sequence-specific features of the DNA structure. The first factor, DNA phosphate backbone-mediated acid catalysis, has been discussed above and cannot be operative since the DNA alkylation reaction is not acid-catalyzed. The second factor, specific features of DNA structure and dynamics, has not yet been completely defined nor systematically investigated.

The implication in the Hurley-Warphoski hypothesis is that the DNA sequence affects the ability to attain the requisite geometry of the transition state, which determines alkylation kinetics. Similarities between the extent of bending of the DNA duplex upon alkylation and DNA bending associated with tracts of three or more consecutive adenine residues have been noted.[16] The hypothesis predicts that any DNA sequence that lacks such flexibility and/or the ability to position an ordered water molecule between the phosphate backbone on the nonalkylated strand and the C6 carbonyl will show low reactivity. However, there is no direct evidence establishing that DNA bonding determines site selectivity, other than DNA alkylation studies for CC-1065 derivatives composed only of the alkylation subunit, which remain controversial. In contrast, there is an extensive series of studies that establish the importance of noncovalent interactions in determining DNA alkylation selectivity and reactivity, all of which are fully consistent with a DNA binding-induced activation mechanism.[4] The new results reported here provide a structural basis for the 20-fold difference in reactivity for an analogue that alters noncovalent binding interactions.

The key element in this in situ activation mechanism for DNA alkylation by these compounds is the binding-induced release of the highly reactive conjugated spirocyclopropylcyclohexadienone that is stabilized by the cross-conjugated vinylogous amide at N12 when the ligand is not bound to DNA. It is important to note that the DNA binding-induced increase in alkylation reactivity far exceeds the effect associated with the increased local concentration of reactants upon binding. By determining the structure of the DSI adduct and comparing it to the structure of the corresponding DSA adduct, we have obtained direct evidence to support the proposal that binding-induced twist between the two alkylation subunits correlates with the rate of DNA alkylation. Although appreciated in studies of enzyme catalysis, the concept of using binding-induced conformational change in DSA/CC-1065 agents to trigger a large increase in reactivity is unique among DNA alkylators known to date[18] but may represent a mode of DNA alkylation catalysis that is more widespread than previously recognized.

Acknowledgment. We thank Jarrod Smith for assistance in setting up the structure calculations and preparing figures, Robert Garbaccio for many insightful discussions, and a referee for

(22) Hurley, L. H.; Lee, C.-S.; McGovern, J. P.; Warpehoski, M. A.; Mitchell, M. A.; Kelly, R. C.; Aristoff, P. A. *Biochemistry* **1988**, *27*, 3886−3892.

helpful suggestions. Support for work in the laboratory of D.L.B. was obtained from the Skaggs Institute for Chemical Biology and the National Institutes of Health (RO1 CA41986). W.J.C. is a Faculty Research Fellow of the American Cancer Society (FRA-436).

Supporting Information Available: A figure showing the chemical shift differences between the DSI and DSA adducts plotted for selected sugar and base protons of the AATTA duplex, a figure showing the variation in the backbone and pseudorotation angles of the DSI adduct, a figure showing the variation in the helicoidal parameters of the DSI adduct, and a table of assigned proton chemical shifts in the DSI adduct. This material is available free of charge via the Internet at http://pubs.acs.org.

JA983556J