Molecular Basis for Sequence-Specific DNA Alkylation by CC-1065†

Laurence H. Hurley,*,‡ Chong-Soon Lee,‡ J. Patrick McGovren,‡,§ Martha A. Warpehoski,∥ Mark A. Mitchell,∥ Robert C. Kelly,∥ and Paul A. Aristoff∥

Division of Medicinal Chemistry, College of Pharmacy, The University of Texas at Austin, Austin, Texas 78712, and Cancer and Viral Diseases Research, The Upjohn Company, Kalamazoo, Michigan 49001

Received January 8, 1988; Revised Manuscript Received March 10, 1988

ABSTRACT: CC-1065 is a potent antitumor antibiotic that binds covalently to N3 of adenine in the minor groove of DNA. The CC-1065 molecule is made up of three repeating pyrroloindole subunits, one of which (the left-hand one or A subunit) contains a reactive cyclopropyl function. The drug reacts with adenines in DNA in a highly sequence-specific manner, overlapping four base pairs to the 5′-side of the covalently modified base. Concomitant with CC-1065 covalent binding to DNA is an asymmetric effect on local DNA structure which extends more than one helix turn to the 5′-side of the covalent binding site. The DNA alkylation, sequence specificity, and biological potency of CC-1065 and a select group of trimeric synthetic analogues were evaluated. The results suggest that (a) noncovalent interactions between this series of compounds and DNA do not lead to the formation of complexes stable enough to be detected by footprinting methods, (b) sequence specificity and alkylation intensity can be modulated by the substituents on the nonreactive middle and right-hand segments, and (c) biological potency correlates well with ability to alkylate DNA. In addition, the extent and the sequence specificity of covalent adduct formation between linear DNA fragments and three analogues comprised of the CC-1065 alkylating subunit linked to zero (analogue A), one (analogue AB), or two (analogue ABC) nonreactive indole subunits were compared. The results suggest that specificity of covalent binding of this analogue series is controlled not by the noncovalent interactions of the B and C subunits with the minor groove but by sequence-dependent reactivity of adenines with the alkylating (A) subunit. However, the B and C subunits markedly increase the apparent rate constant of the reaction with "susceptible" adenines, suggesting that these moieties facilitate noncovalent interactions preceding covalent bond formation. Covalent binding of the analogue consisting only of the alkylating subunit of CC-1065 (analogue A) was associated with the same large asymmetric effect on DNA structure as the entire CC-1065 molecule. This altered local DNA structure could be a consequence of adduct formation. Alternatively, it may be indirect evidence of a particular DNA conformation which existed prior to covalent bond formation and which was "trapped" by the drug. It is proposed that certain adenine-containing sequences have an increased propensity to undergo such a local conformational change and that this is the molecular basis for sequence specificity of these DNA-reactive compounds. These results provide strong experimental evidence for the importance of sequence-dependent site reactivity, rather than noncovalent minor groove interactions, in determining the alkylation specificity of some DNA-reactive molecules.

The molecular mechanisms for sequence-specific recognition of DNA by small molecules and proteins have received considerable attention in the last few years. X-ray crystallographic studies on the *Eco*RI restriction endonuclease (McClarin et al., 1986) and various repressors [see references 1–6 in McClarin et al. (1986)] have revealed that hydrogen bonding is the primary mechanism for sequence recognition in the major groove of DNA. In contrast, the primary mechanism for DNA sequence recognition of minor groove binding sites by small molecules such as the netropsin/distamycin group is proposed to occur by close van der Waals contacts (Kopka et al., 1985), although more recent work has demonstrated that there may also be other as yet unidentified factors to consider (Lown et al., 1986; Marky & Breslauer, 1987). Furthermore, while a considerable effort has gone into determining the noncovalent interactions that lead to sequence-specific recognition of DNA [see Dervan (1986) and Helene and Lancelot (1982) for reviews], much less is known about the origin of sequence specificity of *covalent* DNA modification, although covalent modification of DNA is proposed to increase the intrinsic sequence specificity of *Eco*RI (McClarin et al., 1986). Examples of small molecules exhibiting sequence specificity in covalent adduct formation include aflatoxin B_1 (Muench et al., 1983), *N*-(bromoacetyl)distamycin (Baker & Dervan, 1985), and nitrogen mustards (Mattes et al., 1986). In the latter case, Mattes et al. noted that, at least within some regions of DNA and for some compounds, sequence-dependent variations in the electrostatic potential at the N7 position of guanine correlated with the alkylation intensity. The studies described in this paper provide strong experimental evidence for the importance of sequence-dependent site reactivity in mediating the alkylation specificity of some small molecular weight molecules. Some of these results have appeared in preliminary from (Wierenga et al., 1986; Lee & Hurley, 1986).

CC-1065 is an antitumor antibiotic produced by *Streptomyces zelensis* [see Hanka and Martin (1978) and Martin et al. (1980), and for reviews, see Reynolds et al. (1986) and Hurley and Needham-VanDevanter (1986)]. Its structure, shown in Figure 2, is characterized by three repeating pyrroloindole subunits, one of which (the left-hand one or A subunit) contains a reactive cyclopropyl function. Naturally occurring CC-1065 and analogues sharing the same stereochemical configuration of the cyclopropane ring are termed (+)-enantiomers. (+)-CC-1065 was shown to be modestly

† Supported in part by the Welch Foundation and by Grant CA-30349 from the U.S. Public Health Service.

* To whom correspondence should be addressed.
‡ The University of Texas at Austin.
§ On sabbatical leave from The Upjohn Co., Kalamazoo, MI.
∥ The Upjohn Co.

0006-2960/88/0427-3886$01.50/0

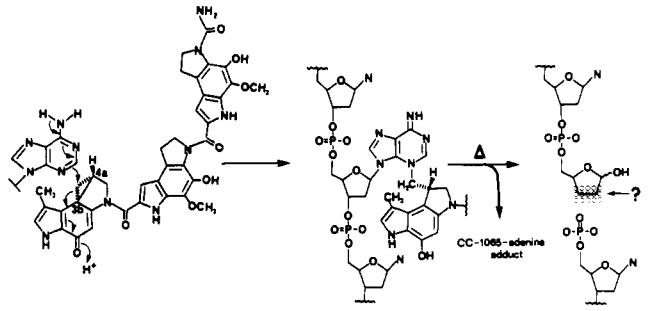

FIGURE 1: Reaction of (+)-CC-1065 with N3 of adenine in DNA and products from thermal cleavage reaction. The extract nature of the species generated on the 5′-side of the strand break is unknown.

efficacious in the treatment of several experimental tumors in mice (Reynolds et al., 1986; Martin et al., 1981). It was also active in the in vitro human tumor cloning assay against a broad spectrum of tumor types with a potency at least 100 times greater than that of doxorubicin (Bhuyan et al., 1982). However, (+)-CC-1065 also produced delayed death in mice at therapeutic doses (McGovren et al., 1984), a finding which precluded its clinical development. Synthetic analogues have recently been prepared which are more efficacious, which do not cause delayed death, and which, therefore, show considerable promise as chemotherapeutic agents (Warpehoski, 1986; Li et al., 1987).

Previous work established that (+)-CC-1065 binds covalently to DNA through N3 of adenine and lies snuggly within the minor groove, covering a four base pair region to the 5′-side of the covalently modified adenine (Hurley et al., 1984). 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 (Reynolds et al., 1985) (Figure 1). Using this strand breakage assay, 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 (Reynolds et al., 1985). The construction of a site-directed (+)-CC-1065-(N3-adenine)-DNA adduct in a 117 base pair fragment (Needham-VanDevanter & Hurley, 1986) allowed us to determine the effect of covalent attachment on local DNA structure. DNase I footprinting and restriction enzyme analysis demonstrated that (+)-CC-1065 adduct formation correlates with an asymmetric effect on DNA structure which extends more than one helix turn to the 5′-side of the covalent binding site (Hurley et al., 1987).

To determine the relative importance of covalent vs. noncovalent interactions in mediating the DNA sequence specificity of (+)-CC-1065, we have examined the degree of covalent adduct formation and sequence specificites of selected synthetic CC-1065 analogues (Warpehoski, 1986; Kelly et al., 1987; Warpehoski et al., 1988; Warpehoski & Bradford, 1988). The structures of the synthetic compounds are shown in Figure 2. Analogues containing the alkylating (cyclopropane ring containing) subunit linked to zero, one, or two indole subunits are designated (+)-A, (+)-AB, and (+)-ABC, respectively. An analogue containing the pyrroloindole moiety linked to two indole subunits but lacking the DNA-reactive cyclopropane ring of CC-1065 is designated des-ABC. An analogue with the structure of CC-1065 but lacking the hydroxyl and methoxyl substituents in the B and C subunits is termed (+)-AB′C′.

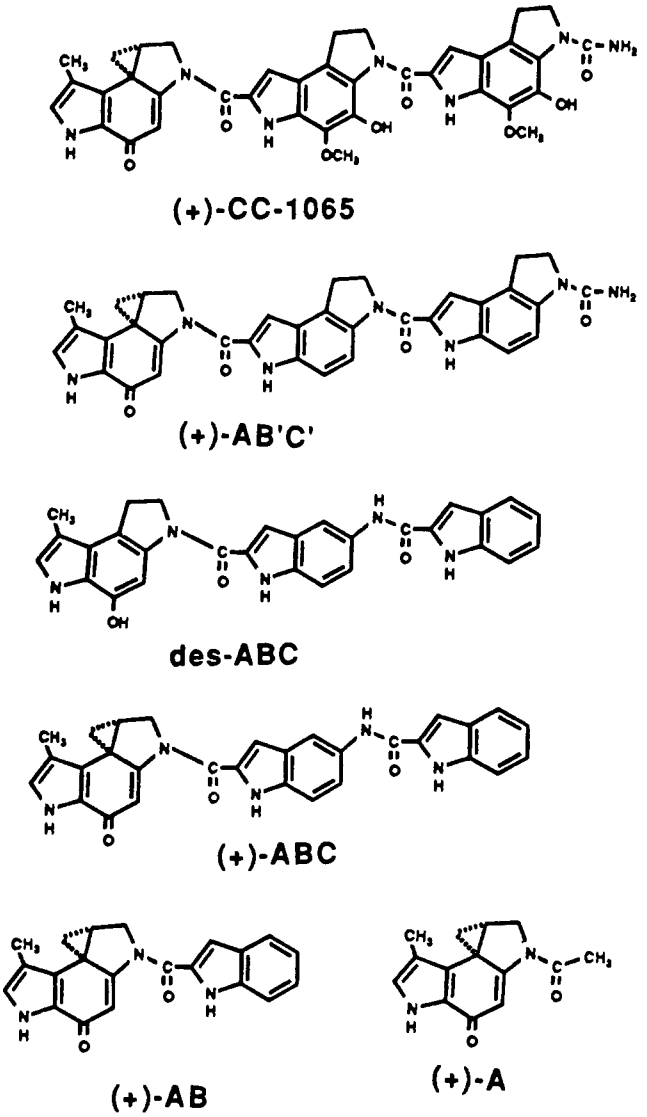

FIGURE 2: Structure of (+)-CC-1065 and its synthetic analogues used in this study.

Experimental Procedures

Materials

The synthetic analogues used in this study were prepared as previously described (Warpehoski, 1986; Kelly et al., 1987; Warpehoski et al., 1988; Warpehoski & Bradford, 1988). Electrophoretic reagents [acrylamide, TEMED,[1] ammonium persulfate, and bis(acrylamide)] were purchased from Bio-Rad. Distamycin, DNase I, BAP, dNTPs, and lysozyme were from Sigma. All other enzymes [restriction endonucleases, T4-PNK, and DNA polymeraseI (large fragment)] were from New England Biolabs. M13mp1 vector was obtained from Professor J. Messing, University of Minnesota. MPE was a generous gift of Professor P. Dervan, California Institute of Technology. M13mp1 RF DNA was isolated from infected *Escherichia coli* JM 101 cells by $CsCl_2$ ultracentrifugation as described by Messing (1983). [γ-^{32}P]- and [α-^{32}P]ATP were from ICN. X-ray film, intensifying screens, and developing chemicals were from Kodak. SV40 DNA was from

[1] Abbreviations: MPE, methidiumpropyl-EDTA; ip, intraperitoneal; O.D., optimum dose; DTT, dithiothreitol; TEMED, *N,N,N′,N′*-tetramethylethylenediamine; BAP, bacterial alkaline phosphatase; DNase I, bovine pancreatic deoxyribonuclease I; dNTP, deoxyribonucleoside triphosphate; T4-PNK, T4 polynucleotide kinase; EDTA, ethylenediaminetetraacetic acid.

Bethesda Research Laboratories.

Methods

Isolation of ^{32}P-Labeled DNA Fragment. 5′- or 3′-^{32}P-labeled M13mp1 117-bp DNA fragment was isolated as described previously (Needham-VanDevanter et al., 1986). 5′-^{32}P-labeled SV40 118-bp DNA fragment was prepared by successive restriction endonuclease digestion with *Mbo*I and *Hin*fI.

Drug Binding and Determination of Covalent Binding Sites. Varying drug concentrations were incubated with labeled DNA fragments, and covalent binding sites were identified by thermal treatment to induce strand breakage (Reynolds et al., 1984).

MPE·Fe(II) and DNase I Footprinting. MPE·Fe(II) and DNase I footprinting were performed as described previously (Hurley et al., 1987).

Polyacrylamide Gel Electrophoresis. Analysis of thermal strand breaks and of MPE·Fe(II) and DNase I footprinting was performed by electrophoresis of products on 8% denaturing polyacrylamide gels adjacent to Maxam–Gilbert DNA sequencing reactions.

Densitometric Analysis of Autoradiograms. Autoradigrams were scanned with a laser densitometer (LKB 2202) coupled to a recording integrator (LKB 2220).

Results and Discussion

Comparison of DNA Binding and Sequence Specificity of (+)-CC-1065 and Des-ABC with Those of Distamycin. We used the MPE footprinting technique to determine if des-ABC, which lacks the alkylating function, could nonetheless bind to DNA. (+)-CC-1065 itself and distamycin were footprinted under the same conditions. The sequence specificity of the distamycin/netropsin group has been previously examined by footprinting techniques (VanDyke & Dervan, 1983; Fox & Waring, 1984).

While distamycin shows four clear MPE·Fe(II) footprints in the interpretable part of the gel (I–IV in lane 5 of Figure 3a), des-ABC did not reveal footprints under the same conditions (compare lanes 1 and 2 of Figure 3a). (+)-CC-1065 also shows four MPE·Fe(II) footprints, of which three (A, C, and D) reflect covalent binding sites on the same strand (arrows A′, C′, and D′). The other footprint (B) corresponds to a covalent binding site on the opposite strand. The relative extents of binding and precise binding sites of distamycin and (+)-CC-1065 are shown in Figure 3b.

Distamycin, which is known to interact noncovalently with DNA through van der Waals contacts, hydrogen bonds, and electrostatic forces (Kopka et al., 1985; VanDyke & Dervan, 1983), yields clear footprints. However, for CC-1065, only those sites which react covalently with DNA are revealed by MPE footprinting. We could find no evidence for noncovalent (reversible) interaction of CC-1065 at other sites on this DNA fragment. If such noncovalent binding is occurring, it evidently is too weak to displace the MPE reagent (in contrast to the noncovalent binding of distamycin). Consistent with this was the absence of any footprinting by des-ABC, which is not capable of covalent adduct formation. While, sterically, this molecule might be expected to fit well in the minor groove in A-T regions which lack the protruding 2-amino group of guanine, it lacks the cationic and hydrogen-bonding capabilities which enhance the binding affinity of distamycin. Des-ABC also lacks significant biological potency in comparison to (+)-CC-1065 and other analogues (Table I).

Biological Activity and DNA Sequence Specificity of (+)-A, (+)-AB, and (+)-ABC. The (+)-A and (+)-AB compounds

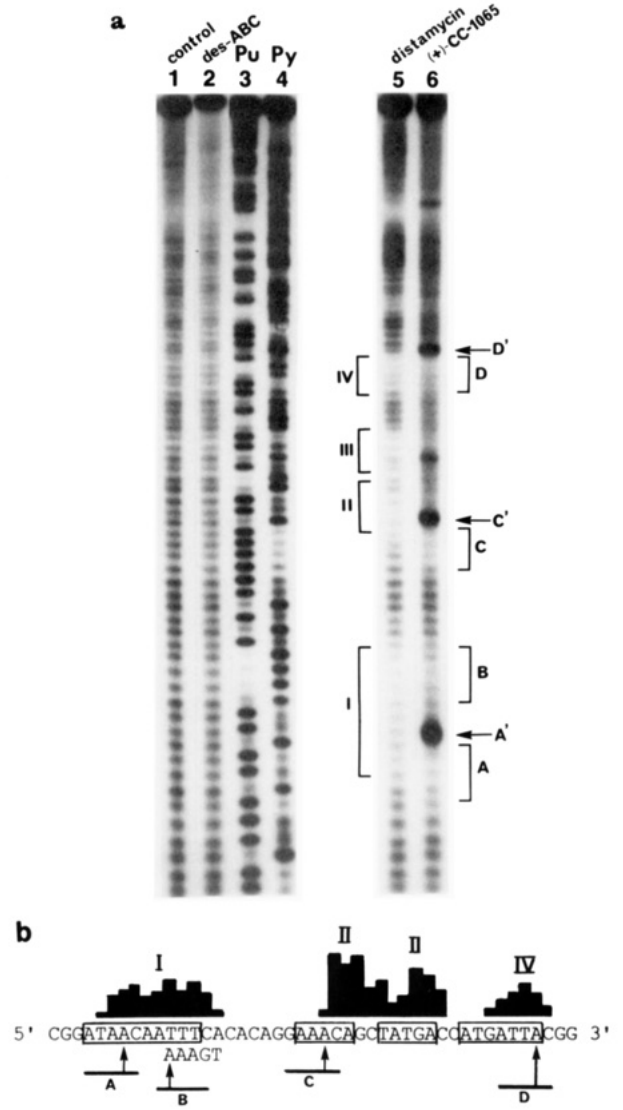

FIGURE 3: (a) MPE·Fe(II) footprinting of des-ABC, distamycin, and (+)-CC-1065-modified 117-bp *Msp*I–*Bst*NI fragments from M13mp1 DNA. Single 5′-^{32}P-labeled aliquots of M13 mp1 117-bp DNA fragments were modified with 2.8 μM solutions of des-ABC (lane 2), distamycin (lane 5), and (+)-CC-1065 (lane 6). For MPE·Fe(II) digestions (lanes 1, 2, 5, and 6), DNA samples were suspended in 16 μL of 50 mM NaCl and 10 mM Tris-HCl, pH 7.4, and 2 μL of 100 μM MPE and 100 μM $Fe(NH_4)_2(SO_4)_2$ and equilibrated for 15 min at room temperature. DTT (2 μL of 20 mM) was added and incubated for 15 min at room temperature, and samples were lyophilized to dryness. DNA residues were suspended in 10 μL of alkaline tracking dye for electrophoresis. Reactions were electrophoresed adjacent to Maxam–Gilbert purine and pyrimidine lanes (lanes 3 and 4). Brackets correspond to the MPE·Fe(II) footprints for distamycin (I–IV) and (+)-CC-1065 (A–D), and arrows (A′, C′, and D′) correspond to the covalent binding sites for (+)-CC-1065. (b) Diagrammatic representation of the distamycin and (+)-CC-1065 binding sites and relative affinities in the readable portion of the sequencing gel shown in (a). The sequences enclosed by boxes and the histogram bars represent the position and relative inhibition of MPE·Fe(II) cutting for the distamycin binding sites on the 117-bp fragment of DNA. Data are taken from lanes 1 and 5 in (a) which were scanned by an integrating laser densitometer. Following normalization of total lane areas, the relative inhibition of cutting was calculated and plotted on the histogram such that maximum inhibition is given the highest histogram rating. The 3′-offset of binding sites (boxes) relative to the inhibition zones (histogram bars) is in accord with the model for asymmetric footprint patterns. Distamycin footprints I–IV correspond to zones of inhibition I–IV in (a). The bars A–D below the DNA sequence correspond to the (+)-CC-1065 alkylation sites shown in (a). The length of the arrows corresponds to the relative sensitivites of sites A–D for (+)-CC-1065, and the arrow heads point to the covalently modified adenines. The longest arrows represent the most sensitive sites. Since the (+)-CC-1065 covalent binding site corresponding to B is on the strand opposite to that which is 5′-^{32}P-labeled, the complementary sequence is shown for this site.

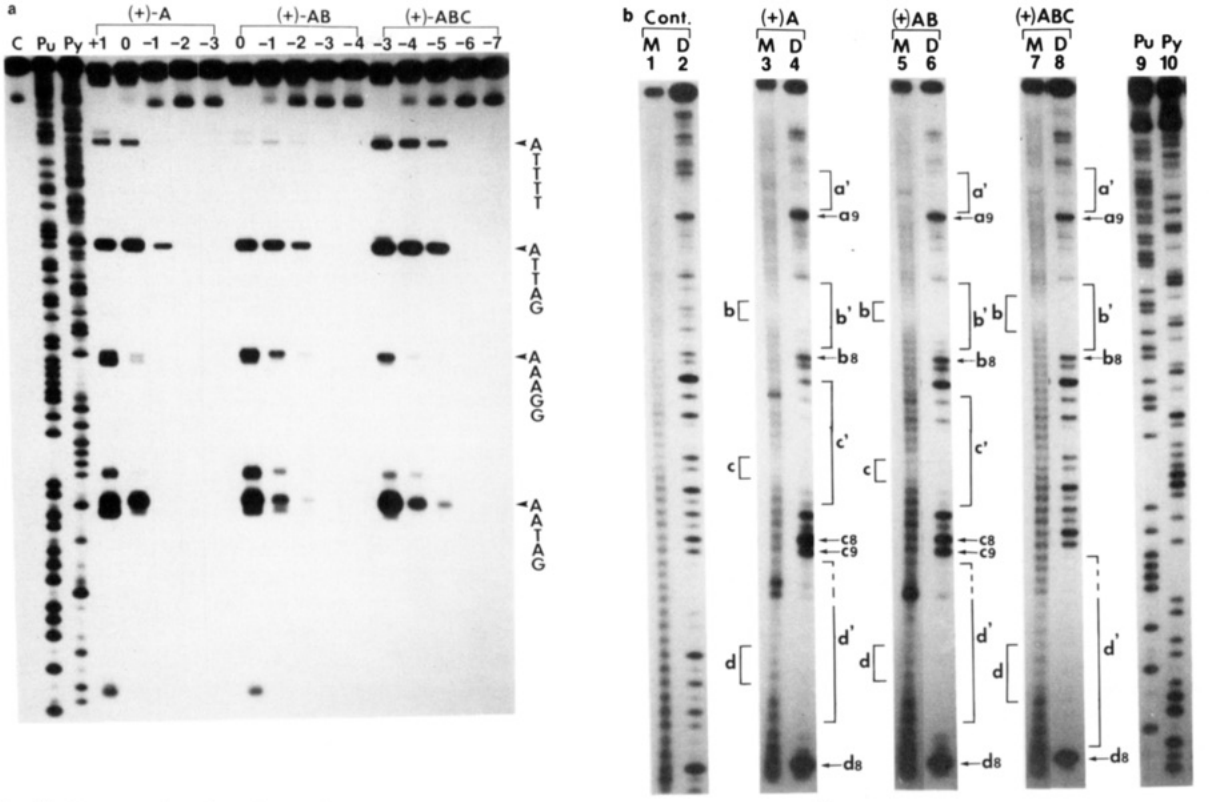

FIGURE 4: (a) Concentration dependency of alkylation by (+)-A, (+)-AB, and (+)-ABC of the 5′-^{32}P-labeled (+) strand in the 117-bp *Msp*I–*Bst*NI fragment of M13 mp1 DNA. Aliquots of single 5′-^{32}P-labeled 117-bp DNA fragments were modified with 1×10^n ($n = +1$ through -7) dilutions (labeled +1 through −7) of 280 μM stock of (+)-A, (+)-AB, or (+)-ABC solution. Modified DNA was heated at 90 °C for 30 min in DSC buffer and electrophoresed adjacent to Maxam–Gilbert purine- and pyrimidine-specific DNA cleavage reactions. The drug alkylation sequences are shown, and arrows indicate adenines modified by (+)-A, (+)-AB, and (+)-ABC. (b) MPE·Fe(II) and DNase I footprinting of (+)-A, (+)-AB, (+)-ABC on the 3′-^{32}P-labeled (−) strand in the 117-bp *Msp*I–*Bst*NI fragment of M13 mp1 DNA. Single 3′-^{32}P-labeled 117-bp DNA fragment was prepared as described under Methods. MPE·Fe(II) digestions (lanes 1, 3, 5, and 7) were carried out as described under Methods. DNase I digestions (lanes 2, 4, 6, and 8) were also performed as described previously (Hurley et al., 1987). 3′-^{32}P-labeled fragments were modified with 2.8 mM (+1) (+)-A (lanes 3 and 4), 280 μM (0) (+)-AB (lanes 5 and 6), and 280 nM (−3) (+)-ABC (lanes 7 and 8). Lanes 1 and 2 are control MPE·Fe(II) and DNase I digestions, and all samples were electrophoresed adjacent to Maxam–Gilbert purine- and pyrimidine-specific DNA cleavage reactions (lanes 9 and 10). M = MPE·Fe(II) and D = DNase I. Brackets b–d correspond to MPE·Fe(II) footprints for (+)-A, (+)-AB, and (+)-ABC. The MPE footprints for the "a" binding site revealed by DNase I cannot be confidently located due to overdigestion by MPE in this portion of the gel. For the MPE footprint labeled d in lanes 3 and 5, the presence of an additional covalent binding site on the 5′-side of the bracket increases the apparent size of the MPE inhibition zone. Brackets a′–d′ correspond to DNase I footprints, and arrows denoted a_9, b_8, c_8, c_9, and d_8 correspond to DNase I enhancement sites for (+)-A, (+)-AB, and (+)-ABC. A broken bracket line is drawn for regions where it is expected that a drug-induced inhibition of DNase I should occur, but where DNase I cleavage sites are absent in the control lane. These drug binding regions are revealed, however, by the presence of their corresponding DNase I cleavage enhancement site.

are 3.3×10^4-fold and 4.0×10^2-fold less potent than (+)-ABC in inhibition of L1210 leukemia cell growth (Table I). Nevertheless, at elevated doses, (+)-A and (+)-AB show moderate levels in vivo activity against P388 leukemia, while (+)-ABC produces multiple long-term survivors in the same model. At 10^4- and 10^3-fold higher drug concentrations, the (+)-A and (+)-AB subunits show levels of covalent binding to DNA comparable to that of the (+)-ABC molecule on the 5′-labeled 117 base pair fragment from M13mp1 (Figure 4a). We were surprised to observe that the DNA sequence specificity of the (+)-A and (+)-AB compounds is virtually the same as that of (+)-ABC. A minor difference noted was the less rigid specificity of (+)-A and (+)-AB for the 3′-adenine in the sequence 5′GATTA* in comparison to that of (+)-ABC. Thus it appears that the (+)-A subunit alone has sufficient structural information to mediate the sequence specificity of the entire (+)-ABC molecule; that is, covalent binding specificity appears to be determined by sequence-dependent reactivity of adenines with the alkylating subunit. However, the B and C subunits serve to markedly increase the apparent rate constant of the reaction with "susceptible" adenines, suggesting that these moieties facilitate noncovalent (presumably hydrophobic and van der Waals) interactions preceding covalent bond formation (proximity effect; Menger, 1985; Bruice, 1976).

Table I: Biological Activities of CC-1065 and Its Analogues

	in vitro L1210 cell growth inhibition, ID_{50}[a]		in vivo P388 leukemia[b]	
	nM	rel. to CC-1065	% ILS	O.D. (μg kg^{-1} day^{-1})
(+)-CC-1065	0.03	1	67	100
(+)-AB′C′	0.05	2	77	100
des-ABC	160	5300	0	6000[d]
(+)-ABC	0.004	0.14	4/6[e]	25
(+)-AB	0.1	3	100	250
(+)-A	12	400	45[c]	3000[c]

[a] 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. [b] 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. O.D. = optimum dose. [c] Tested as racemic mixture of (+)-A and (−)-A. [d] Highest dose tested. [e] Number of day 30 survivors/total number mice in test group.

Importance of Close Contacts of B and C Subunits with the Minor Groove in Determining DNA Sequence Specificity. While the noncovalent interactions between DNA and the indolic B and C subunits evidently do not markedly influence

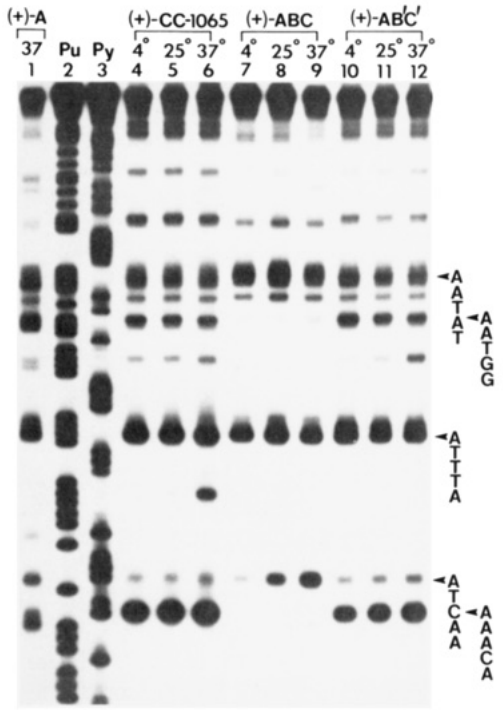

FIGURE 5: Temperature dependency of alkylation by (+)-CC-1065, (+)-ABC, and (+)-AB′C of the 118-bp *Mbo*I–*Hin*fI fragment of SV40 DNA. Aliquots of single 5′-^{32}P-labeled 118-bp fragments were incubated with 2.8 mM (+)-A or 28 μM (+)-CC-1065, (+)-ABC, and (+)-AB′C′ and incubated at 4, 25, or 37 °C for 2 h. Drug-modified DNA was heated at 90 °C for 30 min in DSC buffer and electrophoresed adjacent to Maxam–Gilbert purine- and pyrimidine-specific DNA cleavage reactions. The drug binding sequences are shown, and arrows indicate adenines modified by either (+)-A, (+)-CC-1065, (+)-ABC, or (+)-AB′C′.

alkylation site specificity of the A, AB, and ABC analogue series, the more complex pyrroloindole subunits of CC-1065 have a detectable influence on the selection of certain sites. Figure 5 compares the effect of incubation temperature on the cleavage pattern of (+)-CC-1065, (+)-ABC, (+)-AB′C′, and (+)-A with a 118 base pair fragment from SV40 DNA. While some sites are common to both (+)-CC-1065 and (+)-ABC (5′TATAA, 5′ATTTA), others are unique to (+)-CC-1065 (5′ACAAA, 5′GGTAA). At elevated temperatures (+)-ABC reacts strongly at 5′AACTA, which is a poor binding site for (+)-CC-1065 between 4 and 37 °C. Similar experiments with the 117-mer from M13 mp1 (data not presented) also show more temperature-dependent sequences for (+)-ABC than for (+)-CC-1065. The sequence specificity of (+)-AB′C′, which is intermediate in structure between (+)-CC-1065 and (+)-ABC, was also determined (Figure 5). The sequence specificities of (+)-AB′C′ and (+)-CC-1065 are almost identical. Thus it appears that close contacts (van der Waals or hydrophobic) can fine tune or modulate the sequence specificity of the (+)-A subunit alone (lane 1). In other words, interactions between the inside edge of the drug molecule and the floor of the minor groove may influence the "proximity effect" at certain sites and may cause exclusion from other sites on steric grounds.

DNase I Footprinting Pattern Induced by Covalent Binding of Compounds (+)-A, (+)-AB, and (+)-ABC to DNA. In a previous study we showed that covalent binding of (+)-CC-1065 to DNA is accompanied by an asymmetric effect on local DNA structure of the 5′-side of the covalent binding site (Hurley et al., 1987). The DNase I cleavage pattern around the (+)-CC-1065–DNA adduct shows a 12-nucleotide zone of DNase I inhibition on the *noncovalently modified strand* with a large DNase I enhanced cleavage site at the 3′-terminus of the inhibition zone. Much less pronounced effects were seen on the covalently modified strand. Since the (+)-A subunit alone is sufficient to mediate the sequence specificity of (+)-ABC, it was important to determine whether (+)-A could also mediate the asymmetric effect. A comparison of the MPE·Fe(II) and DNase I footprints of (+)-A, (+)-AB, and (+)-ABC on the noncovalently modified strand of the same 117 base pair fragment seen in Figure 4a is shown in Figure 4b. The MPE·Fe(II) footprints of each covalently bound subunit(s) are shown in the left-hand lanes (odd numbers) while the DNase I footprints are shown in the right-hand lanes (even numbers). Arrows show the position of DNase I enhancement sites. Four (+)-A and (+)-AB binding sites (a′–d′) were revealed by DNase I footprinting with associated enhancement sites (a_8–d_8) in this DNA fragment. A direct comparison of drug binding sites can be made since, although opposite strands of the 117-mer were used, 5′ (Figure 4a) and 3′ (Figure 4b) labels were also used. In accord with the slightly lower band intensity reflecting covalent interaction between (+)-ABC and the sequence 5′GGAAA* (see Figure 4a), MPE·Fe(II) and DNase I footprints were not found at this site. But the other sites (a, b, and d) show the typical DNase I inhibition/enhancement cleavage pattern common to (+)-A, (+)-AB, and (+)-CC-1065 (Hurley et al., 1987). Consequently, (+)-A is not only sufficient to determine the sequence specificity of (+)-ABC but also mediates a similar asymmetric effect on local DNA structure.

Analysis of the Structural Requirements for Covalent Binding and DNA Sequence Recognition by CC-1065. Our findings suggest that the dynamics of the DNA alkylation reaction are of paramount importance in determining the sequence specificity of CC-1065 and its analogues. We suggest further that covalent binding to DNA may require a cost in the form of DNA deformation energy to accommodate the final reaction step and that adenines in sequences which have a low energy barrier to undergo this deformation are favored for reaction with (+)-CC-1065. Therefore, in contrast to noncovalent DNA binding molecules such as repressor proteins and minor groove equilibrium binding agents such as distamycin, certain DNA alkylating molecules may have an alternative or additional mechanism for sequence recognition of DNA. In this regard, the experimentally observed DNA sequence specificity of the pyrrolo[1,4]benzodiazepine antitumor antibiotics (Hertzberg et al., 1986), which bind covalently through N^2 of guanine, has been rationalized by theoretical calculations (Zakrzewska & Pullman, 1986) in which the relative distortion energies of different sequences are predictive of the experimentally determined DNA sequence specificity. Also, *N*-(bromoacetyl)distamycin has been shown to alkylate DNA at only a subset of its noncovalent binding sites (Baker & Dervan, 1985).

While the A subunit plays the primary role in sequence-specific recognition of DNA, the B and C subunits are not completely passive in the DNA recognition and binding processes. First, (+)-ABC and (+)-CC-1065 show subtle differences in sequence specificity which have unknown but possibly important biological consequences, if they occur in vivo.[2] Second, the presence of the B and C subunits in a molecule dramatically increases the apparent reaction rate

[2] In this regard, it is interesting that (+)-CC-1065 and (+)-AB′C′ are only modestly efficacious in tumor-bearing mice (Table I) and both compounds cause delayed death in mice (McGovren et al., 1984; Warpehoski & Bradford, 1988) whereas (+)-ABC is highly efficacious (Table I) and does not cause delayed death (Warpehoski, 1986).

constant for adduct formation. This can best be visualized by means of a set of binding equilibria:

$$[\text{drug}] + [\text{DNA}] \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} [\text{drug·DNA}] \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} [\text{drug·DNA*}] \xrightarrow{k_3} [\text{drug–DNA*}]$$

In this model the k's are rate constants for the forward and reverse reactions, [drug·DNA] symbolizes each noncovalent complex of drug with a five base pair region in the minor groove, [drug·DNA*] indicates each complex in which the local sequence has assumed a conformation favorable for reaction of adenine with drug, and [drug–DNA*] refers to each covalent adduct with a DNA sequence of conformationally altered and "trapped" local structure. Alternatively, [drug·DNA*] may not be an intermediate of finite lifetime but may simply represent a transition state of the covalent reaction. In that case, alkylation would be synchronous with the DNA conformational change, making it irreversible. This model assumes that there are no allosteric interactions between binding sequences.

There are some interesting similarities between this scheme and certain aspects of the proposed mechanism for sequence recognition by the restriction endonuclease *Eco*RI (McClarin et al., 1986). *Eco*RI discriminates between cognate and noncognate sequences by a dissociation of noncognate [protein·DNA] complexes. The noncognate [protein·DNA] complexes are analogous to the [drug·DNA] species, which, because of their nonproductive complexation, can dissociate. McClarin et al. (1986) have suggested that covalently modifying enzymes such as *Eco*RI are intrinsically capable of achieving a higher level of sequence discrimination than a simple binding protein. Our results reported here suggest that this may also apply to certain covalent vs noncovalent DNA-modifying drugs.

Since MPE footprinting experiments with (+)-CC-1065 and des-ABC detected only the [drug–DNA*] species, we may assume that the [drug·DNA] complexes are weak, at least in relation to the MPE–DNA intercalative complex. The B and C moieties apparently increase the k_1/k_{-1} ratio markedly, as judged by the much lower concentrations of (+)-ABC and (+)-CC-1065 required to achieve detectable levels of covalent reaction (strand breakage) relative to (+)-A and (+)-AB. The driving force for this equilibrium binding step (for which spectral evidence exists; Krueger et al., 1985) may be favorable van der Waals contacts between the inside edge of the drug molecule and the floor of the minor groove of DNA, or simply the exclusion of water (hydrophobic effect). That some alkylation sites are different for (+)-CC-1065 and (+)-ABC may reflect steric inhibition to the equilibrium binding of (+)-CC-1065 at some sites. Alternatively, specific favorable noncovalent interactions (van der Waals) may enhance the k_1/k_{-1} ratio for (+)-CC-1065 to a much greater extent than for (+)-ABC at some sites. Certain sequences may have a lower energy barrier for assuming a conformation favoring reaction of adenine with drug. In terms of the model, these sequences would have a relatively high k_2/k_{-2} ratio. A very low k_2/k_{-2} ratio would correspond to sites where covalent reaction does not occur. This model is consistent with data from spectroscopic studies of CC-1065 binding to synthetic DNA polymers (Krueger et al., 1985; Krueger & Prairie, 1987). The model may also be applicable to sequence-specific alkylation by the reactive distamycin derivative *N*-(bromoacetyl)distamycin (Baker & Dervan, 1985).

In summary, three steps are proposed which lead to covalent binding of (+)-CC-1065 to N3 of adenine in DNA. In the first step, by analogy with distamycin, the size and right-hand-twisted banana shape of CC-1065 facilitate minor groove equilibrium binding through the displacement by drug of water molecules along the spine of the groove and through favorable van der Waals contacts. These noncovalently bound complexes, localized in A-T-rich regions, can then undergo the next step. This second step is proposed to involve primarily a conformational change in the local DNA structure, a change which optimizes the distance and geometry of N3 of adenine and the methylene carbon of the cyclopropane. This change may be prerequisite to covalent adduct formation or could occur concomitantly. The conformational change is presumed to be much more facile for some A-T sequences than for others and, therefore, underlies the precise sequence specificity observed. The third and final step is covalent binding.

Conclusions and Biological Significance. The experiments described here provide strong evidence suggesting the prime importance of sequence-dependent reactivity in determining the reaction site specificity of certain covalently binding ligands. The key finding is that the (+)-A subunit of CC-1065 is sufficient to mediate both the sequence specificity and asymmetric effect on DNA structure associated with much larger drug molecules. Since the asymmetric effect of (+)-CC-1065 on local DNA structure extends more than one helix turn to the 5′-side of the covalent binding site (Hurley et al., 1987), we propose that the sequence preference is based upon the relative propensity of certain sequences to undergo a (pre)requisite conformational change.

Covalent adduct formation between drug and naked linear DNA fragments correlated remarkably well with biological potency (Table I), supporting the use of the heat-induced strand breakage assay in assessing the biochemical effects of these compounds.

It is intriguing to speculate that the potent biological effect of (+)-CC-1065 and its analogues may be related either to the propensity of certain biologically important DNA sequences to react or to local effects on DNA structure mimicing a biologically potent form of DNA (Hurley et al., 1987). In this regard, we have previously shown that the most reactive (+)-CC-1065 binding sites so far discovered are adjacent to Sp1 binding sites in the 21 base pair repeats of the SV40 early promoter region (Reynolds et al., 1985).

ACKNOWLEDGMENTS

We gratefully acknowledge Dr. L. H. Li and his staff for providing the in vitro and in vivo testing data.

REFERENCES

Baker, B. F., & Dervan, P. B. (1985) *J. Am. Chem. Soc. 107*, 8266.

Bhuyan, B. K., Newell, K. A., Crampton, S. L., & VonHoff, D. D. (1982) *Cancer Res. 42*, 3532.

Bruice, T. C. (1976) *Annu. Rev. Biochem. 45*, 331.

Dervan, P. B. (1986) *Science (Washington, D.C.) 232*, 464.

Fox, K. R., & Waring, M. J. (1984) *Nucleic Acids Res. 12*, 9271.

Hanka, L. J., Dietz, A., Gerpheide, S. A., Kuentzel, S. L., & Martin, D. G. (1978) *J. Antibiot. 31*, 1211.

Helene, C., & Lancelot, G. (1982) *Prog. Biophys. Mol. Biol. 39*, 1.

Hertzberg, R. P., Hecht, S. M., Reynolds, V. L., Molineaux, I. J., & Hurley, L. H. (1986) *Biochemistry 25*, 1249.

Hurley, L. H., & Needham-VanDevanter, D. R. (1986) *Acc. Chem. Res. 19*, 230.

Hurley, L. H., Reynolds, V. L., Swenson, D. H., Petzold, G. L., & Scahill, T. A. (1984) *Science (Washington, D.C.) 226*, 843.

Hurley, L. H., Needham-VanDevanter, D. R., & Lee, C.-S. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 6412.
Kelly, R. C., Gebhard, I., Wicnienski, N., Aristoff, P. A., Johnson, P. D., & Martin, D. G. (1987) *J. Am. Chem. Soc. 109*, 6837.
Kopka, M. L., Yoon, C., Goodsell, D., Pjura, P., & Dickerson, R. E. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 1376.
Krueger, W. C., & Prairie, M. D. (1987) *Chem.-Biol. Interact. 62*, 281.
Krueger, W. C., Li, L. H., Moscowitz, A., Prairie, M. D., Petzold, G. L., & Swenson, D. H. (1985) *Biopolymers 24*, 1549.
Lee, C.-S., & Hurley, L. H. (1986) *Proc. Am. Assoc. Cancer Res. 27*, 243.
Li, L. H., Wallace, T. L., DeKoning, T. F., Warpehoski, M. A., Kelly, R. C., Prairie, M. D., & Krueger, W. C. (1987) *Invest. New Drugs 5*, 329.
Lown, J. W., Krowicki, K., Bhat, U. G., Skorobogaty, A., Ward, B., & Dabrowiak, J. C. (1986) *Biochemistry 25*, 7408.
Marky, L. A., & Breslauer, K. J. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 4359.
Martin, D. G., Chidester, C. G., Duchamp, D. J., & Mizsak, S. A. (1980) *J. Antibiot. 33*, 902.
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. (1981) *J. Antibiot. 34*, 1119.
Mattes, W. B., Hartley, J. A., & Kohn, K. W. (1986) *Nucleic Acids Res. 14*, 2971.
McClarin, J. A., Frederick, C. A., Wang, B.-C., Green, P., Boyer, H. W., Grable, J., & Rosenberg, J. M. (1986) *Science (Washington, D.C.) 234*, 1526.
McGovren, J. P., Clarke, G. L., Pratt, E. A., & Dekoning, T. F. (1984) *J. Antibiot. 37*, 63.
Menger, F. M. (1985) *Acc. Chem. Res. 18*, 128.
Muench, K. F., Misra, R. P., & Humayun, M. Z. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 6.
Needham-VanDevanter, D. R., & Hurley, L. H. (1986) *Biochemistry 25*, 8430.
Reynolds, V. L., Molineaux, I. J., Kaplan, D. J., Swenson, D. H., & Hurley, L. H. (1985) *Biochemistry 24*, 6228.
Reynolds, V. L., McGovren, J. P., & Hurley, L. H. (1986) *J. Antibiot. 39*, 319.
VanDyke, M. W., & Dervan, P. B. (1983) *Nucleic Acids Res. 11*, 5555.
Warpehoski, M. A. (1986) *Tetrahedron Lett. 27*, 4103.
Warpehoski, M. A., & Bradford, V. S. (1988) *Tetrahedron Lett. 29*, 131.
Warpehoski, M. A., Gebhard, I., Kelly, R. C., Krueger, W. C., Li, L. H., McGovren, J. P., Prairie, M. D., Wicnienski, N., & Wierenga, W. (1988) *J. Med. Chem. 31*, 590.
Wierenga, W., Bhuyan, B. K., Kelly, R. C., Krueger, W. C., Li, L. H., McGovren, J. P., Swenson, D. H., & Warpehoski, M. A. (1986) *Adv. Enzyme Regul. 25*, 141.
Zakrzewska, K., & Pullman, B. (1986) *J. Biomol. Struct. Dyn. 4*, 127.

CORRECTION

Interaction of Synthetic Analogues of Distamycin with Poly(dA-dT): Role of the Conjugated *N*-Methylpyrrole System, by Dipak Dasgupta, Pradipkumar Parrack, and V. Sasisekharan*, Volume 26, Number 20, October 6, 1987, pages 6381–6386.

Page 6384. In column 1, line 23 should read as follows: evidence for a single type of complex formation at $r' < 0.10$.

Page 6386. In column 1, line 41 should read as follows: concomitant with the decrease in the number of contiguous ....