Pergamon

Bioorganic & Medicinal Chemistry, Vol. 2, No. 2, pp. 115-135, 1994

0968-0896/94 $6.00 + .00

0968-0896(93)E0011-C

Molecular Basis for Sequence Selective DNA Alkylation by (+)- and *ent*-(–)-CC-1065 and Related Agents: Alkylation Site Models that Accommodate the Offset AT-rich Adenine N3 Alkylation Selectivity

Dale L. Boger,* Douglas S. Johnson, Weiya Yun and Christine M. Tarby

Department of Chemistry, The Scripps Research Institute, 10666 North Torrey Pines Road, La Jolla, CA 92037, U.S.A.

Abstract—A detailed evaluation of the DNA alkylation selectivity of (+)-CC-1065, *ent*-(–)-CC-1065 and a series of aborted and extended analogs possessing the CPI alkylation subunit is detailed and the refinement of a model that accommodates the offset AT-rich adenine N3 alkylation selectivity of the enantiomeric agents is presented. The natural enantiomers bind in the minor groove in the 3'→5' direction starting from the adenine N3 alkylation site across a 2 base (*N*-BOC-CPI; i.e. 5'-AA), 3.5 base (CPI-CDPI$_1$/CPI-PDE-I$_1$; i.e. 5'-AAA), 5 base (CC-1065/CPI-CDPI$_2$; i.e. 5'-AAAAA) or 6.5 base (CPI-CDPI$_3$; i.e. 5'-AAAAAA) AT-rich site. In contrast, the unnatural enantiomers bind in the reverse 5'→3' direction in the minor groove and the binding site necessarily starts at the first 5' base preceding the adenine N3 alkylation site and extends across the alkylation site to the adjacent 3' bases covering an AT-rich site of 2 bases (*N*-BOC-CPI; e.g., 5'-AA), 5 bases (CC-1065/CPI-CDPI$_2$; eg. 5'-AAAAA), or 6.5 bases (CPI-CDPI$_3$; e.g. 5'-AAAAAA). Notably, the model accommodates the unusual observation that both enantiomers of *N*-BOC-CPI alkylate the same sites within duplex DNA (5'-AA > 5'-TA) and the required reversed binding orientation of the enantiomeric agents. The reversed binding orientation is required to permit access to the electrophilic cyclopropane and the resulting offset AT-rich alkylation selectivity is the natural consequence of the diastereomeric relationship of the adducts. Three dimensional models of the natural and unnatural enantiomer alkylations are presented which clearly illustrate the offset binding sites. A fundamentally simple model for the CC-1065 DNA alkylation reaction, that accommodates the behavior of both enantiomers, is provided in which the sequence selectivity is derived from the noncovalent binding selectivity of the agents preferentially in the narrower, sterically more accessible AT-rich minor groove, the inherent steric accessibility to the adenine N3 alkylation site that accompanies deep penetration of the agent into the minor groove within an AT-rich site, and the 2 base-pair (*N*-BOC-CPI), 3.5 base-pair (CPI-PDE-I$_1$/CPI-CDPI$_1$), 5 base-pair (CC-1065/CPI-CDPI$_2$), or 6.5 base-pair (CPI-CDPI$_3$) site size required to permit agent binding in the minor groove at the alkylation site. Using the models, a simple explanation for the distinguishing DNA alkylation rate, efficiency, and biological potency of enantiomeric pairs of agents is detailed based on the unnatural enantiomer sensitivity to steric bulk at the CPI C7 center.

(+)-CC-1065 (**1**)[1] represents the oldest member of a growing class of potent antitumor antibiotics including the duocarmycins[2–8] that exert their biological effects through a sequence selective alkylation of DNA. Since its initial disclosure, extensive efforts have been devoted to define the structural origin of its sequence selective minor groove adenine N3 alkylation within AT-rich sequences. In the conduct of the efforts, both (+)-CC-1065 and its enantiomer, *ent*-(–)-CC-1065,[26,33] as well as an extensive series of key structural analogs possessing the authentic CPI alkylation subunit or modified alkylation subunits (CI, CBI, C_2BI, DSA) have been shown to participate in the now characteristic adenine N3 alkylation reaction.[9–46] In conjunction with efforts that have resulted in the definition of the DNA alkylation properties of duocarmycin SA including the development of alkylation site models that accommodate the behavior of both enantiomers,[42] we have evaluated in detail the data derived from a study of the natural and unnatural enantiomers of CC-1065 and a key series of aborted and extended structural analogs.[33] Based on the evaluation of this data, herein we detail the further refinement of a fundamentally simple and consistent model for the CC-1065 and CPI DNA alkylation reactions *that accommodates the behavior of both enantiomers.* Central to the model is the continued interpretation that the sequence selectivity is derived from the noncovalent binding selectivity of the agents within the narrower, sterically more accessible AT-rich minor groove, the inherent steric accessibility to the adenine N3 alkylation site that accompanies deep penetration of the agents into the minor groove within an AT-rich alkylation site, and the five base-pair binding site size required to permit full agent (CC-1065) binding in the minor groove at the alkylation site. Comparative three dimensional models of the natural and unnatural enantiomer alkylations are presented and clearly illustrate the offset AT-rich binding sites required to accommodate the enantiomer alkylations. Based on these models, an explanation for the distinguishing properties (DNA alkylation rate, efficiency, adduct stability and biological potency) of enantiomeric pairs of agents is detailed based on the unnatural enantiomer sensitivity to steric interactions surrounding the alkylation subunit (CPI) C7 center.

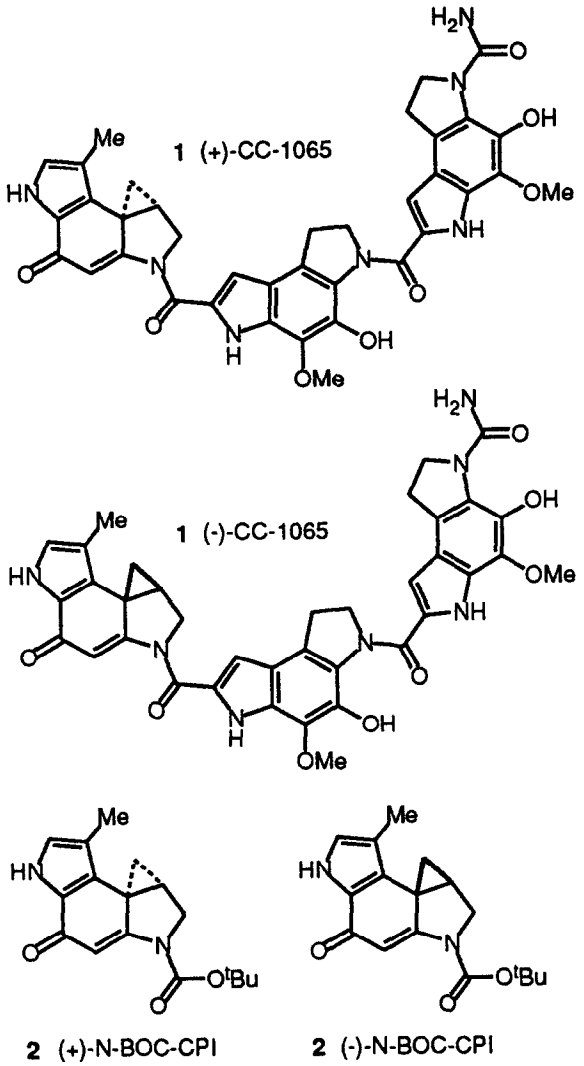
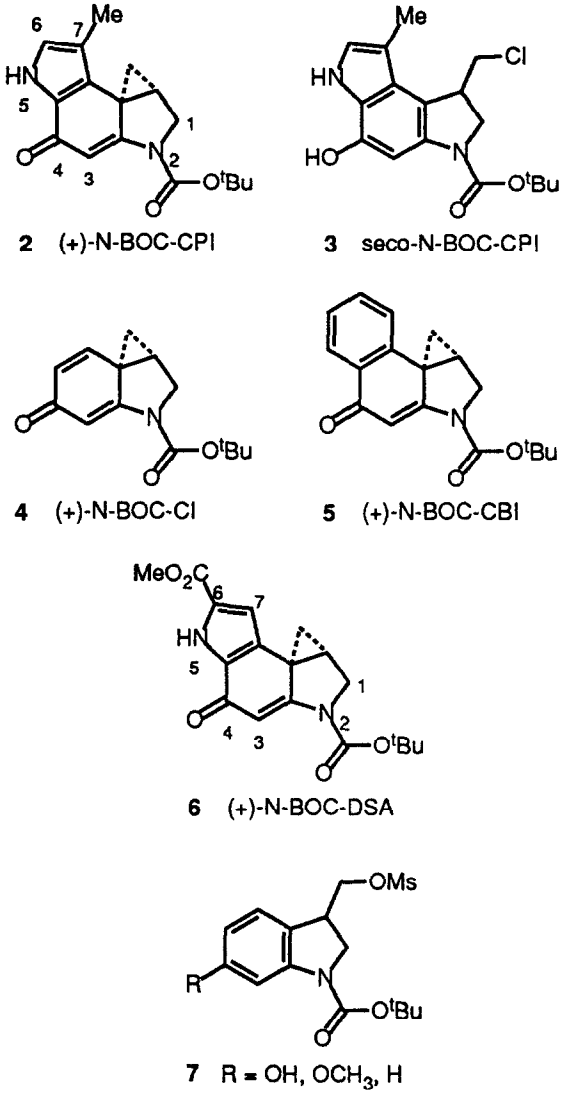


Figure 1.

(+)- and (–)-N-BOC-CPI

In our studies, the DNA alkylation properties of both enantiomers of *N*-BOC-CPI (**2**) proved especially revealing. The *tert*-butyloxycarbonyl (BOC) derivative as well as the *N*-acetyl derivative were examined and extensive comparative data is now available for a series of structurally related agents including *N*-BOC-CI (**4**),[34,37] *N*-BOC-CBI (**5**),[38] and *N*-BOC-DSA (**6**).[42] The choice of the simple derivative (BOC versus acetyl) was shown to subtly affect the relative selectivity among the available alkylation sites and the efficiency of DNA alkylation but does *not* alter the sites of alkylation detected over a 10-fold concentration range.[33] The DNA alkylation by the simple CPI subunit proved substantially less selective than that of (+)- or *ent*-(–)-CC-1065. Within the DNA examined, 40–45% of all adenines were alkylated by *N*-BOC-CPI within a 10-fold concentration range in contrast to the 20–25% alkylation by CC-1065 over a 1000–10,000 fold concentration range. In addition, (+)-*N*-BOC-CPI and *ent*-(–)-*N*-BOC-CPI displayed comparable, essentially identical DNA alkylation selectivities. Both enantiomers of **2** alkylate DNA much less efficiently than CC-1065 and require 10,000x the concentration of (+)-CC-1065 and more vigorous reaction conditions (37 °C versus 4 °C). Moreover, (+)-*N*-BOC-CPI, the natural enantiomer, proved to be 10–100x more efficient than *ent*-(–)-*N*-BOC-CPI at alkylating DNA. Figure 2 and Table 1 summarize the *N*-BOC-CPI alkylation sites observed over a 10-fold concentration range and Table 2 summarizes the composite alkylation consensus sequence.



In all cases, only adenine N3 alkylation was observed and no evidence for even a single guanine alkylation was detected in our protocol. In all but 2 instances of a 5'-CA alkylation, the adenine N3 alkylation was found to occur within a two base AT-rich site. The preference for DNA alkylation proved to be 5'-AA > 5'-TA with the additional weak preference for a purine base at the 3' site preceding the alkylation site. The agents also appear to exhibit a weak preference, but not absolute requirement, for the third 5' base to be A or T versus G or C although this needs to be interpreted with caution since 40–45% of all adenines are reactive over a 10-fold concentration range. In addition, the *seco* agents **3** proved indistinguishable from **2** in the DNA alkylation assays. At least part of the origin of the

5'-AA > 5'-TA preference lies in the equal opportunity for alkylation of the unlabeled complementary strand adenine within the 5'-TA sites competitive with the labeled strand alkylation. Since the occurrence of 5'-TA alkylation is approximately 50% that of 5'-AA alkylation, it is suggestive that this 5'-AA preference is simply a statistical rather than structural preference.

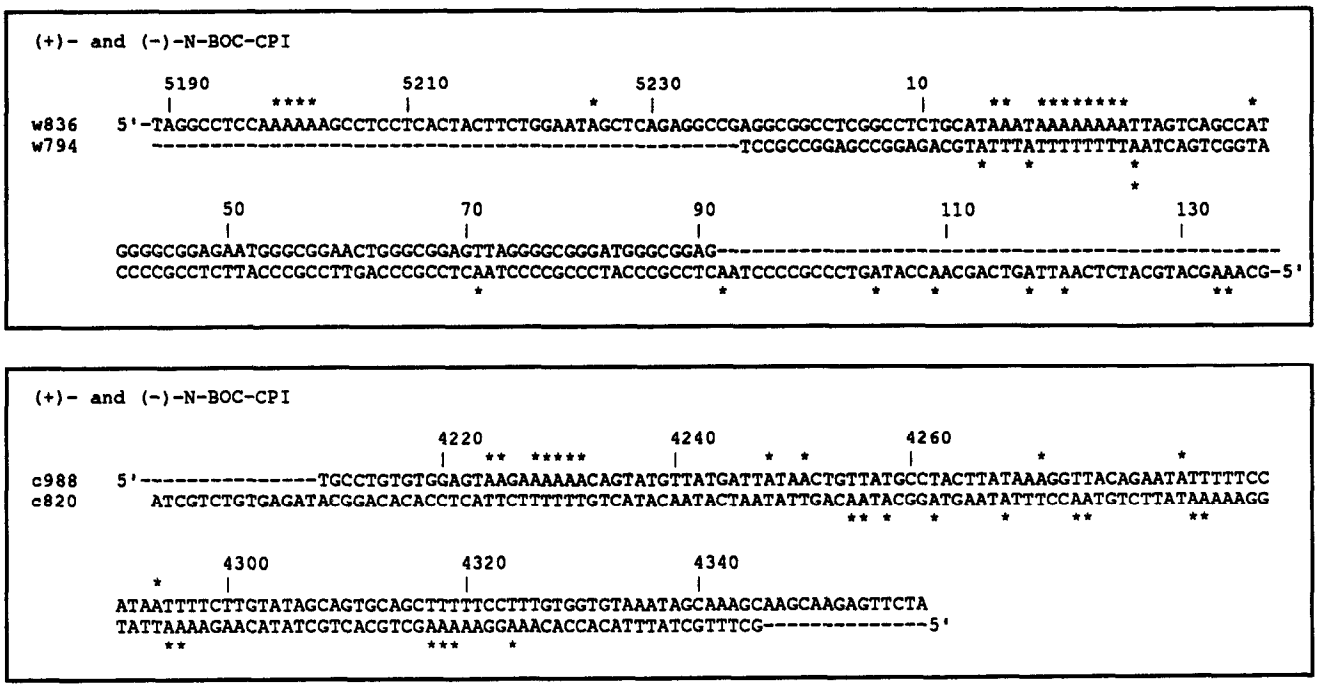


Figure 2. Summary of alkylation and cleavage sites for (+)- and *ent*-(–)-*N*-BOC-CPI, (*) denotes the alkylation and cleavage sites. The SV40 DNA numbering system employs the origin of replication (ORI) as a reference. Two regions of the SV40 DNA are represented. One region spans nucleotides no. 5189-138 and includes the origin of replication and part of the regulatory region. The other includes a segment of the SV40 early genes. The missing terminal regions of the SV40 DNA represented by a line constitute the nonoverlapping regions absent in the complementary clones and are not single-stranded segments of DNA. The data derived from c1346 are not shown

Table 1. Summary of alkylation sites for (+)- and (–)-*N*-BOC-CPI[a]

Sequence	Total Sites	Alkylation Sites	
5′-(NAAN)-3′	83	51	(61%)
5′-(AAAN)-3′	39	30	(77%)
5′-(TAAN)-3′	19	10	(53%)
5′-(GAAN)-3′	11	5	(45%)
5′-(CAAN)-3′	14	6	(43%)
5′-(NAAPu)-3′	54	36	(67%)
5′-(NAAPy)-3′	29	15	(52%)
5′-(NTAN)-3′	53	13	(25%)
5′-(ATAN)-3′	18	7	(39%)
5′-(TTAN)-3′	19	4	(21%)
5′-(GTAN)-3′	7	2	(29%)
5′-(CTAN)-3′	9	0	(00%)
5′-(NTAPu)-3′	29	9	(31%)
5′-(NTAPy)-3′	24	4	(17%)

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

Table 2. (+)- and (–)-*N*-BOC-CPI consensus sequence

base	-1	A	+1	3′
A (30)	71	100	46	
T (26)	26	--	20	
G (21)	0	--	21	
C (23)	3	--	13	
A/T (56)	97	100	66	
Pu (51)			67	
Py (49)			33	
Composite:	A > T	A	Pu > Py	

Two important distinctions in the DNA alkylation profiles of **2** and CC-1065 (**1**) deserve highlighting. Clear from the comparisons are the more selective DNA alkylation properties of (+)-**1** or (–)-**1** versus that of (+)-**2** and (–)-**2**. That is, (+)-**1** and (–)-**1** each alkylate a different subset of sites alkylated by (+)-**2** and (–)-**2**. However, while this simple description is nearly accurate, it is not entirely complete. There exists a subset of sites that either (+)-CC-1065 or *ent*-(–)-CC-1065 alkylates which is not alkylated by *N*-BOC-CPI. For (+)-CC-1065, this subset characteristically contains the 5'-TTA sequence which has proven to be a high affinity site for the natural agent and yet embodies the less preferred sequence for **2**, 5'-AA > 5'-TA. Characteristic of this distinction, (+)-CC-1065 alkylated 63% of all 5'-TTA sites[33,34] while (+)- and (–)-*N*-BOC-CPI alkylated only 21%. As such, (+)-CC-1065 and (–)-CC-1065 are not just more selective among the available alkylation sites, but are distinct in their alkylation profiles as well.

The concentration of (+)-*N*-BOC-CPI (**2**) required for observation of DNA alkylation proved comparable but slightly higher than that required for (+)-*N*-BOC-CBI (**5**)[38] or (+)-*N*-BOC-DSA (**6**) alkylation.[42] In part, this more efficient DNA alkylation reaction by (+)-*N*-BOC-DSA (**6**) > (+)-*N*-BOC-CBI (**5**) > (+)-*N*-BOC-CPI (**2**, *ca.* 5–10x) may be attributed to the enhanced stability of **6** and **5** versus **2** (**6** > **5** > **2**)[38,42] which leads to a more productive participation in the DNA alkylation reaction. However, this undoubtedly constitutes only part of the observed differences. The CPI-based agents participate in the DNA alkylation reaction at a rate which is slower than that of the less reactive agents (CBI and DSA-based agents) and this contributes to the DNA alkylation intensity differences.[38,42] As described in more detail later on, we believe this slower rate of DNA alkylation is derived from the increased steric hindrance surrounding the CPI C7 center.

In contrast to the observation that both enantiomers of *N*-BOC-CPI alkylate DNA with essentially the same selectivity, they do so with apparent differences in efficiency (*ca.* 10–100x). Similar observations have been made in the comparisons of (+)- and (–)-*N*-BOC-CBI (**5**),[38] but they are distinct from the results of the comparisons of (+)- and (–)-*N*-BOC-CI (**4**),[37] or (+)- and (–)-*N*-BOC-DSA (**6**).[42] With the latter agents, no distinctions in the DNA alkylation efficiencies were detected between the natural and unnatural enantiomers. As described in more detail later on, this distinction between the *N*-BOC-CPI enantiomers is the result of the steric hindrance surrounding the CPI C7 center for which the unnatural enantiomer is especially sensitive, *cf.* Figures 3 and 4.

In our comparison of the results derived from the evaluation of *N*-BOC-CPI (**2**), *N*-BOC-CI (**4**),[37] *N*-BOC-CBI (**5**),[38] and *N*-BOC-DSA (**6**),[42] it is clear that both enantiomers of all four classes of agents alkylate the same composite sites in DNA. The distinction among the agents lies not in the number of sites alkylated over an agent 10-fold concentration range, but in the efficiency of DNA alkylation and the relative selectivity among the available alkylation sites. Although the relative selectivity differences are subtle, the least reactive agents not only provide a more efficient DNA alkylation (*N*-BOC-DSA and *N*-BOC-CBI > *N*-BOC-CPI > *N*-BOC-CI), but also exhibit more discrimination among the available alkylation sites. While this unusual behavior of both enantiomers alkylating the same sites within DNA may appear unexpected, it illustrates that the event and selectivity of DNA alkylation are tied to the DNA versus agent structure. This is further reinforced by the observation that the characteristic alkylation profile for the four classes of agents is not uniquely observed with agents containing the activated cyclopropane but that other electrophiles incorporated into structurally related agents including **7** (R = H, OCH_3) behave similarly.[34,36,41] Thus, in addition to being independent of the absolute configuration of the agents, the characteristic alkylation selectivity also appears to be independent of the nature of the electrophile within the limited series of agents examined to date. These seminal observations not only demonstrated that *seco* agents may directly alkylate DNA without closure to the cyclopropane containing agents, but also that the characteristic alkylation selectivity of **2** cannot be the result of a sequence dependent autocatalytic phosphate protonation of the alkylation subunit carbonyl.[9]

The selectivity of DNA alkylation in past studies has been attributed to a sequence-dependent autocatalytic phosphate protonation of the alkylation subunit carbonyl,[9,26,30] the conformational variability of DNA and alkylation at junctions of bent DNA,[27–29] or alkylation within the narrower, sterically more accessible minor groove of AT-rich DNA.[11–17,33–39] Central to the distinctions in the interpretations have been the perceived similarities[26] or distinctions[33] in the alkylation selectivity of simple derivatives of the alkylation subunit versus the natural products themselves. Consistent with the latter interpretation of the origin of the alkylation selectivity, the most prevalent feature of the DNA alkylation sites is the neighboring 5' base-pairs, Tables 1–2. While this requirement for a 5' A or T base adjacent to the adenine alkylation site for both enantiomers of **2** may seem intuitively unlikely, constructed models of the diastereomeric alkylations proved revealing. For the natural enantiomer, this simply requires 3' adenine N3 alkylation and agent binding in the 3'→5' direction from the site of alkylation across the adjacent 5' base, Figures 3 and 4. For the unnatural enantiomer, this similarly occurs by adenine N3 alkylation but with the reversed 5'→3' binding orientation required for adenine N3 access to the activated cyclopropane, Figures 3 and 4. As a consequence of the diastereomeric relationship of the two alkylations, the binding of (–)-**2**, while oriented in the 5'→3' direction, covers the same adjacent 5' base as (+)-**2**. Thus, for simple derivatives of either enantiomer of the alkylation subunit, 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 its steric accessibility to the adenine N3 alkylation site. For simple derivatives of the alkylation subunits including *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. Illustrated in Figure 4 are enlarged models of the central four base-pairs of the natural and unnatural enantiomer *N*-BOC-CPI alkylations of the common 5'-d(GACTAATTTTT) site taken from Figure 3 and their comparisons with an unobserved and hypothetical alkylation at a similar site, 5'-d(GACTGATTTTT), bearing a 5' G base adjacent to the adenine N3 alkylation site. The C2 amine of the 5' G which extends into the minor groove clearly would sterically preclude either a natural or unnatural enantiomer alkylation at such a 5'-d(GA) site. Only when the adjacent 5' base is A or T, is sufficient groove penetration possible to permit the adenine N3 alkylation. This fundamentally simple and consistent interpretation now extended to include the unnatural enantiomer of **2** is identical to that disclosed in our earlier work with the enantiomeric CC-1065 agents[33] but is different from that advanced by others[26] where different polynucleotide recognition features have been suggested to be responsible for the behavior of each of the CC-1065 enantiomers.

The results have further suggested that the distinctions in the relative rate of DNA alkylation by (+)-*N*-BOC-DSA (**6**) versus (+)-*N*-BOC-CPI (**2**) may be interpreted simply in terms of steric accessibility to the alkylation site.[43] 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 DSA subunit which lacks a C7 substituent alkylates DNA at a more rapid rate, Figure 4. As discussed in more detail later on, the models illustrated in Figures 3 and 4 also reveal that the unnatural enantiomers are more sensitive to the C7 substituent destabilizing steric hindrance of the alkylation reaction as a consequence of its proximal positioning next to the alkylation site adjacent 5' base and suggest that this single structural feature is responsible for the diminished DNA alkylation properties of the unnatural enantiomer.

(+)-CC-1065 and (+)-CPI-CDPI$_2$

The DNA alkylation selectivity of (+)-CC-1065 (**1**) and (+)-CPI-CDPI$_2$ (**8**), which proved indistinguishable, has been discussed in detail in prior studies. Summarized in Figure 5 and Tables 3–4 are the alkylation sites for (+)-**1** and (+)-**8** observed in our studies and the composite consensus sequence derived from the observed alkylation sites.

In our protocol, each site of alkylation proved to be adenine and each alkylation site was flanked by two 5' A or T bases. The sequence preference for the three base AT-rich alkylation site was determined to follow the order 5'-AAA = 5'-TTA > 5'-TAA > 5'-ATA. In addition, the two agents exhibited a strong preference but not absolute requirement 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 versus pyrimidine base. No sensitivity to the nature of the sixth 5' base was perceptible. The preferences for the fourth and fifth 5' bases to be A or T versus G or C distinguish many of the high versus low affinity alkylation sites and the preference for a 3' purine versus pyrimidine base preceding the alkylation site was observed most prominently within the lower affinity sites. As detailed elsewhere,[33] we attribute the strict AT preference within the first three base pairs to represent a combination of the initial 3' adenine N3 alkylation site and an adjacent 5' two base AT site required to accommodate the binding of the agent central subunit. The weaker preferences for the fourth and fifth 5' bases to be A or T reflect, as observed with the high affinity alkylation sites, a preferential AT-rich site for the third subunit bound in the minor groove.

The extension of these studies to include agents containing modified alkylation subunits (i.e. CI-CDPI$_2$ (**9**),[34,37] CBI-CDPI$_2$ (**10**))[38] has proven especially revealing. Although the efficiency of DNA alkylation was shown to be inversely related to the agent's reactivity with the chemically more stable agents more effectively alkylating DNA (CBI > CPI > CI)[35] and the relative selectivity among the available alkylation sites was shown to increase with the decreased reactivity (CBI > CPI > CI), identical sites of DNA alkylation were observed with the modified agents. Especially important was the recognition that the characteristic alkylation selectivity of (+)-CC-1065 was observed even with (+)-CI-CDPI$_2$ (**9**) which possesses an exceptionally reactive alkylation subunit ($t_{1/2}$ = 35 s, pH 3; = 5.2 h, pH 7). We believe this study with the

exceptionally reactive CI-based agents represented a clear demonstration of the fundamentally important role that the AT-rich noncovalent binding selectivity of the agents[45,46] played in controlling the DNA alkylation selectivity and that the results of the study are inconsistent with the proposal that it is derived from the inherent selectivity of the CPI alkylation reaction itself.

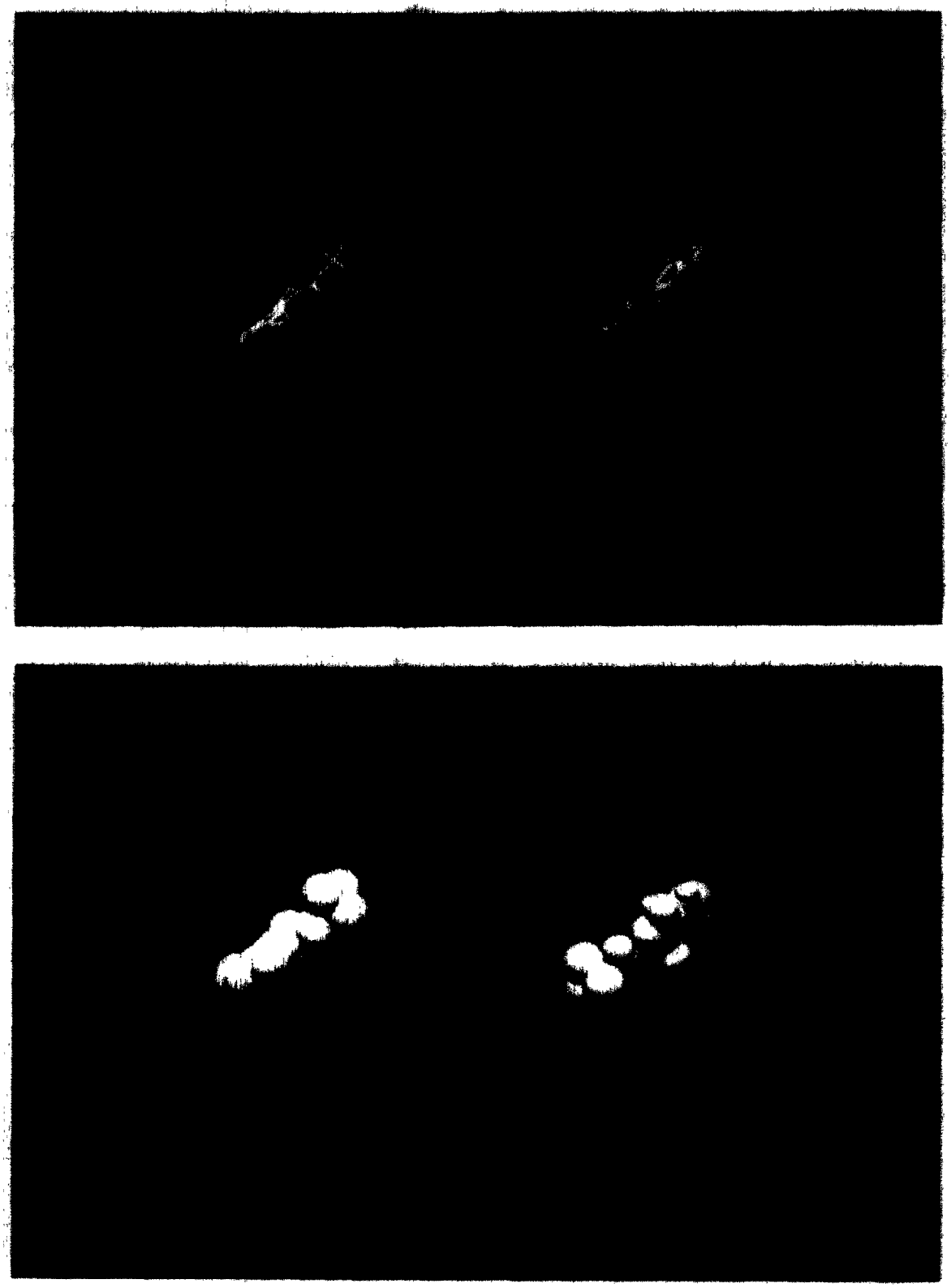


Figure 3. Comparison stick and space-filling models of the (+)-*N*-BOC-CPI (left) and *ent*-(–)-*N*-BOC-CPI (right) alkylation at the same site within w794 DNA: duplex 5'-d(GACTAATTTTT). The natural enantiomer binding extends in the 3'→5' direction from the adenine N3 alkylation site across the 5'-AA site. The unnatural enantiomer binding extends in the 5'→3' direction but binds across the same 5'-AA site. The model complexes were generated with MacroModel (AMBER force field supplemented with agent parameters)

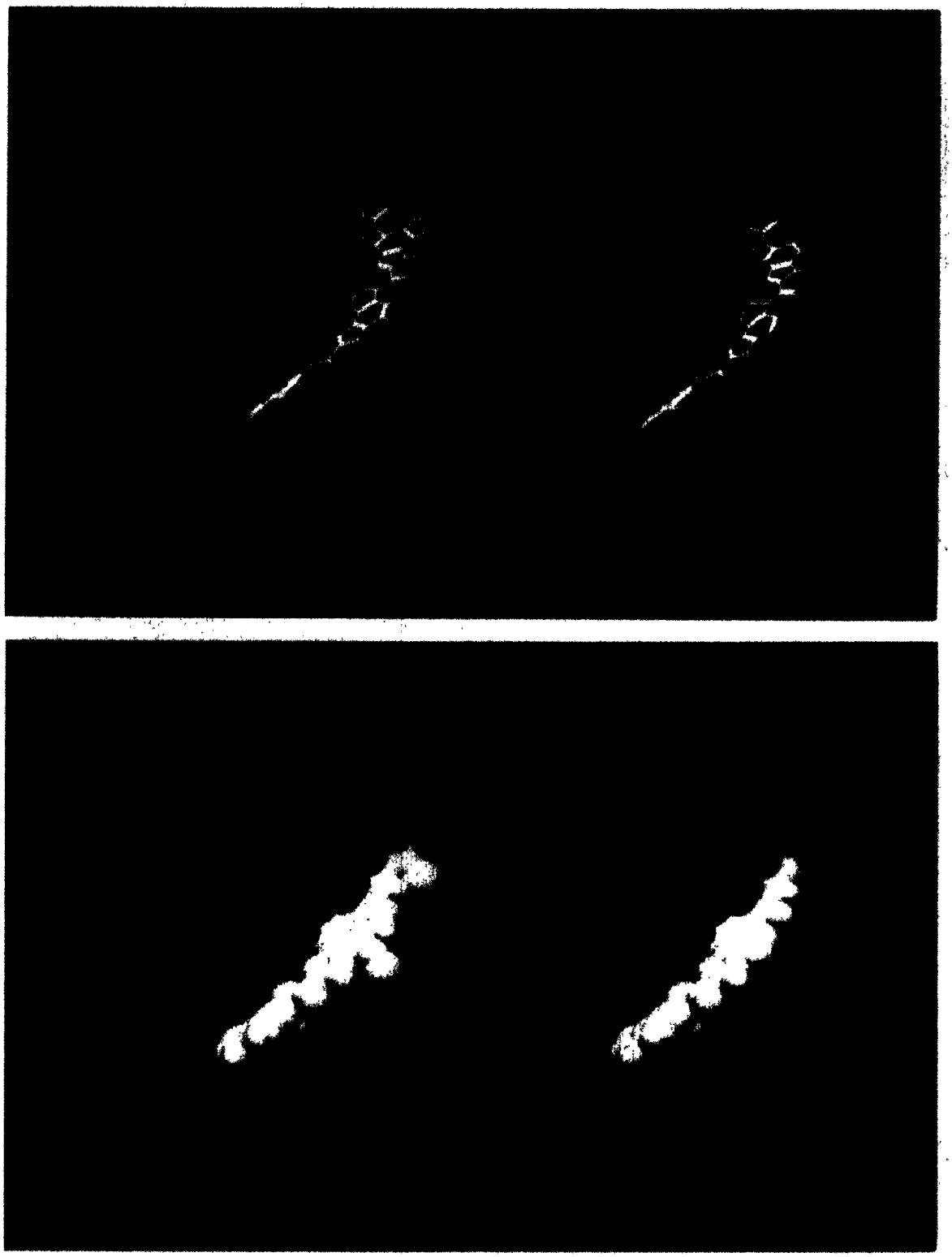


Figure 6. Comparison stick and space-filling models of the (+)-CC-1065 (left) and (+)-CPI-$CDPI_2$ (right) alkylation at the high affinity w794 DNA site: duplex 5'-d(CTCAATTAGTC). The agent binding extends in the 3'→5' direction from the adenine N3 alkylation site across the five base site 5'-AATTA. The model complexes were generated with MacroModel (AMBER force field supplemented with agent parameters)

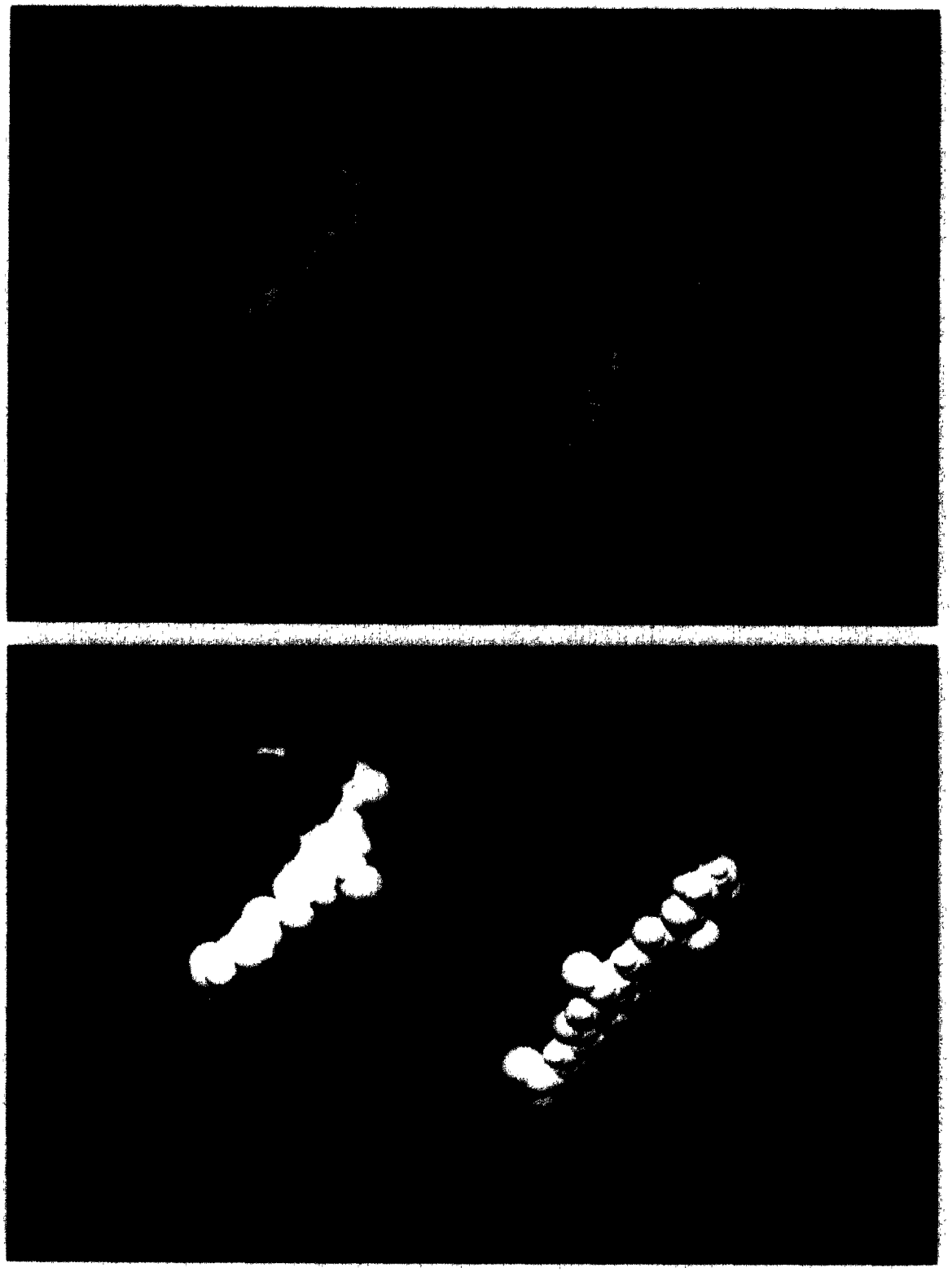


Figure 8. Comparison stick and space-filling models of the (+)-CC-1065 (left) and *ent*-(–)-CC-1065 (right) alkylation at the same site within w794 DNA: duplex 5'-d(GACTAATTTTT). The natural enantiomer binding extends in the 3'→5' direction from the adenine N3 alkylation site across the five base site 5'-ACTAA. The unnatural enantiomer binding extends in the 5'→3' direction across the site 5'-AATTT. The model complexes were generated with MacroModel (AMBER force field supplemented with agent parameters)

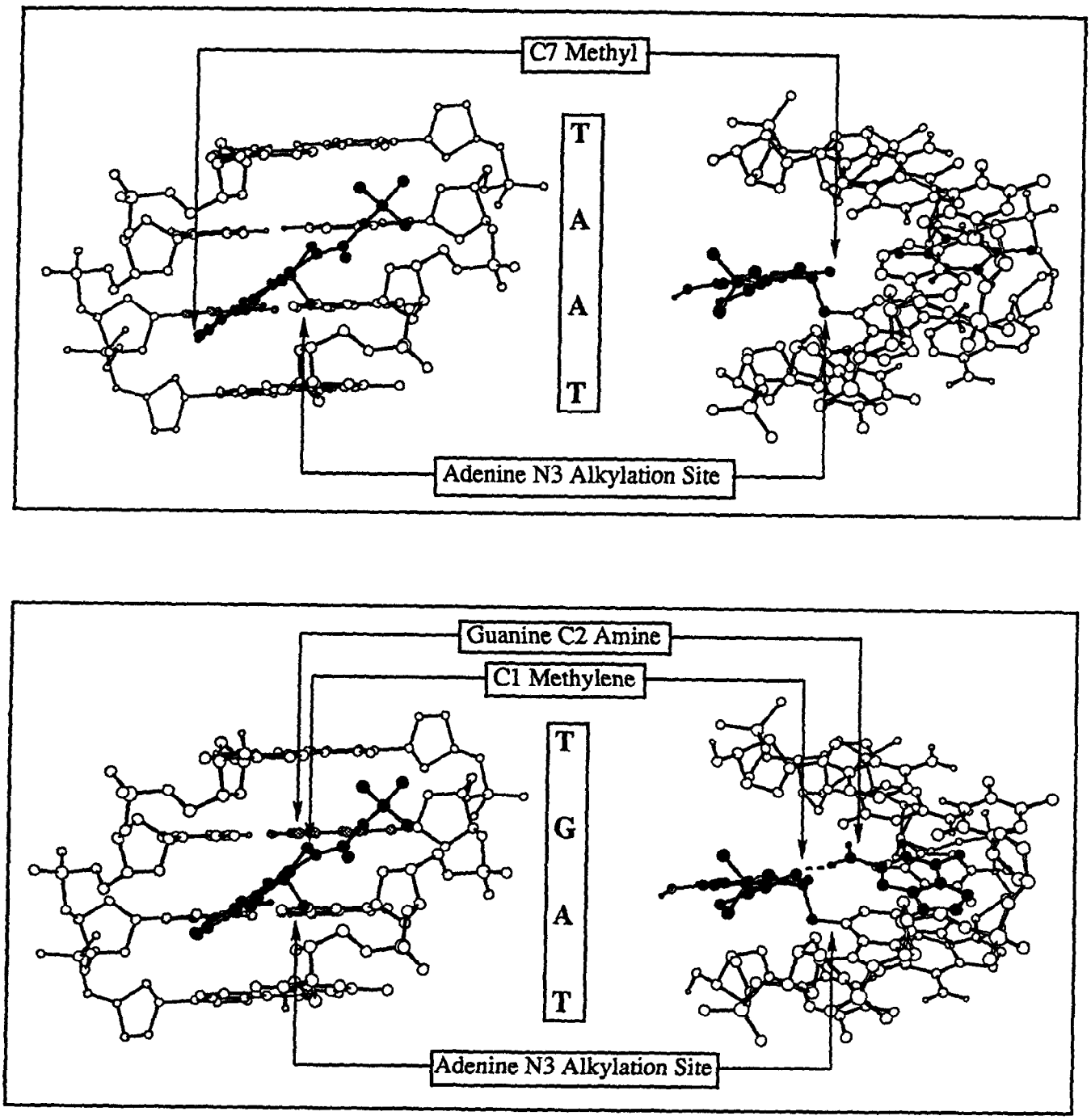


Figure 4-A. Top: Expanded and rotated views of the (+)-*N*-BOC-CPI 5'-AA alkylation taken from Figure 3 [5'-d(GACTAATTTTT) alkylation]. The natural enantiomer binds in the 3'→5' direction from the adenine N3 alkylation site extending across the 5'-AA site. The C7 methyl substituent is proximal to the complementary strand thymine (in black) and produces significant but not serious destabilizing steric interactions, C7-CH_3/thymine C=O distance = 2.99 Å, C1-CH_2/adjacent 5' adenine C2 distance = 3.62 Å. Bottom: Model of a comparable and hypothetical 5'-GA alkylation illustrating the destabilizing guanine C2 amine steric interactions that preclude reaction, C1-CH_2/adjacent 5' guanine C2-NH_2 distance = 1.56 Å

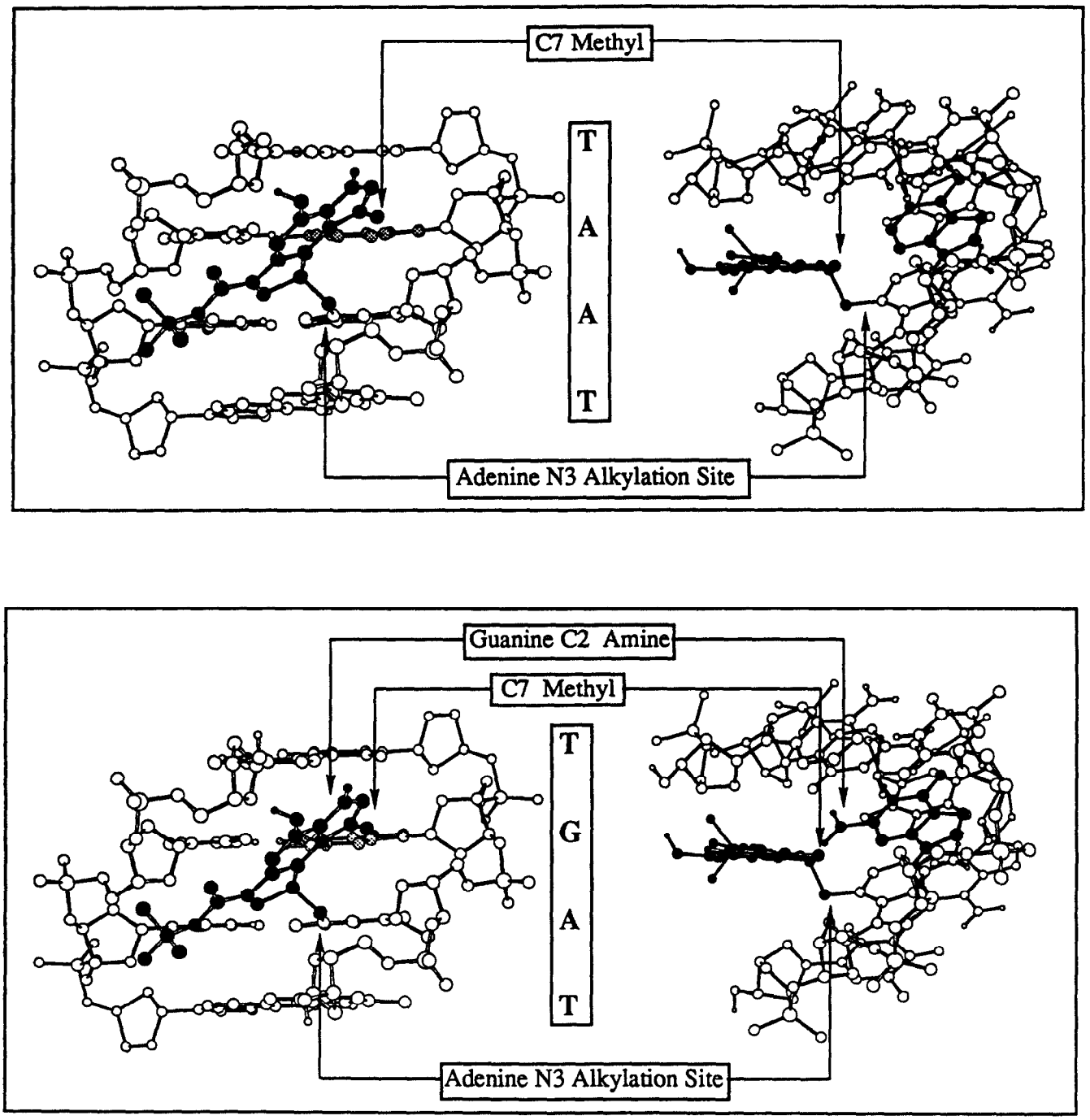
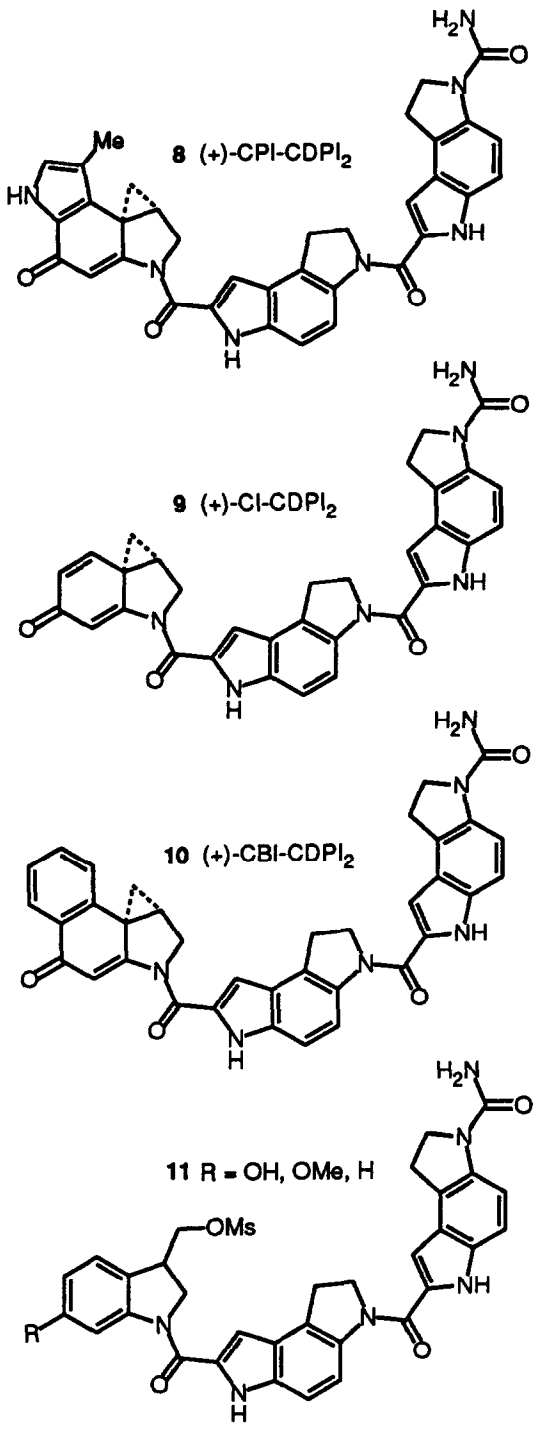


Figure 4-B. Top: Expanded and rotated views of the *ent*-(–)-*N*-BOC-CPI 5'-AA alkylation taken from Figure 3 [5'-d(GACTAATTTTT) alkylation]. The unnatural enantiomer binds in the 5'→3' (versus 3'→5') direction but with binding that covers the same 5'-AA site as the natural enantiomer. The model highlights the unnatural enantiomer sensitivity to the destabilizing steric interactions of the C7 methyl substituent with the alkylation site adjacent 5' adenine (in black) located on the same strand, C7-CH_3/adjacent 5' adenine N3 distance = 2.71 Å. Bottom: Model of a comparable and hypothetical 5'-GA alkylation illustrating the destabilizing guanine C2 amine steric interactions that preclude reaction, C7-CH_3/adjacent 5' guanine C2-NH_2 distance = 1.88 Å



An even clearer demonstration of the role that the noncovalent DNA binding selectivity was playing in the control of the alkylation selectivity was derived from a study of **11** (R = OH, OCH_3, H).[36] The three agents **11** (R = OH, OCH_3, H) exhibited an alkylation selectivity profile that was identical to that of (+)-CC-1065 and (+)-CPI-$CDPI_2$ (**8**). This alkylation selectivity proved readily distinguishable from that of **7** which proved identical to that of **4** (*N*-BOC-CI) and comparable to that observed with **2** (*N*-BOC-CPI). Thus, the characteristic alkylation selectivity of (+)-CC-1065 proved to be independent of the nature of the electrophile within the limited series of agents examined to date. These seminal observations demonstrated that *seco* agents may directly alkylate DNA without closure to the cyclopropane containing agents, that the alkylation selectivity of **11** and (+)-**1** cannot be derived from a sequence dependent phosphate protonation of the alkylation subunit carbonyl[9] and that the noncovalent binding selectivity of the agents contributes prominently to observation of the (+)-CC-1065 characteristic alkylation selectivity.[36]

Comparison models of the (+)-CC-1065 and (+)-CPI-$CDPI_2$ (**8**) alkylations within the high affinity site of w794 DNA, 5'-AATTA, are illustrated in Figure 6. The bound helical conformation of the agents complements the topological curvature and pitch of the minor groove with the agent binding spanning a five base-pair AT-rich site in the 5' direction from the adenine N3 alkylation site. The hydrophobic, concave face of the agent is deeply imbedded in the minor groove and the polar functionality of the agent lies on the outer face of the complex.

We attribute the alkylation selectivity of the agent in part to the AT-rich noncovalent bonding selectivity of the agent preferentially within the narrower, sterically more accessible AT-rich minor groove[45,46] and the five base-pair binding site size required to permit full agent binding. This nicely explains the absolute requirement for the first three base-pairs of the alkylation sites to be A or T that extends less rigidly to the fourth and fifth base-pairs.

Importantly, this interpretation should not be misconstrued to imply that AT-rich binding leads necessarily to productive DNA alkylation. Rather, the noncovalent binding of the agents may serve to further restrict or alter 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 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. As detailed earlier, sufficient minor groove penetration is possible with a single 5' A or T base adjacent to the alkylation site for simple derivatives (i.e. *N*-BOC-CPI) of the alkylation subunit. For (+)-**1** and (+)-**8** sufficient minor groove penetration may be possible only when two or more adjacent 5' bases are A or T and optimal minor groove binding and penetration may be possible when three or four adjacent 5' bases are A or T.

In addition to significantly enhancing the sequence selectivity of the DNA alkylation reaction, the central subunit and right hand subunits of (+)-**1** serve the additional important functional role of stabilizing the inherently reversible nature[31] of the adenine N3 alkylation reaction. We further suggest that it is the simple event of this noncovalent binding stabilization of the readily reversible DNA alkylation reaction that is responsible for the potent biological activity of (+)-**1** and (+)-**8** versus (+)-**2**.

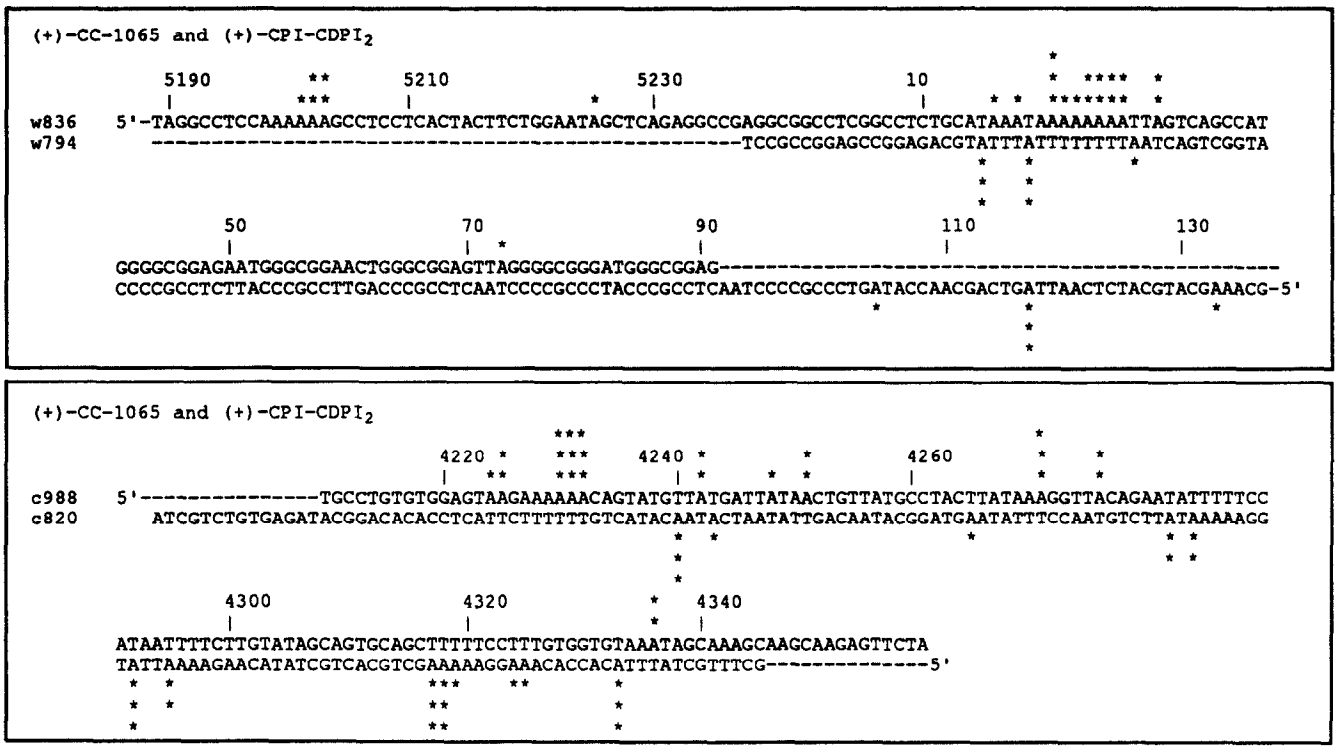


Figure 5. Summary of alkylation and cleavage sites for (+)-CC-1065 and (+)-CPI-CDPI$_2$, (*) denotes the alkylation and cleavage sites. The relative intensity of alkylation and cleavage is represented by the number of symbols above a given site. The SV40 DNA numbering system employs the origin of replication (ORI) as a reference. Two regions of the SV40 DNA are represented. One region spans nucleotides no. 5189-138 and includes the origin of replication and part of the regulatory region. The other includes a segment of the SV40 early genes. The missing terminal regions of the SV40 DNA represented by a line constitute the nonoverlapping regions absent in the complementary clones and are not single-stranded segments of DNA. The data derived from c1346 are not shown

Table 3. (+)-CC-1065 and (+)-CPI-CDPI$_2$ consensus sequence

base	-4	-3	-2	-1	A	+1	3′
A (30)	53	55	56	67	100	31	
T (26)	14	23	38	31	--	24	
G (21)	23	11	2	0	--	31	
C (23)	10	11	4	2	--	14	
A/T (56)	67	78	94	98	100	55	
Pu (51)						62	
Py (49)						38	
Composite:	A/T ≥ G/C	A/T > G/C	A/T	A/T	A	Pu ≥ Py	

ent-(–)-CC-1065 and ent-(–)-CPI-CDPI$_2$

In an initially unexpected observation, *ent*-(–)-CC-1065 (**1**) and *ent*-(–)-CPI-CDPI$_2$ (**8**) were found to be active cytotoxic agents (20 pM, L1210) equipotent with the natural enantiomers. Consistent with this observation, *ent*-(–)-CC-1065 and *ent*-(–)-CPI-CDPI$_2$ and their corresponding *seco* derivatives were found to alkylate DNA at concentrations approximately equivalent to that of the natural enantiomers. The alkylation profiles of (–)-**1** and (–)-**8** were indistinguishable from one another and they were found to be readily distinguishable from the natural enantiomer alkylation profile. Table 6 and Figure 7 summarize the alkylation sites for (–)-**1** and (–)-**8** observed in our studies and Table 5 summarizes their composite alkylation consensus sequence. Notably, each alkylation site proved to be adenine and no minor guanine alkylation was detected in our protocol. Both the rate and intensity of the *ent*-(–)-CC-1065 alkylation were comparable to those of the natural enantiomer.

Table 4. Summary of alkylation sites for (+)-CC-1065 and (+)-CPI-$CDPI_2$

Sequence	No. AS[a]/No. TS		High Affinity[b]		Low Affinity[b]
5′-(NNAAAN)-3′	25/39	(64%)	10	(26%)	15
5′-(NNTTAN)-3′	12/19	(63%)	7	(37%)	5
5′-(NNTAAN)-3′	9/19	(47%)	4	(22%)	5
5′-(NNATAN)-3′	4/18	(22%)	0	(00%)	4
5′-(NNTTAN)-3′					
5′-(NATTAN)-3′	4/4	(100%)	3	(75%)	1
5′-(A/TATTAN)-3′	3/3	(100%)	3	(100%)	0
5′-(G/CATTAN)-3′	1/1	(100%)	0	(00%)	1
5′-(NTTTAN)-3′	5/6	(83%)	4	(67%)	1
5′-(A/TTTTAN)-3′	4/4	(100%)	4	(100%)	0
5′-(G/CTTTAN)-3′	1/2	(50%)	0	(00%)	1
5′-(NG/CTTAN)-3′	3/9	(33%)	0	(00%)	3
5′-(NGTTAN)-3′	3/5	(60%)	0	(00%)	3
5′-(NCTTAN)-3′	0/4	(00%)	0	(00%)	0
5′-(NNTTAPu)-3′	4/6	(67%)	3	(50%)	1
5′-(NNTTAPy)-3′	8/13	(62%)	4	(31%)	4
5′-(NNAAAN)-3′					
5′-(NAAAAN)-3′	17/21	(81%)	8	(38%)	9
5′-(A/TAAAAN)-3′	13/14	(93%)	6	(43%)	7
5′-(G/CAAAAN)-3′	4/7	(57%)	2	(29%)	2
5′-(NTAAAN)-3′	4/6	(67%)	1	(17%)	3
5′-(A/TTAAAN)-3′	3/4	(75%)	1	(25%)	2
5′-(G/CTAAAN)-3′	1/2	(50%)	0	(00%)	1
5′-(NG/CAAAN)-3′	4/12	(33%)	1	(08%)	3
5′-(NGTAAN)-3′	1/5	(20%)	0	(00%)	1
5′-(NCAAAN)-3′	3/7	(43%)	1	(14%)	2
5′-(NNAAAPu)-3′	19/30	(63%)	8	(27%)	11
5′-(NNAAAPy)-3′	6/9	(67%)	2	(22%)	4
5′-(NNTAAN)-3′					
5′-(NTTAAN)-3′	2/3	(67%)	1	(33%)	1
5′-(A/TTTAAN)-3′	1/1	(100%)	1	(100%)	0
5′-(G/CTTAAN)-3′	1/2	(50%)	0	(00%)	1
5′-(NATAAN)-3′	5/9	(56%)	3	(33%)	2
5′-(A/TATAAN)-3′	3/5	(60%)	1	(20%)	2
5′-(G/CATAAN)-3′	2/4	(50%)	2	(50%)	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′	5/9	(56%)	2	(22%)	3
5′-(NNTAAPy)-3′	4/10	(40%)	2	(20%)	2
5′-(NNATAN)-3′					
5′-(NAATAN)-3′	2/5	(40%)	0	(00%)	2
5′-(A/TAATAN)-3′	1/3	(33%)	0	(00%)	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′	2/7	(29%)	0	(00%)	2
5′-(NGATAN)-3′	0/0	--	-	(00%)	-
5′-(NCATAN)-3′	2/7	(29%)	0	(00%)	2
5′-(NNATAPu)-3′	3/14	(21%)	0	(00%)	3
5′-(NNATAPy)-3′	1/4	(25%)	0	(00%)	1

[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/number of total sites (No. TS).

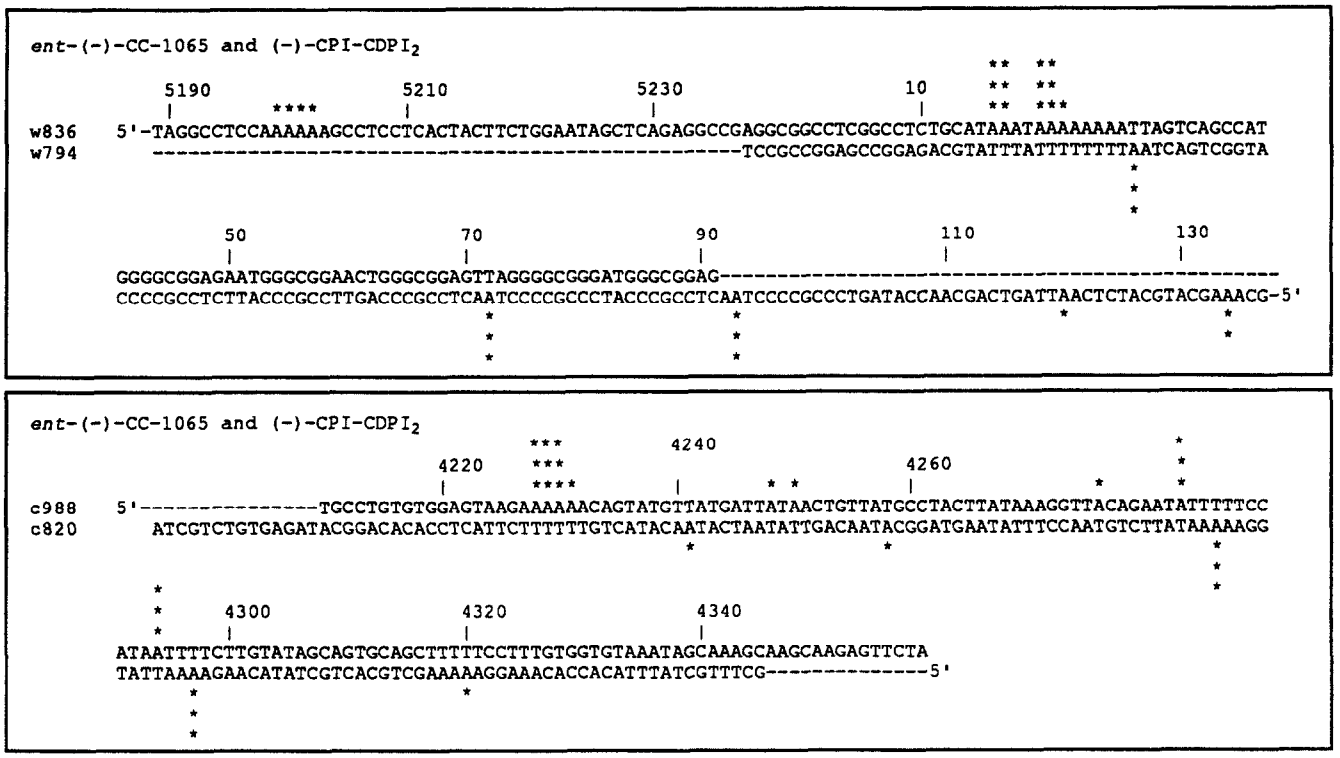


Figure 7. Summary of alkylation and cleavage sites for (–)-CC-1065 and (–)-CPI-CDPI$_2$, (*) denotes the alkylation and cleavage sites. The relative intensity of alkylation and cleavage is represented by the number of symbols above a given site. The SV40 DNA numbering system employs the origin of replicaiton (ORI) as a reference. Two regions of the SV40 DNA are represented. One region spans nucleotides no. 5189-138 and includes the origin of replication and part of the regulatory region. The other includes a segment of the SV40 early genes. The missing terminal regions of the SV40 DNA represented by a line constitute the nonoverlapping regions absent in the complementary clones and are not single-stranded segments of DNA. The data derived from c1346 are not shown

Table 5. *Ent*-(–)-CC-1065 and *ent*-(–)-CPI-CDPI$_2$ consensus sequence

base	-3	-2	-1	A	+1	+2	+3	+4	3′
A (30)	49	40	46	100	77	61	43	30	
T (26)	9	30	42	--	16	23	30	26	
G (21)	15	18	0	--	0	6	18	21	
C (23)	27	12	12	--	7	12	9	23	
A/T (56)	58	70	88	100	93	82	73	56	
Composite:	N	A/T ≥ G/C	A/T	A	A/T	A/T > G/C	A/T > G/C	N	

Each alkylation site proved consistent with 5' adenine alkylation, agent binding in the minor groove in the 5'→3' direction from the alkylation site covering five base-pairs across an AT-rich region. All alkylation sites detected proved to be adenine and nearly all of the 3' and 5' bases flanking the adenine N3 alkylation site proved to be A or T, Tables 5 and 6. The exceptions to this generalization represent a tolerance for a 3' cytosine adjacent to a 5'-T<u>A</u> alkylation site or a 5' cytosine adjacent to a 5'-<u>A</u>A or 5'-<u>A</u>T alkylation site. Like the natural enantiomer, there proved to be a preference for this three base sequence which follows the order 5'-A<u>A</u>A = 5'-T<u>A</u>A > 5'-A<u>A</u>T, 5'-T<u>A</u>T (Table 6) with a pronounced preference for the first two sequences. There also proved to be a substantial preference for the second and third 3' base to be A or T versus G or C. At least part of origin for the preference for the enriched A versus mixed AT-rich sites for *ent*-(–)-CC-1065 (5'-A<u>A</u>A, 5'-T<u>A</u>A > 5'-A<u>A</u>T, 5'-T<u>A</u>T) also observed with *N*-BOC-CPI (5'-A<u>A</u> > 5'-T<u>A</u>) and (+)-CC-1065 (5'-AA<u>A</u> = 5'-TT<u>A</u> > 5'-TA<u>A</u> > 5'-AT<u>A</u>) lies in the statistically relevant opportunity for alkylation of a complementary strand adenine within the mixed AT-rich sequences competitive with the labeled strand site. Presumably, a competitive complementary strand alkylation would preclude an additional labeled strand alkylation within the same overlapping site and lower the apparent alkylation efficiency.

Table 6. Summary of DNA alkylation sites for *ent*-(–)-CC-1065 and (–)-CPI-$CDPI_2$

Sequence	Total Sites	Alkylation Sites	
5′-(NAAN)-3′	53	14	(26%)
5′-(NAAA)-3′	26	12	(46%)
5′-(NAAT)-3′	11	2	(18%)
5′-(NAAG)-3′	8	0	(00%)
5′-(NAAC)-3′	9	0	(00%)
5′-(NTAN)-3′	41	10	(24%)
5′-(NTAA)-3′	14	6	(43%)
5′-(NTAT)-3′	14	2	(14%)
5′-(NTAG)-3′	7	0	(00%)
5′-(NTAC)-3′	6	2	(33%)
5′-(NGAN)-3′	18	0	(00%)
5′-(NGAA)-3′	8	0	(00%)
5′-(NGAT)-3′	2	0	(00%)
5′-(NGAG)-3′	7	0	(00%)
5′-(NGAC)-3′	1	0	(00%)
5′-(NCAN)-3′	30	3	(10%)
5′-(NCAA)-3′	6	2	(33%)
5′-(NCAT)-3′	12	1	(08%)
5′-(NCAG)-3′	8	0	(00%)
5′-(NCAC)-3′	4	0	(00%)

Sequence	No. AS[a]/No. TS		High Affinity[b]		Low Affinity[b]
5′-(NAAAN)-3′	12/26	(46%)	6	(23%)	6
5′-(NTAAN)-3′	6/14	(43%)	4	(29%)	2
5′-(NAATN)-3′	2/11	(18%)	2	(18%)	0
5′-(NTATN)-3′	2/14	(14%)	1	(07%)	1
5′-(NAAAN)-3′					
5′-(NAAAA)-3′	9/16	(56%)	5	(31%)	4
5′-(NAAAT)-3′	1/4	(25%)	1	(25%)	0
5′-(NAAAG)-3′	2/5	(40%)	0	(00%)	2
5′-(NAAAC)-3′	0/1	(00%)	0	(00%)	0
5′-(NTAAN)-3′					
5′-(NTAAA)-3′	2/3	(67%)	2	(67%)	0
5′-(NTAAT)-3′	0/3	(00%)	0	(00%)	0
5′-(NTAAG)-3′	0/2	(00%)	0	(00%)	0
5′-(NTAAC)-3′	4/6	(67%)	2	(33%)	2
5′-(NAATN)-3′					
5′-(NAATA)-3′	0/4	(00%)	0	(00%)	0
5′-(NAATT)-3′	2/5	(40%)	2	(40%)	0
5′-(NAATG)-3′	0/1	(00%)	0	(00%)	0
5′-(NAATC)-3′	0/1	(00%)	0	(00%)	0
5′-(NTATN)-3′					
5′-(NTATA)-3′	1/5	(20%)	0	(00%)	1
5′-(NTATT)-3′	1/4	(25%)	1	(25%)	0
5′-(NTATG)-3′	0/5	(00%)	0	(00%)	0
5′-(NTATC)-3′	0/0	(--)	-	(--)	-

[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/number of total sites (No. TS).

Although this (–)-CC-1065/(–)-CPI-$CDPI_2$ AT-rich alkylation selectivity relative to the adenine N3 alkylation site is offset from that observed with (+)-CC-1065/(+)-CPI-$CDPI_2$, it is the natural consequence of the diastereomeric relationship of the adducts and is fully consistent with a binding orientation that extends in the 5'→3' direction. Thus, the *ent*-(–)-CC-1065 DNA alkylation proved to be analogous to that of (+)-CC-1065 with the exception that the unnatural enantiomer binding direction (5'→3') across an AT-rich region is reversed and offset from that observed with the natural enantiomer (3'→5').

Given the structural conformation of the adenine N3 adduct in which the least substituted cyclopropane carbon was found to be alkylated,[26] the established absolute configuration of the agent, and the AT-rich selectivity of the alkylation which extends in the 5'→3' direction from the adenine alkylation site, a model of the (–)-CC-1065 alkylation of duplex DNA may be constructed. This is illustrated in Figure 8 with the w794 site 5'-(GACTAATTTTT) which constitutes the high affinity site for (–)-CC-1065 and a minor alkylation site for (+)-CC-1065.

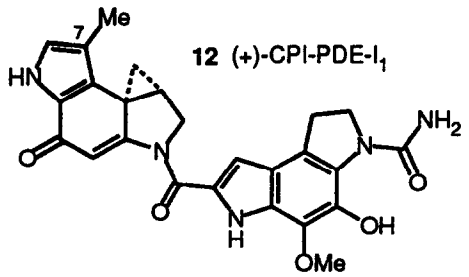
For both enantiomers, the hydrophobic concave face of the agent is bound deeply in the minor groove across an AT-rich region, the polar functionality of the agent lies on the outer face of the complex, and the agents adopt a bound helical conformation that follows the curvature and pitch of the minor groove covering five base-pairs. The difference in the two models is that (–)-CC-1065 binds in the reverse 5'→3' direction from the site of alkylation while (+)-CC-1065 binds in the 3'→5' direction from the same site. For (+)-CC-1065, the binding spans five base-pairs starting with the 3' adenine alkylation site and extending in the 3'→5' direction over the four adjacent 5' base pairs (5'-ACTAA). For (–)-CC-1065, the binding similarly spans a five base-pair AT-rich site but starts at the 5' base adjacent to the alkylation site and extends in the 5'→3' direction to cover the adenine alkylation site and the three adjacent 3' base-pairs (5'-AATTT). Notable is the fact that this offset alkylation selectivity within an AT-rich site is the natural consequence of the diastereomeric relationship of the alkylated DNA and the required reversed binding orientation of the agents in the minor groove to permit adenine N3 access to the electrophilic cyclopropane. Moreover, the model embodies the inherent deep groove penetration of the full bound agent at the appropriate AT-rich five base-pair site surrounding the alkylation site. The (–)-CC-1065 binding illustrated in Figure 8 spans a five base AT-rich site 5'-(AATTT) while the (+)-CC-1065 binding spans a three base AT-rich site 5'-(ACTAA). The relative importance or preference of the fourth base (A/T > G/C) in the binding sequences may be the reason this site constitutes a high affinity site for the unnatural enantiomer but only constitutes a minor site for the natural enantiomer.

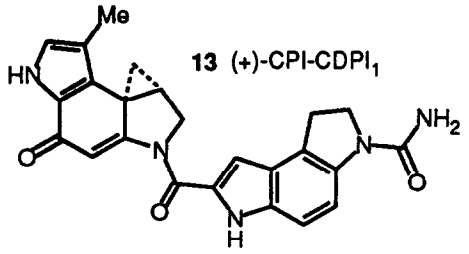
(+)-CPI-PDE-I_1/(+)-CPI-$CDPI_1$ versus ent-(–)-CPI-PDE-I_1 /(–)-CPI-$CDPI_1$: origin of the distinctions in the natural and unnatural enantiomer DNA alkylation capabilities

The comparisons of (+)-CPI-PDE-I_1 (**12**), lacking the right hand subunit of CC-1065, and (+)-CPI-$CDPI_1$ (**13**) with their unnatural enantiomers proved to be especially revealing. Both (+)-**12** and (+)-**13** were found to effectively alkylate duplex DNA although both the rate and efficiency of alkylation proved slightly lower in our protocol[33] than that of (+)-**1** and (+)-**8**. Moreover, the alkylation profiles of (+)-**12** and (+)-**13** were indistinguishable from one another and strikingly similar yet perceptibly different from those of (+)-**1** and (+)-**8**. Table 7 summarizes the composite consensus alkylation sequence derived from our studies. Like (+)-CC-1065 and (+)-CPI-$CDPI_2$ (**8**), each alkylation site detected in our protocol proved to be adenine and each site was flanked by two 5' A or T bases. The same preference for this three base AT alkylation site that was observed with (+)-CC-1065 was observed with (+)-**12**/(+)-**13**: 5'-AAA = 5'-TTA > 5'-TAA > 5'-ATA. The distinctions between (+)-**12**/(+)-**13** and (+)-CC-1065 lie in the extended AT preference within the fourth and fifth 5' bases from the alkylation site. With (+)-**12**/(+)-**13**, a weaker preference for the fourth 5' A or T base was observed and, unlike (+)-CC-1065, no preference for the fifth 5' base was observed for (+)-**12**/(+)-**13**. Importantly, this distinction manifests itself in the relative selectivity of alkylation among the available sites rather than alkylation at a distinguishing set of unique sites. That is, the major and minor alkylation site distinctions observed with (+)-CC-1065/(+)-CPI-$CDPI_2$ attributable to the extended five base AT-rich selectivity were not observed as prominently with (+)-**12**/(+)-**13**.

More interesting was the behavior of the unnatural enantiomers of **12**–**13**. Both (–)-**12** and (–)-**13** failed to alkylate DNA at concentrations of 100–1000x that required for (+)-**12**/(+)-**13** and the DNA alkylation detected at higher concentrations proved to be derived from contaminant natural enantiomer in the resolved samples (≤ 0.1%). Consequently, unlike the unnatural enantiomers of **1** and **8** but like the unnatural enantiomer of **2**, the unnatural enantiomers of **12** and **13** proved to alkylate DNA much more slowly, much less efficiently, and perhaps providing less stable adducts than the natural enantiomers.

The studies have suggested an attractive explanation for this sometimes confusing behavior (relative DNA alkylation efficiency, relative biological potency) of pairs of enantiomeric agents which we believe may be traced to a single structural feature—the relative degree of steric bulk at the site surrounding the CPI C7 center within the alkylation subunit for which the unnatural enantiomers are especially sensitive. The enantiomeric distinctions have proven readily detectable with the simple alkylation subunits themselves (i.e. *N*-BOC-CPI, -CBI), are most prominent with the dimer based agents (i.e. CPI-$CDPI_1$, CPI-PDE-I_1, CI–$CDPI_1$, CBI-$CDPI_1$, duocarmycin SA, duocarmycin A, Table 8), and are less prominent with the larger trimer/tetramer based agents (i.e. CC-1065, CPI-$CDPI_2$, CBI-$CDPI_2$, CI-$CDPI_2$, CPI–$CDPI_3$). In addition, less distinction has been observed with the CI- and duocarmycin SA enantiomeric pairs, both of which lack substituents or steric bulk at this position within the





alkylation subunit. Moreover, the distinctions among enantiomeric CI-based agents which lack the pyrrole ring altogether are generally smaller than those observed with the duocarmycin SA derived agents (DSA > CI). In contrast, the CPI-, CBI-, and duocarmycin A based agents generally exhibit more pronounced distinctions within enantiomeric pairs (CPI > CBI > duocarmycin A) and each possesses a substituent or inherent steric bulk at this position (CPI > CBI > duocarmycin A). We believe the distinguishing behavior of the CPI unnatural enantiomers as well as that of all agents studied to date is derived from a pronounced steric interaction of the CPI C7 substituent with the 5' base adjacent to the adenine N3 alkylation site present with the 5'→3' binding of the unnatural enantiomers, see Figures 3, 4 and 8. The 3'→5' binding model for the natural enantiomers benefits from a less substantial but still significant sensitivity to steric bulk at this same C7 position, Figure 4.

Consistent with the empirical observations to date and in agreement with models of the alkylation adducts, the studies suggest that agents lacking steric bulk surrounding the position occupied by a CPI C7 substituent are, predictably, more effective at alkylating DNA (rate and/or stability), that the unnatural enantiomers of the agents examined to date are especially sensitive to steric bulk at this position, and that as this steric bulk is reduced the distinguishing features of the enantiomers (efficiency of DNA alkylation, biological potency) diminish.

The fact that the monomer (i.e. *N*-BOC-CPI) and dimer (i.e. CPI-$CDPI_1$) based agents are most sensitive to the enantiomeric distinctions provides an additional and indirect verification of the fundamental role of the agent noncovalent binding stabilization of the reversible alkylation adduct formation. The monomer based agents (*N*-BOC-CPI) result in readily reversible adduct formation with the natural CPI enantiomer adduct being more stable as a direct result of the unnatural enantiomer adduct sensitivity to this destabilizing C7 methyl group steric interaction. The natural enantiomer of the dimer based agents (i.e. CPI-$CDPI_1$) are effective DNA alkylating agents and are biologically equipotent with CC-1065 illustrating that a single DNA binding subunit is sufficient to fully stabilize the inherently reversible natural enantiomer alkylation. In contrast, the dimer based unnatural enantiomer which provides an inherently less stable adduct as a consequence of the C7 methyl substituent destabilizing steric interaction is not sufficiently stabilized by the additional DNA binding subunit to provide fully stabilized adduct formation. With the trimer based agents including CC-1065 and CPI-$CDPI_2$, the noncovalent binding stabilization is extensive and sufficient to fully stabilize both the natural and inherently less stable unnatural enantiomer adduct formation.

(+)-CPI-$CDPI_3$ and ent-(–)-CPI-$CDPI_3$

In a final and revealing set of comparisons, the extended analogs of CC-1065, (+)-CPI-$CDPI_3$ (**14**) and its unnatural enantiomer (–)-CPI-$CDPI_3$, were evaluated. Both (+)- and (–)-**14** were found to effectively alkylate duplex DNA.

In addition, the alkylation profiles of (+)- and (–)-**14** were readily distinguishable from one another and were

Table 7. (+)-CPI-PDE-I_1 and (+)-CPI-$CDPI_1$ consensus sequence

base	-3	-2	-1	0	+1	3′
A (30)	43	57	64	100	29	
T (26)	22	43	36	--	29	
G (21)	14	0	0	--	29	
C (23)	21	0	0	--	13	
A/T (56)	65	100	100	100	58	
Pu (51)					58	
Py (49)					42	
Composite:	A/T ≥ G/C	A/T	A/T	A	Pu ≥ Py	

Table 8.

Agent	Configuration	IC_{50} (ng/mL) L1210	rel. IC_{50}	[a]DNA (+)/(-)
(+)-CI-$CDPI_1$	natural	10	1	0.5-2.0
(-)-CI-$CDPI_1$	unnatural	20	0.5	
(+)-duocarmycin SA	natural	0.006	1	10
(-)-duocarmycin SA	unnatural	0.06	0.1	
(+)-duocarmycin A	natural	0.1	1	> 100
(-)-duocarmycin A	unnatural	> 10	< 0.01	
(+)-CBI-$CDPI_1$	natural	0.002	1	> 100
(-)-CBI-$CDPI_1$	unnatural	> 2	< 0.001	
(+)-CPI-PDE-I_1	natural	0.01	1	> 100
(-)-CPI-PDE-I_1	unnatural	> 1.5	< 0.006	
(+)-CPI-$CDPI_1$	natural	0.02	1	> 100
(-)-CPI-$CDPI_1$	unnatural	> 3	< 0.006	

[a]Relative efficiency of the natural(+)/unnatural(-) enantiomer DNA alkyaltion within w794 DNA (25 °C, 24 h).

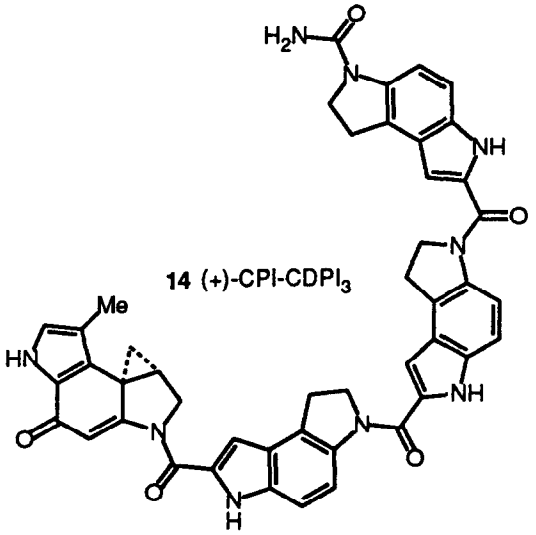
Table 9. (+)-CPI-$CDPI_3$ consensus sequence

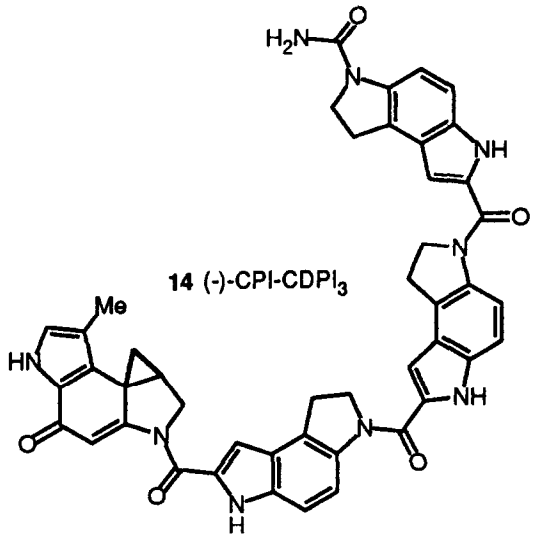
base	-6	-5	-4	-3	-2	-1	A	+1	3′
A (30)	36	39	46	46	38	46	100	15	
T (26)	36	39	23	23	62	54	--	39	
G (21)	18	15	23	16	0	0	--	23	
C (23)	8	8	8	15	0	0	--	23	
A/T (56)	72	77	69	69	100	100	100	54	
Composite:	A/T > G/C	A/T > G/C	A/T > G/C	A/T > G/C	A/T	A/T	A	N	

Table 9. *Continued*

(–)-CPI-CDPI$_3$ consensus sequence

base	-2	-1	A	+1	+2	+3	+4	+5	3′
A (30)	29	57	100	43	29	29	43	14	
T (26)	57	43	--	57	71	57	43	43	
G (21)	14	0	--	0	0	0	0	14	
C (23)	0	0	--	0	0	14	14	29	
A/T (56)	86	100	100	100	100	86	86	57	
Composite:	A/T > G/C	A/T	A	A/T	A/T	A/T > G/C	A/T > G/C	N	





strikingly similar but distinguishable from (+)- and (–)-CC-1065, respectively. Table 9 summarizes the composite consensus sequence derived from the evaluation of our data. Like (+)-CC-1065 and (+)-CPI-CDPI$_2$, each alkylation site detected for the natural enantiomer proved to be adenine and each site was flanked by two 5' A or T bases. The same preferences for the three base AT-rich alkylation site that were detected with the preceding agents were observed with (+)-CPI-CDPI$_3$: 5'-AA<u>A</u> = 5'-TT<u>A</u> > 5'-TA<u>A</u> > 5'-AT<u>A</u>. The distinction of (+)-CPI-CDPI$_3$ from (+)-CC-1065/(+)-CPI-CDPI$_2$ and (+)-CPI-CDPI$_1$ lies in its increasingly extended AT-rich alkylation site selectivity and the accompanying increased selectivity among the available alkylation sites which extends in the 5' direction from the adenine N3 alkylation site. A strong preference for the fourth, fifth, sixth and even seventh 5' base to be A or T was observed and extends two bases beyond that observed with (+)-CC-1065/(+)-CPI-CDPI$_2$. Notably, this extended AT-rich preference for the 5' bases manifests itself in a further restriction of alkylation among the (+)-CC-1065/(+)-CPI-CDPI$_2$ available sites and does not result in the observance of new sites unique to (+)-CPI-CDPI$_3$. Thus, the trend observed in the alkylation selectivity in moving from (+)-*N*-BOC-CPI to (+)-CPI-CDPI$_1$ and (+)-CPI-CDPI$_2$ where an increasingly extended AT-rich selectivity was observed in the 5' direction from the adenine N3 alkylation site (1 base, 2.5 bases, and 4 bases, respectively) continues with (+)-CPI–CDPI$_3$ (6 bases). This trend would seem to verify the prominent contribution of the AT-rich noncovalent binding selectivity to the agent DNA alkylation selectivity.

Similarly, (–)-CPI-CDPI$_3$ exhibited an alkylation selectivity comparable to that of (–)-CC-1065/(–)-CPI-CDPI$_2$ but increasingly more selective. Each alkylation site proved to be adenine and each alkylation site detected proved to be flanked by a 5' A or T bases and two 3' A or T bases. An additional and extended A/T preference for the fourth and fifth 3' bases to be A or T versus G or C was also observed and this 3' AT-rich preference extends beyond that observed with (+)-CC-1065. Thus, each alkylation site proved consistent with 5' adenine alkylation and agent binding in the minor groove in the 5'→3' direction from the site of alkylation across a six–seven base-pair AT-rich site.

Like (–)-CC-1065 and (–)-CPI-$CDPI_2$, this 5'→3' AT-rich alkylation selectivity relative to the adenine N3 alkylation site is both reversed and offset from that of the natural enantiomer (3'→5') and is the natural consequence of the diastereomeric relationship of the adducts. Thus, the (–)-CPI-$CDPI_3$ DNA alkylation proved analogous to that of *ent*-(–)-CC-1065 with the exception of the more extended 3' AT-rich selectivity.

Conclusions

A detailed evaluation of the DNA alkylation selectivity of (+)-CC-1065 and *ent*-(–)-CC-1065 and a series of aborted and extended analogs possessing the CPI alkylation subunit was detailed and the refinement of a model that accommodates the offset AT-rich adenine N3 alkylation selectivity of the enantiomeric agents was presented. The natural enantiomers bind in the minor groove in the 3'→5' direction starting from the adenine N3 alkylation site across a 2 base (*N*-BOC-CPI; i.e. 5'-A<u>A</u>), 3.5 base (CPI-$CDPI_1$/CPI-PDE-I_1; i.e. 5'-AA<u>A</u>), 5 base (CC-1065/CPI-$CDPI_2$; i.e. 5'-AAAA<u>A</u>), or 6.5 base (CPI-$CDPI_3$; i.e. 5'-AAAAA<u>A</u>) AT-rich site. In contrast, the unnatural enantiomers bind in the reverse 5'→3' direction in the minor groove and the binding site of the agents necessarily starts at the first 5' base preceding the adenine N3 alkylation site and extends across the alkylation site to the adjacent 3' bases covering an AT-rich alkylation site of 2 bases (*N*-BOC-CPI; e.g. 5'-A<u>A</u>), 5 bases (CC–1065/CPI-$CDPI_2$; e.g. 5'-A<u>A</u>AAA), or 6.5 bases (CPI-$CDPI_3$; e.g. 5'-A<u>A</u>AAAA). Notably, the model accommodates the unusual observation that both enantiomers of *N*-BOC-CPI alkylate the same sites within duplex DNA (5'-A<u>A</u> > 5'-T<u>A</u>) and the required reversed binding orientation of the enantiomeric agents. The reversed binding orientation is required to permit access to the electrophilic cyclopropane and the resulting offset AT-rich alkylation selectivity is the natural consequence of the diastereomeric relationship of the adducts. Notably, the DNA alkylation selectivities of (+)-*N*-BOC-CPI and *ent*-(–)-*N*-BOC-CPI are nearly indistinguishable from one another as well as being substantially less selective and readily distinguishable from (+)- and *ent*-(–)-CC-1065. Three dimensional models of the natural and unnatural enantiomer alkylations were presented which clearly illustrate the offset binding sites. A fundamentally simple and consistent model for the CC-1065 DNA alkylation reaction, that accommodates the behavior of both enantiomers, was provided in which the sequence selectivity is derived from the noncovalent binding selectivity of the agents preferentially in the narrower, sterically more accessible AT-rich minor groove, the inherent steric accessibility to the adenine N3 alkylation site that accompanies deep penetration of the agent into the minor groove within an AT-rich site, and the 2 base-pair (*N*-BOC-CPI), 3.5 base-pair (CPI-PDE-I_1/CPI-$CDPI_1$), 5 base-pair (CC-1065/CPI-$CDPI_2$), or 6.5 base-pair (CPI-$CDPI_3$) site size required to permit agent binding in the minor groove at the alkylation site. Furthermore, the results of the study are inconsistent with interpretations in which the DNA alkylation selectivities of the natural enantiomers may be attributed to the inherent selectivity of the alkylation event independent of noncovalent binding selectivity which have been proposed to be derived from a sequence dependent phosphate protonation of the alkylation subunit carbonyl[9,26,30] or alkylation at A/T junctions of bent DNA.[27–29] Using the models, a simple explanation for the distinguishing DNA alkylation rate and efficiency as well as biological potency of enantiomeric pairs of agents was detailed based on the unnatural enantiomer sensitivity to steric bulk at the CPI C7 center.

Acknowledgments

We gratefully acknowledge the financial support of the National Institutes of Health (CA41986) and the award of an ACS Organic Division fellowship sponsored by Zeneca Pharmaceuticals to D. S. Johnson.

References

1. Chidester, C. G.; Krueger, W. C.; Mizsak, S. A.; Duchamp, D. J.; Martin, D. G. *J. Am. Chem. Soc.* **1981**, *103*, 7625.

2. Duocarmycin A: Takahashi, I.; Takahashi, K.; Ichimura, M.; Morimoto, M.; Asano, K.; Kawamoto, I.; Tomita, F.; Nakano, H. *J. Antibiot.* **1988**, *41*, 1915.

3. Yasuzawa, T.; Iida, T.; Muroi, K.; Ichimura, M.; Takahashi, K.; Sano, H. *Chem. Pharm. Bull.* **1988**, *36*, 3728.

4. DC89-A1 (duocarmycin C_1) and DC88-A (duocarmycin A): Nakano, H.; Takahashi, I.; Ichimura, M.; Kawamoto, I.; Asano, K.; Tomita, F.; Sano, H.; Yasuzawa, T.; Morimoto, M.; Fujimoto, K. PCT. Int. Appl. W087 06265; *Chem. Abstr.* **1988**, *108*, 110858s.

5. DC89-Al: Ichimura, M.; Muroi, K.; Asano, K.; Kawamoto, I.; Tomita, F.; Morimoto, M.; Nakano, H. *J. Antibiot.* **1988**, *41*, 1285.

6. Duocarmycin B_1 and B_2: Ogawa, T.; Ichimura, M., Katsumata, S.; Morimoto, M.; Takahashi, K. *J. Antibiot.* **1989**, *42*, 1299.

7. Pyrindamycin A and B: Ohba, K.; Watabe, H.; Sasaki, T.; Takeuchi, Y.; Kodama, Y.; Nakazawa, T.; Yamamoto, H.; Shomura, T.; Sezaki, M.; Kondo, S. *J. Antibiot.* **1988**, *41*, 1515; Ishii, S.; Nagasawa, M.; Kariya, Y.; Yamamoto, H.; Inouye, S.; Kondo, S. *J. Antibiot.* **1989**, *42*, 1713.

8. Duocarmycin SA: Ichimura, M.; Ogawa, T.; Takahashi, K.; Kobayashi, E.; Kawamoto, I.; Yasuzawa, T.; Takahashi, I.; Nakano, H. *J. Antibiot.* **1990**, *43*, 1037; Ichimura, M.; Ogawa, T.; Katsumata, S.; Takahashi, K.; Takahashi, I.; Nakano, H. *J. Antibiot.* **1991**, *44*, 1045.

9. Warpehoski, M. A.; Hurley, L. H. *Chem. Res. Toxicol.* **1988**, *1*, 315; Warpehoski, M. A. In *Advances in DNA Sequence Specific Agents*, Vol. 1, p. 217, Hurley, L. H., Ed.; JAI Press Inc.; Greenwich, CT, 1992.

10. Hurley, L. H.; Needharn-VanDevanter, D. R. *Acc. Chem. Res.* **1986**, *19*, 230; Hurley, L. H.; Draves, P. H. In *Molecular Aspects of Anticancer Drug–DNA Interactions*, Vol. 1, p. 89, Neidle, S.; Waring, M., Eds.; CRC Press; Ann Arbor, 1993.

11. Coleman, R. S.; Boger, D. L. In *Studies in Natural Product Chemistry*, Vol. 3, p. 301, Rahman, A.-u., Ed.; Elsevier; Amsterdam, 1989.

12. Boger, D. L. *Chemtracts: Org. Chem.* **1991**, *4*, 329.

13. Boger, D. L. In *Heterocycles in Bioorganic Chemistry*, p. 103, Bergman, J.; van der Plas, H. C.; Simonyl, M., Eds.; Royal Society of Chemistry; Cambridge, 1991.

14. Boger, D. L. In *Proceed. R. A. Welch Found. Conf. on Chem. Res., XXXV. Chem. at the Frontiers of Medicine* **1991**, *35*, 137.

15. Boger, D. L. In *Advances in Heterocyclic Natural Products Synthesis*, Vol. 2, pp. 1–188, Pearson, W. H.; Ed.; JAI Press; Greenwich, CT, 1992.

16. Boger, D. L. *Pure Appl. Chem.* **1993**, *65*, 1123.

17. Boger, D. L. *Pure Appl. Chem.* in press.

18. Hurley, L. H.; Reynolds, V. L.; Swenson, D. H.; Petzold, G. L.; Scahill, T. A. *Science* **1984**, *226*, 843.

19. Reynolds, V. L.; Molineux, I. J.; Kaplan, D. J.; Swenson, D. H.; Hurley, L. H. *Biochemistry* **1985**, *24*, 6228.

20. Wierenga, W.; Bhuyan, B. K.; Kelly, R. C.; Krueger, W. C.; Li, L. H.; McGovren, J. P.; Swenson, D. H.; Warpehoski, M. A. *Adv. Enzyme Regul.* **1986**, *25*, 141; 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.

21. Needham-Van Devanter, D. R.; Hurley, L. H. *Biochemistry* **1986**, *25*, 8430.

22. Hurley, L. H.; Needham-Van Devanter, D. R.; Lee, C.-S. *Proc. Natl Acad. Sci. U.S.A.* **1987**, *84*, 6412.

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. Scahill, T. A.; Jensen, R. M.; Swenson, D. H.; Hatzenbuhler, N. T.; Petzold, G.; Wierenga, W.; Brahme, N. D. *Biochemistry* **1990**, *29*, 2852.

25. Lin, C. H.; Hurley, L. H. *Biochemistry* **1990**, *29*, 9503.

26. 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.

27. Lin, C. H.; Sun, D.; Hurley, L. H. *Chem. Res. Toxicol.* **1991**, *4*, 21.

28. Lin, C. H.; Hill, G. C.; Hurley, L. H. *Chem. Res. Toxicol.* **1992**, *5*, 167.

29. Sun, D.; Lin, C. H.; Hurley, L. H. *Biochemistry* **1993**, *32*, 4487.

30. Lin, C. H.; Beale, J. M.; Hurley, L. H. *Biochemistry* **1991**, *30*, 3597.

31. Warpehoski, M. A.; Harper, D. E.; Mitchell, M. A.; Monroe, T. J. *Biochemistry* **1992**, *31*, 2502.

32. Powers, R.; Gorenstein, D. G. *Biochemistry* **1990**, *29*, 9994.

33. 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; Boger, D. L.; Coleman, R. S. *J. Am. Chem. Soc.* **1988**, *110*, 4796 and 1321.

34. Boger, D. L.; Munk, S. A.; Zarrinmayeh, H.; Ishizaki, T.; Haught, J.; Bina, M. *Tetrahedron* **1991**, *47*, 2661.

35. Boger, D. L.; Munk, S. A.; Ishizaki, T. *J. Am. Chem. Soc.* **1991**, *113*, 2779.

36. Boger, D. L.; Munk, S. A.; Zarrinmayeh, H. *J. Am. Chem. Soc.* **1991**, *113*, 3980.

37. CI-based analogs: Boger, D. L.; Zarrinmayeh, H.; Munk, S. A.; Kitos, P. A.; Suntornwat, O. *Proc. Natl Acad. Sci. U.S.A.* **1991**, *88*, 1431. See also ref. 34; Synthesis: Boger, D. L.; Wysocki, Jr R. J. *J. Org. Chem.* **1989**, *54*, 1238; Boger, D. L.; Wysocki, Jr R. J.; Ishizaki, T. *J. Am. Chem. Soc.* **1990**, *112*, 5230; Drost, K. J.; Jones, R. J.; Cava, M. P. *J. Org. Chem.* **1989**, *54*, 5985; Tidwell, J. H.; Buchwald, S. L. *J. Org. Chem.* **1992**, *57*, 6380; Lown, J. W.; Wang, Y. *Heterocycles* **1993**, *36*, 1399; Sundberg, R. J.; Baxter, E. W. *Tetrahedron Lett.* **1986**, *27*, 2687.

38. CBI-based analogs: Boger, D. L.; Munk, S. A. *J. Am. Chem. Soc.* **1992**, *114*, 5487; Synthesis: Boger, D. L.; Ishizaki, T.; Wysocki, Jr R. J.; Munk, S. A.; Kitos, P. A.; Suntornwat, O. *J. Am. Chem. Soc.* **1989**, *111*, 6461; Boger D. L.; Ishizaki, T.; Kitos, P. A.; Suntornwat, O. *J. Org. Chem.* **1990**, *55*, 5823; Boger, D. L.; Ishizaki, T. *Tetrahedron Lett.* **1990**, *31*, 793; Boger, D. L.; Ishizaki, T.; Zarrinmayeh, H.; Kitos, P. A.; Suntornwat, O. *BioMed. Chem. Lett.* **1991**, *1*, 55; 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; Drost, K. J.; Cava, M. P. *J. Org. Chem.* **1991**, *56*, 2240; Boger, D. L.; Yun, W.; Teegarden, B. R. *J. Org Chem.* **1992**, *57*, 2873; Aristoff, P. A.; Johnson, P. D. *J. Org Chem.* **1992**, *57*, 6234; Aristoff, P. A.; Johnson, P. D.; Sun, D. *J. Med. Chem.* **1993**, *36*, 1956.

39. C_2BI-based analogs: Boger, D. L.; Palanki, M. S. S. *J. Am. Chem. Soc.* **1992**, *114*, 9318; Boger, D. L.; Johnson, D. S.; Palanki, M. S. S.; Kitos, P. A.; Chang, J.; Dowell, P. *BioMed. Chem.* **1993**, *1*, 27.

40. Bizelesin: Lee, C.-S.; Gibson, N. W. *Biochemistry* **1993**, *32*, 9108; Sun, D.; Hurley, L. H. *J. Am. Chem. Soc.* **1993**, *115*, 5925.

41. Duocarmycin A, C_1–C_2: Boger, D. L.; Ishizaki, T.; Zarrinmayeh, H.; Kitos, P. A.; Suntornwat, O. *J. Org. Chem.* **1990**, *55*, 4499; Boger, D. L.; Ishizaki, T.; Zarrinmayeh, H.; Munk, S. A.; Kitos, P. A.; Suntornwat, O. *J. Am. Chem. Soc.* **1990**, *112*, 8961; Boger, D. L.; Ishizaki, T.; Zarrinmayeh, H. *J. Am. Chem. Soc.* **1991**, *113*, 6645; Boger, D. L.; Yun, W. *J. Am. Chem. Soc.* **1993**, *115*, 9872.

42. Duocarmycin SA and *ent*-(–)-duocarmycin SA: Boger, D. L.; Johnson, D. S.; Yun, W. *J. Am. Chem. Soc.* in press; Synthesis: Boger, D. L.; Machiya, K.; Hertzog, D. L.; Kitos, P. A.; Holmes, D. *J. Am. Chem. Soc.* **1993**, *115*, 9025; Boger, D. L.; Machiya, K. *J. Am. Chem. Soc.* **1992**, *114*, 10056.

43. *epi*- and *ent*-Duocarmycin A: Boger, D. L.; Yun, W.; Terashima, S.; Fukuda, Y.; Nakatani, K.; Kitos, P. A.; Jin, Q. *BioMed. Chem. Lett.* **1992**, *2*, 759.

44. Duocarmycin A: Sugiyama, H.; Hosoda, M.; Saito, I.; Asai, A.; Saito, H. *Tetrahedron Lett.* **1990**, *31*, 7197; Lin, C. H.; Patel, D. J. *J. Am. Chem. Soc.* **1992**, *114*, 10658; Sugiyama, H.; Ohmori, K.; Chan, K. L.; Hosoda, M.; Asai, A.; Saito, H.; Saito, I. *Tetrahedron Lett.* **1993**, *34*, 2179; Yamamoto, K.; Sugiyama, H.; Kawanishi, S. *Biochemistry* **1993**, *32*, 1059.

45. Boger, D. L.; Invergo, B. J.; Coleman, R. S.; Zarrinmayeh, H.; Kitos, P. A.; Thomspon, S. C.; Leong, T.; McLaughlin, L. W. *Chem. Biol. Interact.* **1990**, *73*, 29; Boger, D. L.; Coleman, R. S.; Invergo, B. J. *J. Org. Chem.* **1987**, *52*, 1521; Boger, D. L.; Coleman, R. S. *J. Org. Chem.* **1984**, *49*, 2240.

46. Boger, D. L.; Sakya, S. M. *J. Org. Chem.* **1992**, *57*, 1277.

(*Received* 28 *September* 1993; *accepted* 2 *December* 1993)