VOLUME 114, NUMBER 14
JULY 1, 1992

JOURNAL
OF THE
AMERICAN
CHEMICAL
SOCIETY

DNA Alkylation Properties of Enhanced Functional Analogs of CC-1065 Incorporating the 1,2,9,9a-Tetrahydrocyclopropa[1,2-*c*]benz[1,2-*e*]indol-4-one (CBI) Alkylation Subunit

Dale L. Boger* and Stephen A. Munk

Contribution from the Department of Chemistry, Purdue University, West Lafayette, Indiana. *Received April 29, 1991. Revised Manuscript Received December 27, 1991*

Abstract: Full details of a study of the DNA alkylation properties of analogs of (+)-CC-1065 possessing the 1,2,9,9a-tetrahydrocyclopropa[1,2-*c*]benz[1,2-*e*]indol-4-one (CBI) alkylation subunit are described. Both enantiomers of *N*-BOC-CBI, a simple derivative of the alkylation subunit, exhibit an identical profile of DNA alkylation that is not distinguishable from that of (+)-*N*-BOC-CPI, the authentic alkylation subunit of (+)-CC-1065. In addition, the intensity of DNA alkylation as judged by the concentration of agent required for SV40 DNA fragment alkylation was determined to follow the following order: (+)-*N*-BOC-CBI > (−)-*N*-BOC-CBI (ca. 5–10×) and (+)-*N*-BOC-CBI > (+)-*N*-BOC-CPI (ca. 3–6×). A total of 40–45% of the adenines in the double-stranded DNA examined were found to be alkylated over a 10-fold agent concentration range. With three exceptions, the alkylation was found to occur at adenine flanked by a single 5′ A or T base with a preference for 5′-d(A*A*)-3′ > 5′-d(T*A*)-3′. In contrast, (+)-CBI-$CDPI_2$ was found to alkylate DNA in a more selective fashion, and a total of 20–25% of the adenines were alkylated over a 1000–10000-fold agent concentration range. The DNA alkylation was found to proceed with greater intensity (ca. 10000×) and with a readily distinguishable profile of selectivity in which the adenine *N*-3 alkylation site proved to be flanked by two 5′ A or T bases: 5′-d(AA*A*)-3′ > 5′-d(TT*A*)-3′ > 5′-d(TA*A*)-3′ ≥ 5′-d(AT*A*)-3′. With (+)-CBI-$CDPI_2$ there proved to be a strong preference for the fourth 5′ base to be A or T, 5′-d(A/TXX*A*)-3′ > 5′-d(G/CXX*A*)-3′, a weaker preference for the fifth 5′ base to be A or T, 5′-d(A/TXXX*A*)-3′ > 5′-d(G/CXXX*A*)-3′, and a weak preference for the 3′ base preceding the alkylation site to be a purine base, 5′-d(NNXX*A*Pu)-3′ ≥ 5′-(NNXX*A*Py)-3′ where X = A or T. Potentially contributing to their enhanced cytotoxic potency, the CBI-based agents were found to alkylate DNA with a greater intensity and a faster relative rate than the corresponding CPI-based agent (e.g. k((+)-CBI-$indole_2$)/k((+)-CPI-$indole_2$) = 14). A discussion of the potential origin of the alkylation selectivity and the origin of the distinctions in the CBI and CPI-based agents is provided.

CC-1065 (**1**, NSC 298223), an antitumor antibiotic isolated from *Streptomyces zelensis*,[1,2] has been shown to possess exceptionally potent in vitro cytotoxic activity,[3] confirmed in vivo antitumor activity,[4] and broad spectrum antimicrobial activity.[1] The mechanism of the (+)-CC-1065 antitumor activity has been related to its sequence-selective DNA minor groove alkylation which has been demonstrated to proceed by 3′-adenine *N*-3 alkylation of the electrophilic cyclopropane present in the CPI subunit of the natural product.[5,6] The observation that (+)-*N*-acetyl-CPI

* Address correspondence to this author at the Department of Chemistry, The Scripps Research Institute, 10666 North Torrey Pines Road, La Jolla, California 92037.

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

(2) Rachelmycin, isolated from *Streptomyces* strain C-329, has been shown to be identical to CC-1065: Nettleton, D. E., Jr.; Bush, J. A.; Bradner, W. T. U.S. Patent 4,301,248; *Chem. Abstr.* **1982**, *96*, 33362e.

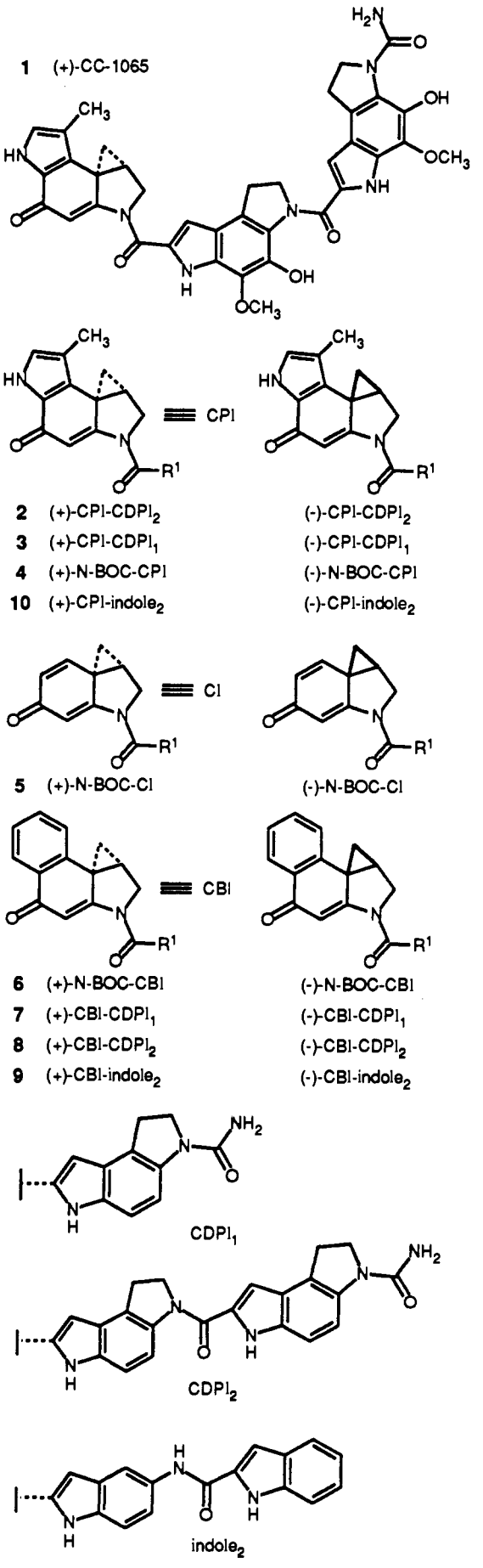
(3) Bhuyan, B. K.; Newell, K. A.; Crampton, S. L.; Von Hoff, D. D. *Cancer Res.* **1982**, *42*, 3532.

0002-7863/92/1514-5487$03.00/0

exhibits a comparable albeit substantially less intense (ca. 10000×) DNA alkylation within some segments of double-stranded DNA led to the suggestion that the alkylation subunit of (+)-CC-1065 plays the dominant role in controlling the properties of the agent. This has been further interpreted to be derived from a rate-determining and sequence-dependent autocatalytic activation of the alkylation step through carbonyl protonation by a strategically located phosphate in the DNA backbone two base-pairs removed from the alkylation site although more recent studies have addressed the additional contributing role of DNA conformational variability.[7]

Complementary to these latter studies, recent investigations have demonstrated that simplified agents including $CDPI_3$ methyl ester[8] exhibit a substantial preference for DNA minor groove binding within A-T rich versus G-C rich DNA[9] attributable to the preferential stabilization of the noncovalent complex within the narrower, sterically more accessible A-T rich minor groove.[10] In these efforts, the comparable profile of DNA alkylation by (+)-*N*-BOC-CPI (**4**) and (+)-*N*-acetyl-CPI proved to be readily distinguished from that of (+)-CC-1065,[11] and an autocatalytic phosphate activation of the alkylation step was shown not to be uniquely responsible for the DNA alkylation selectivity.[12] In addition, the electrophilic cyclopropane proved not to be obligatory to observation of the (+)-CC-1065 characteristic alkylation selectivity, and the noncovalent binding selectivity of the agents was shown to exhibit a pronounced effect independent of the nature of the electrophile.[12] In the course of the continuing investigations, we have prepared analogs of (+)-CC-1065 incorporating both enantiomers of the 1,2,9,9a-tetrahydrocyclopropa[1,2-*c*]benz-[1,2-*e*]indol-4-one (CBI) alkylation subunit including *N*-BOC-CBI (**6**), $CBI\text{-}CDPI_1$ (**7**), $CBI\text{-}CDPI_2$ (**8**), and $CBI\text{-(indole)}_2$ (**9**).[15,16] Herein, we report the results of a study of the DNA alkylation properties of the agents.

Chart I



(4) Martin, D. G.; Chidester, C. G.; Duchamp, D. J.; Mizsak, S. A. *J. Antibiot.* **1980**, *33*, 902. Chidester, C. G.; Krueger, W. C.; Mizsak, S. A.; Duchamp, D. J.; Martin, D. G. *J. Am. Chem. Soc.* **1981**, *103*, 7629.

(5) Hurley, L. H.; Reynolds, V. L.; Swenson, D. H.; Petzold, G. L.; Scahill, T. A. *Science* **1984**, *226*, 843. Reynolds, V. L.; Molineux, I. J.; Kaplan, D. J.; Swenson, D. H.; Hurley, L. H. *Biochemistry* **1985**, *24*, 6228. 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.

(6) For recent reviews see: Rawal, V. H.; Jones, R. J.; Cava, M. P. *Heterocycles* **1987**, *25*, 701. Warpehoski, M. A.; Hurley, L. H. *Chem. Res. Toxicol.* **1988**, *1*, 315. Coleman, R. S.; Boger, D. L. In *Studies in Natural Products Chemistry*; Atta-ur-Rahman, Ed.; Elsevier: Amsterdam, 1989; Vol. 3, p 301. Boger, D. L. *Advances in Heterocyclic Natural Products Synthesis*; Pearson, W. H., Ed.; JAI Press: Greenwich, CT, 1992; Vol. 2, pp 1–177.

(7) Hurley, L. H.; Warpehoski, M. A.; Lee, C.-S.; McGovren, J. P.; Scahill, T. A.; Kelly, R. C.; Mitchell, M. A.; Wicnienski, N. A.; Gebhard, I.; Johnson, P. D.; Bradford, V. S. *J. Am. Chem. Soc.* **1990**, *112*, 4633. Lin, C. H.; Sun, D.; Hurley, L. H. *Chem. Res. Toxicol.* **1991**, *4*, 21. Lee, C.-S.; Sun, D.; Kizu, R.; Hurley, L. H. *Chem. Res. Toxicol.* **1991**, *4*, 203.

(8) Boger, D. L.; Coleman, R. S.; Invergo, B. J. *J. Org. Chem.* **1987**, *52*, 1521.

(9) Boger, D. L.; Invergo, B. J.; Coleman, R. S.; Zarrinmayeh, H.; Kitos, P. A.; Thompson, S. C.; Leong, T.; McLaughlin, L. W. *Chem.-Biol. Interact.* **1990**, *73*, 29. Boger, D. L.; Sakya, S. M. *J. Org. Chem.* **1992**, *57*, 1277.

(10) Nelson, H. C. M.; Finch, J. T.; Luisi, B. F.; Klug, A. *Nature (London)* **1987**, *330*, 221. Coll, M.; Frederick, C. A.; Wang, A. H.-J.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 8385. Coll, M.; Aymami, J.; van der Marel, G. A.; van Boom, J. H.; Rich, A.; Wang, A. H.-J. *Biochemistry* **1989**, *28*, 310.

(11) Boger, D. L.; Coleman, R. S.; Invergo, B. J.; Sakya, S. M.; Ishizaki, T.; Munk, S. A.; Zarrinmayeh, H.; Kitos, P. A.; Thompson, S. C. *J. Am. Chem. Soc.* **1990**, *112*, 4623.

(12) Boger, D. L.; Zarrinmayeh, H.; Munk, S. A.; Kitos, P. A.; Suntornwat, O. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 1431. Boger, D. L.; Munk, S. A.; Zarrinmayeh, H.; Ishizaki, T.; Haught, J.; Bina, M. *Tetrahedron* **1991**, *47*, 2661. Boger, D. L.; Munk, S. A.; Zarrinmayeh, H. *J. Am. Chem. Soc.* **1991**, *113*, 3980.

(13) 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.

(14) Boger, D. L.; Ishizaki, T. *Tetrahedron Lett.* **1990**, *31*, 793. Boger, D. L.; Munk, S. A.; Ishizaki, T. *J. Am. Chem. Soc.* **1991**, *113*, 2779.

(15) Boger, D. L.; Ishizaki, T.; Sakya, S. M.; Munk, S. A.; Kitos, P. A.; Jin, Q.; Besterman, J. M. *BioMed. Chem. Lett.* **1991**, *1*, 115.

(16) Boger, D. L.; Ishizaki, T. *J. Org. Chem.* **1990**, *55*, 5823. Boger, D. L.; Ishizaki, T.; Wysocki, R. J., Jr.; Munk, S. A.; Kitos, P. A.; Suntornwat, O. *J. Am. Chem. Soc.* **1989**, *111*, 6461.

The DNA alkylation properties of the agents were examined within five 150 base-pair segments of double-stranded SV40 DNA since comparative results were available for related agents including (+)-CC-1065.[11,12] From a library of phage M13mp10 harboring SV40 nucleosomal DNA, five clones (750 base-pairs) containing the following inserts were selected for study: w794 (nucleotide no. 5238-138) and its complement w836 (nucleotide no. 5189-91); clone c988 (nucleotide no. 4359-4210) and its complement c820 (nucleotide no. 4201-4356), and c1346 (nucleotide no. 1632-1782).[17] The demonstration of the sites and selectivity (relative intensity) of DNA alkylation was obtained from the thermally-induced strand cleavage of singly 5′ end-labeled double-stranded DNA after exposure to the agents.[5,7,13] Thus, a range of concentrations of the agent was incubated with the singly end-labeled double-stranded DNA for 24 h (4 or 37 °C). The unbound agent was removed by ethanol precipitation of the DNA, and the agent–DNA complex was redissolved in aqueous buffer and warmed at 100 °C (30 min) to induce strand cleavage at the sites of alkylation. Electrophoresis of the resulting DNA adjacent to Sanger dideoxynucleotide sequencing standards[18] under denaturing conditions followed by autoradiography permitted the identification of the sites of DNA alkylation.[12]

Although the details of this protocol have been described elsewhere,[12] a few pertinent features of the procedure and the treatment of the results deserve elaboration. The preparation of the singly 5′ end-labeled duplex DNA simply entails annealing chemically synthesized 5′ end-labeled universal primer to the single-strand template DNA containing the clones of interest and extending the primer with the Klenow fragment of DNA polymerase I (Figure 1). Subsequent cleavage of the extended primer template with *Eco* RI at a site immediately following the inserted (cloned) DNA provides an end-labeled DNA of homogeneous length. Although there are a number of technical advantages to this procedure over the conventional use of end-labeled restriction fragments,[12] the most prominent is that it takes less than 24 h to prepare large quantities of the end-labeled DNA for binding studies. Concurrent with the extension reactions, Sanger[18] sequencing reactions may be performed on the template. This avoids the necessity of conducting Maxam–Gilbert chemical cleavage reactions for the preparation of sequencing standards as is required in the use of restriction fragments.[18] In addition, moderate care in reproducing the conditions of the primer-template extension reaction with Klenow fragment provides sufficient unlabeled duplex DNA liberated with the final restriction enzyme cleavage to reproducibly serve as carrier duplex DNA (ca. 10× the amount of labeled DNA).

The thermal treatment of the agent-alkylated DNA results in depurination and strand cleavage at the site of alkylation while the Sanger sequencing reactions result in base incorporation but halted chain elongation at the sequenced site. Consequently, cleavage at nucleotide N (sequencing) represents alkylation at nucleotide N + 1. The gel legends included herein identify the alkylation sites. The 3′ end heterogeneity in the gels (double bands) observed with the 5′ end-labeled DNA constitute a single alkylation site. The band observed at higher molecular weight (lower gel mobility) constitutes a thermal cleavage product containing a modified sugar, presumably a 2,3-dehydro-2,3-dideoxyribose or subsequent derivative. The second band at lower molecular weight (higher gel mobility) constitutes the subsequent β-elimination product with complete loss of the 3′ modified sugar which results in DNA containing a 3′-phosphate at the 5′ side of the strand break. This end heterogeneity may be removed by piperidine treatment providing a single cleavage product constituting the lower molecular weight band which comigrates with DNA derived from the Maxam–Gilbert adenine reaction. In the conduct of the work it was further demonstrated that the thermal treatment of the agent-alkylated DNA for longer than 30 min at 100 °C does not lead to additional detectable cleavage.[19]

Electrophoresis of the agent-treated DNA was conducted under conditions that maximize the resolution of the SV40 DNA examined. Not shown on the gels is the remaining 5′ end 55 base-pairs derived from the M13mp10 DNA. This carefully chosen 55 base-pair region contains a single alkylation site for the agents adjacent to the 5′ end label (Figure 1). The apparent complete consumption of radioactivity on the gels at high agent concentrations constitutes complete cleavage of the duplex DNA at this site. The competitive cleavage at this site does not affect the relative selectivity of the examined agent within the SV40 DNA, especially at the relevant low agent concentrations constituting single alkylation events on the DNA, but it does ensure that multiple alkylations at high agent concentrations result in cleavage to a single short fragment that is not examined on the gel. This ensures that the commonly encountered gel fade to a single short fragment of 5′ end-labeled DNA at high agent concentrations is not misrepresented as increasing selectivity. Thus, built in to the protocol are features that permit the identification of high affinity alkylation sites detected at the lowest agent concentrations and lower affinity alkylation sites detected at the higher agent concentrations. Multiple strand alkylation leads to complete consumption of the duplex DNA and production of a short fragment of DNA not examined on the gel. The use of a 10-fold amount of carrier unlabeled DNA in the binding assays permits the use of easily measured substantial changes in the agent concentration (10-fold) to effect small changes in the relative concentrations of agent and labeled DNA.

In addition, the statistical treatment of the alkylation sites provided in Tables I and II proved more revealing than a simple and conventional analysis of the observed alkylation sites. That is, the evaluation that included the consideration of sites *not* alkylated helped distinguish the composite consensus sequence and highlighted subtle features not apparent from a simple examination of the alkylated sites.

The DNA alkylation reaction selectivities observed under the incubation conditions of 37 °C (24 h) for the BOC derivatives of the alkylation subunits and 4 °C (24 h) for the more advanced analogs have proven identical to the alkylation selectivities observed with shorter or extended reaction periods (37 or 25 °C, 0.5–7 d). In addition, the selectivity of the DNA alkylation observed under conditions of 25 or 37 °C (24–48 h) for the advanced analogs and (+)-CC-1065 has not proven distinguishable from that observed at 4 °C.[20]

(+)-*N*-BOC-CPI (4) and (+)- and (−)-*N*-BOC-CBI (6). The examination of the DNA alkylation properties of the BOC derivative of the authentic alkylation subunit of (+)-CC-1065 and both enantiomers of *N*-BOC-CBI (**6**) proved revealing. The *tert*-butyloxycarbonyl (BOC) derivative of CBI was selected for study since comparative data were available for related agents, *N*-BOC-CPI and *N*-BOC-CI. Importantly and as detailed elsewhere,[12] the choice of the simple derivative, i.e. *N*-BOC-CBI versus *N*-acetyl-CBI, may affect the relative selectivity observed among the available alkylation sites but does *not* alter the number of observed alkylation sites detected over a 10-fold agent concentration range. The alkylation of DNA by the simple subunits proved substantially less selective than that of (+)-CC-1065 or (+)- and (−)-CBI-CDPI$_n$ (n = 1, 2). Within the five 150 base-pair fragments of DNA examined, 40–45% of the adenines were alkylated by the agents within a 10-fold concentration range. This is in contrast to the 20–25% alkylation of the total adenines by (+)-CC-1065 and (+)-CBI-CDPI$_2$ over a 1000–10000-fold concentration range. (+)-*N*-BOC-CPI, (+)-*N*-BOC-CBI, and (−)-*N*-BOC-CBI proved to display comparable DNA alkylation

(17) Ambrose, C.; Rajadhyaksha, A.; Lowman, H.; Bina, M. *J. Mol. Biol.* **1989**, *210*, 255.

(18) Sanger, F.; Nicklen, S.; Coulsen, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 5463. Maxam, A. M.; Gilbert, W. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 560.

(19) In a study of an isolated oligo-agent adduct, a related adenine *N*-3 adduct exhibits a $t_{1/2}$ = 134 h (37 °C) and $t_{1/2}$ = 1.1 h (60 °C). See: Sugiyama, H.; Hosoda, M.; Saito, I.; Asai, A.; Saito, H. *Tetrahedron Lett.* **1990**, *31*, 7197.

(20) For a published comparison example of a 4 °C (24 h) and 37 °C (24 h) incubation of (+)-CC-1065 with duplex DNA, see ref 11.

(21) Two exceptions to this generalization have been observed and may be found in Figure 4.

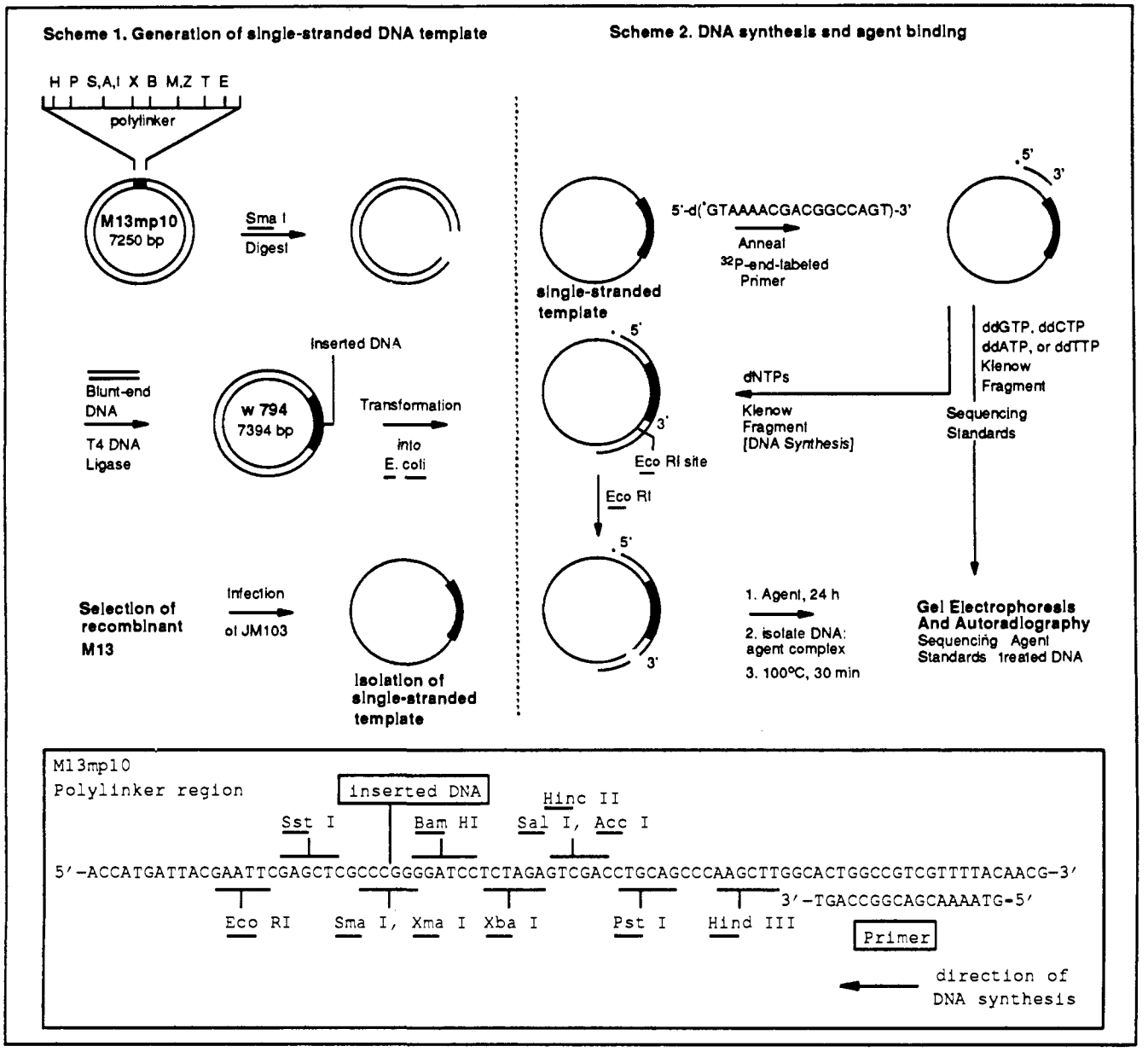


Figure 1.

profiles, and both enantiomers of *N*-BOC-CBI alkylate DNA at concentrations approximately 10000× that required for (+)-CC-1065 or CBI-CDPI$_n$ (Figure 2). The intensity of DNA alkylation by (+)-*N*-BOC-CBI exceeded that of (+)-*N*-BOC-CPI (ca. 5×) as judged by the relative concentration at which the DNA alkylation is observed. This result proved unexpected in that (+)-*N*-BOC-CBI is approximately 4× less reactive than (+)-*N*-BOC-CPI and yet participates in the DNA alkylation more efficiently. In part, this may be attributed to the enhanced stability of **6** which leads to a more productive participation in the DNA alkylation. However, as described in detail with the more advanced CBI-based agents, this constitutes only a part of the observed differences. In addition, the CBI-based agents undergo DNA alkylation at a rate that exceeds that of the more reactive CPI-based agents, and this contributes to the DNA alkylation intensity differences of (+)-*N*-BOC-CBI and (+)-*N*-BOC-CPI.

In addition, the two enantiomers of *N*-BOC-CBI alkylate DNA with essentially the same selectivity. However, the intensity of alkylation by (+)-*N*-BOC-CBI exceeds that of (−)-*N*-BOC-CBI (ca. 5–10×). These observations are comparable to those made in the examination of (+)- and (−)-*N*-acetyl-CPI[7] in which the natural enantiomer alkylates DNA with an identical selectivity and a greater intensity (ca. 10×) than the unnatural enantiomer. Similarly, the examination of the exceptionally reactive agents (+)- and (−)-*N*-BOC-CI revealed that the two enantiomers exhibit an indistinguishable profile *and* intensity of DNA alkylation.[12] All three classes of agents alkylate the same composite sites in DNA, but do so with distinctions in the relative selectivity among the available sites, and the least reactive CBI agents are the most discriminant. Thus, results of the combined studies suggest that the event and selectivity of the DNA alkylation may be tied closely to the DNA versus agent structure. These observations in conjunction with the demonstration that the near-identical DNA alkylation selectivity of the two enantiomers of the simple CI alkylation subunit does not require the cyclopropane electrophile[12] and thus cannot be uniquely related to an autocatalytic phosphate activation of the alkylation reaction by carbonyl protonation suggest that the origin of this selectivity has not yet been well established.[7] Although subtle distinctions may be made among the classes of agents studied to date, the examination of the DNA alkylation properties of both enantiomers of *N*-BOC-CBI detailed herein, along with the results of a comparable study of both enantiomers of *N*-acetyl-CPI,[7] and *N*-BOC-CI[12] and related agents including those lacking the electrophile cyclopropane suggests that the inherent selectivity of the alkylation event is not unique to the CPI structure or the absolute configuration of the alkylation

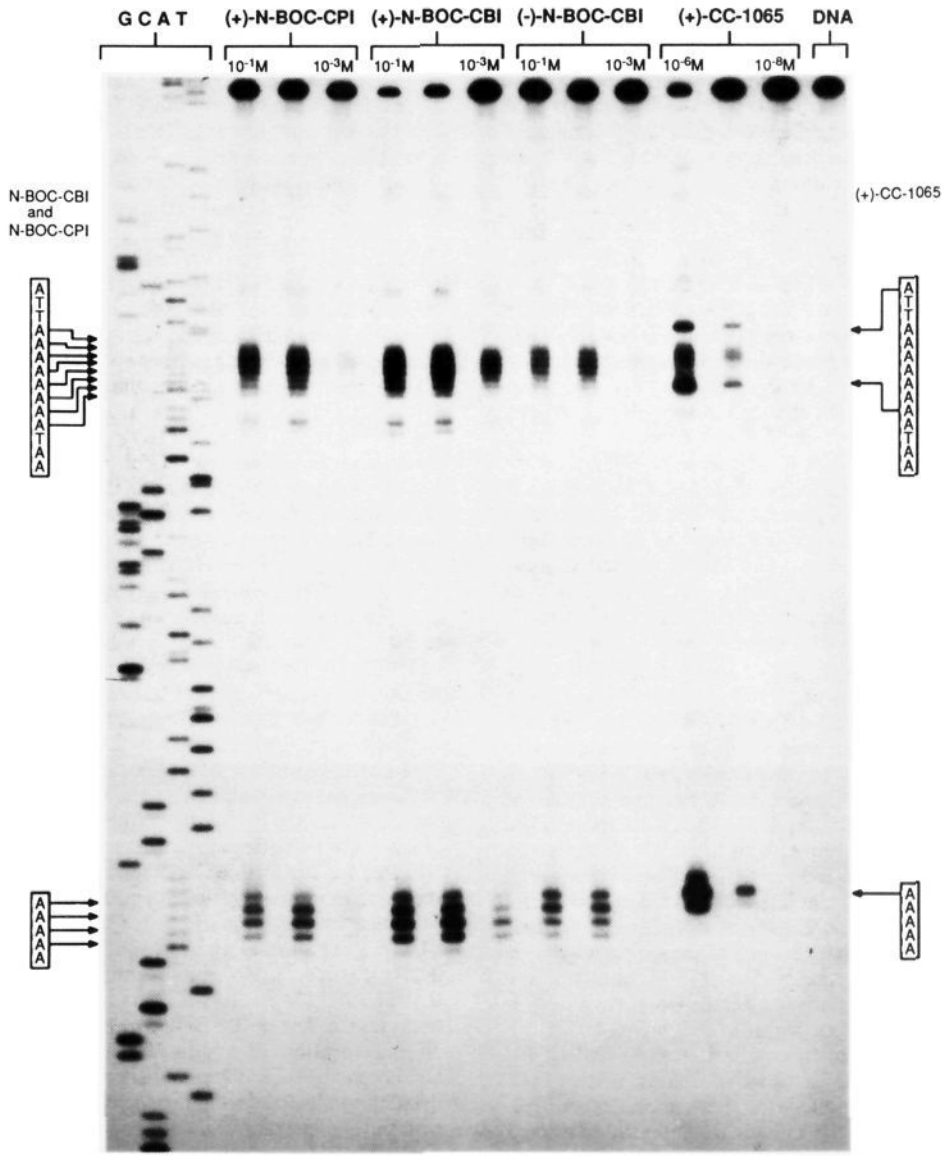


Figure 2. Thermally-induced strand cleavage of double-stranded DNA (SV40 DNA fragment, 146 b.p.; nucleotide no. 5189-91, clone w836) after 24 h incubation of agent–DNA at 37 °C followed by removal of unbound agent and 30 min incubation at 100 °C; denaturing 8% polyacrylamide gel and autoradiography. Lanes 1–4, Sanger G, C, A, and T reactions; lanes 5–7, (+)-*N*-BOC-CPI ((+)-**4**, 2.5×10^{-1}–2.5×10^{-3} M); lanes 8–10, (+)-*N*-BOC-CBI ((+)-**6**, 2.5×10^{-1}–2.5×10^{-3} M); lanes 11–13, (−)-*N*-BOC-CBI ((−)-**6**, 2.5×10^{-1}–2.5×10^{-3} M); lanes 14–16, (+)-CC-1065 ((+)-**1**, 2.5×10^{-6}–2.5×10^{-8} M); lane 17, control DNA.

Tabel I. Summary of Alkylation Sites for (+)- and (−)-*N*-BOC-CBI and (+)-*N*-BOC-CPI[a]

sequence	total sites	alkylated sites
5′-(NNAAN)-3′	83	51 (61%)
5′-(NAAAN)-3′	39	30 (77%)
5′-(NTAAN)-3′	19	10 (53%)
5′-(NGAAN)-3′	11	5 (45%)
5′-(NCAAN)-3′	14	6 (43%)
5′-(NNAAPu)-3′	54	36 (67%)
5′-(NNAAPy)-3′	29	15 (52%)
5′-(NNTAN)-3′	53	13 (25%)
5′-(NATAN)-3′	18	7 (39%)
5′-(NTTAN)-3′	19	4 (21%)
5′-(NGTAN)-3′	7	2 (29%)
5′-(NCTAN)-3′	9	0 (00%)
5′-(NNTAPu)-3′	29	9 (31%)
5′-NNTAPy)-3′	24	4 (17%)

[a] Alkylation sites observed over a 10-fold agent concentration range.

subunit. In addition, the most effective DNA alkylating agents within the series examined to date are the CBI-based agents.

Table I summarizes the alkylation sites observed for (+)- and (−)-*N*-BOC-CBI over a 10-fold agent concentration range. In all but three cases [5′-d(CA)-3′], the alkylation was found to occur at adenine within a two base A-T rich region of DNA and the preference for DNA alkylation proved to be 5′-d(AA)-3′ > 5′-d(TA)-3′ with an additional weak preference for a purine base at the 3′ site preceding the alkylation site. Most interesting is the fact that the high affinity 5′-d(TTA)-3′ sites for (+)-CC-1065 and (+)-CBI-CDPI$_n$ proved to be the least preferred sequence for *N*-BOC-CBI and *N*-BOC-CPI. Characteristic of this distinction, (+)-CC-1065 (Figure 2) as well as (+)-CBI-CDPI$_2$ (not shown) alkylate the 5′-d(AATTAG)-3′ site within w836 as a high affinity alkylation site while (+)-*N*-BOC-CPI and (+)- and (−)-*N*-BOC-CBI show no evidence for alkylation at this site. This site also proved to be a high affinity alkylation site for (+)-CPI-CDPI$_n$,[11] (+)-CBI-CDPI$_1$, (+)- and (−)-CBI-CDPI$_2$, as well as (+)- and (−)-CI-CDPI$_n$.[12] As such, the alkylation selectivity of (+)-CC-1065 and the more advanced CBI-based analogs of (+)-CC-1065 is not just more selective among the available sites but also distinct in the observed alkylation profile as well.

In past studies, the selectivity of the DNA alkylation by agents

related to CC-1065 has been attributed to a sequence-dependent autocatalytic phosphate protonation of the alkylation subunit,[7] the conformational variability of DNA and alkylation at junctions of bent DNA,[7] or alkylation within the narrower, sterically more accessible A-T rich DNA.[9,11,12] Central to the distinction in the interpretations has been the perceived similarities[5,7] or distinctions[11,12] in the alkylation selectivity of simple derivatives of the alkylation subunit versus more advanced analogs or (+)-CC-1065 itself. Consistent with the latter interpretation of the origin of the alkylation selectivity,[9,11,12] the most prevalent feature of the DNA alkylation sites is the neighboring 5′ base-pairs. For simple derivatives of the alkylation subunits, a single adjacent 5′ A or T base is sufficient for observation of DNA alkylation. This has suggested to us that the simple event governing DNA alkylation is the depth of minor groove penetration by the agent and steric accessibility to the adenine *N*-3 alkylation site. For simple derivatives of the alkylation subunits including *N*-BOC-CBI and *N*-BOC-CPI, this requires a single 5′ A or T base adjacent to the adenine alkylation site in order to permit sufficient access for alkylation. The results have further suggested that the distinctions in the relative rate of DNA alkylation by (+)-*N*-BOC-CBI versus (+)-*N*-BOC-CPI may be interpreted simply in terms of steric accessibility to the alkylation site. That is, the C7 methyl group of the CPI subunit sterically decelerates the rate of DNA alkylation to the extent that the less reactive but more accessible CBI subunit alkylates DNA at a more rapid rate.

(+)- and (−)-CBI-CDPI$_1$ (7) and (+)- and (−)-CBI-CDPI$_2$ (8). A comparison of the profiles of DNA alkylation for (+)- and (−)-CBI-CDPI$_1$ (**7**), (+)- and (−)-CBI-CDPI$_2$ (**8**), and (+)-CC-1065 within clone w794 is presented in Figure 3. The comparison is representative of that found with clones w794, w836, c988, c820, and c1346, and Figure 4 summarizes the alkylation sites observed for (+)- and (−)-*N*-BOC-CBI and (+)- and (−)-CBI-CDPI$_2$. There are several striking features about the DNA alkylation by this set of agents. Both enantiomers of CBI-CDPI$_2$ and the natural enantiomer (+)-CBI-CDPI$_1$ alkylate DNA in a comparable sequence-selective fashion at concentrations as low as 10^{-7} M, four orders of magnitude below that required of the simple alkylation subunits. Although the alkylation profiles of (+)-CBI-CDPI$_1$ and (+)-CBI-CDPI$_2$ are only subtly distinguishable, that of (−)-CBI-CDPI$_2$ is more readily distinguished. (−)-CBI-CDPI$_2$ alkylates the high affinity alkylation sites found for (+)-CBI-CDPI$_2$ albeit with a different order of selectivity. Characteristic of this difference, (−)-CBI-CDPI$_2$ alkylates the same sites as (+)-CBI-CDPI$_2$ within w794 DNA (Figure 3) but with an altered preference. Presumably this may be attributed to the noncovalent binding of the agents restricting the available sites and subsequent distinctions in the facility with which the alkylation may occur within the binding sites. In contrast, (−)-CBI-CDPI$_1$ alkylates DNA in a more distinct fashion and with an intensity that is 10–100× less than that observed for (+)-CBI-CDPI$_1$. This behavior is essentially identical to the observations made with CPI-CDPI$_n$[11] and CI-CDPI$_n$.[12] Although subtle, both (+)-CBI-CDPI$_1$ and (+)-CBI-CDPI$_2$ alkylate DNA in a perceptibly more selective fashion than (+)-CC-1065. As detailed in Table II, the alkylation sites for (+)-CBI-CDPI$_2$ proved to be adenine flanked by two 5′ A or T bases,[21] and there proved to be a preference for the three base-pair sequence that follows the order of 5′-d(AAA)-3′ > 5′-d(TTA)-3′ > 5′-d(TAA)-3′ ≥ 5′-d(ATA)-3′. This proved to be only slightly different than that of (+)-CC-1065 where the observed preference proved to be 5′-d(AAA)-3′ = 5′-d(TTA)-3′ > 5′-d(TAA)-3′ > 5′-d(ATA)-3′.[19] There was also a strong preference for the fourth 5′ base to be A or T, a weaker preference for the fifth 5′ base to be A or T, and a weak preference for the 3′ base preceding the alkylation site to be a purine base. Thus, consistent with the relative reactivities of the two agents, the alkylation selectivity of (+)-CBI-CDPI$_2$ is comparable to that of (+)-CC-1065, but slightly more discriminate among the available alkylation sites.

As with (+)-CC-1065, (+)-CI-CDPI$_n$, and structurally related agents, we have attributed the strict A-T preference for the first three base-pairs to represent a combination of the 3′-adenine alkylation site and the binding selectivity of the agent central subunit and its two base-pair A-T binding selectivity. Similarly, the strong preference for the A or T fourth base may be attributed to the 3.5 base-pair binding site size required to accommodate the minor groove binding of the first two subunits of the agent and the A-T binding preference that extends less strongly to the fourth base-pair. As discussed elsewhere,[9,11] the third subunit of the agent may not be required to be bound in the minor groove and consequently the A-T selectivity of the fifth base-pair proved to be weaker and may reflect, as in the high affinity sequences, the preferential but not required binding of the third subunit within the minor groove. Importantly, this interpretation should not be misconstrued to imply that A-T rich binding necessarily leads to productive DNA alkylation. Rather, the noncovalent binding of the agents may serve to further restrict the number of available alkylation sites. Additional distinctions in the facility or rate of DNA alkylation within the available binding sites will be expectedly observed due to the steric accessibility to the alkylation center and stereoelectronic effects imposed on the alkylation reaction. A simple interpretation of the results to date suggests that the depth of minor groove penetration by the agent and steric accessibility to the alkylation site are important features contributing to the observed efficiency or intensity of DNA alkylation. For simple derivatives of the alkylation subunit, sufficient minor groove penetration is possible with a single 5′ A or T base adjacent to the alkylation site. For more advanced analogs including **2–3** and **7–9**, sufficient minor groove penetration may be possible only when two or more adjacent 5′ bases are A or T.

Intensity and Rate of DNA Alkylation. In contrast to the similarities in the selectivity of DNA alkylation between the CBI-based and CPI-based agents, there proved to be a general but unexpected distinction in the intensity of DNA alkylation observed with the two classes of agents. The observed intensity of DNA alkylation as judged by the relative concentration of agent required for observable thermal cleavage of DNA reflects both the relative rate and efficiency of DNA alkylation. This is exemplified by the comparison of (+)-*N*-BOC-CBI with (+)-*N*-BOC-CPI (Figure 2) and (+)-CBI-(indole)$_2$ with (+)-CPI-(indole)$_2$ (Figure 5). A visual examination of Figure 5 illustrates that at 4 °C (24 h), (+)-CPI-(indole)$_2$ alkylates DNA at concentrations as low as 10^{-5} M and (+)-CBI-(indole)$_2$ alkylates DNA at concentrations as low as 10^{-6} M. Within the high affinity alkylation site of w794 DNA, the measured relative rate of alkylation conducted during the course of a 4 d incubation (37 °C) with the two agents was determined to be 14 (k_{rel}(**9**/**10**) (Figure 6).[15] Under the conditions of the experiment, excess double-stranded DNA was present and the agents were effectively sequestered from the aqueous reaction medium by the double-stranded DNA. Under such conditions, the observed distinctions reflect the relative rate of alkylation rather than competitive nonproductive agent solvolysis. Similarly, the relative intensity (4 °C, 24 h; 5×) and relative rate (4 °C, k_{rel}(**8**/**2** = 20) of DNA alkylation by (+)-CBI-CDPI$_2$ versus (+)-CPI-CDPI$_2$ within the high affinity site of w794 DNA were found to be greatest with the more stable CBI-based agent (Table III).[14] Although this more productive covalent modification of DNA by the more stable agent may be attributed in part to the enhanced agent availability, the pronounced stability of both classes of agents and the comparable intensities of alkylation derived with incubation under more vigorous conditions suggest this accounts for only a small part of the distinctions observed in the cell free assays.[22] The comparisons additionally reflect the relative rates of DNA alkylation by the two classes of agents which follow the order of (+)-CBI-(indole)$_2$ > (+)-CPI-(indole)$_2$, (+)-CBI-CDPI$_2$ > (+)-

(22) Both the CPI-based and CBI-based agents are stable to conditions of assay in the absence of DNA. In the case of **2** and **8**, the adenine *N*-3 alkylation has been shown to account for ≥85% of the consumption of agent in the presence of DNA.

(23) The profile of DNA alkylation selectivity exhibited by (−)-CBI-CDPI$_1$ is distinct from that of (+)-CBI-CDPI$_1$ (cf. Figure 3), clearly illustrating that the DNA alkylation properties are attributable to the unnatural enantiomer and not due to contaminant natural enantiomer.

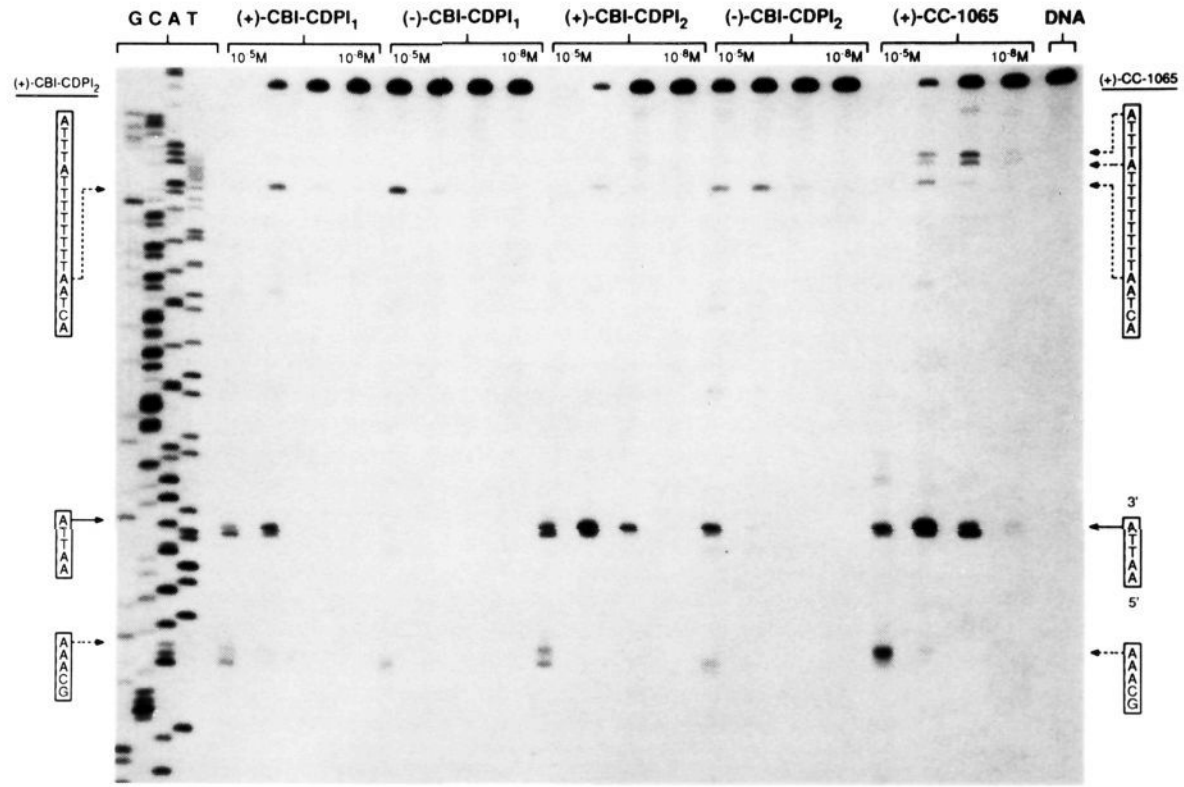
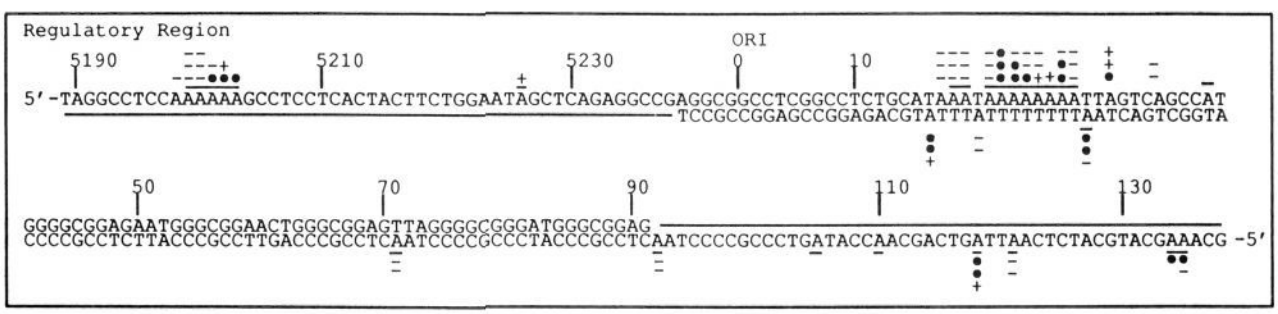


Figure 3. Thermally-induced strand cleavage of double-stranded DNA (SV40 DNA fragment, 144 b.p.; nucleotide no. 138-5238, clone w794) after 24 h incubation of agent–DNA at 4 °C followed by removal of unbound agent and 30 min incubation at 100 °C; denaturing 8% polyacrylamide gel and autoradiography. Lanes 1–4, Sanger G, C, A, and T reactions; lanes 5–8, (+)-CBI-CDPI$_1$ ((+)-**7**, 5.6×10^{-5}–5.6×10^{-8} M); lanes 9–12, (−)-CBI-CDPI$_1$ ((−)-**7**, 5.6×10^{-5}–5.6×10^{-8} M); lanes 13–16, (+)-CBI-CDPI$_2$ ((+)-**8**, 5.6×10^{-5}–5.6×10^{-8} M); lanes 17–20, (−)-CBI-CDPI$_2$ ((−)-**8**, 5.6×10^{-5}–5.6×10^{-8} M); lanes 21–24, (+)-CC-1065 ((+)-**1**, 5.6×10^{-5}–5.6×10^{-8} M); lane 25, control DNA. The observation and origin of double bands for a single alkylation reaction with 5′ end-labeled DNA have been detailed elsewhere (refs 7 and 12).



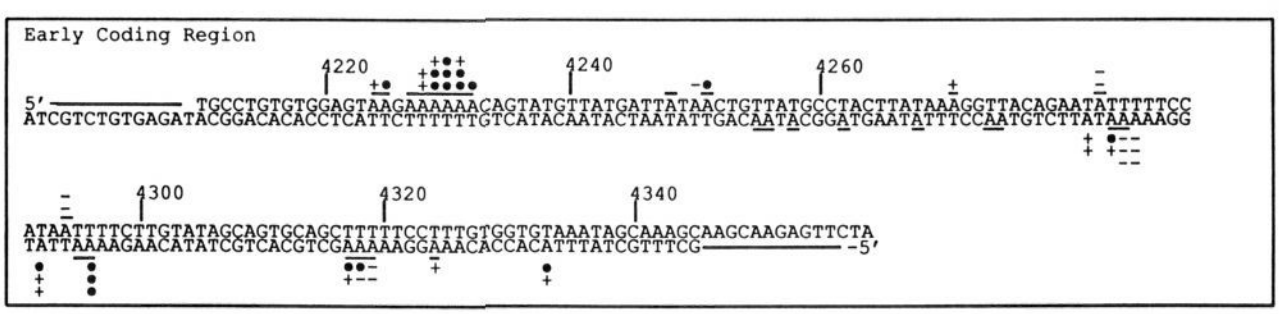


Figure 4. Summary of cleavage sites for (+)-CBI-CDPI$_2$ ((+)-**8**), (−)-CBI-CDPI$_2$ ((−)-**8**), and (+)-*N*-BOC-CPI and (+)- and (−)-*N*-BOC-CBI ((+)- and (−)-**6**): (●) cleavage by (+)-CBI-CDPI$_2$ and (−)-CBI-CDPI$_2$; (+) cleavage by (+)-CBI-CDPI$_2$; (−) cleavage by (−)-CBI-CDPI$_2$; underlined adenines represent cleavage by (+)-*N*-BOC-CPI and (+)- and (−)-*N*-BOC-CBI. The relative intensity of cleavage is represented by the number of symbols above a given site. The SV40 numbering system employs the origin of replication (ORI) as a reference. Two regions of the SV40 genome are presented: one region spans nucleotide no. 5189-138 and includes the origin of replication and part of the SV40 regulatory sequences, the other includes a segment of the SV40 early genes and contains sequences representing strong agent cleavage sites. The missing terminal regions of the SV40 DNA represented by a line constitute the overlapping regions not present in the complementary clones for examination and is not a single-stranded segment of DNA. The data derived from c1346 are not shown.

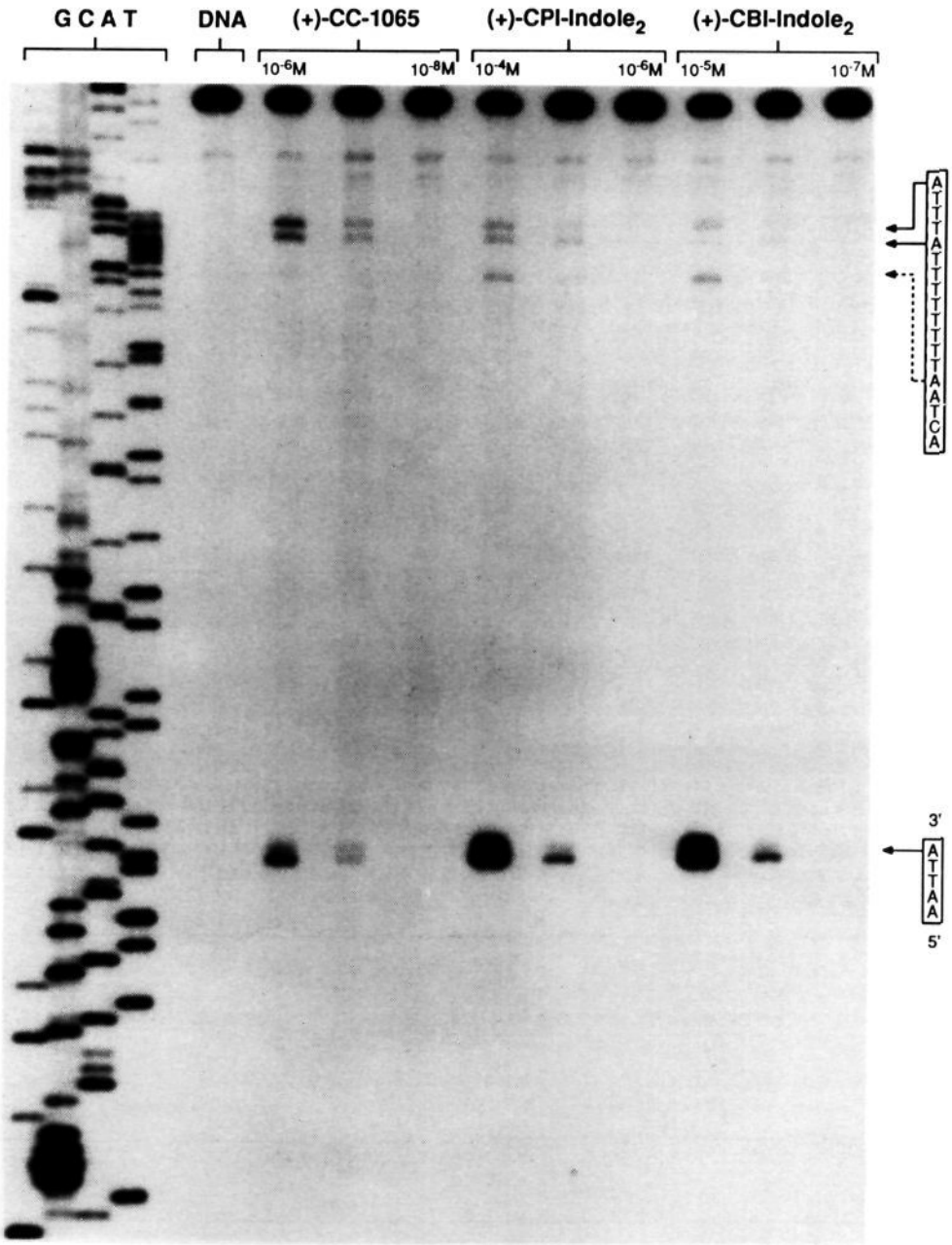


Figure 5. Thermally-induced strand cleavage of double-stranded DNA (SV40 DNA fragment, 144 b.p.; nucleotide no. 138-5238, clone w794) after 24 h incubation of agent–DNA at 4 °C followed by removal of unbound agent and 30 min incubation at 100 °C; denaturing 8% polyacrylamide gel and autoradiography. Lanes 1–4, Sanger G, C, A, and T reactions; lane 5, control DNA; lanes 6–8, (+)-CC-1065 ((+)-**1**, 1×10^{-6}–1×10^{-8} M); lanes 9–11, (+)-CPI-$(indole)_2$ ((+)-**10**, 1×10^{-4}–1×10^{-6} M); lanes 12–14, (+)-CBI-$(indole)_2$ ((+)-**9**, 1×10^{-5}–1×10^{-7} M). The observation and origin of double bands for a single alkylation reaction with 5′ end-labeled DNA have been detailed elsewhere (refs 7 and 12).

CPI-$CDPI_2$. This more rapid (rate) and more productive (intensity) DNA alkylation by the chemically less reactive CBI-based agents is most likely attributable to their steric accessibility to the alkylation site. It would appear reasonable to suggest that the CPI C7 methyl substituent sterically decelerates the rate of DNA alkylation to the extent that the less reactive CBI agents alkylate DNA at a more rapid rate.

Although the results of the in vitro cytotoxic evaluation[16] of the CBI-based agents and the efficacious antitumor activity of **9** have been presented elsewhere,[15] a short discussion of their relationship to the results of the studies detailed herein is merited. The cytotoxic activity of the CBI enantiomers corresponding to the absolute configuration of (+)-CC-1065 proved approximately four times more potent than corresponding CPI-based analogs (Table IV). The cytotoxic potencies of both (+)-CBI-$CDPI_1$ and (+)-CBI-$CDPI_2$ proved remarkably high and indistinguishable from one another. The CBI agent possessing the natural configuration proved slightly more potent than its corresponding unnatural enantiomer. (+)-CBI-$CDPI_1$ is two orders of magnitude more potent than (−)-CBI-$CDPI_1$, and this difference corresponds to the differences in concentrations at which the DNA alkylation is observed.[23] Unlike the pairs of enantiomers (+)- and (−)-CC-1065/(+)- and (−)-CPI-$CDPI_2$, which proved indistinguishable in cytotoxic assays, (+)-CBI-$CDPI_2$ proved 8× more potent than (−)-CBI-$CDPI_2$. Nonetheless, (−)-CBI-$CDPI_2$ proved quite potent (0.5× (+)-CC-1065). Similar to the trends observed in the DNA alkylation studies, (+)-CBI-$(indole)_2$ proved to be 4× more potent than (+)-CPI-$(indole)_2$ in in vitro cytotoxic assays and was found to possess efficacious antitumor activity at doses substantially lower than that required of (+)-CPI-$(indole)_2$.[15]

The natural enantiomers of the CBI-based agents proved to be 4× more potent than their CPI-based counterparts, and they

Table II. Summary of Alkylation Sites for (+)-CBI-CDPI$_2$

sequence	no. AS/no. TS[a]	high affinity[b]	low affinity[b]
5′-(NNAAAN)-3′	22/39 (56%)	11 (28%)	11
5′-(NNTTAN)-3′	7/19 (37%)	5 (26%)	2
5′-(NNTAAN)-3′	4/19 (21%)	1 (5%)	3
5′-(NNATAN)-3′	3/18 (17%)	1 (6%)	2
5′-(NNAAAN)-3′			
5′-(NAAAAN)-3′	15/21 (71%)	7 (33%)	8
5′-(A/TAAAAN)-3′	11/14 (79%)	6 (43%)	5
5′-(G/CAAAAN)-3′	4/7 (57%)	1 (14%)	3
5′-(NTAAAN)-3′	3/6 (50%)	1 (17%)	2
5′-(A/TTAAAN)-3′	3/4 (75%)	1 (25%)	2
5′-(G/CTAAAN)-3′	0/2 (00%)	0 (00%)	0
5′-(NG/CAAAN)-3′	4/12 (33%)	3 (25%)	1
5′-(NGAAAN)-3′	3/5 (60%)	3 (60%)	0
5′-(NCAAAN)-3′	1/7 (14%)	0 (00%)	1
5′-(NNAAAPu)-3′	18/30 (60%)	9 (30%)	9
5′-(NNAAAPy)-3′	4/9 (44%)	2 (22%)	2
5′-(NNTTAN)-3′			
5′-(NATTAN)-3′	3/4 (75%)	3 (75%)	0
5′-(A/TATTAN)-3′	3/3 (100%)	3 (100%)	0
5′-(G/CATTAN)-3′	0/1 (00%)	0 (00%)	0
5′-(NTTTAN)-3′	3/6 (50%)	2 (33%)	1
5′-(A/TTTTAN)-3′	3/4 (75%)	2 (50%)	1
5′-(G/CTTTAN)-3′	0/2 (00%)	0 (00%)	0
5′-(NG/CTTAN)-3′	1/9 (11%)	0 (00%)	1
5′-(NGTTAN)-3′	0/5 (00%)	0 (00%)	0
5′-(NCTTAN)-3′	1/4 (25%)	0 (00%)	1
5′-(NNTTAPu)-3′	3/6 (50%)	3 (50%)	0
5′-(NNTTAPy)-3′	4/13 (31%)	3 (23%)	1
5′-(NNTAAN)-3′			
5′-(NTTAAN)-3′	1/3 (33%)	0 (00%)	1
5′-(A/TTTAAN)-3′	1/1 (100%)	0 (00%)	1
5′-(G/CTTAAN)-3′	0/2 (00%)	0 (00%)	0
5′-(NATAAN)-3′	1/9 (11%)	1 (11%)	0
5′-(A/TATAAN)-3′	1/5 (20%)	1 (20%)	0
5′-(G/CATAAN)-3′-	0/4 (00%)	0 (00%)	0
5′-(NG/CTAAN)-3′	2/7 (29%)	0 (00%)	2
5′-(NGTAAN)-3′	1/4 (25%)	0 (00%)	1
5′-(NCTAAN)-3′	1/3 (33%)	0 (00%)	1
5′-(NNTAAPu)-3′	3/9 (33%)	1 (11%)	2
5′-(NNTAAPy)-3′	1/10 (10%)	0 (00%)	1
5′-(NNATAN)-3′			
5′-(NAATAN)-3′	3/5 (60%)	1 (20%)	2
5′-(A/TAATAN)-3′	2/3 (66%)	1 (33%)	1
5′-(G/CAATAN)-3′	1/2 (50%)	0 (00%)	1
5′-(NTATAN)-3′	0/6 (00%)	0 (00%)	0
5′-(A/TTATAN)-3′	0/4 (00%)	0 (00%)	0
5′-(G/CTATAN)-3′	0/2 (00%)	0 (00%)	0
5′-(NG/CATAN)-3′	0/7 (00%)	0 (00%)	0
5′-(NGATAN)-3′			
5′-(NCATAN)-3′	0/7 (00%)	0 (00%)	0
5′-(NNATAPu)-3′	3/14 (21%)	1 (07%)	2
5′-(NNATAPy)-3′	0/4 (00%)	0 (00%)	0

[a] Number of alkylated sites (no. AS)/number of total sites (no. TS). [b] Intensity of alkylation at the sites: high = high affinity site observed at lowest agent concentration; low = low affinity site observed at higher agent concentrations. The expressed percent is that of the number of high affinity alkylation sites (no. HAAS)/number of total sites (no. TS).

also proved to be 4× less reactive toward aqueous acid-catalyzed solvolysis (Table IV). This inverse relationship between cytotoxic potency and solvolytic reactivity not only holds for the four classes of CBI-based agents and CPI-based agents discussed above but also includes the exceptionally reactive CI-based agents.[12,14] This has suggested that the direct relationship between agent stability and cytotoxic potency may constitute a relevant feature in the design of functional analogs. That is, within the range of agents possessing sufficient reactivity to alkylate DNA, the more stable

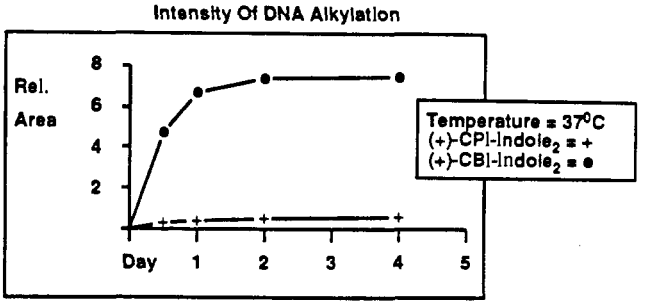


Figure 6. Plot of the relative area (optical density) of 5′ end-labeled DNA derived from the alkylation and thermally-induced cleavage at the high affinity alkylation site [5′-d(AATTA)-3′] within clone w794 versus time at 37 °C (10^{-5} M agent). The optical density determinations were conducted using a scanning laser densitometer. The results at day 4 represent 50% consumption of the end-labeled DNA by (+)-CBI-(indole)$_2$ and full consumption of the available agent.

Table III

	(+)-CBI-CDPI$_2$	(+)-CPI-CDPI$_2$
rel intensity of DNA alkylation, 37 °C (24 h)[a]	1.1 (100%)	1.0 (100%)
rel intensity of DNA alkylation, 25 °C (24 h)	1.0 (90%)	0.2 (20%)
k_{rel},[b] 4 °C	1.0	0.05

[a] Relative intensity of alkylation (thermally-induced strand cleavage) at the high affinity alkylation site (5′-d(AATTA)-3′) within w794 DNA determined with a scanning densitometer. Percent reaction is expressed as the percentage of the total alkylation at this site observed when the alkylation is taken to >90% completion (37 °C, 24–48 h). The relative intensities of alkylation detectable at 10^{-8} M at this point are compared to (+)-CBI-CDPI$_2$ (4 °C, 23 h) = 1.0. [b] Relative first-order rate constants for DNA alkylation at the high affinity alkylation site taken from plots of the intensity of DNA cleavage versus time (4 °C, 10^{-7} M agent; 12, 24, 48, 96, and 192 h).

Table IV. In Vitro Cytotoxic Activity, L1210[a]

agent	configuration	IC_{50}, pM
(+)-*N*-BOC-CPI[b]	natural	330000
(+)-*N*-BOC-CBI[b]	natural	77000
(−)-*N*-BOC-CBI[b]	unnatural	940000
(+)-CBI-CDPI$_1$[b]	natural	5
(−)-CBI-CDPI$_1$[b]	unnatural	≥380
(+)-CPI-CDPI$_1$[e]	natural	40
(−)-CPI-CDPI$_1$[c]	unnatural	≥630
(+)-CBI-CDPI$_2$[b]	natural	5
(−)-CBI-CDPI$_2$[b]	unnatural	40
(+)-CC-1065[c]	natural	20
(+)-CPI-CDPI$_2$[c]	natural	20
(−)-CC-1065[c]	unnatural	20
(−)-CPI-CDPI$_2$[c]	unnatural	20
(+)-CPI-(indole)$_2$[d]	natural	40
(+)-CBI-(indole)$_2$[d]	natural	10

[a] IC_{50} = inhibitory concentration for 50% cell growth relative to untreated controls, L1210 mouse lymphocytic leukemia cell culture, see: Boger, D. L.; Yasuda, M.; Mitscher, L. A.; Drake, S. D.; Kitos, P. A.; Thompson, S. C. *J. Med. Chem.* **1987**, *30*, 1918. [b] Taken from ref 14. For *N*-BOC-CPI, $t_{1/2}$ = 36.7 h (pH = 3); for *N*-BOC-CBI, $t_{1/2}$ = 133 h (pH = 3). [c] Taken from ref 11. [d] Taken from ref 15.

agents may constitute the more potent agents. Presumably this may be the consequence of the enhanced stability of the agent and its relative ability to reach its biological target effectively providing a more productive covalent modification of duplex DNA. Potentially contributing to the distinctions in the agents examined to date may be the observation of the more rapid (rate), more discriminate (selectivity), and more productive (intensity or efficiency) DNA alkylation by the more stable CBI-based agents.

Experimental Section

Agents **6–9** were prepared as detailed and were >99% enantiomerically pure.[14–16] General procedures, the preparation of singly 5′ end-labeled double-stranded DNA, and the agent binding studies were con-

ducted following procedures described in full detail elsewhere.[12] Eppendorf tubes containing the end-labeled DNA[12] (9 μL) in TE buffer (10 mM Tris, 1 mM EDTA, pH = 7.6) were treated with the agent in a solution of DMSO (1 μL at the specified concentration). Agent concentrations were measured by UV using known extinction coefficients. The reaction was mixed by vortexing and brief centrifugation and subsequently incubated at 4 or 37 °C for the specified time (generally 24 h). The DNA was separated from unbound agent by ethanol precipitation and resuspended in TE buffer (20 μL). The solution in an Eppendorf tube sealed with Teflon tape was warmed to 100 °C for 30 min to induce cleavage at the alkylation sites, allowed to cool to room temperature, and centrifuged. Formamide dye was added (10 μL) to the supernatant. Prior to electrophoresis, the samples were warmed at 100 °C for 5 min, placed in an ice bath, and centrifuged and the supernatant was loaded onto the gel. Sanger dideoxynucleotide sequencing reactions were run as standards adjacent to the agent treated DNA. Gel electrophoresis was carried out using an 8% sequencing gel (19:1 acrylamide:*N,N*-methylene bisacrylamide, 8 M urea). Formamide dye contained xylene cyanol FF (0.03%), bromophenol blue (0.03%), and aqueous Na_2EDTA (8.7%, 250 mM). Electrophoresis running buffer (TBE) contained Tris base (100 mM), boric acid (100 mM), and $Na_2EDTA{\cdot}H_2O$ (0.2 mM). Gels were prerun for 30 min with formamide dye prior to loading the samples. Autoradiography of dried gels was carried out at −70 °C using Kodak X-Omat AR film and a Picker Spectra intensifying screen. The relative intensity of DNA alkylation at the high affinity alkylation site within clone w794 [5′-d(AATTA)-3′] versus time was determined by measuring the optical density from the autoradiogram at that site using an LKB UltroScan XL scanning laser densitometer interfaced to LKB 2400 Gelscan XL integration software.

Acknowledgment. This work was supported by the National Institutes of Health (Grant No. CA41986 to D.L.B.) and the American Cancer Society (Grant No. PF-3311 to S.A.M.). We gratefully acknowledge Professor M. Bina for the gift of M13mp10 clones.

Nonenzymatic Template-Directed Synthesis on Hairpin Oligonucleotides. 2. Templates Containing Cytidine and Guanosine Residues

Taifeng Wu and Leslie E. Orgel*

Contribution from The Salk Institute for Biological Studies, P.O. Box 85800, San Diego, California 92186-5800. *Received December 6, 1991*

Abstract: We have prepared hairpin oligonucleotides in which a 5′-terminal single-stranded segment contains cytidylate (C) and guanylate (G) residues. When these hairpin substrates are incubated with a mixture of cytidine 5′-phosphoro(2-methyl)imidazolide (2-MeImpC) and guanosine 5′-phosphoro(2-methyl)imidazolide (2-MeImpG), the 5′-terminal segment acts as a template to facilitate sequence-specific addition of G and C residues to the 3′-terminus of the hairpin. If an isolated G residue is present at the 3′-end of the template strand, it is copied regiospecifically in the presence of 2-MeImpC and 2-MeImpG to give a product containing an isolated C residue linked to its G neighbors by 3′–5′-internucleotide bonds. However, if only 2-MeImpC is present in the reaction mixture, very little reaction occurs. Thus, the presence of 2-MeImpG catalyzes the incorporation of C. If the template strand contains a short sequence of G residues, it is copied in the presence of a mixture of 2-MeImpC and 2-MeImpG. If only 2-MeImpC is present in the reaction mixture, efficient synthesis occurs to give a final product containing one fewer C residue than the number of G residues in the template.

Introduction

Template-directed reactions on oligonucleotides, oligodeoxynucleotides, and polynucleotides that contain cytidylic acid (C) and guanylic acid (G) residues have been studied extensively.[1-12] It has been established that random copolymer templates that contain a substantial excess of riboC residues over riboG residues, when incubated with a mixture of cytidine 5′-phosphoro(2-methyl)imidazolide (2-MeImpC) and guanosine 5′-phosphoro-(2-methyl)imidazolide (2-MeImpG), are able to direct an efficient synthesis of copolymers rich in G residues.[7] However, copolymers containing more or less equal numbers of G and C residues are not effective templates, presumably because of their stable intramolecular self-structure. Poly(G) fails to act as a template for a different reason. It forms a very stable tetrahelical self-complex,[13] which prevents binding by C residues. On the basis of our studies with random copolymer templates we speculated that the failure to copy G-rich random poly(CG) templates was probably due to the intra- and intermolecular secondary structures of the templates rather than to the inability of the underlying condensation reaction of activated monomers to proceed on templates containing single-stranded G-rich sequences.[7]

Studies of reactions on short templates of defined sequence are consistent with this interpretation, but do not confirm it unambiguously. Isolated G residues in a ribo[4] or deoxyribo[6,8,10] template direct the efficient incorporation of C residues, but short oligomers containing the sequences -CCGGCC- and -CCCGCGCC- do not function as templates. We attributed this failure to the formation of double-helical regions stabilized by four strong G:C base pairs since, in our experiments, we were obliged to use a template concentration of at least 0.01 M, a concentration high enough to stabilize a CGCG or CGGC double-helical segment at 0 °C.[8]

Recently we introduced a novel method for studying template-directed reactions, namely the extension of a hairpin oligonucleotide in which the 5′-end of the oligomer functions as a template for the extension of the 3′-end (Figure 1a).[14] A ribo-

(1) Inoue, T.; Orgel, L. E. *J. Am. Chem. Soc.* **1981**, *103*, 7666-7667.
(2) Inoue, T.; Orgel, L. E. *J. Mol. Biol.* **1982**, *162*, 201-217.
(3) Inoue, T.; Orgel, L. E. *Science* **1983**, 219, 859-862.
(4) Inoue, T.; Joyce, G. F.; Grzeskowiak, K.; Orgel, L. E.; Brown, J. M.; Reese, C. B. *J. Mol. Biol.* **1984**, *178*, 669-676.
(5) Chen, C. B.; Inoue, T.; Orgel, L. E. *J. Mol. Biol.* **1985**, *181*, 271-279.
(6) Haertle, T.; Orgel, L. E. *J. Mol. Biol.* **1986**, *188*, 77-80.
(7) Joyce, G. F.; Orgel, L. E. *J. Mol. Biol.* **1986**, *188*, 433-441.
(8) Haertle, T.; Orgel, L. E. *J. Mol. Evol.* **1986**, *23*, 108-112.
(9) Grzeskowiak, K.; Orgel, L. E. *J. Mol. Evol.* **1986**, *23*, 287-289.
(10) Acevedo, O. L.; Orgel, L. E. *J. Mol. Biol.* **1987**, *197*, 187-193.
(11) Kanavarioti, A.; White, D. H. *Origins Life* **1987**, *17*, 333-349.
(12) Kanavarioti, A.; Bernasconi, C. F.; Doodokyan, L. D.; Alberas, D. J. *J. Am. Chem. Soc.* **1989**, *111*, 7247-7257.
(13) *Basic Principles in Nucleic Acid Chemistry*; Ts'o, P. O. P., Ed.; Academic Press: New York and London, 1974; Vol. 1.

0002-7863/92/1514-5496$03.00/0