Sequence Specificity of DNA Alkylation by the Unnatural Enantiomer of CC-1065 and Its Synthetic Analogues

Laurence H. Hurley,*[1a] Martha A. Warpehoski,*[1b] Chong-Soon Lee,[1a] J. Patrick McGovren,[1b] Terrence A. Scahill,[1b] Robert C. Kelly,[1b] Mark A. Mitchell,[1b] Nancy A. Wicnienski,[1b] Ilse Gebhard,[1b] Paul D. Johnson,[1b] and V. Susan Bradford[1b]

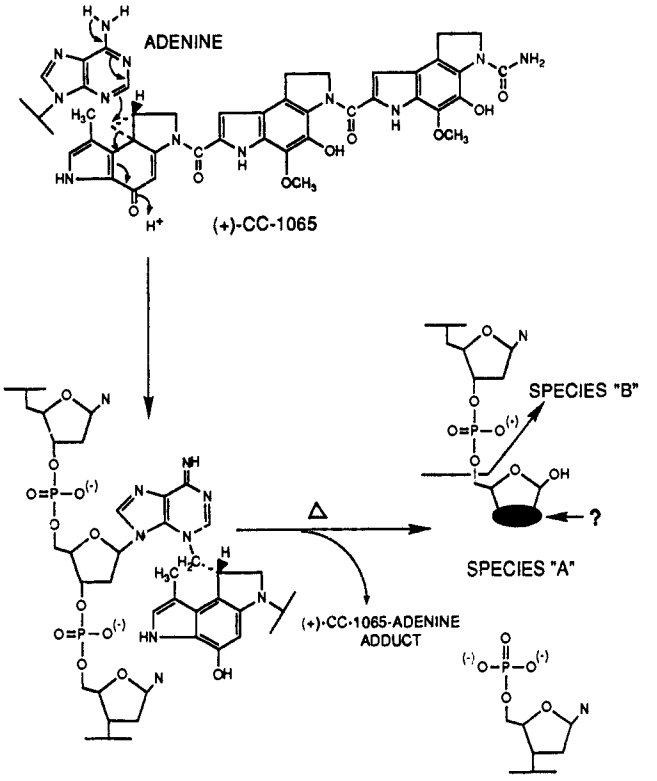
Contribution from the Drug Dynamics Institute, College of Pharmacy, The University of Texas at Austin, Austin, Texas 78712, and Research Laboratories, The Upjohn Company, Kalamazoo, Michigan 49001. *Received December 1, 1989*

Abstract: (−)-CC-1065, the unnatural enantiomer of the potent and sequence-selective, DNA-reactive antibiotic, (+)-CC-1065, was prepared by synthesis and its covalent reaction with DNA was studied and compared to that of the natural product. Although (−)-CC-1065 also formed covalent adducts in which the cyclopropyl carbon was bonded to the N3 atom of adenine, and the thermal strand breakage that it produced paralleled that seen for (+)-CC-1065, it lay in the opposite direction along the minor groove and exhibited a markedly different sequence requirement for the covalently modified adenine. While (−)-CC-1065 and its full carbon framework analogue, (−)-AB′C′, reacted readily at adenines near to, but generally distinct from, (+)-CC-1065-reactive adenines and exhibited potent cytotoxicity, their simpler analogues did not alkylate DNA under the conditions employed and were biologically nonpotent. At relatively high concentrations, the smallest such analogue, (−)-A, reacted detectably only at the same sites selected by (+)-CC-1065. An analysis of the reactivity patterns of (+)- and (−)-CC-1065 and their analogues with DNA restriction fragments supported the conclusion that the mode of sequence recognition for (−)-CC-1065 adduct formation is fundamentally different from that of (+)-CC-1065 and is primarily controlled by specific minor groove, AT-selective binding interactions, rather than by sequence requirements of the covalent step, as occurs for (+)-CC-1065 and the (+)-CPI analogues. Models are proposed comparing the interactions of the enantiomeric alkylating moieties variously oriented in the minor groove at potential reaction sites. The evolutionary significance of both the alkylating moiety and the minor groove binding segments of the natural product is discussed.

The concept of DNA as a target of drug action has undergone considerable refinement in recent years, as information accrues on the range of structural and dynamic diversity available to this critical biomacromolecule.[2] In particular, drugs and carcinogens that react with nucleic acid bases in DNA to form covalent adducts have been found to exhibit sometimes remarkable discrimination toward the sequence contexts of the bases undergoing modification.[3-7] The mechanisms mediating this recognition of specific sequences of DNA by small molecules may include sequence-dependent molecular electrostatic potential,[4a] conformational flexibility,[5b] catalytic functional group juxtapositions,[6] and pre-covalent binding affinities.[7]

Our previous work with CC-1065, an extremely potent antitumor antibiotic produced by *Streptomyces zelensis*,[8,9] revealed a DNA alkylating agent extraordinary both for the base-heteroatom specificity and the sequence selectivity of its covalent reaction with double-helical DNA.[10] The CC-1065 molecule consists of three repeating pyrroloindole subunits, one of which (the CPI, or cyclopropylpyrroloindole subunit) contains a potentially reactive cyclopropyl function.[11] Naturally occurring

Scheme I. Reaction of (+)-CC-1065 with N3 of Adenine in DNA and Products from Thermal (Δ) and Piperidine (Pip) Cleavage Reactions (ref 10)

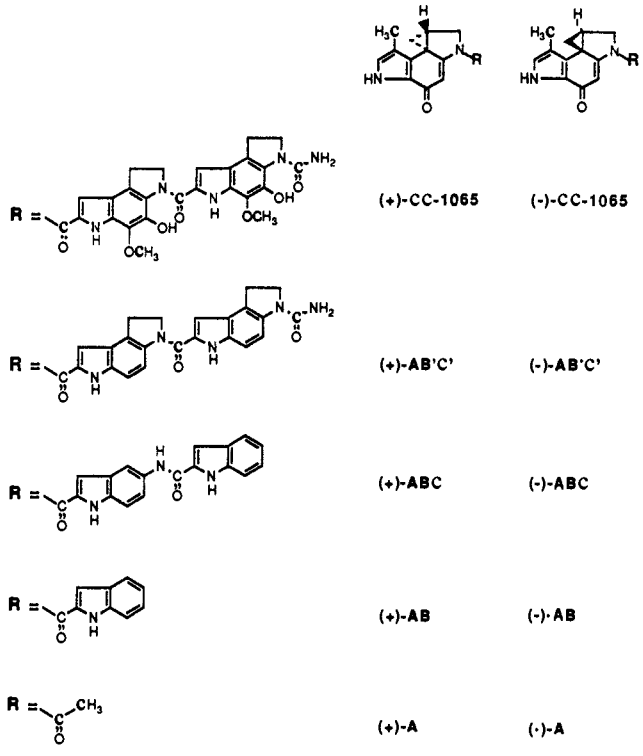


CC-1065 has the 7b*R*,8a*S* stereochemical configuration of the cyclopropane ring[12] and is dextrarotatory at the sodium D line.[13] (+)-CC-1065 binds strongly and reversibly in the minor groove of A–T regions in double-stranded DNA or duplex oligomers[14]

(1) (a) University of Texas. (b) The Upjohn Co.
(2) (a) Dickerson, R. E. *J. Mol. Biol.* **1983**, *166*, 419. (b) Calladine, C. R. *J. Mol. Biol.* **1982**, *161*, 343. (c) Dickerson, R. E.; Kopka, M. L.; Pjura, P. E. In *DNA–Ligand Interactions. From Drugs to Proteins*; Guschlbauer, W., Saenger, W., Eds.; Plenum Press: New York, 1987; pp 45–62.
(3) Warpehoski, M. A.; Hurley, L. H. *Chem. Res. Toxicol.* **1988**, *1*, 315.
(4) (a) Kohn, K. W.; Hartley, J. A.; Mattes, W. B. *Nucleic Acids Res.* **1987**, *15*, 10531. (b) Benasutti, M.; Ejadi, S.; Whitlow, M. D.; Loechler, E. L. *Biochemistry* **1988**, *27*, 472.
(5) (a) Hertzberg, R. P.; Hecht, S. M.; Reynolds, V. L.; Molineux, I. J.; Hurley, L. H. *Biochemistry* **1986**, *25*, 1249. (b) Zakrzewska, K.; Pullman, B. *J. Biomol. Struct. Dyn.* **1986**, *4*, 127.
(6) Buckley, N. *J. Am. Chem. Soc.* **1987**, *109*, 7918.
(7) Baker, B. F.; Dervan, P. B. *J. Am. Chem. Soc.* **1989**, *111*, 2700.
(8) (a) Hanka, L. J.; Dietz, A.; Gerpheide, S. A.; Kuentzel, S. L.; Martin, D. G. *J. Antibiot.* **1978**, *31*, 1211. (b) Martin, D. G.; Chidester, C. G.; Duchamp, D. J.; Mizsak, S. A. *J. Antibiot.* **1980**, *33*, 902.
(9) Reviews: (a) Reynolds, V. L.; McGovren, J. P.; Hurley, L. H. *J. Antibiot.* **1986**, *39*, 319. (b) Hurley, L. H.; Needham-VanDevanter, D. R. *Acc. Chem. Res.* **1986**, *19*, 230.
(10) (a) Hurley, L. H.; Reynolds, V. L.; Swenson, D. H.; Petzold, G. L.; Scahill, T. A. *Science* **1984**, *226*, 843. (b) Reynolds, V. L.; Molineux, I. J.; Kaplan, D. J.; Swenson, D. H.; Hurley, L. H. *Biochemistry* **1985**, *24*, 6228.
(11) Chidester, C. G.; Krueger, W. C.; Mizsak, S. A.; Duchamp, D. J.; Martin, D. G. *J. Am. Chem. Soc.* **1981**, *103*, 7629.
(12) Martin, D. G.; Kelly, R. C.; Watt, W.; Wicnienski, N.; Mizsak, S. A.; Nielsen, J. W.; Prairie, M. D. *J. Org. Chem.* **1988**, *53*, 4610.
(13) Kelly, R. C.; Gebhard, I.; Wicnienski, N.; Aristoff, P. A.; Johnson, P. D.; Martin, D. G. *J. Am. Chem. Soc.* **1987**, *109*, 6837.

0002-7863/90/1512-4633$02.50/0

Chart I. Structures of (+)- and (-)-Cyclopropylpyrroloindole (CPI) Compounds Used in This Study



and then bonds covalently to the N3 of adenine in target sequences, with opening of the cyclopropyl ring (Scheme I).[10] The reacted molecule lies snugly within the minor groove covering a four base pair region to the 5'-side of the covalently modified adenine.[10,15] Upon thermal treatment of (+)-CC-1065-(N3-adenine)-DNA adducts, cleavage of the N-glycosidic linkage and subsequent backbone breakage occurs to the 3'-side of the covalently modified adenine to leave a 5'-phosphate on the 3'-side of the break and, we presume, a modified deoxyribose on the 5'-side.[10a] Using a heat-induced strand breakage assay based on this observation, we determined that the most reactive adenines in a set of different DNA fragments were generally found in two sequences, 5'AAAAA* or 5'PuNTTA*, where * indicates the covalently modified adenine and N indicates any of the four bases in DNA.[10b] The construction of a site-directed (+)-CC-1065-(N3-adenine)-DNA adduct in a 117 base pair fragment[16] allowed us to determine the effect of covalent bonding on local DNA structure. Analysis of DNase I footprinting experiments and restriction enzyme cutting patterns using the DNA fragment containing the site-directed adduct demonstrated that (+)-CC-1065 adduct formation correlates with an asymmetric effect on DNA structure that extends more than one helix turn to the 5'-side of the covalent bonding site.[17]

Despite this novel mechanism of action of (+)-CC-1065, and its highly potent in vitro[18] and in vivo[19] antitumor activity, clinical development of the natural product was precluded by an unusual toxicity, which led to delayed death in mice receiving therapeutic doses.[20] Synthetic analogues have been prepared that are more efficacious, do not cause delayed death, and that, therefore, show considerable promise as chemotherapeutic agents.[21] Chirally resolved synthetic analogues are shown in Chart I.

Like (+)-CC-1065, analogues of the 7b*R*,8a*S* cyclopropyl configuration [(+)-CPI agents] react with double-helical DNA to produce adducts that, on thermal treatment, cause strand breaks at alkylated adenines.[22] We recently described the unexpected finding that an acetyl-substituted (+)-CPI [(+)-A, Chart I], which shows little or no evidence of the noncovalent binding to DNA so characteristic of (+)-CC-1065, nevertheless has sufficient structural information to mediate the sequence specificity of the entire drug molecule, i.e., (+)-A preferentially alkylates the same sequences in DNA as does (+)-CC-1065.[23] This suggests that the sequence specificity of (+)-CC-1065 depends, not on the relative noncovalent binding affinity of the different sequences as originally thought,[10] but on sequence-dependent site reactivity (i.e., upon covalent bonding dynamics). It was suggested that the prime function of the middle and right-hand segments in the (+)-CPI series is to increase the equilibrium constant for nonspecific reversible interaction with the minor groove of DNA where alkylation can occur, accounting for the differences in absolute reactivity among the analogues. Several sequences in an SV40 fragment that reacted with (+)-CC-1065 and (+)-AB'C' failed to be alkylated by (+)-ABC, indicating that specific reversible interactions between the inside edge of the B and C subunits and the floor of the minor groove of DNA can "modulate or fine tune" the sequence specificity of the (+)-CPI subunit.[23]

One remarkable feature of this strand breakage assay as applied to the (+)-CC-1065 analogues was that the absolute reactivity, gauged by the concentration of drug required to produce detectable cleavage under given incubation conditions, correlated closely with biological potency (in vitro cytotoxicity and in vivo toxic and therapeutic doses).[23] We have previously reported the preparation of chirally resolved synthetic CC-1065 analogues (+)-ABC and (-)-ABC, along with the interesting observation that, while (+)-ABC was a highly potent and efficacious antitumor agent, (-)-ABC was essentially devoid of biological activity.[21a] We have also disclosed the first synthesis of the unnatural enantiomer of CC-1065, namely (-)-CC-1065, and the surprising observation [in view of the inactivity of (-)-ABC] of its extremely potent cytotoxicity.[13] Other investigators have recently confirmed the potent cytotoxicity of synthetic (-)-CC-1065 and, further, have reported that the didehydroxy, didemethoxy compound (-)-AB'C' is likewise a potent cytotoxin.[24]

We felt that these observations warranted a systematic study of the interactions of (-)-CC-1065 and its (-)-CPI analogues with DNA. The object of the present study was to experimentally determine [by methods previously applied to (+)-CC-1065] the following: (1) the base specificity and structure of the (-)-CC-1065-DNA adduct(s); (2) the nature of the products of DNA strand cleavage resulting from these adducts; (3) the orientation of covalently bonded (-)-CC-1065 in the DNA minor groove; (4) the sequence specificity of (-)-CC-1065 adduct formation; and (5) the influence of the middle and right-hand subunits on the reactivity and sequence selectivity of (-)-CPI alkylation of DNA.[25] Full experimental details for the preparation of (+)- and (-)-CC-1065 and their analogues, and pertinent biological data for

(14) (a) Swenson, D. H.; Li, L. H.; Hurley, L. H.; Rokem, J. S.; Petzold, G. L.; Dayton, B. D.; Wallace, T. L.; Lin, A. H.; Krueger, W. C. *Cancer Res.* **1982**, *42*, 2821. (b) Krueger, W. C.; Li, L. H.; Moscowitz, A.; Prairie, M. D.; Petzold, G. L.; Swenson, D. H. *Biopolymers* **1985**, *24*, 1549.

(15) Scahill, T. A.; Jensen, R. M.; Swenson, D. H.; Hatzenbuhler, N. T.; Petzold, G. L.; Wiereng, A. W.; Brahme, N. M. *Biochemistry* **1990**, *29*, 2852.

(16) Needham-VanDevanter, D. R.; Hurley, L. H. *Biochemistry* **1986**, *25*, 8430.

(17) Hurley, L. H.; Needham-VanDevanter, D. R.; Lee, C.-S. *Proc. Natl. Acad Sci. U.S.A.* **1987**, *84*, 6412.

(18) (a) Li, L. H.; Swenson, D. H.; Schpok, S.; Kuentzel, S. L.; Dayton, B. D.; Krueger, W. C. *Cancer Res.* **1982**, *42*, 999. (b) Bhuyan, B. K.; Newell, K. A.; Crampton, S. L.; Von Hoff, D. D. *Cancer Res.* **1982**, *42*, 3532.

(19) Martin, D. G.; Biles, C.; Gerpheide, S. A.; Hanka, L. J.; Krueger, W. C.; McGovren, J. P.; Mizsak, S. A.; Neil, G. L.; Stewart, J. C.; Visser, J. *J. Antibiot.* **1981**, *34*, 1119.

(20) McGovren, J. P.; Clarke, G. L.; Pratt, E. A.; DeKoning, T. F. *J. Antibiot.* **1984**, *37*, 63.

(21) (a) Warpehoski, M. A. *Tetrahedron Lett.* **1986**, *27*, 4103. (b) Warpehoski, M. A.; Gebhard, I.; Kelly, R. C.; Krueger, W. C.; Li, L. H.; McGovren, J. P.; Prairie, M. D.; Wicnienski, N.; Wierenga, W. *J. Med. Chem.* **1988**, *31*, 590. (c) DeKoning, T. F.; Kelly, R. C.; Wallace, T. L.; Li, L. H. *Proc. Am. Assoc. Cancer Res.* **1989**, *30*, 491.

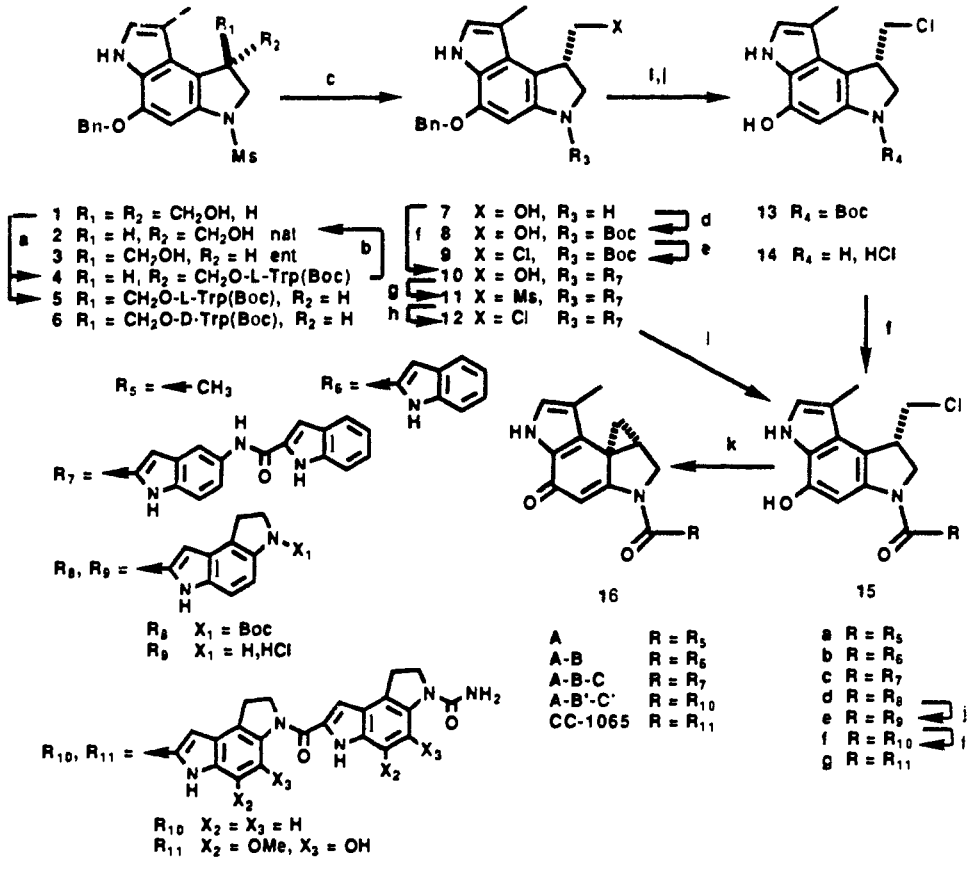
(22) Swenson, D. H.; Petzold, G. L. Unpublished results.

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

(24) Boger, D. L.; Coleman, R. S. *J. Am. Chem. Soc.* **1988**, *110*, 4796.

(25) Some of these results have been presented in preliminary form: McGovren, J. P.; Lee, C.-S.; Hurley, L. H.; Li, L. H.; Kelly, R. C. *Proc. Am. Assoc. Cancer Res.* **1987**, *28*, 262.

Scheme II. Synthetic Routes to Compounds Used in This Study



(a) L-Trp(Boc), EDC, DMAP; (b) NaOH; (c) Red-Al; (d) Boc-ON; (e) 1) MsCl, pyr; 2) LiCl, DMF; (f) EDC, RCOOH; (g) MsCl, pyr; (h) LiCl, DMF; (i) $NH_4^+HCO_2^-$, Pd/C; (j) HCl (gas); (k) Et_3N, H_2O, CH_3CN.

these compounds, are provided.

Results

Synthesis. The starting material (**1**, Scheme II) for the synthesis has been described previously in a full paper,[21b] and the details of the conversions to some of the final products have been disclosed in preliminary reports[13,21a] as has the resolution of **1** to its natural and unnatural enantiomers, **2** and **3**, respectively.[21a]

In our previously described route to CC-1065 analogues through mesylate (e.g., **11**) without the intermediacy of the chloride,[21] problems were encountered in the debenzylation reaction. Attempted hydrogenolysis of the benzyl ether under mild conditions often produced no reaction and under more forcing conditions resulted in numerous side reactions, including hydrogenolysis of the mesylate. Alternatively, cleavage of the benzyl ether with in situ generated trimethylsilyl iodide[26] proceeded more dependably but always generated a multitude of unknown byproducts in addition to the desired spirocyclopropyl compounds. These difficulties in the benzyl ether cleavage reaction were avoided by conversion of the mesylate to the chloride (e.g., **11** to **12**). The chlorides in general are both more soluble and more stable and these properties combine to allow phase-transfer hydrogenolysis of the benzyl ether to proceed in reproducible, high yield. The stability of the chlorides carried through to the next step and they gave no or only partial closure in the presence of a tertiary amine in anhydrous solvents, conditions that readily cyclized the mesylates. Stronger bases such as NaH, KO-*t*-Bu, or NaOMe gave substantial decomposition when attempted in the cyclization. Happily, it was found that large excesses of tertiary amine in highly aqueous solvents produce rapid, reproducible, and high-yielding spirocyclization.

The above procedure readily produced a wide range of analogues in excellent yield (40–60% from **2** and **3**) but suffered from two major shortcomings. It required five steps for each analogue (e.g., **2** → **7** → **10** → **11** → **15** → ABC) and it did not work with all the desired B and C subunits, including the PDE-I dimer of CC-1065. The failure with PDE-I dimer was principally due to the dehydration of the urea group during the mesylation or clorination reaction. Both of these problems were avoided by first protecting **7** as its Boc derivative **8** and carrying this through the mesylation, chlorination, and debenzylation chemistry to **13**. The stability and improved solubility of **8** and subsequent intermediates resulted in improved yields and ease of handling in all those steps. From **13**, each analogue required only removal of the Boc and coupling with the desired side-chain carboxylic acid, all of which was done in one pot, followed by the spirocyclopropyl ring closure. The overall yield from **12** for these conversions ranged from 60 to 85% with the lower yield being for CC-1065 itself and its enantiomer. In this manner, all the natural and enantiomeric materials used in this study were prepared with the exception of (+)-ABC, which was synthesized by the five-step procedure through **12**. The (+)-CC-1065 used in the experiments with DNA was isolated as described previously.[8,19]

Scheme III summarizes our route to the dihydropyrroloindoles **23** and **24** used in the synthesis of (+)- and (−)-AB'C' as previously reported.[27,28]

(26) Olah, G. A.; Narang, S. C.; Gupta, G. B.; Malhorta, R. *J. Org. Chem.* **1979**, *44*, 1247.

(27) Warpehoski, M. A.; Bradford, V. S. *Tetrahedron Lett.* **1986**, *27*, 2735.
(28) Warpehoski, M. A.; Bradford, V. S. *Tetrahedron Lett.* **1988**, *29*, 131.
(29) The consensus sequence is a composite sequence derived by determining which of the nucleotide bases or base combinations occur with the highest frequency at a given position relative to the alkylated adenine.

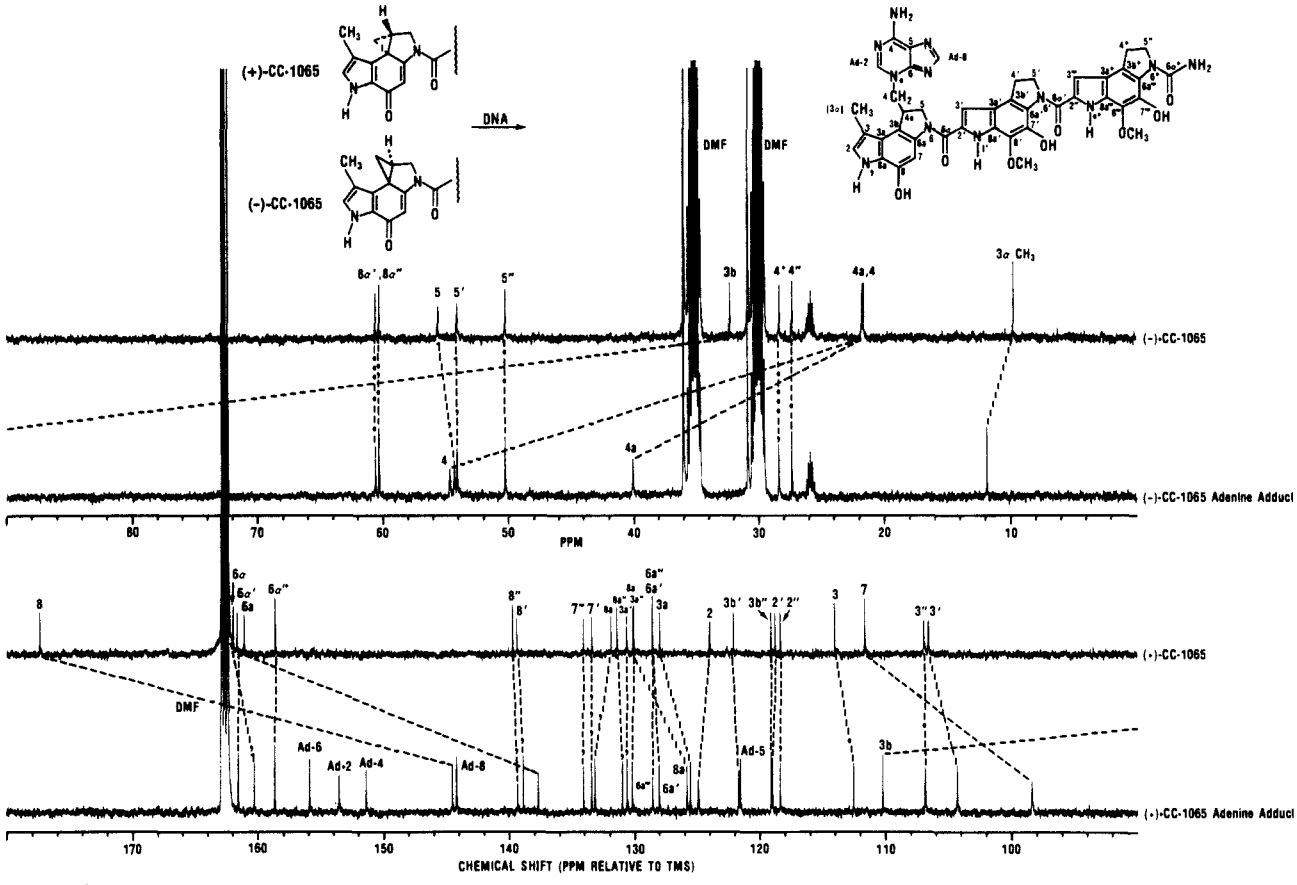
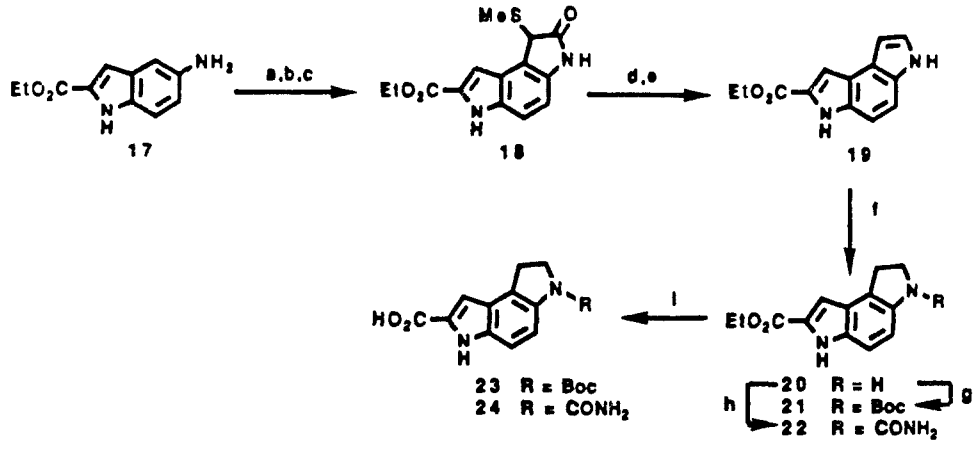


Figure 1. ^{13}C NMR of (−)-CC-1065 (upper spectrum) and the (−)-CC-1065–N3-adenine adduct (lower spectrum) (ppm relative to TMS).

Scheme III. Synthetic Route to B′ and C′ Segments



(a) $CH_3SCH_2CO_2Et$, SO_2Cl_2, ProtonSponge, CH_2Cl_2, -78°; (b) Et_3N; (c) HOAc; (d) Lawesson's reagent; (e) Raney nickel; (f) BH_3·SMe_2, TFA, THF, 0°; (g) Boc-ON; (h) KOCN, HOAc; (i) LiOH

DNA Base Specificity of (−)-CC-1065 Alkylation and Structure of the Adduct. As with (+)-CC-1065, when (−)-CC-1065–DNA adducts were heated at 100 °C for 30 min, a molecule with the CC-1065-like chromophore was extractable into butanol. Each of the bases in L1210 mouse leukemia cell DNA was labeled separately by ^3H-precursor incorporation and then the isolated DNA was reacted with either (+)- or (−)-CC-1065, subjected to thermal treatment, and butanol-extracted. The DNA that released the predominant amount of tritium had been prelabeled with

Table I. Reaction of (+)-CC-1065 and (−)-CC-1065 with DNA Prelabeled in Specific Bases and Extraction of Labeled Adducts[a]

labeled base	total dpm added	dpm recovered in hot butanol extract		
		no drug	(+)-CC-1065	(−)-CC-1065
A	165 000	31	2781	6891
T	153 000	18	0	0
G	34 000	0	11	126
C	146 000	11	0	3

[a] In preliminary experiments, ca. 85% of DNA-adducted (+)-CC-1065 and (−)-CC-1065 absorbance equivalents were extracted into butanol after heating at 100°, 45 min.

(30) For most DNA fragments, under the conditions of DNA–drug incubation used, cleavage sites appearing at low drug concentrations remained visible at higher drug concentration, with additional cleavage sites appearing. An exception was the (−) strand of the 117 base pair M13mp1 DNA fragment in its reaction with (+)-CC-1065. At 2.8–28 nM (+)-CC-1065, three cleavage sites were observed (see Chart II, bottom strand). At 280 nM and higher concentrations of (+)-CC-1065, these cleavage sites disappeared and instead there appeared cleavage sites 25–30 base pairs to the 3′-side of the initially alkylated adenines, suggesting that the initially alkylated fragments were hyperreactive to subsequent drug-induced cleavage at the distant sites.

[^{3}H]adenine (Table I). A very small amount of excess tritium was also released from the sample labeled in the guanine nucleus, which, in the case of (−)-CC-1065, was 2% of the amount of the radioactivity released from the [^{3}H]adenine-labeled DNA. [The comparable figure for (+)-CC-1065 was 0.5%.] The significance of these results is unknown.

In previous work with (+)-CC-1065 bound covalently to calf thymus DNA, a product was purified from a hot butanol extract and, on the basis of ^{1}H and ^{13}C NMR, was assigned the CC-1065-(N3-adenine) adduct structure (Scheme I).[10a] A similar experiment with (−)-CC-1065 gave a product that eluted identically upon reversed-phase HPLC with that previously found with (+)-CC-1065 (data not presented). The ^{13}C NMR spectra of the purified product obtained from (−)-CC-1065 (Figure 1) was in all respects identical with that obtained from the analogous product with (+)-CC-1065. Five adenine signals in the aromatic region (121.5–155.9 ppm) correspond closely to the chemical shift values for 3-methyladenine.[10a] The resonance signals for C4 and C4a are downfield-shifted in accord with a structure resulting from opening of the cyclopropyl ring. Arguments similar to that made previously distinguish between the assigned structure and the alternate one in which ring opening of the cyclopropane ring occurs at the more substituted carbon (C4a).[31] Chemical shift changes for C3b, C6a, C7, C8, and C8a in the CC-1065–(N3-adenine) adduct are in accord with conversion of the cyclohexadienone to a phenol. ^{1}H NMR spectra for (+)- and (−)-CC-1065 were also identical (data not presented).

Identical adduct structure but with opposite stereochemistry was strongly supported by mirror-image circular dichroism spectra of the (+)- and (−)-CC-1065 adducts with adenine (data not presented).

Thermal and Piperidine Treatment of CC-1065-Modified DNA Restriction Fragments. Electrophoretic analysis of 3′- and 5′-^{32}P-labeled DNA reacted with (+)-CC-1065 and then heat-treated has demonstrated that strand breakage occurs on the 3′-side of the covalently modified adenine. The products of this reaction appear to be a 5′-phosphate on the 3′-side of the break site and a modified sugar on the 5′-side. Consequently, thermal treatment of (+)-CC-1065–DNA adducts on 3′-labeled DNA leads directly to a product that migrates at a similar position to the Maxam and Gilbert adenine sequencing reaction product.[10b] However, with 5′-labeled DNA fragments, the heat cleavage product migrates at a higher molecular weight than an adenine sequencing reaction product due to the presumed presence of the modified deoxyribose moiety (species A in Scheme I). Upon subsequent piperidine treatment, a second reaction (perhaps a β-elimination) leads to loss of the modified sugar, resulting in a product (species B in Scheme I) which then comigrates with the Maxam and Gilbert adenine reaction.[10]

A comparison of thermal in aqueous buffer and thermal in piperidine treatments of (+)- and (−)-CC-1065-modified 5′- and 3′-labeled DNA fragments is shown in Figure 2. The same DNA strand of the 117 base pair fragment was either 3′- or 5′-^{32}P-labeled. Although (+)- and (−)-CC-1065 treatment led to strand breakage at different adenines, the appearance and behavior of the bands were identical for each enantiomer. Using 5′-labeled DNA, thermal cleavage led to pairs of closely migrating bands (lanes 1 and 3). We believe that the major band observed at higher molecular weight is the thermal cleavage product still containing the modified sugar, while the minor product is a small amount of the presumed second β-elimination product. Quantitative conversion of the higher molecular weight product to the minor product took place upon piperidine treatment (lanes 2 and 4). Using 3′-labeled DNA, thermal treatment led directly to a product comigrating with an adenine sequencing reaction fragment (lanes 5 and 7). Migration of this fragment was unchanged upon subsequent piperidine treatment (lanes 6 and 8). Thus, the chemistry of strand breakage of both (+)- and (−)-CC-1065–(N3-adenine)–DNA adducts seems to be purely a function of N3 alkylation and is independent of other variables such as the stereochemistry at the linkage site and orientation in the minor groove of DNA.

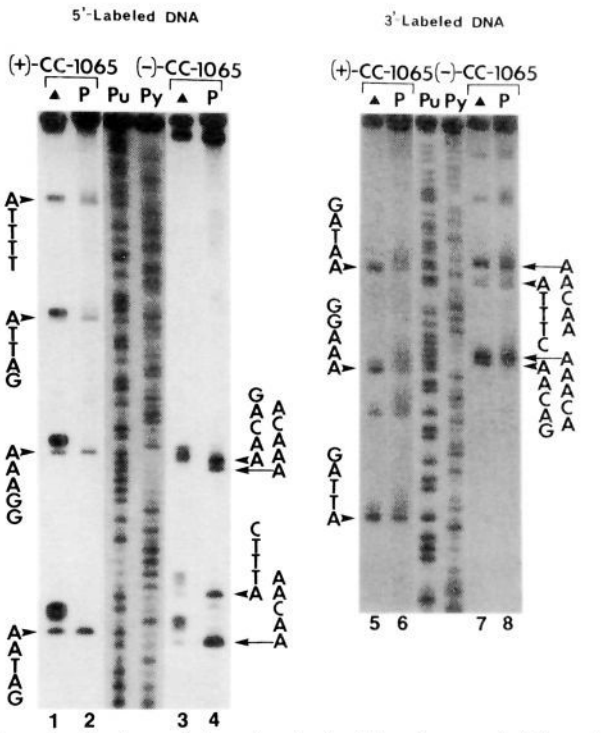


Figure 2. Products of thermal and piperidine cleavage of (+)- and (−)-CC-1065–DNA adducts on the 117-bp DNA fragment from M13mp1. DNA fragments were modified with 28 μM of either (+)- or (−)-CC-1065 subjected to either thermal treatment (▲) or thermal followed by piperidine treatment (P) as described. The center lanes are Maxam and Gilbert purine (Pu) and pyrimidine (Py) sequencing lanes. The arrow heads point to the covalently modified adenines, and the remainder of the 5-mer sequences are the drug overlap region. The larger amount of product in lane 1 vs lane 2 is due to a decreased recovery of radioactivity for the sample in lane 2 following the additional manipulations associated with piperidine treatment.

Orientation of (+)- and (−)-CC-1065 Molecules Following N3-Adenine Alkylation in DNA. Molecular modeling of the (+)-CC-1065-DNA adduct had led us to predict that the middle and right-hand subunits would be oriented to the 5′-side of the covalently modified adenine.[10] Footprinting of a site-directed adduct in a 117 base pair DNA fragment from M13mp1[16] and one-dimensional difference NOE experiments[15] were in accord with this prediction. The DNA alkylation sequence specificity of (+)-CC-1065 also lies to the 5′-side of the covalent adduct, although this is a less reliable indicator of orientation specificity.[10b,23] Modeling of the (−)-CC-1065-DNA adduct suggested that it would have the opposite orientation; i.e., lying to the 3′-side of the covalent adduct. Footprinting of the covalently bound (+)- and (−)-CC-1065 molecules on 5′- and 3′-labeled DNA fragments, respectively, using MPE·Fe(II) is shown in Figure 3. A comparison of control (untreated) DNA lanes (1 and 3) with (+)- and (−)-CC-1065–pretreated DNA lanes (2 and 4, respectively) shows four (+)-CC-1065 bonding sites and five (−)-CC-1065 bonding sites within the interpretable areas of the autoradiogram. In lane 4 the unidentified footprint in the center of the gel reflects drug covalently bound on the nonlabeled strand and consequently the results of this experiment do not reveal its orientation. In all other cases for both (+)- and (−)-CC-1065, each footprint site is associated with a strand breakage site at the high molecular weight end of the footprint. Since the footprinting experiments for (+)- and (−)-CC-1065 were carried out with 5′- and 3′-labeled DNA fragments, respectively, this result fixes the orientations of (+)- and (−)-CC-1065 to the 5′- and 3′-sides, respectively, of the covalent-bonding sites, in agreement with molecular modeling predictions.

(31) Martin, D. G.; Mizsak, S. A.; Krueger, W. C. *J. Antibiot.* **1985**, *38*, 746.

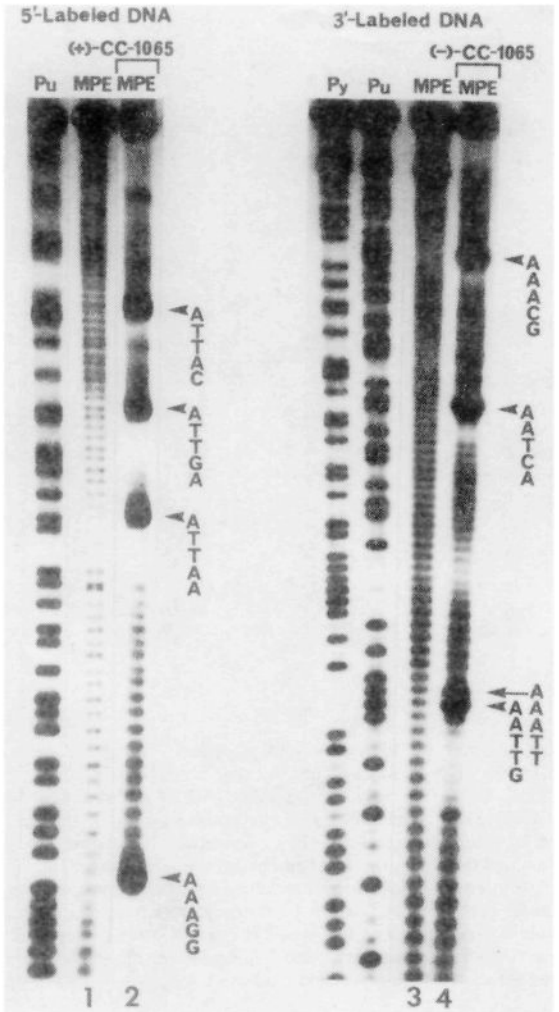


Figure 3. MPE·Fe(II) footprinting of (+)- and (−)-CC-1065 on a 5′-labeled 106-bp fragment from SV40 DNA [(+)-CC-1065] or on a 3′-labeled 117-bp DNA fragment from M13mp1 [(−)-CC-1065]. DNA fragments were modified with 28 μM of either (+)- or (−)-CC-1065 and subjected to MPE·Fe(II) footprinting as described in the Experimental Section. Samples were denatured by heat (90 °C, 2 min) in alkaline tracking dye (80% formamide and 10 mM NaOH) and then electrophoresed adjacent to Maxam–Gilbert purine (Pu) and pyrimidine (Py) sequencing lanes. Lanes 1 and 3 are control MPE·Fe(II) digestions and lanes 2 and 4 (+)- and (−)-CC-1065-treated samples, respectively. The arrow heads point to the covalently modified adenines, and the remainder of the 5-mer sequence is the drug overlap site. The pronounced bands at the 3′- and 5′-ends of the footprints in lanes 2 and 4, respectively, are due to heat-induced cleavage (90 °C for 2 min) at (+)- or (−)-CC-1065 alkylation sites. MPE does not detect noncovalent binding sites[23] except as described in ref 40.

Comparison of (+)- and (−)-CC-1065 Sequence Selectivity. The heat-induced strand breakage assay with (−)-CC-1065 was carried out on two DNA restriction fragments from SV-40 (one strand each of a 276 base pair fragment and a 118 base pair fragment), and three restriction fragments from M13mp1 (one strand of a 180 base pair fragment and both strands of a 117 base pair fragment). Exclusion of base pairs further from the end label than the largest cleavage fragment identified on each gel, and closer to the end label than the smallest identified cleavage fragment (i.e., experimentally unevaluable regions), leaves about 470 base pairs in these five restriction fragments that can be analyzed for sequence context of (−)-CC-1065 alkylation. This sampling of DNA is A–T rich, with 34% adenine, 32% thymine, and about 17% each of guanine and cytosine bases. The DNA was incubated with the drug at a range of concentrations. Figure 4 illustrates the gel results obtained with one of the DNA fragments. The appearance of each band on the polyacrylamide gel

Table II. Most Reactive (−)-CC-1065 Sequences[a]

				*					
5′	A	C	A	A	A	T	T	A	3′
	G	A	A	A	A	A	A	C	
	C	A	T	A	A	T	T	G	
	G	T	A	A	A	A	C	G	
	T	G	A	A	A	T	T	G	
	G	A	T	A	A	C	A	A	
	C	G	T	A	A	T	C	A	
	A	C	A	A	T	T	C	C	

[a] *, reactive (covalently modified) adenine.

Table III. Analysis of the DNA Sequences Flanking the Covalent Bonding Sites of (−)-CC-1065

		frequency of occurrence,[a] (%)							
		−3	−2	−1	A*	+1	+2	+3	+4
A	(34)[b]	34	26	53	100	70	43	30	34
T	(32)	30	42	42		28	32	36	32
G	(17)	23	17	4			11	11	17
C	(17)	13	15	2		2	13	23	17
A/T	(66)	64	68	94	100	98	75	66	66
consensus[c]		N	N	A/T	A*	A/T	N	N	N

[a] A* represents the adenine to which (−)-CC-1065 is covalently bonded. Columns to the left and right of the A* column represent bases to the 5′- and 3′-sides, respectively, of the alkylated adenine. The frequencies of occurrence of individual bases or base combinations in the region adjacent to A* are computed from a total of 53 sites on SV40 and M13 DNA, which were shown by thermal cleavage assay to be (−)-CC-1065 alkylation sites at 28 μM or lower drug concentrations. Individual sequences are shown in supplementary material. [b] Random frequency of occurrence of each base in the evaluable region (ca. 470 base pairs) of the five DNA fragments used in this study. [c] The consensus sequences were determined by χ^2 analysis ($\alpha = 0.0005$) using the raw frequency values observed.

Table IV. Frequency Analysis of (−)-CC-1065 Consensus Sequences[a]

triplet	no. in DNA	no. reacted $c < 28$ μM	no. reacted $c < 0.28$ μM	patterns correlated to lower reactivity
5′AA*A	23	19	14	5′AA*AG
5′TA*A	20	15	10	5′TA*AG; TA*AAG
5′AA*T	13	8	7	5′A/T AA*TT
5′TA*T	19	7	0	none

[a] A* identifies potentially reactive adenine.

indicates that a detectable level of heat-induced cleavage has occured at that adenine (Scheme I). Since relative band intensity is not sensitive to the heat-induced strand scission variables under the conditions employed, we infer that this intensity is proportional to the rate of alkylation of the adenine at which subsequent cleavage occurs. Under a given set of drug–DNA incubation conditions, the cleavage bands produced at the lowest concentration of a drug identify the most reactant adenines for that drug. Each reactive adenine, *when placed in the context of its consensus sequence* (vide infra) identifies a reactive *alkylation* site for the drug under consideration.

At 2.8–28 nM (−)-CC-1065, only eight heat-induced cleavage bands (and thus alkylation sites) were detected. These are listed in Table II. At 280 nM (−)-CC-1065, 24 additional alkylation sites are produced. These 32 reactive sites include 14 5′AA*A, 10 5′TA*A, and 7 5′AA*T sequences, and a single 5′AA*C sequence (A* refers to the covalently modified adenine). Thus, while the alkylated adenine can be flanked by one thymine in this set of sequences, either on the 5′- or 3′-side, there is always at leat one adjacent adenine. When the DNA is incubated with 2.8–28 μM (−)-CC-1065, 21 additional alkylation sites are generated, one-third of which now contain the 5′TA*T sequence not seen at lower drug concentrations. The composite tabulated base frequency for all 53 reaction sites is given in Table III, with the resulting consensus sequence[29] 5′A/T A* A/T.

Of the 53 alkylated sequences, 49 contain 5′AA*A, TA*A, AA*T, or TA*T. If the DNA sequences evaluable in this study

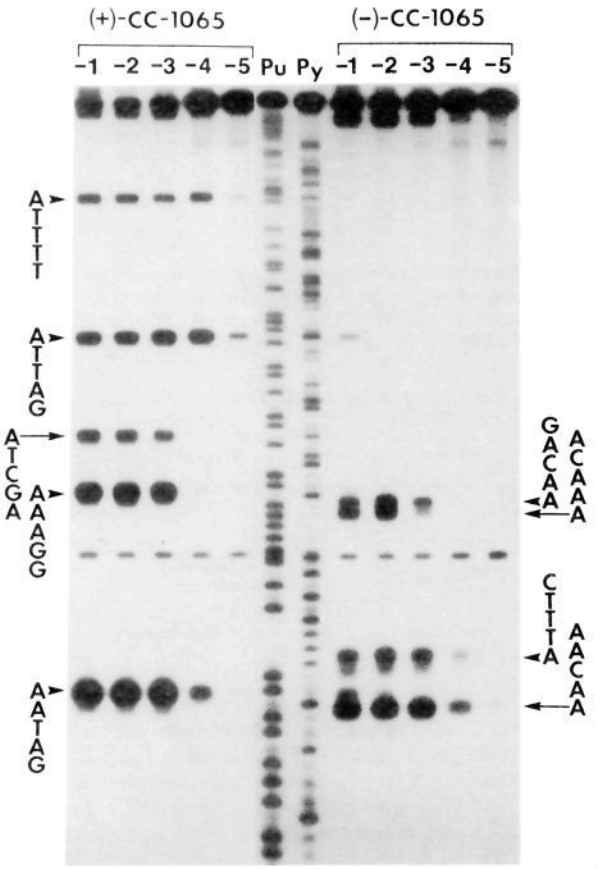


Figure 4. Concentration dependence of alkylation by (+)-CC-1065 and (−)-CC-1065 of the 117-bp *Msp*I–*Bst*NI fragment of M13mp1 DNA. Single 5′-^{32}P-labeled (+)-strand of the 117-bp DNA fragment was isolated and labeled as described in the Experimental Section. Fragments were modified with 2.8 nM (lanes labeled −5), 28 nM (−4), 280 nM (−3), 2.8 μM (−2), or 28 μM (−1) of either (+)-CC-1065 (left-hand lanes) or (−)-CC-1065 (right-hand lanes). Thermal cleavage was as described in the Experimental Section, and samples were electrophoresed adjacent to Maxam–Gilbert purine- and pyrimidine-specific DNA cleavage reactions. The drug alkylation sequences are shown and arrows indicate adenines modified by either (+)- or (−)-CC-1065.

are examined for these four triplets (supplementary material), it is found that they occur 75 times. Thus, about a third of the "consensus sequences" are not alkylated by 28 μM (−)-CC-1065. Slightly more than a third are alkylated by only 0.28 μM (−)-CC-1065. Table IV describes features of the reactivity of each of the four sets of consensus triplets. It can be seen that most of the sequences containing the first three triplets react to some extent, with the frequency of alkylation roughly in proportion to the frequency of occurrence. The question arises as to why the several consensus sequences in each category fail to alkylate. Several apparent patterns emerge from analysis of individual sequences. For example, 5′AAG sequences frequently are associated with weak or unreactive sites. In 5′NAATT sequences, the 3′A is reactive when N is C, but not generally when it is A or T, since then the 5′A is able to react instead. The 5′TAT sequence is clearly less reactive, since a much smaller percentage of TAT sequences are alkylated, and then only at higher drug concentrations.

Previous analysis of (+)-CC-1065 alkylation sites on restriction fragments from SV 40 and T7 DNA representing a somewhat larger sampling of base pairs had revealed that the sequences 5′PuNTTA* and 5′AAAAA* were most reactive (i.e., were cleaved at 1.4–0.14 mM (+)-CC-1065, the lowest drug concentrations leading to detectable cleavages).[10a] They also represented the consensus sequences for all cleavages produced by 140 nM or less (+)-CC-1065 in these DNA fragments under the conditions employed. These conditions differed from those of the present

Table V. Most Reactive (+)-CC-1065 Sequences[a]

						*			
5′	G	G	A	T	A	A	C	A	3′
	T	G	A	T	T	A	C	G	
	G	T	T	T	T	A	C	A	
	T	T	G	T	A	A	A	A	
	G	T	A	A	A	A	C	G	
	T	T	G	T	T	A	T	C	
	G	A	T	T	T	A	A	A	
	G	A	C	A	A	A	C	T	

[a] *, reactive (covalently modified) adenine.

Table VI. Analysis of the DNA Sequences Flanking the Covalent Bonding Sites of (+)-CC-1065

		frequency of occurrence,[a] %						
		−4	−3	−2	−1	A*	+1	+2
A	(34)[b]	27	36	31	44	100	36	42
T	(32)	45	27	58	55		27	24
G	(17)	18	29	4			9	24
C	(17)	9	7	7	2		27	11
A/T	(66)	73	64	89	98	100	64	65
consensus[c]		N	N	A/T	A/T	A	N	N

[a] A* represents the adenine to which (+)-CC-1065 is covalently bonded. Columns to the left and right of the A* column represent bases to the 5′- and 3′-sides, respectively, of the alkylated adenine. The frequencies of occurrence of individual bases or base combinations in the region adjacent to A* are computed from a total of 55 sites on SV40 and M13 DNA, which were shown by thermal cleavage assay to be (+)-CC-1065 alkylation sites at 28 μM or lower drug concentrations. Individual sequences are shown in supplementary material. [b] Random frequency of occurrence of each base in the evaluable region (ca. 470 base pairs) of the five DNA fragments used in this study. [c] The consensus sequences were determined by χ^2 analysis ($\alpha = 0.0005$) using the raw frequency values observed.

Table VII. Frequency Analysis of (+)-CC-1065 Consensus Sequences[a]

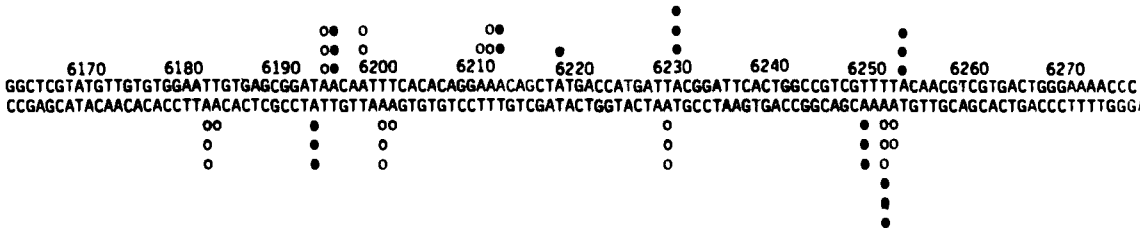
triplet	no. in DNA	no. reacted $c < 28$ μM	no. reacted $c < 0.28$ μM	patterns correlated to lower reactivity
5′TTA*	20	18	13	5′CTTA*
5′TAA*	20	14	13	5′TAA*T; 5′A/C NTAA*
5′AAA*	23	9	7	5′AAA*A/T/G
5′ATA*	14	7	2	5′ATA*A

[a] A* identifies potentially reactive adenines.

study in that the DNA–drug incubations were carried out for 24 h at 4 °C rather than for 2 h at 37 °C. To allow a more direct comparison, the alkylation sites for (+)-CC-1065 in the present set of DNA fragments, under conditions identical with those used for (−)-CC-1065, were determined. The eight most reactive (28–2.8 nM drug) sequences for (+)-CC-1065 are listed in Table V. At 280 nM (+)-CC-1065, 30 additional alkylation sites were detected.[30] Three triplets comprised the great majority of these 38 reactive sites: 5′TTA* (13 sites), TAA* (13 sites), and AAA* (7 sites). The high frequency of 5′TAA* in these experiments is a featue not found in earlier work[10b] and may possibly reflect the higher temperature of drug incubation used in the present study. At still higher (+)-CC-1065 concentrations, 17 additional alkylation sites are produced. Table VI tabulates the frequency of occurrence of the bases at each position surrounding these 55 alkylated adenines, giving the consensus sequence 5′A/T A/T A*.

Of these 55 alkylated sequences, 48 contain the triplets 5′AAA*, TTA*, TAA*, and ATA*. Examination of the evaluable regions of these DNA fragments shows these triplet sequences to occur 77 times. More than a third of these "consensus sequences" are not alkylated by 28 μM (+)-CC-1065. Table VII summarizes the reactivity of each of the four sets of consensus triplets. The 5′TTA* sequence appears to be intrinsically reactive: virtually all such sites are alkylated by (+)-CC-1065. Interestingly, the only two such sequences not found to be alkylated in this work were the only two 5′CTTA* sequences present in these DNA regions. In earlier sequencing work on (+)-CC-1065, however, at least one 5′CTTA* sequence was found to be highly reactive

Chart II. Diagrammatic Presentation of the (+)- and (-)-CC-1065 Bonding Sequences on the 117 Base Pair *MspI–BstNI* Fragment of M13mp1[a]

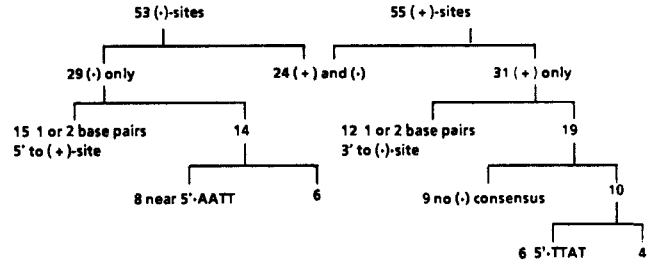


[a] The data are taken from concentration-dependency experiments in which 10-fold dilutions of a standard (+)- or (-)-CC-1065 stock solution (280 μM) were incubated for 2 h at 37 °C and then heated under standard conditions to produce DNA strand breakage. The O show the (-)-CC-1065 bonding sites and the ● show the (+)-CC-1065 bonding sites. The number of circles reflects the relative reactivity toward alkylation as indicated by the minimum concentration of drug producing detectable strand breakage at that site: ●●●, 2.8–28 nM; ●●, 280 nM; ●, 2.8–28 μM.

to alkylation.[10b] The subset of 5′TAA* sequences that are not alkylated by 28 μM (+)-CC-1065 include three of the four 5′TAA*T sequences present and, independently, five of the six 5′A/C NTAA* sequences present in this DNA. Only 40% of the 5′AAA* sequences available were alkylated by 28 μM (+)-CC-1065 under these conditions. All six such sequences containing a 3′C (5′AAA*C) were highly reactive to (+)-CC-1065. Any other base in this position was correlated with weak or no alkylation. In particular, of seven 5′AAA*A sequences present, only one was alkylated, and only at the higher drug concentrations. No correlation of reactivity of 5′AAA* with the identity of the bases lying to the 5′side was discerned. The 5′ATA* sequences showed limited reactivity, mainly at higher drug concentrations. While half of these sequences in the DNA examined contained a 3′A (5′ATA*A), only one such sequence was alkylated; in most cases the resulting 5′TAA* site is more reactive.

It is clear that (+)-CC-1065 and (-)-CC-1065 have distinct sequence requirements, even though both favor AT-rich DNA. The most striking difference is that (+)-CC-1065 strongly prefers two A/T base pairs in the 5′-direction from its adenine alkylation site, where (-)-CC-1065 requires A/T flanking on both sides of the reactive adenine. Comparison of these two drugs, not merely with respect to flanking sequence requirements, but with respect to the location of reaction sites along the DNA strands as illustrated in Figure 4 for one DNA fragment, leads to several intriguing observations. Chart II shows a 117 base pair fragment from M13mp1 on which are indicated for both strands the adenines alkylated by (+)- or (-)-CC-1065. In this DNA fragment, one adenine reacts with both enantiomers of CC-1065. It must be emphasized that the sequence context (i.e., consensus sequence) for (+)- and (-)-CC-1065 is different, and hence, while the specific adenine that is alkylated by (+)- or (-)-CC-1065 is the same, the alkylation *sites* themselves are not identical. (We use the term "alkylation site" to include both the alkylated base and its sequence context.) All of three favorable 5′TTA* (+)-CC-1065 sites are accompanied by a favorable 5′TA*A (-)-CC-1065 reaction site directly opposite on the complementary strand. Another common motif, occurring five times in this sampling, is the occurrence of a (-)-CC-1065 alkylated adenine one or two base pairs in the 5′-direction from a (+)-CC-1065 reactive adenine on the same strand. Since the footprinting experiments discussed above show that (+)-CC-1065 extends in the 5′-direction from its modified adenine, while (-)-CC-1065 extends in the 3′-direction from its covalent link, this indicates that alkylation with either enantiomer leads to occupation of the same or overlapping spaces in the minor groove in many instances. A similar analysis of the set of (+)- and (-)-CC-1065 reaction sites on all of the five restriction fragments used in the (-)-CC-1065 consensus studies (supplementary material) is summarized in Chart III. Of the 53 adenines that react with (-)-CC-1065, 24 can also react with (+)-CC-1065 [i.e., these adenines belong to both the set of (+)-CC-1065 alkylation sites and the set of (-)-CC-1065 alkylation sites]. Of the remaining adenines targeted by (-)-CC-1065, 15 occur one or two base pairs 5′ to a (+)-CC-1065 alkylated base. Of the 14 (-)-CC-1065 reactive adenines that do not appear to be near (+)-CC-1065 alkylation sites (in some cases it is possible that such reaction would occur on the complementary strand), 8 are in or

Chart III. Relative Locations of Adenines That Are Alkylated by (+)- or (-)-CC-1065



adjacent to 5′AATT sequences. Four of these are represented in the most reactive (-)-CC-1065 sites listed in Table II. In this sequence the (-)-CC-1065 molecule prefers to react at the 5′-adenine (5′A*ATT) whenever the 5′ flanking A/T requirement is honored. If not, it will react at the 3′-adenine (5′AA*TT).

Of the 55 (+)-CC-1065 alkylation sites identified in this DNA sample at 28 μM drug, nearly two-thirds occur at adenines that also react with (-)-CC-1065 or are one or two base pairs 3′ to adenines that react with (-)-CC-1065. Of the 19 exceptions, 9 do not contain the consensus sequence (A/T A* A/T) for (-)-CC-1065. Six of the remaining 10 are 5′TTA*T sequences, which are predicted to have reactive (-)-CC-1065 sites on the opposite strand, leaving only four (+)-CC-1065 sites that do not have an obvious reason for failing to react with (-)-CC-1065. One of these sites, 5′CTA*TG, reacts with (+)-CC-1065 only at relatively high concentrations. The other three sites contain 5′AAG sequences adjacent to, or incorporating, the target adenine.

In summary, within the constraints of their distinct consensus sequence requirements, (+)- and (-)-CC-1065 tend to react within the same A/T-rich sequences such that much of the same space in the minor groove is accessible to either enantiomer. Nearly half of the alkylation sites for each of the enantiomers permit covalent bond formation with the same adenines. Another large fraction of sequences contains a (-)-CC-1065 alkylated adenine one or two base pairs 5′ to an adenine targeted by (+)-CC-1065 on the same strand, or directly across on the complementary strand such that the consensus sequences overlap. The relatively small number of exceptions to this appear to fall mainly into three categories: (1) 5′AATT sequences, which are excellent (-)-CC-1065 alkylation sites, but which react poorly or not at all with (+)-CC-1065; (2) sequences that can accommodate (+)-CC-1065, but not the stringent 5′A/T A* A/T (-)-CC-1065 consensus; and (3) nominally consensus (-)-CC-1065 sequences in which the target adenine is within or directly 5′ to a 5′AAG sequence. While these are usually highly reactive sequences for (+)-CC-1065 when A/T base pairs lie in the 5′-direction, they are generally poorly reactive or unreactive with (-)-CC-1065. Indeed three of the five 5′A*AG cleavage sites that appear at the higher concentrations of (-)-CC-1065 are highly susceptible (+)-CC-1065 sites and could conceivably arise at least partly from traces of enantiomeric contamination.

Comparison of (+)- and (-)-CPI Analogues of CC-1065: A. Sequence Specificity and Reactivity of DNA Alkylation. The (+)- and (-)-CPI analogues listed in Chart I were subjected to the

Table VIII. Reactivity of (+)- and (−)-CC-1065 Alkylation Sites

agent	concn of earliest detectn, μM	sequences alkylated at A* at earliest detectn, 5′ to 3′
	(+)-CC-1065 Alkylation Sites	
(+)-ABC	0.0028	TTTTA*, GATTA*, GATAA*
(+)-CC-1065	0.0028	GATTA*
	0.028	TTTTA*, GATAA*
(+)-AB′C′	0.0028	GATTA*
	0.028	TTTTA*, GGAAA*, GATAA*
(+)-AB	2.8	GATTA*, GGAAA*, GATAA*
(+)-A	28	TTTTA*, GATTA*, GATAA*
(−)-AB′C′	28[a]	GATTA*, GGAAA*, GATAA*
(−)-CC-1065	28[a]	GATTA*, GATAA*
(−)-A	280	TTTTA*, GATTA*, GATAA*
(−)-AB	2800[a]	GATTA*
(−)-ABC	280[b]	no (+)-CC-1065 sites
	(−)-CC-1065 Alkylation Sites	
(−)-AB′C′	0.028	ATA*AC, CAA*TT
(−)-CC-1065	0.028	ATA*AC
	0.28	CAA*TT, GAA*AC
(+)-ABC	0.28	ATA*AC, CAA*TT
(−)-ABC	28	CAA*TT, GAA*AC
(+)-AB	28	ATA*AC, CAA*TT
(+)-A	2800	(GAA*TT)ATA*AC, CAA*TT, GAA*AC
(−)-AB	2800	CAA*TT, GAA*AC
(−)-A	2800[b]	no (−)-CC-1065 sites
(+)-AB′C′	28[b]	no (−)-CC-1065 sites
(+)-CC-1065	28[b]	no (−)-CC-1065 sites

[a] Detectable cleavage may be due to contamination by (+). [b] Maximum concentration studied.

thermally induced strand breakage assay using the 5′-labeled 117 base pair (+ strand) M13mp1 DNA fragment, with a range of concentrations for each drug. This fragment contains five identified (+)-CC-1065 alkylation sites, and four identified (−)-CC-1065 sites (Figure 4 and Chart II, top strand). Figure 5 presents the results for (+)- and (−)-A, -AB, and -ABC. The gels obtained for (+)- and (−)-AB′C′ were virtually identical with those of (+)- and (−)-CC-1065 in Figure 4 (data not presented). Table VIII summarizes these data.

In confirmation of our earlier results,[23] Table VIII shows that the same set of sequences that reacts with (+)-CC-1065 also reacts with all of the (+)-CPI agents (although consensus studies were not carried out for each analogue, the consistent and exact correspondence of covalently reactive adenines strongly suggests that the alkylation sites themselves correspond as well).[23] While the 10-fold dilutions provide only approximate quantitation, the overall reaction rate at all of the sensitive sequences increases markedly as the molecular structure of the drug is lengthened from (+)-A to (+)-ABC, since correspondingly lower concentrations of drug are needed to produce detectable cleavage. With (+)-ABC this effect reaches a plateau, and no further increase in alkylation reactivity results from further elaboration of the CC-1065 structure for these sequences. At very high concentrations [1000-fold higher than required for detection of cleavage by their respective (+)-CPI enantiomers], (−)-CC-1065, (−)-AB′C′, and (−)-AB produce detectable cleavage at these (+)-CC-1065 sites. This is most likely due to traces of enantiomeric contaminant in these synthetic agents, since as little as 0.1% (+)-enantiomer could produce the observed bands. A very different situation obtains for (−)-A, however, whose reactivity with the set of adenines in (+)-CC-1065 sites is within an order of magnitude of the reactivity of (+)-A. Since all of the (−)-CPI compounds are prepared from the same resolved intermediate **3**, (−)-A cannot contain sufficient (+)-enantiomer contaminant to account for this result.

Some interesting differences emerge for the (−)-CC-1065 alkylation sites. While these distinct alkylation sites are highly reactive toward (−)-CC-1065 and its close analogue (−)-AB′C′, they are somewhat less reactive than the (+)-CC-1065 sites are toward the corresponding (+)-CPI agents, judging by the higher concentrations required to produce detectable cleavage. A second difference is that (−)-ABC is very poorly reactive with adenines at these (−)-CC-1065-sensitive sequences, requiring 100- to 1000-fold greater concentrations than the more structurally elaborated (−)-CC-1065 and (−)-AB′C′. This large rate discrepancy continues with the smaller (−)-AB, and no adenine reactions corresponding to (−)-CC-1065 alkylation sites could be detected for (−)-A at the highest concentration tested. Indeed, (−)-A reacts at (+)-CC-1065 sites much more readily than it would appear to react at (−)-CC-1065 sites. The third intriguing observation from this part of Table VIII is that the less elaborate (+)-series analogues, (+)-ABC, (+)-AB, and (+)-A, react to a small extent at adenines associated with (−)-CC-1065 sites. To be sure, reaction at these adenines requires 10- to 100-fold higher concentrations of each drug than reaction at adenines in (+)-CC-1065 sites. However, this still allows these (+)-CPI analogues to be far more reactive than their respective (−)-CPI enantiomers with these same adenines. Indeed, the adenine (A*) in the 5′CAA*TT sequence, which is one of the few (−)-CC-1065 alkylated sequences not near to a (+)-CC-1065 bonding site, appears to react as readily with (+)-ABC as it does with (−)-CC-1065. A 5′GAA*TT sequence rather near the labeled end of this DNA fragment, which would be expected to have a good bonding site for (−)-CC-1065, also provides an alkylation site for (+)-A and (+)-AB at 2800 and 280 μM, respectively. [The observation that intense bands from heat-treated, 5′-end-labeled DNA are often accompanied by traces of slower running cleavage product presumed to result from β-elimination of the modified deoxyribose, as shown in Figure 2, might make it somewhat difficult to unambiguously identify weak bands resulting from reaction at adenines in (−)-CC-1065 sites, which frequently occur immediately 5′ to (+)-CC-1065 sites. Such a complication does not occur, however, for the 5′CAA*TT and GAA*TT sequences in this DNA fragment.] At the highest concentrations tested, (+)-AB′C′ and (+)-CC-1065 did not produce detectable cleavage at adenines associated with (−)-CC-1065 sites. Because of the lower reactivity of (−)-CC-1065 sites, traces of (−)-enantiomeric contaminant in synthetic (+)-AB′C′ might not be detected at these concentrations.

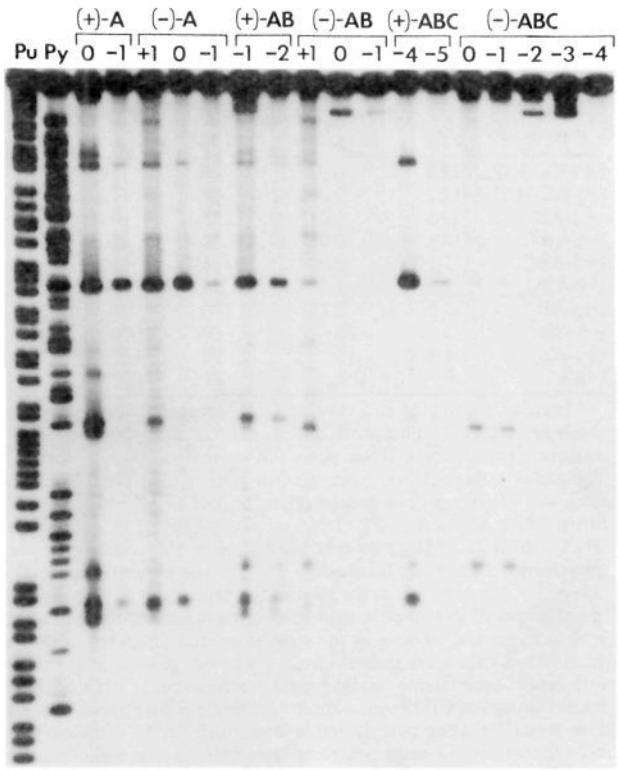


Figure 5. Concentration dependence of alkylation by (+)- and (−)-A, (+)- and (−)-AB, and (+)- and (−)-ABC of the 117-bp *MspI–BstNI* fragment of M13mp1. Aliquots of 5′-labeled 117-bp DNA fragment were incubated with appropriate concentration of (+)- or (−)-A, (+)- or (−)-AB, or (+)- or (−)-ABC as described in the Experimental Section. +1, 0, −1, −2, −3, and −4 denote 2.8 mM and 280, 28, 2.8, 0.28 and 0.028 μM, respectively. Thermal cleavage conditions are described in the Experimental Section. Pu, purine; Py, pyrimidine.

Table IX. Physical Binding Interaction and Biological Activities of CC-1065 and Analogues

		in vitro	in vivo P388 leukemia[c]		
	ΔT_m[a]	L1210 leukemia ID_{50},[b] nM	% ILS	OD, μg/kg-day	delayed death[d]
(+)-CC-1065	29.8	0.03	60	50	yes
(−)-CC-1065	12.7	0.007	50	12	no
(+)-AB′C′	22.8	0.01	75	100	yes
(−)-AB′C′	12.4	0.01	50	50	no
(+)-ABC	19.1	0.004	(4/6)[f]	25	no
(−)-ABC	7.5	1.0	55	200	no[e]
(+)-AB	13.3	0.1	100	250	no[e]
(−)-AB	3.3	20	80	10000	no[e]
(+)-A	1.4	10	90	1200	no
(−)-A	0	100	75	50000[g]	nt

[a] Thermal melting of calf thymus DNA compared with or without the drug in 0.01 M phosphate buffer, pH 7.2, after the drug–DNA mixture, drug alone, or DNA alone was incubated at 25 °C for 24 h. The molar ratio of DNA (measured as nucleic acid phosphate) and drug was 13 with a drug concentration of 0.85×10^{-5} M. ΔT_m, T_m (drug–DNA mixture) − T_m (DNA alone). The T_m of calf thymus DNA is 65.6 °C. [b] Drug concentration causing 50% inhibition of cell growth with 3-day drug incubation. L1210 cells were at a concentration of 5×10^3 cells/mL at the start of the experiment. [c] Mice bearing ip-implanted P388 leukemia with ip drug treatment on days 1, 5, and 9; % ILS, percent increase in life span of treated compared to control mice. Most values are rounded averages of two or more experiments with repeat experiments showing good reproducibility. OD, optimum dose. [d] Groups of CD2F1 mice received a single iv dose over a range of dose levels including acutely lethal doses and deaths were recorded daily for 90 days. Deaths occurring later than 20 days were classified as "delayed". [e] Tested as racemic mixture of (+)- and (−)-enantiomers. [f] Number of day 30 survivors/total. [g] Highest dose tested.

In summary, those sequences in the M13mp1 DNA strand described as (+)-CC-1065 alkylation sites appear to be intrinsically more reactive and more generally chirally discriminating than those sequences described as (−)-CC-1065 alkylation sites. Their relative reactivity is largely independent of the structural appendages attached to the (+)-CPI subunit, although binding enhancement by those appendages accelerates reaction at all susceptible sequences.[23] Only the smallest (−)-CPI agent, (−)-A, can react at these sequences, and this occurs with a 10-fold loss of rate relative to (+)-A. The sequences identified as (−)-CC-1065 alkylation sites are only chirally discriminatory to the structurally elaborated (+)-CC-1065 and (+)-AB′C′. The react, although more slowly than (+)-CC-1065 sites, with other (+)-CPI analogues [presumably with the bonded (+)-CPI lying to the 5′-side of the alkylated adenine, opposite to the 3′-orientation of (−)-CC-1065]. The (−)-CPI analogues in general appear to be intrinsically less reactive alkylating agents for DNA. The smaller analogues, (−)-A, (−)-AB, and (−)-ABC, are less reactive at any sequence than their (+)-enantiomers. This low reactivity of the (−)-CPI subunit is overcome, however, only by (−)-CC-1065 and (−)-AB′C′, which contains the full carbon skeleton of (−)-CC-1065, and may react at adenines that react relatively poorly with smaller (+)-CPI analogues, and not at all with (+)-CC-1065 and (+)-AB′C′.

B. DNA Thermal Melting. Table IX lists the change in the melting temperature of calf-thymus DNA in the presence of (+)- and (−)-CC-1065 and their analogues. This experimental parameter can reflect the density of binding (covalent or noncovalent) sites on the DNA, as well as the average strength of the noncovalent interactions at these sites and the stabilization of the helix organization resulting from the bound or bonded complexes. All of these factors probably contribute to the high ΔT_m of (+)-CC-1065. At the other extreme, (+)-A and (−)-A have little or no measurable effect when incubated with calf thymus DNA under the same conditions. There is a monotonic decrease in ΔT_m for the (+)-CPI analogues as the right-hand portion of the molecule is progressively truncated, consistent with the expected contribution of this structure to hydrophobic interactions within the minor groove.[21b] (−)-CC-1065, as well as its analogues, has a strikingly lower net stabilizing effect on calf thymus DNA than do the corresponding (+)-CPI analogues under these conditions of temperature and stoichiometry.

C. Biological Activity. Table IX summarizes some important biological characteristics of the (+)- and (−)-CPI compounds described above. These include the cytotoxicity against L1210 tumor cells in suspension culture, antitumor efficacy against P388 leukemia cells in mice and the dose required to achieve it, and the ability to cause delayed death in nontumored mice. In comparing the activities of enantiomers, three types of relationships are evident from Table IX: (1) (+)- and (−)-CC-1065, and also (+)- and (−)-AB′C′, show comparable, high potency both in vitro and in vivo, as well as comparable efficacy against P388 leukemia in mice. However, the (+)- and (−)-compounds are strikingly differentiated in their biological action by the ability of only the (+)-agents to cause delayed death in nontumored mice. (2) (+)- and (−)-ABC, and also (+)- and (−)-AB, in sharp contrast to the more complex structures, show very disparate potencies between (+) and (−) counterparts. Since these potency differences approach 2 orders of magnitude or more, it becomes experimentally impossible to distinguish the contribution of a minor, potent contaminant from that of the bulk agent. Thus, the ID_{50} value for (−)-ABC might be due to the possible presence of 0.5% (+)-ABC. Similarly, the 55% increased life span observed with 200 μg/kg (−)-ABC could in part be due to contamination by (+)-ABC. Thus, the cytotoxicities and efficacies reported for (−)-ABC and (−)-AB must be interpreted with caution. What can be said with assurance is that these (−)-CPI agents are dramatically less active biologically than are their (+)-enantiomers. (3) (+)-A and (−)-A show yet another relationship. They differ in in vitro cytotoxic potency by only an order of magnitude. Since all of the various enantiomers shown in Chart I were prepared from the same diasteriomerically resolved intermediate, and since the observed potencies of (−)-ABC and (−)-AB are inconsistent with high (10%) levels of (+) contamination, it follows that the cytotoxic potency observed for (−)-A is real and does not reflect a significant contribution by (+)-A contamination.

Discussion

A recent analysis of sequence-dependent features of the DNA major and minor grooves noted that the narrower minor groove contains only half of the hydrogen-bonding information available in the wide major groove.[2c] A ligand in the minor groove cannot distinguish, through hydrogen bonding, an A–T base pair from a T–A base pair, or a G–C from a C–G base pair. Despite this reduced discrimination, groove-binding drugs almost invariably prefer the deeper, narrower cleft of the minor groove, most likely due to the powerful driving force of hydrophobic interaction as the tightly held spine of hydration in this groove is displaced into the aqueous medium.[2c,32] Natural products such as distamycin and netropsin further refine this hydrophobic interaction to include specific favorable van der Waals contacts with AT bases lining the minor groove.[33] Interestingly, this binding specificity does not seem to preclude binding of these drugs in both possible orientations (e.g., 5′ to 3′ or 3′ to 5′) for most sequences, according to affinity cleavage studies.[34] A large and diverse class of linearly linked aryl or heteroaryl oligomeric molecules, generally bearing a positive charge, all bind noncovalently in the minor groove, and all have a more or less pronounced preference for stretches of A–T and T–A base pairs.[35] This AT preference has been described as a GC avoidance:[2c] the minor groove cleft is interrupted by the protruding 2-amino group of guanine. The bias for AT regions of the groove is strong, and synthetic "lexitropsins" designed to "recognize" GC bases[36] may simply be more tolerant of those bases than the noncovalent, groove-binding agents designed by Nature.[2c]

(32) Marky, L. A.; Breslauer, K. J. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 4359.
(33) Kopka, M. L.; Yoon, C.; Goodsell, D.; Pjura, P.; Dickerson, R. E. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 1376.
(34) Youngquist, R. S.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 2565.
(35) Braithwaite, A. W.; Baguley, B. C. *Biochemistry* **1980**, *19*, 1101.
(36) Lown, J. W.; Krowicki, K.; Bhat, U. G. *Biochemistry* **1986**, *25*, 7408.

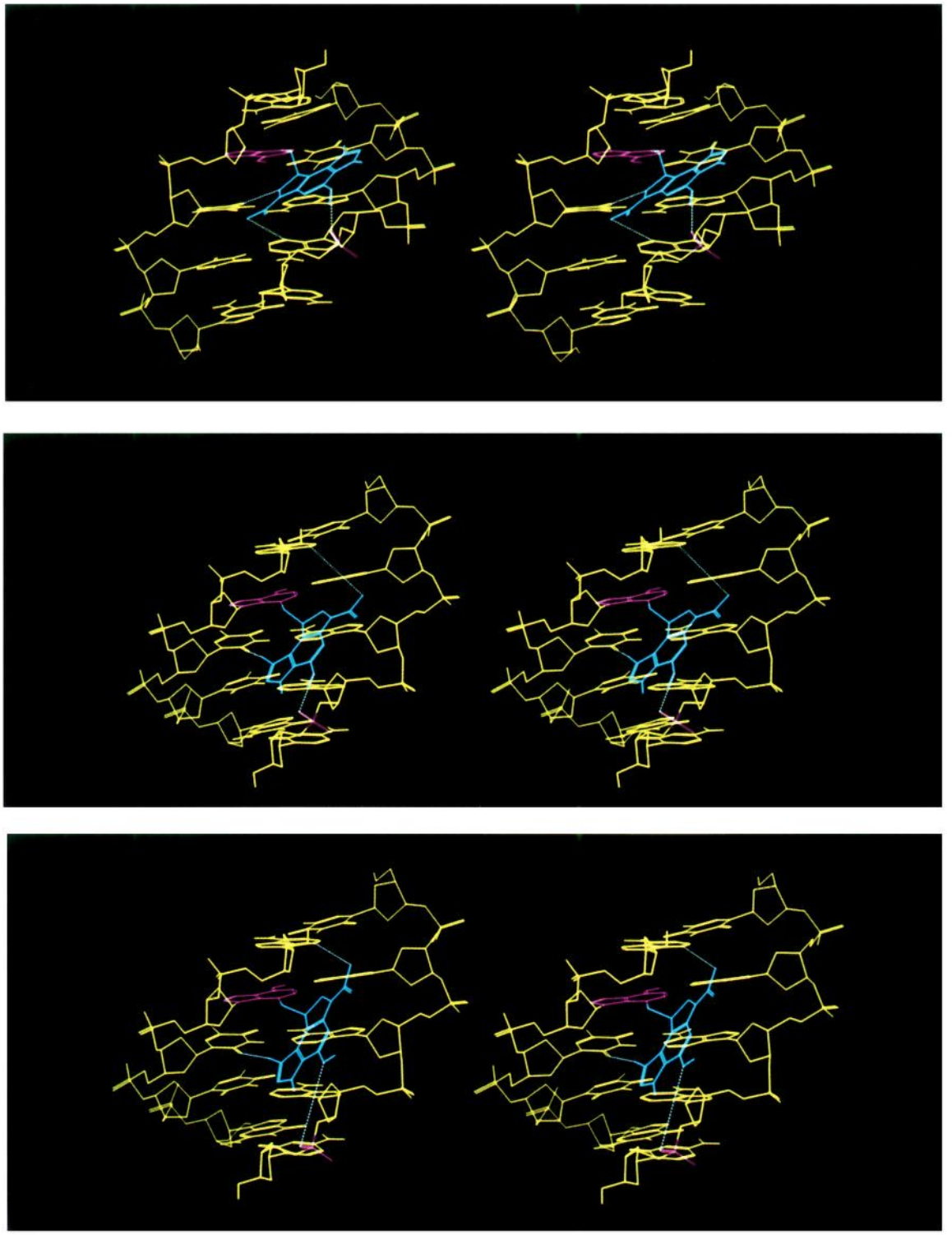


Figure 6. Stereoscopic diagrams of the adducts of (+)-A (I, top) and (−)-A (II, middle, and III, bottom) with the duplex 5′CGTTA*ACG (showing only 5′GTTA*A) generated as described in the text. The drug moiety is shown in blue, and the covalently modified adenine (A*) and the phosphate group on the opposite strand, two base pairs away, are shown in magenta. White dotted lines depict some distances that are discussed in the text.

Footprinting and affinity cleavage studies confirm that drugs such as netropsin and distamycin bind strongly to DNA fragments or oligomers at sites of four or five contiguous A–T or T–A base pairs.[2c,34,37] As discussed above, there appears to be no clear distinction between A–T and T–A pairs.[2c] Recently, this AT binding selectivity was exploited in the design of *N*-bromoacetyldistamycin, which can covalently modify DNA at adenine N3 in the minor groove, at some of the AT-rich sites that tightly hold the ligand.[37]

(+)-CC-1065 is also a natural product that binds strongly to the DNA minor groove in AT-rich regions, but unlike distamycin, it reacts with DNA to form a covalent adduct.[8-11] The irreversibility of this molecular lesion undoubtedly contributes to the 6 orders of magnitude greater cytotoxic potency of (+)-CC-1065[8b] relative to distamycin.[38] However, (+)-CC-1065 is not simply

(37) Baker, B. F.; Dervan, P. D. *J. Am. Chem. Soc.* **1989**, *111*, 2700.

a natural product version of *N*-bromoacetyldistamycin, in which sequence- (AT-) selective minor groove binding constrains a reluctant alkylating center to react with nucleophilic atoms in proximity to the binding site.[3] If this were the case, it would be expected that truncated (+)-CPI analogues of (+)-CC-1065, whose properties, such as the ability to affect the thermal melting of DNA (Table IX), indicate much weaker binding, would be unable to alkylate DNA sequence selectively, if at all. In fact, while the overall rates of the alkylation reactions with DNA are lower for the less strongly binding (+)-CPI agents, it has been shown that the set of sequences in which adenines are alkylated is the same for (+)-CPI analogues regardless of binding ability.[23] We have interpreted this result as demonstrating that the sequence selectivity of (+)-CPI compounds is primarily determined, not by sequence-recognizing noncovalent binding, but by the sequence requirements inherent in the covalent-bonding step itself.[23] This might be postulated to arise from a sequence-dependent conformational flexibility allowing adenine in certain environments to approach the electrophilic center sufficiently closely for bond formation. It might also be postulated to involve sequence-dependent electrophilic activation of the (+)-CPI moiety simultaneously with adenine N3 nucleophilic attack, in a process similar to the bifunctional catalysis observed in many enzyme–substrate reactions.[39] From a purely structural perspective, such a process resembles suicide substrate–enzyme inactivation.

Consensus sequence analysis of covalent adduct formation by (+)-CC-1065,[10] as well as comparative footprinting experiments,[23] also indicates that (+)-CC-1065 and distamycin recognize the minor groove through different mechanisms. In the work described here, in which incubations of drug and DNA were carried out at 37 °C, an almost universal reaction with 5′TTA* sequences was observed (Table VII). While 5′TAA* sequences were more often reactive than not, interchange of the base pair more distant from the adduct, as in 5′AAA* and 5′ATA*, had a pronounced effect on the probability of reaction at these sequences, i.e., only a small fraction of such sequences were reactive toward (+)-CC-1065. Thus, (+)-CC-1065, through its covalent reaction, is able to discriminate between A–T and T–A base pairs in the minor groove.

The strand breakage assay used in these studies can only detect covalent reaction sites of (+)-CC-1065.[40] Recent spectroscopic studies of (+)-CC-1065 with various synthetic oligomers reveal that the drug may bind strongly to certain sequences with which it does not react covalently.[41] The most dramatic example of this behavior is provided by the strong noncovalent complex between (+)-CC-1065 and the duplex of 5′CGCGAATTCGCG. This oligomer also provides an excellent binding site for netropsin, and indeed, the first crystal structure of a netropsin–oligomer complex was obtained with a 5-bromocytosine-modified duplex of this sequence.[33] Such evidence of reversible binding of (+)-CC-1065 to the same minor groove locations selected by netropsin or distamycin suggests that the purely noncovalent recognition mechanisms of (+)-CC-1065 are not greatly different from those of other minor groove binding agents. It also emphasizes the overlying, powerful A vs T discriminating ability of the (+)-CPI covalent reaction.

$$\text{drug} + \text{DNA} \overset{K_b}{\rightleftharpoons} (\text{drug-DNA})_{\text{noncov}} \xrightarrow{k_r} (\text{drug-DNA})_{\text{cov}} \quad (1)$$

In terms of the generalized kinetic scheme in eq 1, a drug may preferentially react with a particular sequence through either a very high binding affinity (K_b) at that sequence, or a covalent reaction whose activation energy is especially favorable for that sequence.[3] If k_r is not sequence-dependent, the sites of covalent reaction will be those having the highest on-rate of noncovalent binding (if k_r is fast relative to dissociation) or the highest K_b (if k_r is slow). (+)-CC-1065, although it has a very high binding affinity for AT-rich sequences, is also constrained by a sequence-dependent k_r. Simple (+)-CPI analogues with low binding affinity to DNA still react, though at a lower overall rate, at the same sequences that react with (+)-CC-1065, sequences that allow a favorable k_r for covalent reaction between adenine and the (+)-CPI moiety.

Implicit in the sequence recognition capability of the (+)-CPI alkylating segment is that alterations to the structure of that moiety could have significant effects on its absolute reactivity and relative reactivity (sequence selectivity) with DNA. The accessibility, through synthesis, of the unnatural enantiomer of the alkylating segment,[21a] and the methodology to prepare (−)-CC-1065 and other (−)-CPI analogues,[13] have provided an ideal system with which to explore these implications. The identical physical and chemical properties of the (−)-CPI and (+)-CPI agents focus comparison on the differences of their interactions with the chiral DNA helix. Indeed, we have found that (−)-CC-1065, while it superficially resembles (+)-CC-1065 in reacting covalently with certain adenines in the minor groove of DNA in AT-rich regions, shows subtle but significant differences. It lies in the opposite direction in the groove, it alkylates a nonidentical set of adenines, it reveals a different consensus requirement for the covalent adduct, and, finally, its simplified (−)-CPI analogues fail to react with DNA. The sole exception is (−)-A, which reacts at the (+)-CC-1065 sites rather than at the (−)-CC-1065 sites. How can these observations be rationalized in terms of structural models of drug–DNA bonding? The following discussion describes, not a rigorous modeling study, but a qualitative attempt to illustrate some of the possible interactions that might underlie the experimental observations described in this work.

Figure 6 presents stereo views of (+)-CPI (I) and (−)-CPI (II or III) adducts with the duplex oligomer of 5′CGTTA*ACG-3′ (only the 5′GTTA*A-3′ portion is shown for clarity). These structures were generated by the modeling program MOSAIC, an Upjohn-developed program based on Macromodel.[42] Using the AMBER set of force field parameters,[43,44] we performed energy minimizations starting from idealized DNA helices to which a ring-opened CPI moiety had been attached. These structures are intended to help illustrate possible interactions of these agents with recognition elements in the minor groove of DNA. The adduct is assumed to resemble the activated complex in the transition state for the alkylation reaction. The covalently modified adenine (A*) is in a good consensus bonding sequence for both (+)-CC-1065 (5′TTA*) and (−)-CC-1065 (5′TA*A). In structure I, the bonded (+)-CPI is oriented to place the acyl appendage [for simplicity the adduct of (+)-A was used for the modeling] in the minor groove, lying to the 5′-side of the covalently modified adenine. The energy-minimized structure places the phenolic hydroxyl group within hydrogen-bonding distance of the phosphate oxygen (O–O distance = 2.6 Å, I; Figure 6) on the opposite strand, two base pairs down from the covalently bonded adenine. While the proximity of these groups in the energy-minimized model most likely reflects the magnification of such dipolar interactions in the gas-phase calculations employed, it also suggests that such an interaction is possible and thus might play a role in protic activation of the (+)-CPI moiety. This orientation also positions the pyrrolidine methylene carbon for a favorable van der Waals contact (C–C distance = 3.3 Å) with the C2 of the adenine one base pair down, on the opposite strand, and close to the O2 of the thymine of that base pair (C–O distance = 3.1 Å, I; Figure 6). There is a favorable van der Waals distance between the carbon atom of the acyl appendage (methyl, in I) and the C2 of

(38) Lown, J. W.; Krowicki, K.; Balzarini, J.; DeClercq, E. *J. Med. Chem.* **1986**, *29*, 1210.
(39) Fersht, A. In *Enzyme Structure and Function*, 2nd ed.; W. H. Freeman and Co.: New York, 1985.
(40) Footprinting studies with MPE·Fe(II) were initially thought to reveal only covalent bonding sites for (+)-CC-1065 (ref 23); however, the footprint labeled B in Figure 3 of that reference may possibly reflect a strong noncovalent complex, particularly in view of the spectroscopic studies of ref 41. The attribution of that footprint to a covalent adduct on the opposite strand is an error; there is no evidence of a covalent adduct at that site on the opposite strand.
(41) Thériault, N. Y.; Krueger, W. C.; Prairie, M. D. *Chem.-Biol. Interact.* **1988**, *65*, 187.
(42) Still, C. W., Columbia University.
(43) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S., Jr.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765.
(44) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230.

the adenine on the opposite strand, two base pairs in the 5′-direction from the bonded adenine (C–C distance = 3.8 Å, I; Figure 6).

All of the observed contacts in I lie on the 5′-side of the adduct. Even the indole methyl substituent, while it faces the 3′-side of the adduct, lies far from the base pair to the 3′-side of the alkylated adenine. Thus, this model is consistent with the (+)-CC-1065 consensus data, which only show a general sequence requirement in the 5′-direction from the adduct, and with footprinting, as well as NMR data.[15] It further suggests that a favorable dipolar interaction between the oxo group of (+)-CPI and a phosphate group in the backbone of the helix is compatible with favorable van der Waals interactions of this alkylating moiety.

With such a significant number of possible recognition points for (+)-CPI at this reactive sequence it is not surprising that chiral discrimination is high. In fact, the greater surprise is that (−)-A reacts as well as it does at the set of adenines corresponding to (+)-CC-1065 reactive sites. One possible model for this reaction is shown in II, the diastereomer of I, in which the (−)-CPI moiety has been turned ca. 180° around the axis running through the original cyclohexadienone carbonyl group. This orientation allows covalent bonding to N3 of the same adenine bonded in I, while still maintaining the close phenolic hydroxy group interaction with the phosphate two base pairs in the 5′-direction (O–O distance = 2.7 Å, II; Figure 6). This minimized structure undoubtedly benefits in the modeling from the additional hydrogen bond of the indole N–H to the phosphate (H–O distance = 1.8 Å). The indole methyl carbon has favorable contacts with the 5′ T–A base pair (methyl carbon to thymine O2 distance = 3.1 Å, II; Figure 6). This orientation of (−)-CPI places the acyl appendage in the 3′-direction from the adduct site, but in order to maintain the phenol phosphate interaction, the molecule is skewed in a manner that pulls this appendage out from the groove. In II, where this group is methyl, it is quite distant from the 3′ A–T base pair (carbon to adenine N3 distance = 5.4 Å, II; Figure 6). It is this model that we suggest to explain the observed reaction of (−)-A at the adenines of (+)-CC-1065 sites. The 10-fold lower reactivity of (−)-A relative to (+)-A indicated by the concentration dependence of the cleavage assay would suggest that II represents only a slightly higher energy pathway than I. The lack of reactivity of (−)-CPI compounds with acyl appendages larger than methyl at these (+)-CC-1065 site adenines would indicate, according to this model, that (1) leaving such large hydrophobic appendages protruding out of the groove is not a favorable situation, and (2) reorientation to optimize groove binding of these appendages could rotate the (−)-CPI moiety away from the favorable orientation for alkylation depicted in II.

The 10-fold lower alkylation reactivity of (−)-A at the (+)-CC-1065 site adenines correlated with its 10-fold lower in vitro cytotoxic potency (Table IX). We had previously noted a positive, rank order correlation between relative DNA reactivity (as reflected by the drug concentration required in the cleavage assay) of a series of (+)-CPI analogues and their in vitro cytotoxic potency and in vivo potency.[23] The observed lack of potent cytotoxicity of (−)-AB and (−)-ABC (Table IX) is consistent with their observed lack of reactivity with the (+)-CC-1065 site adenines (Table VIII) and further extends the correlation of reactivity at these sites on naked DNA with cytotoxic potency for these compounds.

Model II can account for the observed (+)-CC-1065 site reactivity of (−)-A. While alternative models might be considered, including one in which the acyl methyl group of the (−)-A adduct lies in the 5′-direction from the covalently modified adenine (not shown), II appears to be the more attractive conformation for the (−)-A adduct. A different model (III, Figure 6) is required to account for the observed reactivity of (−)-CC-1065 and (−)-AB′C′ at the (−)-CC-1065 alkylation sites. [Recall that we are using the term "alkylation site" to refer not only to the specific adenine involved in bonding but also to the sequence context as summarized in the consensus sequence for the drug. Hence, while models I and II depict possible (+)- and (−)-CPI adduct interactions at a 5′TTA* "(+)-CC-1065 alkylation site", model III describes a possible (−)-CPI adduct ((−)-A is shown for simplicity) at the overlapping but distinct 5′TA*A "(−)-CC-1065 alkylation site".]

In III, the acyl appendage lies in the minor groove to the 3′-side of the covalently modified adenine. This model accommodates the footprinting findings for (−)-CC-1065, the recovery of essentially only adenine N3 adducts, and the high degree of minor grooves overlap between (+)-CC-1065 sites and (−)-CC-1065 sites on the DNA fragments examined in this work (Chart III). In the energy-minimized structure III, the acyl methyl group is brought into favorable contact with the adenine lying 3′ to the adduct, particularly with its anomeric carbon (C–C distance = 3.6 Å) and its N3 atom (C–N distance = 3.8 Å, III; Figure 6). This proximity suggests that larger acyl appendages can also enjoy close interaction with the groove in the 3′-direction, to a much greater extent than is possible in structure II. The indole methyl group enjoys contact with the thymine O2 (C–O distance = 3.3 Å, III; Figure 6) and the adenine C2 (C–C distance = 3.4 Å) and adenine N3 (C–N distance = 3.8 Å) in the base pair immediately 5′ to the adduct. This 5′ base pair recognition of III provides a modeling rationale for the observed 5′ flanking AT requirement of (−)-CC-1065, even though the acyl appendage lies in the 3′-direction from the adduct. To achieve the favorable 3′-contacts of the adduct structure III, the phenol hydroxyl group interaction with the phosphate group (O–O distance = 5.3 Å, III; Figure 6) has been sacrificed. If indeed the phosphate proximity in I and in II is important to provide electrophilic activation of the cyclopropylcyclohexadienone moiety (nucleophilic ring opening of CPI compounds in solution is strongly acid catalyzed),[21b,31,45] then (−)-CPI compounds would be expected to suffer a severe loss of reactivity with DNA when oriented as in III.[46] The observations (Table VIII) that (−)-A, (−)-AB, and (−)-ABC are extremely unreactive at (−)-CC-1065 sites is consistent with this model. Particularly in the case of (+)- and (−)-A, where differential binding interactions are minimized, it is apparent that (−)-A reacts at (−)-CC-1065 sites (III) at least 2 orders of magnitude more slowly than (+)-A reacts at (+)-CC-1065 sites (I). Indeed, (−)-A appears better able to react at (+)-CC-1065 sites (II, vide supra). Clearly the (−)-CPI alkylating structure by itself, in contrast to (+)-CPI, does not enjoy uniquely facilitated alkylation.

Why then do (−)-CC-1065 and (−)-AB′C′ readily alkylate DNA at the (−)-CC-1065 sites? The answer must lie with the interaction of the large, bis(pyrroloindole) acyl appendage with the minor groove when the drug is oriented as depicted in III. The low reactivity of (−)-AB and (−)-ABC at these sites suggests that the less specific hydrophobic binding interactions that these molecules allow (as indicated, for example, by their measurable stabilization of DNA toward thermal melting, see Table IX) are simply inadequate to facilitate the less favorable covalent reaction. The dramatic alkylation rate enhancement resulting from attachment of the full carbon skeleton of the natural products to the (−)-CPI moiety strongly implies that specific AT recognizing interactions, such as favorable van der Waals contacts involving the rigidly held framework of pyrrolidine methylene groups, play a dominant role in both the absolute and relative reactivities of (−)-CC-1065 and (−)-AB′C′ with DNA.

That the pyrrolidine methylene groups of the B′ and C′ segments can play a decisive role in minor groove binding is supported by the observation of an induced circular dichroism of this dimeric fragment [from the alkaline hydrolysis of (+)-CC-1065] in the presence of calf thymus DNA.[31] In contrast, the amide linked bis(indole-2-carboxylate) from the alkaline hydrolysis of (+)-ABC shows no induced circular dichroism with DNA.[47] Important synthetic[48] and computational[49] studies by Boger have contributed

(45) Warpehoski, M. A.; Harper, D. E. Unpublished results.

(46) Another feature distinguishing model I from the diastereomeric model III is the steric conformation around the newly formed N–C bond. Viewed along the axis of this bond, the adenine ring in structure I is almost fully staggered relative to the nearest carbon–carbon bond of the drug moiety. In structure III, in contrast, the adenine ring is nearly eclipsed by this carbon–carbon bond. An interaction such as this might significantly add to the instability of a transition state resembling III.

(47) Warpehoski, M. A.; Krueger, W. C. Unpublished results.

to an appreciation of the strong minor groove affinity of B′–C′ oligomers. Indeed, these studies are highly relevant to the DNA sequence recognition characteristics of (−)-CC-1065.

Interestingly, (−)-CC-1065 does not discriminate significantly among 5′AA*A, 5′TA*A, and 5′AA*T sequences (Table IV), as expected from an agent whose primary recognition mode fundamentally resembles that of noncovalent minor groove binders.[2c] There is, however, clearly a lower degree of reactivity at 5′TA*T sequences, and this may reflect a particular bonding limitation in III at that sequence. It is intriguing that the sequencing studies show (Chart III) that of the 14 (−)-CC-1065 sites not overlapping (+)-CC-1065 alkylation sites, 8 are within or adjacent to 5′AATT sequences. Four of these are among the most reactive (−)-CC-1065 sites (Table II). This sequence, as discussed above, is a highly favorable binding site for netropsin and related reversible binding agents and is an exceptionally stable, reversible (noncovalent) binding site for (+)-CC-1065. The implication of these observations is that AT-specific minor groove binding interactions attributable to the B′ and C′ segments appear to be the main determinant of the reactivity of (−)-CC-1065 at its target adenines. In terms of eq 1, k_r for (−)-CC-1065 is not greatly facilitated by favorable transition-state stabilization at particular sequences. Rather, high binding affinity for certain sequences is necessary for significant alkylation to occur there.

The similar level of cytotoxic potency (ID_{50} values) of (+)- and (−)-CC-1065 against several cell lines observed by us (Table IX) and others[24] is somewhat fortuitous and misleading. Cytotoxicity, of course, may be the result of diverse lethal biochemical processes. While (−)-CC-1065 and (−)-AB′C′ are, if anything, slightly more potent in the biological systems tested than are (+)-CC-1065 and (+)-AB′C′ (Table IX), they are about an order of magnitude less reactive than their (+) counterparts toward DNA under the reaction conditions used in the alkylation experiments (Table VIII). Thus, the rank order correlation of potency and reactivity observed for compounds reacting at (+)-CC-1065 sites must be considered separately from a similar correlation for compounds reacting at (−)-CC-1065 sites. This is expected if different biochemical mechanisms for cytotoxicity result from the different molecular lesions on DNA. Because it seems reasonable that the observed low level of reaction of the simpler (+)-CPI agents at adenines in (−)-CC-1065 sites (Table VIII) produces adducts lying to the 5′-side of the alkylated adenine, rather than to the 3′-side, these (+)-CPI agents should not be included in correlations of reactivity and potency of (−)-CPI agents at (−)-CC-1065 sites. Further delineation of the biological and biochemical differences between (+)- and (−)-CC-1065 is in progress. One striking difference already noted (Table IX) is that (−)-CC-1065 does not cause delayed death in mice.

The full carbon structure right-hand appendage has a somewhat more subtle, but nevertheless, significant effect on the properties of (+)-CPI as well.[23] At some sequences, it appears that the specific noncovalent interactions of the two pyrrolidine rings with the minor groove can either promote an otherwise less favorable bonding reaction [operationally similar to (−)-CC-1065] or inhibit, perhaps through a competitive binding mode, a bonding reaction that is available to simpler (+)-CPI analogues. While these specific binding interactions are clearly not responsible for the large majority of (+)-CC-1065-reactive alkylation sites on DNA restriction fragments,[23] they appear to have profound biological significance. Only (+)-CC-1065 and (+)-AB′C′ [and not the simpler (+)-CPI agents] cause delayed death in mice at therapeutic doses.[20,21,28] This difference could conceivably reflect adduct formation at unique sites with these agents. Alternatively, unique structural and functional consequences to DNA caused by these specific binding interactions, available only to (+)-CC-1065 and (+)-AB′C′ covalent adducts, may be critical for the expression of this toxicity.

In earlier biosynthesis studies,[50] we had concluded that *S. zelensis* incorporated a single carbon atom from methionine to form the cyclopropyl group of (+)-CC-1065. Was it strictly fortuitous that the 7b*R*,8a*S* enantiomer was formed, or was this the result of selective pressure? Only the (+)-CPI moiety appears to recognize the minor groove through sequence-dependent covalent reactivity. The synthetically prepared unnatural enantiomer of the natural product, while it is also a potent cytotoxin, has a fundamentally different ultimate interaction with double-helical DNA and at least some different resulting biological properties. It is speculative but conceivable that these different biological properties would not have given the microorganism a particular survival advantage. Likewise, it is interesting to note that the B′C′ pyrrolidine carbon atoms are not incorporated through tyrosine, as might, a priori, be supposed. Instead, two carbon atoms from serine are incorporated into each of the tyrosine-derived indoles of the B′C′ segment.[50] It is this additional structural elaboration that leads to the subtle modulation of the (+)-CPI covalent reaction site or adduct structure, which in turn appears to underlie delayed toxicity in rodents. Conceivably, this structural elaboration might also result from selective pressure. An exciting and general implication of these considerations is that the very biosynthesis of natural product toxins may provide important clues about their essential reactivity features. If these often elaborate molecules have evolved through multiple levels of structural refinement to optimize key biological interactions, then a knowledge of the biosynthetic origins of the structural components could serve as a valuable inspiration to structure–activity investigations.

The synthetic toxin (−)-CC-1065 appears to have "pirated" an AT-specific minor groove binding ligand that was exquisitely designed through evolution to optimize the critical (to the producing organism) DNA interaction of (+)-CC-1065. It uses this structural key to gain entry to AT regions generally favorable to (+)-CC-1065 (and other minor groove binders). Once there, however, the rules governing the critical covalent reaction are quite different than for the natural product. One might also regard the synthetic (+)-CPI analogues, such as (+)-ABC, as having "pirated" the sequence-discriminating alkylating moiety of the natural product. Separated from the natural binding ligand, and efficiently delivered to the reactive DNA site by less specific binding ligands, these analogues are proving to be highly effective and potent antitumor agents in animal testing.[21]

It has been suggested that the predilection of natural product defense molecules for the minor groove is an evolutionary adaptation to the accessibility of this groove, which is not occupied by endogenous proteins.[2c] The extraordinary and highly evolved DNA binding and bonding interactions of (+)-CC-1065 invite even further intriguing speculations. The "recognition" by (+)-CC-1065 of its critical reactive sequences in DNA may involve both cognitive and response features, i.e., these DNA sequences may be pharmacological receptors.[51] They might, for example, be normal binding regions for regulatory proteins, and modifications of those sequences could affect protein binding, and ultimately the processes that those proteins control. Alternatively, unusual DNA local structures that are sequence dependent could be receptors for selective drug action. Our clearer understanding, resulting from the present work, of the structural basis for the binding and bonding selectivity of (+)-CC-1065 and related molecules will help to focus the study of the consequences of those interactions and the efforts to exploit them for therapeutic ends.

Experimental Section

Instrumentation and Reagents. Melting points were taken on a Kofler hot-stage or a Thomas-Hoover Unimelt apparatus and are uncorrected. Combustion analysis and IR and mass spectra were obtained by the Physical and Analytical Chemistry Research Department of The Upjohn Co. Fast atom bombardment (FAB) mass spectra were run on a JG ZAB 2F instrument, and electron-impact (EI) mass spectra were run at 70 eV on a CEC 21-110B mass spectrometer. Proton NMR spectra were

(48) Boger, D. L.; Coleman, R. S.; Invergo, B. J. *J. Org. Chem.* **1987**, *52*, 1521.
(49) Coleman, R. S.; Boger, D. L. In *Studies in Natural Products Chemistry*; Atta-ur-Rahman, Ed.; Elsevier: Amsterdam, 1988; Vol. 2.
(50) Hurley, L. H.; Rokem, J. S. *J. Antibiot.* **1983**, *36*, 383.
(51) Hurley, L. H. *J. Med. Chem.* **1989**, *32*, 2027.

recorded on a Varian 390 (90-MHz) or a Brüker Aspect 3000 (300-MHz) instrument, and ^{13}C NMR spectra were recorded on a Varian CFT-20 (20-MHz) instrument. Chemical shifts are reported in parts per million relative to internal tetramethylsilane. Ultraviolet spectra were recorded on a Perkin-Elmer Lambda 3 spectrophotometer. Analtech silica gel GF (0.25-mm) plates were used for TLC, with visualization by UV light. Column chromatography was performed on silica gel (70–230 mesh, gravity; 230–400 mesh, low–medium pressure) supplied by E. Merck, Darmstadt. Tetrahydrofuran (THF) was dried by distillation under nitrogen from sodium/benzophenone ketyl. Acetonitrile was distilled from calcium hydride. Dry pyridine was obtained by distillation from sodium hydroxide. All other solvents were reagent grade or reagent grade distilled from glass (Burdick & Jackson). Ammonium formate was dried in vacuo, in the presence of H_2SO_4, for 24 h. All other reagents were used as purchased and were reagent grade where available. In order to minimize the health risks posed by these potent cytotoxic agents to analytical service personnel outside of our own laboratory and to allow preparation of only the very limited quantities needed for testing, combustion elemental analyses were not obtained on the final analogues, except for ABC. These substances were, however, homogeneous by TLC. Experimental details of the resolution of **1** to **2** and **3** are provided in the supplementary materials.

(*S*)-1,6-Dihydro-1-(hydroxymethyl)-8-methyl-5-(phenylmethoxy)-benzo[1,2-*b*:4,3-*b'*]dipyrrole-3(2*H*)-carboxylic Acid 1,1-Dimethylethyl Ester [(*S*)-8]. Under a nitrogen atmosphere, a solution of 1.0 g (2.6 mmol) of **2** in 50 mL of anhydrous THF and 50 mL of toluene was treated dropwise, with stirring, with 5.0 mL (17 mmol) of a 3.4 M solution of REDAL in toluene. The solution was quickly heated under a nitrogen flow, allowing the escape of THF until the temperature reached 85 °C. After 0.5 h, the reaction mixture was cooled in an ice bath and carefully treated with 50 mL of 15% aqueous potassium carbonate. The mixture was partitioned between water and ethyl acetate, which had previously been purged with nitrogen. The aqueous layer was separated and twice more extracted with nitrogen-purged ethyl acetate. The combined organic layers were dried over sodium sulfate and concentrated in vacuo to give (*S*)-**7** as a tan foam. This material is exquisitely sensitive to air and light and was used directly in the next reaction. Under a nitrogen atmosphere, crude (*S*)-**7** was dissolved in 10 mL of anhydrous THF and the resultant solution treated with 0.38 mL (5.1 mmol) of triethylamine and 0.70 g (2.8 mmol) of BOC-ON. After 90 h, the reaction mixture was concentrated in vacuo and the residue chromatographed over 180 g of silica gel, eluting with 10% ethyl acetate in hexane to 50% ethyl acetate in hexane. This afforded 0.75 g (71%) of (*S*)-**8**: ^1H NMR ($CDCl_3$) δ 8.5 (s, 1 H), 7.8–7.2 (m, 6 H), 6.9 (s, 1 H), 5.15 (s, 2 H), 4.4–3.4 (m, 6 H), 2.3 (s, 3 H), 1.6 (s, 9 H); MS calcd for $C_{24}H_{28}N_2O_4$ *m/e* 408.2049, found 408.2051. Analogously, (*R*)-**8** was prepared from **3**.

(*S*)-1-(Chloromethyl)-1,6-dihydro-8-methyl-5-(phenylmethoxy)-benzo[1,2-*b*:4,3-*b'*]dipyrrole-3(2*H*)-carboxylic Acid 1,1-Dimethylethyl Ester [(*S*)-9]. A 120-mg (0.29-mmol) quantity of alcohol (*S*)-**8** was dissolved in 2 mL of pyridine under a nitrogen atmosphere and the solution cooled to 0 °C and treated with 0.10 mL (0.145 g, 1.27 mmol) of methanesulfonyl chloride. The reaction was allowed to warm to room temperature and stand for 6 h. It was then recooled to 0 °C and treated with a few drops of water. After about 10 min, the reaction was partitioned between methylene chloride and 5% aqueous sodium bisulfate. The aqueous layer was separated and reextracted with methylene chloride. The combined organic solutions were dried over sodium sulfate and concentrated under vacuum leaving the mesylate: ^1H NMR ($CDCl_3$) δ 8.4 (s, 1 H), 7.7–7.2 (m, 6 H), 6.9 (s, 1 H), 5.1 (s, 2 H), 4.6–3.6 (m, 5 H), 2.8 (s, 3 H), 2.4 (s, 3 H), 1.6 (s, 9 H); R_f 0.31 (20% ethyl acetate in hexane). The crude mesylate was dissolved in 3 mL of dry DMF and treated with 40 mg (0.95 mmol) of lithium chloride. The solution was heated to 80 °C under nitrogen and stirred at that temperature for 25 min. The reaciton was then partitioned between methylene chloride and water. The aqueous layer was separated and reextracted with methylene chloride. The combined organic solutions were dried over sodium sulfate and evaporated to dryness in vacuo. The crude solids were chromatographed over 10 g of silica gel, eluting with 20% ethyl acetate in hexane. There was obtained 105 mg (85%) of (*S*)-**9**: ^1H NMR ($CDCl_3$) δ 8.3 (s, 1 H), 7.7–7.3 (m, 6 H), 7.0 (s, 1 H), 5.2 (s, 2 H), 4.5–3.2 (m, 5 H), 2.4 (s, 3 H), 1.6 (s, 9 H); MS calcd for $C_{24}H_{27}ClN_2O_3$ *m/e* 426.1710, found 426.1721. A similar conversion of (*R*)-**8** afforded (*R*)-**9**.

(*S*)-1-(Chloromethyl)-1,6-dihydro-5-hydroxy-8-methylbenzo[1,2-*b*:4,3-*b'*]dipyrrole-3(2*H*)-carboxylic Acid 1,1-Dimethylethyl Ester [(*S*)-13]. To 110 mg (0.26 mmol) of (*S*)-**9** dissolved in 2 mL of THF and 2 mL of methanol was added 120 mg of 10% palladium on carbon and 130 mg (2.1 mmol) of ammonium formate. The mixture was stirred at room temperature for 20 min. The reaction mixture was then filtered and the collected solids washed with THF. The combined filtrates were partitioned between water and ethyl acetate. The aqueous layer was separated and reextracted with ethyl acetate. The combined organic layers were dried over sodium sulfate and concentrated in vacuo. This left 69 mg of tan solid (78%), which was pure enough to use direclty in the next step. Purified (*S*)-**13** was obtained with 94% recovery after chromatography over silica gel eluted with 30:60:10 ethyl acetate–toluene–chloroform followed by the same solvent mixture in the proportions of 50:40:10: IR (film) 3383, 2961, 2853, 2867, 1681, 1500, 1458, 1415, 1377, 1369 cm^{-1}; UV (ethanol) 245 nm (23 500), 277 (9000), 285 (sh, 8200), 307 (5130), 320 (4890), 345 (3630); ^1H NMR ($CDCl_3$ + acetone-d_6) δ 9.7 (s, 1 H), 8.9 (s, 1 H), 7.4 (s, 1 H), 7.1 (s, 1 H), 4.4–3.3 (m, 5 H), 2.4 (s, 3 H), 1.6 (s, 9 H); MS calcd for $C_{17}H_{21}ClN_2O_3$ *m/e* 336.1241, found 336.1230; $[\alpha]^{25}_{Hg} = -20.0° \pm 1.2$ (DMF, 0.170 g/100 mL). In like manner, (*R*)-**9** was converted to (*R*)-**13**: $[\alpha]^{25}_{Hg} = +19.3° \pm 0.6$ (DMF, 0.187 g/100 mL).

(*S*)-1-(Chloromethyl)-1,6-dihydro-5-hydroxy-8-methylbenzo[1,2-*b*:4,3-*b'*]dipyrrole [(*S*)-14]. Since **14** is unstable, it was prepared directly before use for each analogue. A 50-mg (0.15-mmol) quantity of (*S*)-**13** was dissolved in 2 mL of ethyl acetate saturated with hydrogen chloride gas. After 40 min, the reaction was evaporated to dryness in vacuo and the residue reevaporated three times with methylene chloride to remove traces of HCl. The residue was then used directly in the coupling reactions. The enantiomer (*R*)-**14** was prepared analogously.

(*S*)-6-Acetyl-8-(chloromethyl)-3,6,7,8-tetrahydro-1-methylbenzo[1,2-*b*:4,3-*b'*]dipyrrol-4-ol [(*S*)-15a]. The hydrochloride salt (*S*)-**14** from a 100-mg (0.30-mmol) conversion of (*S*)-**13** was dissolved in 3 mL of anhydrous dimethylacetamide (DMA) and the solution stirred under nitrogen in the dark. To this was added 33 mg (0.3 mmol) of acetyl imidazolide and the resultant solution stirred at 45 °C for 2.5 h and at room temperature for 16 h. The reaction was then partitioned between water and ethyl acetate. The aqueous layer was separated and reextracted with ethyl acetate. The combined organic layers were dried over sodium sulfate and concentrated in vacuo. The residue was chromatographed over 20 g of silica gel, eluting with 50% acetone in hexane, to afford 63 mg (77%) of (*S*)-**15a**: ^1H NMR (DMSO-d_6) δ 10.61 (s, I H), 9.67 (s, 1 H), 7.61 (s, 1 H), 7.0 (s, 1 H), 4.18 (dd, J = 11, 8 Hz, 1 H), 4.10–4.04 (d, J = 11 Hz, 1 H), 3.94 (br t, 1 H), 3.85 (dd, J = 11, ~3 Hz, 1 H), 3.54 (t, J = 10 Hz, 1 H), 2.32 (s, 3 H), 2.16 (s, 3 H); MS *m/e* 279, 278, 229, 199, 187; R_f 0.28 (40% acetone, 60% hexanes, silica). Likewise, (*R*)-**15a** was prepared from (*R*)-**13**.

(+)-(7b*R*,8a*S*)- and (−)-(7b*S*,8a*R*)-A. Compound (*S*)-**15a** (0.23 mmol) was dissolved in 14 mL of acetonitrile, 3.5 mL of triethylamine, and 3.5 mL of water and the resultant mixture stirred for 30 min at ambient temperature, under nitrogen. The reaction was complete by TLC. The mixture was partitioned between ethyl acetate and water. The phases were separated and the aqueous layer was reextracted with ethyl acetate. The combined organic phases were dried (Na_2SO_4) and concentrated to afford 54 mg (96%) of (+)-A (**16**) as a solid with spectral properties identical with those reported previously for the racemic compounds[20b] and exhibiting in addition a CD spectrum as follows: CD (methanol) 318 nm, br (−7700) 306 (0), 281 (61 000), 257 (0). Analogously, (−)-A was prepared from (*R*)-**15a** in 79% yield: CD (DMF) 329 nm, br (19 000), 311 (0), 284 (−71 000).

(*S*)-6-(1*H*-Indol-2-ylcarbonyl)-8-(chloromethyl)-3,6,7,8-tetrahydro-1-methylbenzo[1,2-*b*:4,3-*b'*]dipyrrol-4-ol [(*S*)-15b]. The hydrochloride salt (*S*)-**14** from a 50-mg (0.15-mmol) conversion of (*S*)-**13**, was dissolved in 2 mL of anhydrous DMA and the solution stirred under nitrogen in the dark. To this was added 0.15 mmol of indole-2-carboxylic acid and 29 mg (0.15 mmol) of EDC as a solid in one portion. After stirring about 45 min at room temperature, an additional 0.04 mmol of indole-2-carboxylic acid and 7 mg (0.04 mmol) of EDC were added. After an additional 1.5 h, the reaction mixture was partitioned between water and ethyl acetate. The aqueous layer was separated and reextracted with ethyl acetate. The combined organic layers were dried over sodium sulfate and concentrated in vacuo. Chromatography on silica gel, eluting with 50% acetone–50% hexane, afforded 50 mg (88%) of (*S*)-**15b**: UV (methanol) 292 nm (26 000), 316 (sh, 18 000), 330 (sh, 15 000); ^1H NMR (DMSO-d_6) δ 11.68 (s, 1 H), 10.74 (s, 1 H), 9.8 (br s, 1 H), 7.71–7.63 (m, 2 H), 7.50 (d, J = 8 Hz, 1 H), 7.24 (t, J = 8 Hz, 1 H), 7.1–7.05 (m, 3 H), 4.65 (br t, J = 8 Hz, 1 H), 4.53 (d, J = 11 Hz, 1 H), 4.02 (br t, 1 H), 3.90 (dd, J = 11, 2 Hz, 1 H), 3.59 (t, J = 10 Hz, 1 H), 2.36 (s, 3 H); MS (FAB) calcd for $C_{21}H_{18}ClN_3O_2$ *m/e* 379.1087, found 379.1070; R_f 0.47 (40% acetone, 60% hexane, silica). Likewise (*R*)-**15b** was obtained from (*R*)-**13**.

(+)-(7b*R*,8a*S*)- and (−)-(7b*S*,8a*R*)-AB. A 0.11-mmol quantity of (*S*)-**15b** was dissolved in 9 mL of acetonitrile and the solution treated with 2.4 mL of triethylamine and 2.4 mL of water. The resultant solution was stirred 15 min at room temperature and then partitioned between ethyl acetate and water. The water layer was separated and reextracted several times with ethyl acetate. The combined organic layers were dried

over sodium sulfate and concentrated in vacuo leaving (+)-AB in quantitative yield, homogeneous by TLC, and identical in spectral properties with previously described racemic material:[20b] ^{1}H NMR (DMSO-d_6) δ 11.81 (s, 1 H), 11.55 (s, 1 H), 7.68 (d, J = 8 Hz, 1 H), 7.48 (d, J = 8 Hz, 1 H), 7.26 (m, 1 H), 7.17 (s, 1 H), 7.08 (m, 1 H), 6.88 (s, 1 H), 6.70 (s, 1 H), 4.52 (dd, J = 10, 5 Hz, 1 H), 4.42 (d, J = 10 Hz, 1 H), 3.16 (m, 1 H), 2.0 (s, 3 H), 1.97 (m, 1 H), 1.39 (t, J = 4 Hz, 1 H); CD (methanol) 370 nm, br (−4000), 357 (0), 329 (31 000), 303 (0), 284 (25 000), 263 (0).

In like manner, (−)-AB was obtained: CD (DMF) 359 nm, br (9500), 339 (0), 320 (−25 000), 285 (−20 000).

(+)-(7b*R*,8a*S*)-ABC. [While (+)- and (−)-ABC have also been prepared according to the route described above, the actual samples used in this study were prepared by the earlier and longer routes as indicated in Scheme II. Experimental details up to the preparation of intermediate **12** are available in the supplementary material.] To (*S*)-**12** (5.0 g, 8.0 mmol) dissolved in 125 mL of freshly distilled THF and 125 mL of dry methanol, under nitrogen, at 0–5 °C, and protected from light, was added 5.0 g (80 mmol) of dry ammonium formate and 5.0 g of 10% palladium on charcoal. The mixture was stirred for 30 min and then quickly filtered through Celite, rinsing the filter cake with two 25-mL portions of freshly distilled THF. The filtrate was eluted with 500 mL of ethyl acetate, washed twice with 200 mL of water and once with 200 mL of brine, dried (Na_2SO_4), and concentrated to 4.1 g (95%) of (*S*)-**15c**, a pale yellow solid. This solid was dissolved in 350 mL of acetonitrile and 300 mL of freshly distilled triethylamine, under nitrogen, and 500 mL of pure water was added. The mixture was stirred, at room temperature, protected from light, for 30 min, and at 0–5 °C for an additional 30 min. The yellow solid that precipitated was collected by vacuum filtration and dried in vacuo to afford 3.5 g (88%) of (+)-ABC. (CAUTION! Extremely toxic substance. Potential carcinogen.) This substance gave IR, UV, ^{1}H NMR, and mass spectra identical with those of the previously prepared racemic material[20b] and, in addition, gave the following: CD (methanol) 335 nm (28 000), 327 (0), 314 (−60 000), 297 (0), 283 (53 000), 266 (0); $[\alpha]^{25}_D$ = +175° (DMF, 0.54 g/100 mL). Anal. Calcd for $C_{30}H_{23}N_5O_3$: C, 71.84; H, 4.62; N, 13.96. Found: C, 72.02; H, 4.85; N, 14.08.

(−)-(7b*S*,8a*R*)-ABC. (−)-ABC was prepared from **3** by the procedure described previously for its racemic compound.[21b] Its spectra were identical with those of (+)-ABC and racemic ABC, but with the following differences: CD (methanol) 335 nm (−33 000), 326 (0), 314 (51 000), 297 (0), 283 (−54 000), 264 (0). The identical compound was also prepared via (*R*)-**8** (see below).

(+)-(7b*R*,8a*S*)- and (−)-(7b*S*,8a*R*)-AB′C′. Details of the preparation of acids **23** and **24** are available in the supplementary material. Compound (*S*)-**13** (56 mg, 0.165 mmol) was converted to (*S*)-**14** in 5 mL of HCl-saturated ethyl acetate. This was taken up in a solution of distilled DMA containing the dissolved acid **23** (50 mg, 0.165 mmol). Solid EDC (63 mg, 2 equiv) was added and the coupling was allowed to proceed for 3 h. The reaction mixture was cooled in ice, and water was added to precipitate the product. The solid was washed twice with water and dried. Extraction of the aqueous portions with 1:1 ethyl acetate–THF gave a mixture of product and starting acid. The combined crude product was chromatographed on silica gel (80:17:3 methyl ethyl ketone–acetone–water) to yield 50 mg (58%) of (*S*)-**15d**: ^{1}H NMR (acetone-d_6) δ 10.75 (br s, 1 H), 10.07 (br s, 1 H), 7.85 (s, 2 H), 7.45 (d, 1 H, J = 9 Hz), 7.10 (m, 2 H), 4.8–3.0 (m, CH_2CH_2 + H_2O), 2.40 (s, 3 H), 1.57 (s, 9 H); MS (FAB) calcd for $C_{28}H_{29}ClN_4O_4$ m/e 521.1955, found 521.1911; (M + H$^+$) 520, 465, 421, 285, 237, 236, 229, 201. Compound (*S*)-**15d** (19 mg, 0.0366 mmol) was deprotected in 1.5 mL of HCl-saturated ethyl acetate. The salt **15e** was dissolved in a solution of 0.9 mL of DMA containing 9 mg (0.037 mmol) of **24**. Solid EDC (21 mg, 3 equiv) was added and the mixture was stirred under nitrogen for 2 h. (The reaction was followed by TLC densitometry.) Product (*S*)-**15f** was precipitated with water, collected by centrifugation, and washed with water twice. Extraction of the aqueous portions with ethyl acetate provided further (*S*)-**15f** and small amounts of **24**. The combined precipitate and extract were dissolved in DMF, adsorbed onto silica gel, and placed on a silica gel column. DMF eluted 16 mg of a mixture of (*S*)-**15f** and (+)-AB′C′. The mixture was suspended in 2.0 mL of acetonitrile, 0.5 mL of triethylamine, and 0.5 mL of water for 30 min under nitrogen in the dark. The reaction mixture was cooled on ice, water was added, and the cyclized product was collected by centrifugation. The pellet was washed with water and dried to give 12.6 mg of (+)-AB′C′ as a brown solid. Extraction of the aqueous portions with 1:1 ethyl acetate–THF provided an additional 3.4 mg to give a 60% yield of (+)-AB′C′ from (*S*)-**15f**: IR (KBr) 3398, 1622, 1612, 1582, 1505, 1431, 1394, 1363, 1336, 1264 cm^{-1}; UV (methanol) 318 nm (40 000), 355 (34 000); ^{1}H NMR (DMF-d_7, verified by 500-MHz COSY) δ 11.82 (s, 1 H, NH), 11.52 (s, 2 H, NH), 8.39 (br s, ArH), 8.14 (d, 1 H, J = 9 Hz), 7.48 (d, 1 H, J = 9 Hz), 7.38 (d, 1 H, J = 9 Hz), 7.26 (s, 1 H), 7.08 (s, 1 H), 6.99 (s, 1 H), 6.79 (s, 1 H), 6.15 (s, 1 H), 4.74 (t, J = 8 Hz, 2 H), 4.57 (m, 2 H), 4.15 (t, J = 8 Hz, 2 H), 3.53 (t, J = 9 Hz, 2 H), 3.40 (t, J = 9 Hz, 2 H), 3.24 (m, 1 H), 2.07 (s, 3 H), 2.04 (m, 1 H), 1.49 (m, 1 H); MS (FAB) calcd for $C_{35}H_{29}N_7O_4$ m/e 612.2359, found 612.2385; R_f 0.32 (80% methyl ethyl ketone, 17% acetone; 3% H_2O, silica gel), R_f 0.61 (70% DMF, 30% saline LKC18D Linear K reversed phase). (+)-AB′C′: $[\alpha]^{25}_D$ = +90° (DMA, 0.1 g/100 mL); CD (DMF) 385–335 nm, br (7000), 328 (0), 315 (−20 000), 302 (0), 287 (35 000). An analogous route afforded (−)-AB′C′: $[A]^{25}_D$ = −80° (DMA, 0.1 g/100 mL); CD (methanol) 385–335 nm, br (−5000), 331 (0), 315 (44 000), 298 (0), 287 (−44 000).

In a separate preparation, (*S*)-**15f** was isolated: ^{1}H NMR (DMF-d_7) δ 11.69 (br s, 1 H), 11.51 (br s, 1 H), 10.85 (br s, 1 H), 8.33 (br s, 1 H), 8.09 (d, 1 H, J = 9 Hz), 7.72 (br s, 1 H), 7.44 (d, 1 H, J = 9 Hz), 7.34 (d, 1 H, J = 9 Hz), 7.12 (s, 1 H), 7.11 (s, 1 H), 7.05 (s, 1 H), 6.15 (s, 2 H), 4.1–3.3 (m, 4 H), 2.38 (s, 3 H); MS (FAB) calcd for $C_{35}H_{30}ClN_7O_4$ m/e 648.2126, found 648.2115; R_f 0.28 (70% DMF, 30% saline, LKC18-D Linear K reversed phase).

(*S*)-15g and (*R*)-15g. The PDE(I) dimer carboxylic acid[31] obtained from the alkaline hydrolysis of the fermentation product, (+)-CC-1065 (36.5 mg, 0.07 mmol) was dissolved in 1.5 mL of dry DMA and the solution added to the amine salt (*S*)-**14** from a 24-mg (0.07-mmol) conversion of (*S*)-**13**. To this solution was added EDC (13.3 mg, 0.07 mmol) and the reaction was stirred at room temperature under nitrogen for 2 h. The reaction was then quenched with water and centrifuged. The pellet was washed with water and dried overnight under vacuum. The resultant solid was adsorbed onto 1 g of silica gel from DMF and placed on top of a 7-g silica gel column. The column was eluted with (8:17:75) DMF–acetone–methylene chloride and the product, (*S*)-**15g**, was isolated: 28 mg, 54% yield; IR (mull) 3341, 3208, 1638, 1610, 1566, 1516, 1331, 1169, 1024 cm^{-1}; UV (ethanol) 260 nm (31 000), 300 (26 000), 360 (41 000); ^{1}H NMR (DMSO-d_6) δ 12.93 (s, 1 H), 11.37 (m, 2 H), 10.93 (s, 1 H), 10.72 (s, 1 H), 9.80 (s, 1 H), 7.6 (br s, 1 H), 7.1–6.88 (m, 5 H), 4.79–4.41 (m, 4 H), 4.12–3.93 (m, 3 H), 3.90 (s, 3 H), 3.93–3.85 (m, 1 H), 3.83 (s, 3 H), 3.58 (t, 1 H), 3.45–3.27 (m, 4 H), 2.36 (d, 3 H); MS (FAB) calcd for $C_{37}H_{34}ClN_7O_8$ m/e 740.2235, found 740.2204; R_f 0.54 (20% DMF–80% toluene, silica); R_f 0.41 (8% DMF, 17% acetone, 75% methylene chloride, silica). Analogously, (*R*)-**15g** was prepared from (*R*)-**13** and showed ^{1}H NMR, UV, and mass spectra and chromatographic properties identical with (*S*)-**15g**.

(+)-CC-1065 and (−)-CC-1065. A 58-mg (0.078-mmol) quantity of (*S*)-**15g** was dissolved in 14 mL of acetonitrile and 4 mL of freshly distilled triethylamine and the solution treated with 4 mL of water. The reaction was kept in the dark and stirred 60 min at room temperature. The resultant precipitate was collected by filtration and dried under vacuum, giving 37 mg (67% yield) of (+)-CC-1065 as a yellow solid. The filtrate from above was diluted with ethyl acetate and washed twice with water and once with saturated brine. The organic layer was dried ($MgSO_4$) and concentrated under vacuum giving an additional 11 mg of (+)-CC-1065. The ^{1}H NMR and UV spectra and chromatographic retention were identical with those of natural CC-1065 obtained from fermentation.[19] In addition, $[\alpha]^{25}_D$ = +98° for synthetic (+)-CC-1065 and +97° for natural CC-1065 (DMF, 0.2 g/100 mL). Natural and synthetic (+)-CC-1065 also gave identical CD spectra (DMF): 385 nm, br (9000), 365 (0), 330 (−25 000), 317 (0), 295 (40 000). Reaction of (*R*)-**15g** as above afforded (−)-CC-1065 in comparable yield: $[\alpha]^{25}_D$ = −96° (DMF, 0.2 g/100 mL); CD (DMF) 385 nm, br (−2000), 377 (0), 330 (32 000), 317 (0), 295 (−43 000). This material had ^{1}H NMR and UV spectra and chromatographic properties identical with (+)-CC-1065.

Compounds Used in Alkylation Sequence Specificity Studies. (+)-CC-1065 isolated from the fermentation broth of *S. zelensis*[8,19] was used in the sequencing experiments. All of the other compounds were synthetic analogues (Chart I) synthesized at The Upjohn Co. as described above.

DNA Fragments. The 117 base pair *Msp*I–*Bst*NI fragment of M13mp1 bacteriophage DNA was isolated as described previously.[16] The 180 base pair fragment of M13mp1 DNA was isolated by successive restriction endonuclease treatments with *Hinp*I and *Alu*I. Similar methods were used to isolate 118 and 276 base pair fragments from SV40 DNA using *Mbo*I and *Hinf*I.

3′- and 5′-^{32}P-End Labeling. 3′- and 5′-^{32}P-end-labeling methods were as described previously.[16,23]

Modification of DNA with CC-1065 and Analogues. Various 10-fold dilutions of a 280 μM stock solution of CC-1065 or analogues were added to aliquots of 3′- or 5′-^{32}P-labeled DNA fragments and incubated at 37 °C for 2 h, followed by ethanol precipitation to terminate the reaction.

MPE-Fe(II) Footprinting and Thermal Breakage of CC-1065-Modified Fragments. MPE·Fe(II) footprinting and heat-induced breakage of DNA fragments containing CC-1065–DNA adducts were performed as described.[10b,16]

Characterization of Base Specificity of Alkylation by (+)- and (−)-CC-1065. CHO (J-116) cells were grown in Ham's F-10 cell culture medium supplemented with glutamine and 15% fetal calf serum. Specific base-labeled DNA was obtained by growing the cells in the presence of (2,8-$^3H_2$)adenine, (*methyl*-$^3H_3$)thymidine, deoxy(5-^{3}H)cytidine, or (8-^{14}C)guanosine hydrochloride (Amersham Corp). DNA from the cells was isolated, after addition of carrier calf thymus DNA, by phenol extraction and ribonuclease treatment.[52] Separate portions of DNA were then reacted with (+)-CC-1065 or (−)-CC-1065. Specific base-labeled DNA with or without bound drug was extracted at 100 °C in a butanol emulsion as described previously.[10a] The amount of base label extracted into the organic layer for drug-treated samples was compared to extracts of labeled unreacted DNA.

In preliminary experiments, the recovery of drug–DNA chromophore in butanol extracts was determined by reacting 0.1 μmol of (−)-CC-1065 with 1.0 μmol of calf thymus DNA at 37 °C overnight in pH 7.4 phosphate buffer. The adducted DNA was precipitated with high salt and ethanol and rinsed with 70% ethanol to remove unreacted drug. The pellet was resuspended in 1.0 mL of pH 7.4 phosphate buffer and absorbance at 365 nm was read. Water-saturated butanol (1 mL) was added to the aqueous solution and this emulsion was heated at 100 °C for 45 min. After cooling and centrifugation to clear the two phases, each phase was transferred to a cuvette and absorbance at 365 nm read. Results showed that 85% of the total drug-associated absorbance equivalents were recovered in the butanol extract. To assure that the (−)-CC-1065 adduct had a similar extinction coefficient in an aqueous phosphate buffer as in organic butanol, 6 μg of the adduct was dissolved in 1 mL of phosphate-buffered saline and centrifuged to precipitate any adduct not in solution. The supernatant was drawn off and absorbance read at 365 nm. Extinction coefficients were not significantly different.

Isolation and Characterization of Adduct. (+)-CC-1065 and (−)-CC-1065 (52 mg, 72 μmol) in 26 mL of dimethylformamide (DMF) were reacted separately with calf thymus DNA (1 g) in 500 mL of 10 mM sodium phosphate buffer, pH 7.4. After precipitation with sodium acetate and ethanol, the DNA was dissolved in 500 mL of 10 mM sodium phosphate buffer, pH 7.4, and heated at 100 °C with 500 mL of 1-butanol. The sample was shaken frequently to emulsify. At 25 min, the phases were separated and the aqueous phase was further heated with 400 mL of 1-butanol for 25 min. The butanol extracts were combined and analyzed for CC-1065 chromophore.

The combined butanol extracts were dried at 40 °C in vacuo and the residue was dissolved in 7.7 mL of 70% DMF and applied to a 2 × 30 cm open column of C-18 stationary phase in 70% DMF. After the residue was washed into the column with 4 mL of 70% DMF, the column was eluted at about 0.45 mL/min with 70% DMF. Fractions (9 mL) were taken, and aliquots of those with significant absorbance at 365 nm were examined by chromatography on Whatman CS5 Multi-K reversed-phase thin-layer plates in 70% DMF. High-pressure liquid chromatography was then employed to isolate the adduct for NMR analysis.

All NMR spectra were obtained on a Bruker AM-500 spectrometer equipped with an Aspect-3000 computer and digital phase shifters. Proton spectra (both 1-D and 2-D) were obtained with the following typical parameters. 1-D: sweep width, 10000 Hz; pulse width, 3 μs, data table, 64K; relaxation delay, 1 s. All FIDs were resolution-enhanced with a Gaussian function, then Fourier transformed. All ^{13}C spectra (1-D and DEPT) were obtained at 125.76 MHz with Waltz decoupling.[53] 2-D: correlated spectroscopy (COSY) was obtained by using the sequence of Ave et al.[54]

A total of 512 FIDs (64 scans each) were collected by using 2K of complex points in T_2 while incrementing t_1 from 2 to 1024 μs in steps of 2 μs. The time domain data were treated with sine-bell windows and Fourier transformed to yield a final real matrix of 1K × 1K data points.

Heterocorrelated spectroscopy was performed by use of the pulse sequence of Bodenhausen and Freemen[55] with a final matrix of 4K × 1K data points. A total of 256 FIDs (96 scans each) were collected in t_2. The time domain spectra were multiplied in both dimensions by a sine-bell window (phase shifted by $\pi/2$ in t_2 and $\pi/8$ in t_1).

Long-range heterocorrelated spectra (COLOC) were collected by use of the sequence of Kessler et al.[56]

DNA Melting Temperature. DNA melting temperature experiments were performed on a Perkin-Elmer Lambda 7 equipped with a Digital temperature programmer and thermoelectric cells. The solvent for the interaction studies was aqueous buffer, 0.01 M sodium phosphate, pH 7.2. Mixtures for spectral analysis were prepared by injecting the appropriate amount of a concentrated solution of (+)-CC-1065 or one of its analogues dissolved in dimethylformamide or dimethylacetamide into a predetermined volume of aqueous buffer containing the calf thymus DNA. The final drug concentration was 0.85 × 10^{-5} M, and the concentration of DNA, calculated from the absorbance spectra by using E_{260} = 6600, was 11 × 10^{-5} M, giving a molar DNA to drug ratio of 13.

Acknowledgment. Supported in part by a grant from the U.S. Public Health Service (CA-49751) and the Robert A. Welch Foundation (F890). We gratefully acknowledge Upjohn colleagues L. H. Li for conducting in vitro and in vivo antitumor studies, G. L. Petzold and M. G. Williams for carrying out the DNA base-specific alkylation experiments and for isolating and purifying the (−)-CC-1065–(N^3-adenine)–DNA adduct, W. C. Krueger for conducting the DNA melting temperature studies, D. G. Martin for the PDE(1) dimer carboxylic acid used in the synthesis of (+)- and (−)-CC-1065, and P. A. Aristoff, M. F. Lipton, and A. Scott for contributions to the syntheses.

Supplementary Material Available: DNA alkylation sites for (+)- and (−)-CC-1065 (as presented in Chart II) on the three other restriction fragments included in this analysis and experimental details for the syntheses of **4**, **2**, (*S*)-**10**, (*S*)-**12**, **18**, **19**, **23**, and **24** (12 pages). Ordering information is given on any current masthead page.

(52) Frei, J. V.; Swenson, D. H.; Warren, W.; Lawley, P. D. *Biochem. J.* **1978**, *174*, 1031.

(53) Levitt, M. H.; Freeman, R.; Frenkiel, T. *Adv. Magn. Reson.* **1983**, *11*, 47.

(54) Ave, W. P.; Bartholdi, E.; Ernst, R. R. *J. Chem. Phys.* **1976**, *64*, 2229.

(55) Bodenhausen, G.; Freeman, R. *J. Magn. Reson.* **1977**, *28*, 471.

(56) Kessler, H.; Griesinger, C.; Zarbok, J.; Loosli, H. R. *J. Magn. Reson.* **1984**, *57*, 331.