Reversed and Sandwiched Analogs of Duocarmycin SA: Establishment of the Origin of the Sequence-Selective Alkylation of DNA and New Insights into the Source of Catalysis

Dale L. Boger,* Bernd Bollinger, Donald L. Hertzog, Douglas S. Johnson, Hui Cai, Philippe Mésini, Robert M. Garbaccio, Qing Jin, and Paul A. Kitos

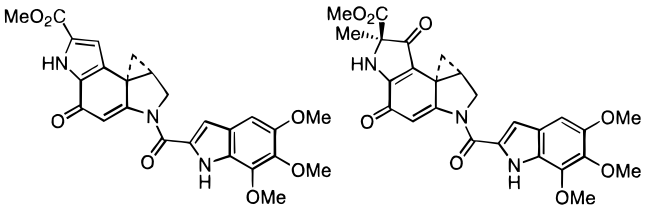
Contribution from the Department of Chemistry, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037

Received October 25, 1996[⊗]

Abstract: The synthesis and examination of two unique classes of duocarmycin SA analogs are described which we refer to as reversed and sandwiched analogs. Their examination established both the origin of the DNA alkylation selectivity and that both enantiomers of this class of natural products are subject to the same polynucleotide recognition features. The most beautiful demonstration of this is the complete switch in the enantiomeric alkylation selectivity of the reversed analogs which is only consistent with the noncovalent binding model and incompatible with alkylation site models of the origin of the DNA alkylation selectivity. In addition, dramatic alterations in the rates of DNA alkylation were observed among the agents and correlate with the presence or absence of an extended, rigid N^2 amide substituent. This has led to the proposal of a previously unrecognized source of catalysis for the DNA alkylation reaction which was introduced in the preceding paper of this issue (*J. Am. Chem. Soc.* **1997**, *119*, 4977−4986).

In the preceding paper[1] we described the evaluation of a series of conventional analogs of the duocarmycins (**1** and **2**, Figure 1) and CC-1065 (**3**).[2−4] Herein, we report two unconventional classes of agents that we refer to as reversed[5] and sandwiched analogs of duocarmycin SA. In addition to their importance as structurally novel analogs of duocarmycin SA,[6] their examination alongside the agents in the accompanying paper[1] has established the origin of the sequence selective DNA alkylation and has provided a unique insight into its catalysis.

Three proposals have been advanced to account for the DNA alkylation sequence selectivity and these have been discussed in detail (Figure 2).[2−4] Two are based on the premise that **3** and **6** exhibit identical alkylation selectivities.[4,7] One proposes a sequence-dependent phosphate protonation of the C4 carbonyl which activates the agent for DNA alkylation[4,7−9] and the other invokes alkylation at junctions of bent DNA equivalently expressed for both **3** and **6** without addressing the source of catalysis.[10] A third proposal, which is based on the premise that **3** and **6** exhibit distinct alkylation selectivities[2,11] that are controlled by the AT-rich noncovalent binding selectivity of the agents and their steric accessibility to the adenine N3



1, (+)-duocarmycin SA **2**, (+)-duocarmycin A

Figure 1.

[⊗] Abstract published in *Advance ACS Abstracts,* May 1, 1997.

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

(2) Boger, D. L.; Johnson, D. S. *Angew Chem., Int. Ed. Engl.* **1996**, *35*, 1439. Boger, D. L.; Johnson, D. S. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 3642. Boger, D. L. *Acc. Chem. Res.* **1995**, *28*, 20. Boger, D. L. In *Advances in Heterocyclic Natural Product Synthesis*; Pearson, W. H., Ed.; JAI: Greenwich, CT, 1992; Vol. 2, p 1. Boger, D. L. *Chemtracts: Org. Chem.* **1991**, *4*, 329. Boger, D. L. In *Proc. R. A. Welch Found. Conf. Chem. Res., XXXV., Chem. Frontiers Med.* **1991**, *35*, 137. Boger, D. L. In *Heterocycles in Bioorganic Chemistry*; Bergman, J., van der Plas, H. C., Simonyl, M., Eds.; Royal Society of Chemistry: Cambridge, 1991; p 103. Coleman, R. S.; Boger, D. L. In *Studies in Natural Product Chemistry*; Rahman, A. u., Ed.; Elsevier: Amsterdam, The Netherlands, 1989; Vol. 3, p 301.

(3) Hurley, L. H.; Needham-VanDevanter, D. R. *Acc. Chem. Res.* **1986**, *19*, 230. Hurley, L. H.; Draves, P. H. In *Molecular Aspects of Anticancer Drug-DNA Interactions*; Neidle, S., Waring, M., Eds.; CRC: Ann Arbor, MI, 1993; Vol. 1, p 89. Warpehoski, M. A. In *Advances in DNA Sequence Specific Agents*; Hurley, L. H., Ed.; JAI: Greenwich, CT, 1992; Vol. 1, p 217. Aristoff, P. A. In *Advances in Medicinal Chemistry*; Maryanoff, B. E., Maryanoff, C. E., Eds.; JAI: Greenwich, CT, 1993; Vol. 2, p 67. Warpehoski, M. A.; McGovren, J. P.; Mitchell, M. A.; Hurley, L. H. In *Molecular Basis of Specificity in Nucleic Acid-Drug Interactions*; Pullman, B., Jortner, J., Eds.; Kluwer: Dordrecht, The Netherlands, 1990; p 531. Reynolds, V. L.; McGovren, J. P.; Hurley, L. H. *J. Antibiot.* **1986**, *39*, 319.

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

(5) Boger, D. L.; Johnson, D. S. *J. Am. Chem. Soc.* **1995**, *117*, 1443.

(6) Kobayashi, E.; Okamoto, A.; Asada, M.; Okabe, M.; Nagamura, S.; Asai, A.; Saito, H.; Gomi, K.; Hirata, T. *Cancer Res.* **1994**, *54*, 2404. Ichimura, M.; Tatsuhiro, O.; Keiichi, T.; Akira, M.; Isami, T.; Hirofumi, N. *Oncol. Res.* **1993**, *5*, 165. Gomi, K.; Kobayashi, E.; Miyoshi, K.; Ashizawa, T.; Okamoto, A.; Ogawa, T.; Katsumata, S.; Mihara, A.; Okabe, M.; Hirata, T. *Jpn. J. Cancer Res.* **1992**, *83*, 113.

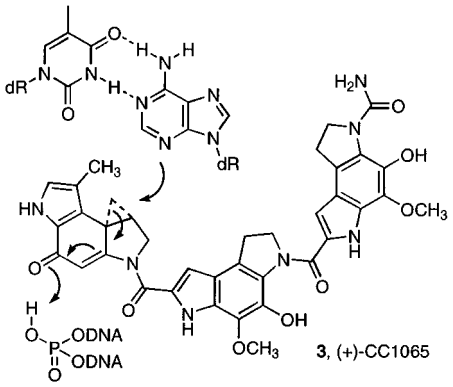
(7) Hurley, L. H.; Reynolds, V. L.; Swenson, D. H.; Petzold, G. L.; Scahill, T. A. *Science* **1984**, *226*, 843. Reynolds, V. L.; Molineux, I. J.; Kaplan, D. J.; Swedson, D. H.; Hurley, L. H. *Biochemistry* **1985**, *24*, 6228. Hurley, L. H.; Lee, C.-S.; McGovren, J. P.; Warpehoski, M. A.; Mitchell, M. A.; Kelly, R. C.; Aristoff, P. A. *Biochemistry* **1988**, *27*, 3886. Scahill, T. A.; Jensen, R. M.; Swenson, D. H.; Hatzenbuhler, N. T.; Petzold, G.; Wierenga, W.; Brahme, N. D. *Biochemistry* **1990**, *29*, 2852.

(8) 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.

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

(10) 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. Lin, C. H.; Hill, G. C.; Hurley, L. H. *Chem. Res. Toxicol.* **1992**, *5*, 167. Ding, Z.-M.; Harshey, R. M.; Hurley, L. H. *Nucleic Acids Res.* **1993**, *21*, 4281. Sun, D.; Lin, C. H.; Hurley, L. H. *Biochemistry* **1993**, *32*, 4487. Thompson, A. S.; Sun, D.; Hurley, L. H. *J. Am. Chem. Soc.* **1995**, *117*, 2371.

S0002-7863(96)03721-3 CCC: $14.00



Alkylation Site Model

- (+)-CPI and (+)-CC-1065 selectivities identical selectivity inherent in CPI subunit and nature of alkylation reaction
- Alkylation selectivity independent of noncovalent binding selectivity
- Different features control *ent*-(-)-CC-1065
- Alkylation at junctions of bent DNA or
- Sequence dependent phosphate activation C4 carbonyl required

Noncovalent Binding Model

- (+)-CPI and (+)-CC-1065 selectivities different modest electrophile superimposed on 5 base-pair AT-rich DNA binding agent
- Alkylation selectivity is controlled by noncovalent binding selectivity
- Same features control *ent*-(-)-CC-1065
- Alkylation within AT-rich binding sequences
- Phosphate protonation not involved C4 carbonyl not required

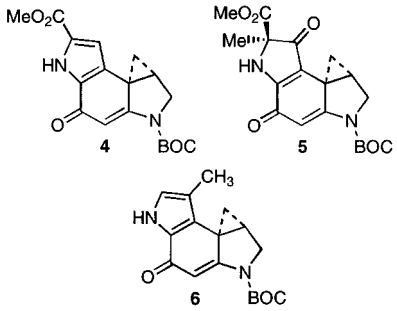


Figure 2.

alkylation site, is fully consistent with a growing set of experimental observations.[2,11−16] It accommodates and explains the reverse and offset 5 or 3.5 base-pair AT-rich adenine N3 alkylation selectivities of the natural and unnatural enantiomers of **1**,[11] or **2** and **3**[12,13] and requires that **4**−**6** exhibit alkylation selectivities distinct from the natural products. It offers a beautiful explanation for the identical alkylation selectivities of both enantiomers of simple derivatives including **4**−**6** (5′-A<u>A</u> > 5′-T<u>A</u>), and the more extended AT-rich selectivity of the advanced analogs of **1**−**3** corresponds nicely to the length of the agent and the size of the required binding region surrounding the alkylation site. Further support of this model includes the demonstrated AT-rich noncovalent binding of the agents,[14] their preferential noncovalent binding coincidental with DNA alkylation,[15] the demonstration that the characteristic DNA alkylation is also observed with isomeric alkylation subunits, and that it does not require the presence of the C4 carbonyl or even the activated cyclopropane.[16] The comparisons of the DNA alkylation selectivities of the reversed and sandwiched analogs of duocarmycin SA detailed herein further establish the accuracy of this model.

In previous studies, the issue of catalysis with the noncovalent binding model has not been addressed. The chemical stability of **1**−**3** and the acid catalysis requirement for addition of typical nucleophiles has led to the assumption that the DNA alkylation must also be an acid-catalyzed reaction. Although efforts have gone into supporting the extent and role of this acid catalysis,[4−10,17] it remains largely undocumented for the DNA alkylation reaction.[1] At pH 7.4, the DNA phosphate backbone is fully ionized (0.0001−0.00004% protonated). Consequently, it is unlikely that catalysis is derived from a phosphate backbone delivery of a proton to the C4 carbonyl as advanced in the alkylation site model. Although increases in the local hydronium ion concentrations surrounding "acidic domains" of DNA have been invoked to explain DNA mediated acid catalysis,[18] nucleotide reactivity,[18] and extrapolated in studies with **3** to alkylation site catalysis,[17] the remarkable stability of **1**−**3** even at pH 5 suggests that it is unlikely to be the source of catalysis. Consistent with this, the rate of the DNA alkylation reaction for **1** exhibits only a very modest pH dependence below pH 7 and essentially no dependence in the more relevant pH 7−8 range.[1]

Herein, we disclose studies which are inconsistent with this simple role of acid catalysis. Our observations suggest this rate acceleration is derived from a DNA binding-induced conformational change in the agent which substantially alters its inherent reactivity.[1,19] This ground state destabilization of the substrate, which we suggest results from a binding-induced twist in the linking N^2 amide, disrupts the vinylogous amide stabilization of the alkylation subunit and increases its reactivity toward nucleophiles. We further suggest that this activation is not uniquely alkylation site dependent, but rather a general consequence of the forced adoption of a helical conformation upon AT-rich minor groove binding. As such, the DNA minor groove binding cocks the pistol but does not pull the trigger for further reaction. The consequences of this proposal are diametrically opposite to those of the alkylation site model where a sequence-dependent DNA reactivity[10] is proposed to be responsible for the purportedly identical alkylation selectivities of **3** and **6**.[7,8] Rather, any sequence-dependent activation derived from this binding-induced conformational change that might contribute to catalysis of the DNA alkylation reaction would lead to distinctions, not similarities, in the DNA alkylation profiles of **3** and **6**.

(11) Boger, D. L.; Johnson, D. S.; Yun, W.; Tarby, C. M. *Bioorg. Med. Chem.* **1994**, *2*, 115. Boger, D. L.; Munk, S. A.; Zarrinmayeh, H.; Ishizaki, T.; Haught, J.; Bina, M. *Tetrahedron* **1991**, *47*, 2661. 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.; 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.

(13) Boger, D. L.; Johnson, D. S.; Yun, W. *J. Am. Chem. Soc.* **1994**, *116*, 1635.

(14) Boger, D. L.; Coleman, R. S.; Invergo, B. J.; 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.

(15) Boger, D. L.; Zhou, J.; Cai, H. *Bioorg. Med. Chem.* **1996**, *4*, 859.

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

(17) Warpehoski, M. A.; Harper, D. E. *J. Am. Chem. Soc.* **1994**, *116*, 7573. Warpehoski, M. A.; Harper, D. E. *J. Am. Chem. Soc.* **1995**, *117*, 2951.

(18) Jayaram, B.; Sharp, K. A.; Honig, B. *Biopolymers* **1989**, *28*, 975. Lamm, G.; Pack, G. R. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 9033. Lamm, G.; Wong, L.; Pack, G. R. *J. Am. Chem. Soc.* **1996**, *118*, 3325.

(19) Boger, D. L.; Garbaccio, R. M. *Bioorg. Med. Chem.* **1997**, *5*, 263.

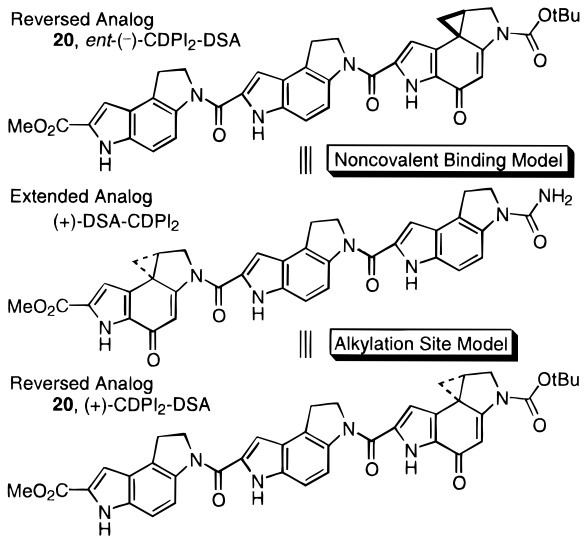
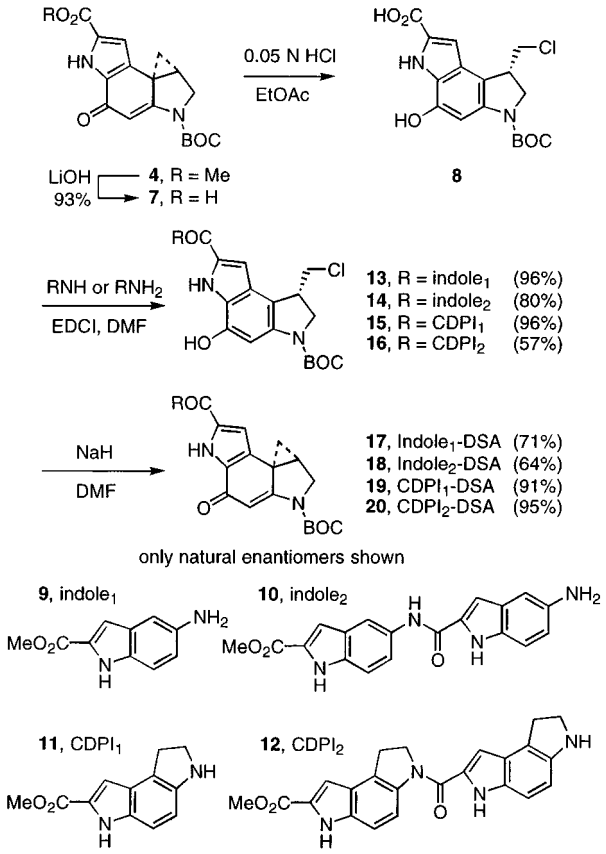


Figure 3.

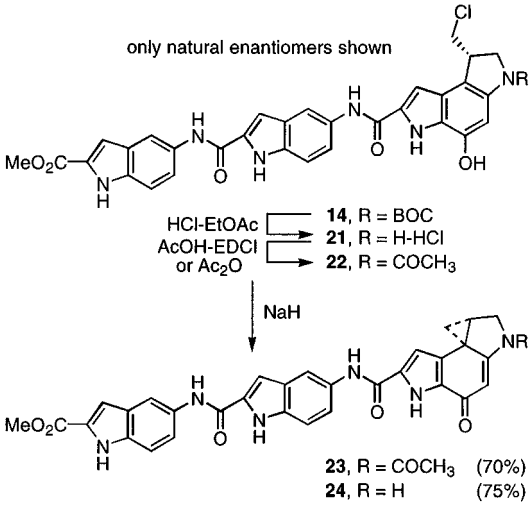
Rationale for the Evaluation of the Reversed Analogs: Test of the Origin of the DNA Alkylation Selectivity. Duocarmycin SA contains a C6 methyl ester that complements the right-hand side linking amide. This provides the opportunity to introduce DNA-binding subunits on either side of the alkylation subunit. Thus, coupling of a DNA-binding subunit through the C6 carboxylate provides a novel class of agents we refer to as reversed analogs. From the noncovalent binding model, the reversed agents were projected to exhibit an AT-rich alkylation selectivity that extends in the atypical reverse direction from an alkylation site. The predicted alkylation sites for the natural enantiomers coincide with those of the unnatural enantiomers of the typical extended agents. Similarly, the predicted alkylation sites for the unnatural enantiomers of the reversed agents coincide with those of the natural enantiomers of the extended agents. Thus, a complete switch in the enantiomeric alkylation selectivity would be observed with the reversed analogs if it is controlled by the AT-rich binding selectivity.[5] In contrast, the alkylation site model would require that natural enantiomers of both the extended and reversed agents alkylate the same sites rather than exhibit this switch. Thus, the agents provide a rigorous test of the two models with a definitive resolution (Figure 3).

Synthesis of the Reversed Analogs of Duocarmycin SA. The preparation of **17**−**20** was accomplished through coupling of the free amine of **9**−**12** (0.9 equiv, 2−4 h, 25 °C) with the C6 carboxylic acid of **8**. While this was conducted as outlined in Scheme 1, the manner in which this could be accomplished was not straightforward. Hydrolysis (1−1.3 equiv of LiOH) of *N*-BOC-DSA (**4**) cleanly provided **7** (93%) without competitive hydrolysis of the carbamate or addition to the cyclopropane. Prolonged reaction times (72 h) at 25 °C under typical conditions (LiOH, THF/CH_3OH/H_2O) provided predominately recovered starting material, and the conversion to **7** was observed only upon warming (60 °C). Even under these conditions, only methyl ester hydrolysis was observed. Both the direct coupling of **7** (1,3-dicyclohexylcarbodiimide (DCC), 1-(3-(dimethylamino)propyl)-3-ethylcarbodiimide hydrochloride (EDCI), diethyl pyrocarbonate (DEPC)) in the presence or absence of added $NaHCO_3$ or the use of preformed active esters (*e.g.*, imidazolide) provided only low yields (10−30%) of the desired agent with the more soluble coupling partners and failed altogether with the insoluble $CDPI_2$. Consequently, the preparation of **17**−**20** was more effectively accomplished in an indirect manner. Treatment of **7** with dilute HCl (0.05 N HCl/EtOAc, 25 °C, 30 min) provided **8** without *N*-BOC deprotection, and no trace of the ring expansion HCl addition product was detected. Subsequent coupling of **8** with the **9**−**12** (0.9 equiv, 2 equiv of EDCI, DMF, 2−4 h, 25 °C) proceeded in high yields (57−96%), and spirocyclization was effected by NaH (DMF, 0 °C, 30 min, 64−95%).

Scheme 1



Scheme 2



Modifications in the Terminal N^2 Acyl Substituent of the Reversed Analogs. To insure that the behavior of **17**−**20** was not substantially influenced by the nature of the N^2 substituent, both enantiomers of **23** and **24** terminating with a N^2 acetyl group or the free amine were prepared (Scheme 2). Thus, *N*-BOC deprotection of **14** (3.4 N HCl/EtOAc, 25 °C, 30 min) followed by spirocyclization or N^2 acetylation and spirocyclization afforded **24** and **23**, respectively.

Synthesis of the Sandwiched Analogs of Duocarmycin SA. An important complement to the extended and reversed analogs

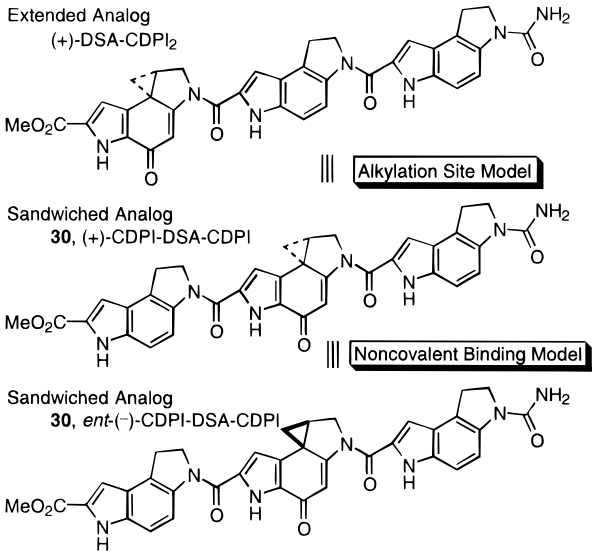


Figure 4.

are the agents **30** (CDPI-DSA-CDPI) and **31** (CDPI-DSA-TMI) which we refer to as sandwiched analogs. Their examination was more informative than their initial consideration might suggest. The noncovalent binding model led us to predict that both enantiomers of **30** and **31** would alkylate the same sites independent of their absolute configuration and that their sites of DNA alkylation would be distinct from either enantiomer of both the extended and reversed analogs. Such a demonstration would further distinguish the noncovalent binding model from the alkylation site model which would require that the natural enantiomers of the sandwiched analogs alkylate the same sites as (+)-**1**−**6** (Figure 4).

N-BOC deprotection of **15** (4 N HCl/EtOAc, 25 °C, 30 min) followed by coupling of $CDPI_1$ (**26**, 68%) or **27** (0.95 equiv, 2 equiv of EDCI, DMF, 25 °C, 6−15 h, 70%) and spirocyclization (NaH, DMF, 0 °C, 30 min) cleanly provided **30** (70%) and **31** (67%), Scheme 3.

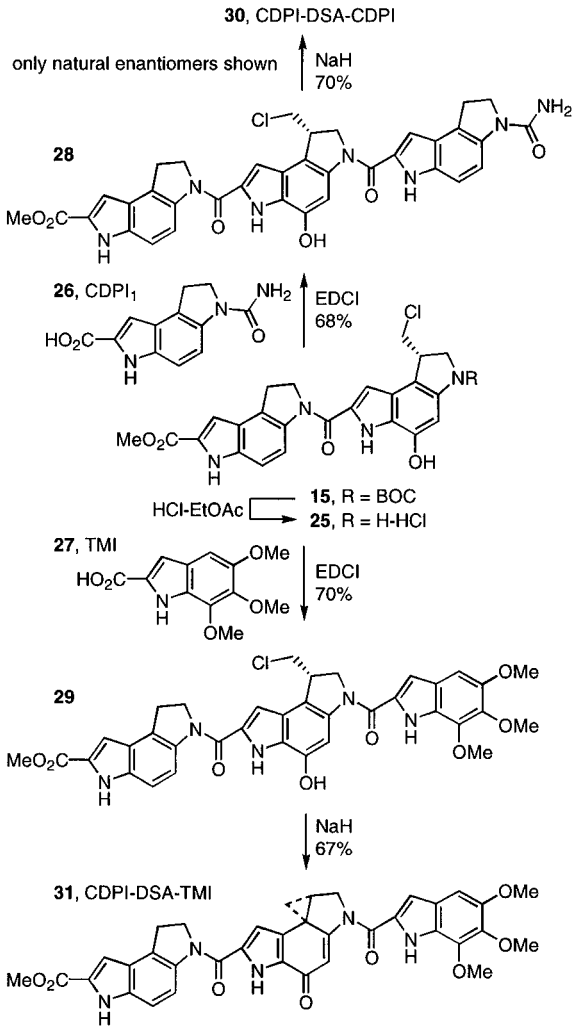
DNA Alkylation Studies. The DNA alkylation properties of the agents were examined within five segments of duplex DNA as described in the accompanying paper.[1]

DNA Alkylation Properties of Reversed versus Extended Analogs. With both enantiomers of the reversed analogs $CDPI_2$-DSA (**20**) in hand, their comparison with the enantiomers of DSA-$CDPI_2$ was addressed. The key question was whether the enantiomeric alkylation selectivity of **20** would switch with the simple reversal of the orientation of the DNA-binding subunits (noncovalent binding model) or whether the two natural enantiomers would behave in an identical or comparable manner (alkylation site model). The comparisons proved unambiguous and are illustrated in Figures 5 and 6.

The natural enantiomer of the reversed agent, (+)-$CDPI_2$-DSA (**20**), was found to alkylate the same sites and to exhibit the same sequence selectivity as the unnatural enantiomer of CC-1065 and the extended agent *ent*-(−)-DSA-$CDPI_2$. Thus, a complete reversal of the enantiomeric alkylation selectivity was observed with (+)-**20**, and these results are only consistent with the noncovalent binding model. This is illustrated nicely in Figure 5 where the natural enantiomer of the reversed analog (+)-$CDPI_2$-DSA (**20**) alkylated the same, single high-affinity site in w794 (5′-AATTT) as the unnatural enantiomers (−)-CC-1065 and (−)-DSA-$CDPI_2$ without detectable alkylation at the single, high-affinity site observed with the natural enantiomers (+)-CC-1065 and (+)-DSA-$CDPI_2$ (5′-AATTA).

This is also highlighted beautifully in Figure 6 with w836 DNA where several additional elements of the DNA alkylation

Scheme 3



reaction are illustrated. Within the 6-base A-rich sequence illustrated, (+)-duocarmycin SA alkylated each of the four 3′-adenines (5′-AAAAAA) corresponding nicely to 3′ → 5′ binding across a 3.5 base-pair AT-rich site. The natural enantiomer of the extended analog (+)-DSA-$CDPI_2$ also alkylates only the 3′-adenines (5′-AAAAAA) in this sequence corresponding to the same 3′ → 5′ binding but across a more extended 5 base-pair AT-rich sequence restricting alkylation to the first three versus four 3′-adenines. In this same sequence, the unnatural enantiomer of the extended agent, *ent*-(−)-DSA-$CDPI_2$, alkylates the 5′ terminal adenines in accordance with 5′ → 3′ binding across a 5 base-pair AT-rich binding site (5′-AAAAAA). Consistent with the offset AT-rich binding site of the unnatural enantiomers due to the diastereomeric nature of the adducts, *ent*-(−)-DSA-$CDPI_2$ does not alkylate the terminal 5′-adenine, but does alkylate the following four 5′-adenines. Indicative of the complete switch in the enantiomeric alkylation selectivity of the reversed agents, the natural enantiomer of **20** alkylated the same sites as *ent*-(−)-DSA-$CDPI_2$ while the unnatural enantiomer of **20** alkylated the same sites as the natural enantiomer (+)-DSA-$CDPI_2$. In addition to establishing that the noncovalent binding selectivity of the agents is controlling the DNA alkylation selectivity, the comparisons also establish that the two enantiomers are subject to the same polynucleotide recognition features.[11−13] In related studies, it is maintained that the two enantiomers are subject to different polynucleotide recognition elements.[3,4]

The consensus alkylation sequence of (+)-$CDPI_2$-DSA (**20**) is summarized in Table 1. A table of the statistical treatment

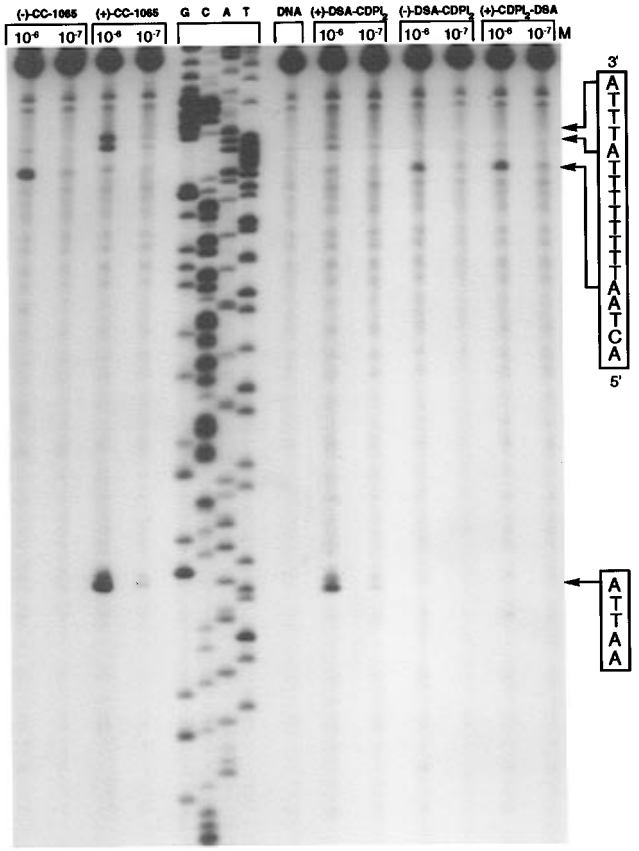


Figure 5. Thermally-induced strand cleavage of w794 DNA (144 bp, nucleotide no. 5238-138) after DNA−agent incubation with DSA-$CDPI_2$ (25 °C, 24 h) and $CDPI_2$-DSA (37 °C, 5 days), removal of unbound agent by EtOH precipitation and 30 min thermolysis (100 °C), followed by denaturing 8% PAGE and autoradiography. Lanes 1 and 2, *ent*-(−)-CC-1065 (25 °C, 1×10^{-6} to 1×10^{-7} M); lanes 3 and 4, (+)-CC-1065 (25 °C, 1×10^{-6} to 1×10^{-7} M); lanes 5−8, Sanger G, C, A, and T sequencing standards; lane 9, control DNA; lanes 10 and 11, (+)-DSA-$CDPI_2$ (25 °C, 1×10^{-6} to 1×10^{-7} M); lanes 12 and 13, *ent*-(−)-DSA-$CDPI_2$ (37 °C, 1×10^{-6} to 1×10^{-7} M); lanes 14 and 15, (+)-$CDPI_2$-DSA ((+)-**20**, 37 °C, 1×10^{-6} to 1×10^{-7} M).

of the available alkylation sites that established the relative selectivity of (+)-**20** among the available alkylation sites is provided in the Supporting Information. All alkylation sites were found to be adenine and essentially each adenine N3 alkylation site was flanked by a 5′ and 3′ A or T base. The preference for this 3-base sequence followed the order: 5′-A$\underline{A}$A > 5′-A$\underline{A}$T ≥ 5′-T$\underline{A}$A > 5′-T$\underline{A}$T (Supporting Information). There was also a strong preference for the second and third 3′ base to be A or T, and the preferences distinguished the high-affinity versus lower-affinity alkylation sites. Thus, (+)-$CDPI_2$-DSA exhibits a 5 base-pair AT-rich alkylation selectivity that corresponds to 5′ → 3′ binding in the minor groove starting at the 5′ base preceding the alkylation site and extending over the alkylated adenine in the 3′ direction covering the 3 adjacent 3′ bases (*e.g.* 5′-A$\underline{A}$AAA).

The consensus alkylation selectivity of (−)-$CDPI_2$-DSA is summarized in Table 1, and a table of the statistical treatment of the available alkylation sites that established the relative selectivity of (−)-**20** among the available sites is provided in the Supporting Information. All alkylation sites were found to be adenine, and essentially each adenine N3 alkylation site was flanked by two 5′ A or T bases. The preference for this 3-base sequence followed the order: 5′-AA$\underline{A}$ > 5′-TA$\underline{A}$ > 5′-TT$\underline{A}$ > 5′-AT$\underline{A}$ (Supporting Information). There was also a preference for the third and fourth 5′ bases to be A or T and this distinguished the high-affinity versus low-affinity sites. Thus, (−)-$CDPI_2$-DSA exhibits a 5 base-pair AT-rich alkylation selectivity starting at the alkylated adenine and extending in the 3′ → 5′ direction across the 4 adjacent 5′ bases (*e.g.*, 5′-AAAA$\underline{A}$).

Examination of **17**−**19** revealed the same characteristics except that they exhibited either a 5 (**18**) or 3.5 (**17** and **19**) base-pair AT-rich selectivity corresponding to their sizes and lengths. This is summarized in Table 1 for $CDPI_1$-DSA (**19**) which was found to behave analogous to duocarmycin SA albeit with the reversed enantiomeric alkylation selectivity. This is illustrated beautifully in Supporting Information Figures 1 and 2 especially with w794 DNA where the natural enantiomers of **17**−**20** preferentially alkylate the single high-affinity site of the

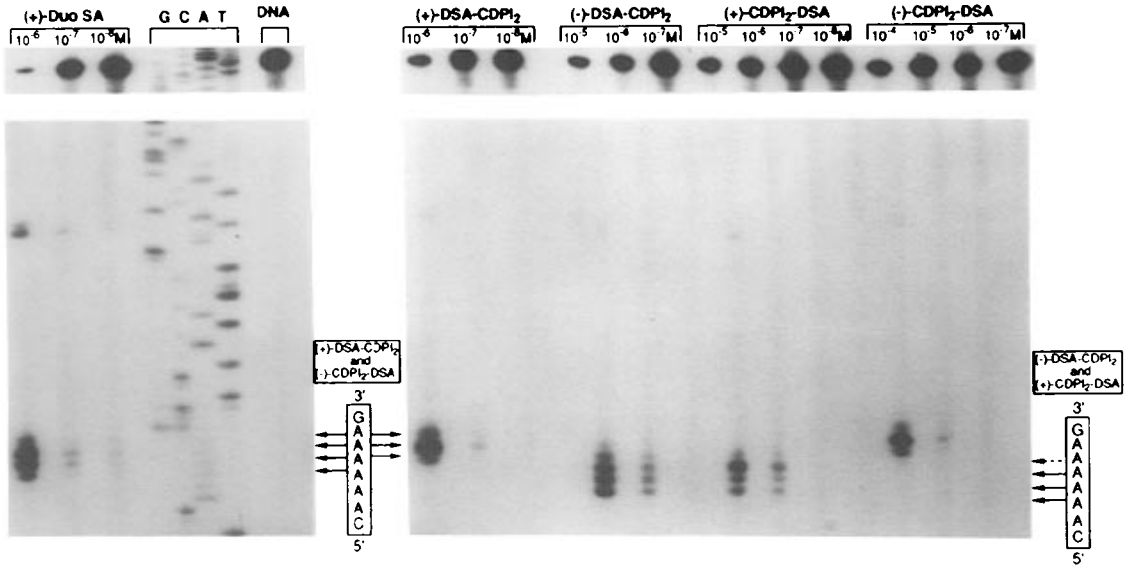


Figure 6. Thermally-induced strand cleavage of w836 DNA (146 bp, nucleotide no. 5189-91) after DNA−agent incubation at 25 °C (24 h, duocarmycin SA and DSA-$CDPI_2$) or 37 °C (5 days, $CDPI_2$-DSA), removal of unbound agent by EtOH precipitation and 30 min thermolysis (100 °C), followed by denaturing 8% PAGE and autoradiography. Lanes 1-3, (+)-duocarmycin SA ((+)-**1**, 1×10^{-6} to 1×10^{-8} M); lanes 4−7, Sanger G, C, A, and T sequencing standards; lane 8, control DNA; lanes 9−11, (+)-DSA-$CDPI_2$ (1×10^{-6} to 1×10^{-8} M); lanes 12−14, *ent*-(−)-DSA-$CDPI_2$ (1×10^{-5} to 1×10^{-7} M); lanes 15−18, (+)-$CDPI_2$-DSA ((+)-**20**, 1×10^{-5} to 1×10^{-8} M); lanes 19−22, *ent*-(−)-$CDPI_2$-DSA ((−)-**20**, 1×10^{-4} to 1×10^{-7} M).

Table 1. Consensus DNA Alkylation Sequences

agent	base[a]	5′	4	3	2	1	0	−1	−2	−3	−4	3′
Typical Agents: Natural Enantiomers												
(+)-CC-1065 (**3**)	A/T (56)[b]		67	78	94	98	100	55				
	consensus		A/T > G/C	A/T > G/C	A/T	A/T	A	Pu ≥ Py				
(+)-DSA-CDPI$_2$	A/T (56)		71	85	100	100	100	63				
	consensus		A/T > G/C	A/T > G/C	A/T	A/T	A	Pu > Py				
(+)-DSA-CDPI$_1$	A/T (56)			65	100	100	100	58				
	consensus			A/T ≥ G/C	A/T	A/T	A	Pu ≥ Py				
(+)-duocarmycin SA (**1**)	A/T (56)			79	100	100	100	69				
	consensus			A/T > G/C	A/T	A/T	A	Pu > Py				
(+)-*N*-BOC-DSA (**4**)	A/T (56)					95	100	65				
	consensus					A/T	A	Pu > Py				
Typical Agents: Unnatural Enantiomers												
(−)-*N*-BOC-DSA (**4**)	A/T (56)					95	100	65				
	consensus					A/T	A	Pu > Py				
(−)-duocarmycin SA (**1**)	A/T (56)					93	100	96	73	56		
	consensus					A/T	A	A/T	A/T > G/C	N		
(−)-DSA-CDPI$_2$	A/T (56)					100	100	100	90	73	58	
	consensus					A/T	A	A/T	A/T > G/C	A/T > G/C	N	
(−)-CC-1065 (**3**)	A/T (56)					88	100	93	82	73	56	
	consensus					A/T	A	A/T	A/T > G/C	A/T > G/C	N	
Reversed Analogs: Natural Enantiomers												
(+)-CDPI$_2$-DSA (**20**)	A/T (56)					95	100	98	85	70	55	
	consensus					A/T	A	A/T	A/T > G/C	A/T > G/C	N	
(+)-CDPI$_1$-DSA (**19**)	A/T (56)					92	100	94	73	60		
	consensus					A/T	A	A/T	A/T > G/C	N		
Reversed Analogs: Unnatural Enantiomers												
(−)-CDPI$_1$-DSA (**19**)	A/T (56)			70	98	98	100	53				
	consensus			A/T > G/C	A/T	A/T	A	Pu > Py				
(−)-CDPI$_2$-DSA (**20**)	A/T (56)		69	81	98	98	100	59				
	consensus		A/T > G/C	A/T > G/C	A/T	A/T	A	Pu > Py				
Sandwiched Analogs: Natural Enantiomers												
(+)-CDPI-DSA-CDPI (**30**)	A/T (56)			68	87	95	100	95	74	68		
	consensus				A/T > G/C	A/T	A	A/T	A/T > G/C			
Sandwiched Analogs: Unnatural Enantiomers												
(−)-CDPI-DSA-CDPI (**30**)	A/T (56)			64	86	94	100	94	78	68		
	consensus				A/T > G/C	A/T	A	A/T	A/T > G/C			

[a] Percentage of the indicated base located at the designated position relative to the adenine-N3 alkylation site. [b] Percentage composition within the DNA examined.

unnatural enantiomers of **1**−**3** and do not alkylate the typical natural enantiomer site. The smaller and shorter reversed analogs also proved less efficient at alkylating DNA. This efficiency followed the predictable trends of **20** > **18** > **19** > **17** and was more sensitive to the size of the agents than the typical extended agents. Moreover, the full set of reversed analogs **17**−**20** alkylated DNA at substantially slower rates and ultimately with lower efficiencies. Typically, the reactions for the reversed analogs were conducted at 37 °C for 5 days (versus 25 °C, < 12 h) and even then were 10−1000 times less efficient than **1**−**3** and related agents. In fact, their behavior proved more similar to the simple derivatives **4**−**6** than to **1**−**3**.

DNA Alkylation Properties of Agents Containing Modifications in the Terminal N^2 Substituent of the Reversed Analogs. The agents **23** and **24** were examined in efforts to determine whether the terminal N^2-BOC group of the reversed analogs was contributing to their reduced rate and efficiency of DNA alkylation. The three agents exhibited no differences in their DNA alkylation selectivity and only small differences in both the rate and efficiency of DNA alkylation. Moreover, the magnitude of these differences was much smaller than might be anticipated. The acetyl derivative **23** was found to be 2−5 times more efficient than the BOC derivative **18**, and the terminal NH agent **24** was about an order of magnitude less efficient than **18**. While these trends nicely follow the expected relative reactivities of the agents (**23** > **18** ≫ **24**), all three were still substantially slower (*ca.* 1000 times) and much less efficient (*ca.* 100 times) at alkylating DNA than the typical extended analog DSA-indole$_2$ (Figure 7). The unnatural enantiomers of **23** and **24**, like that of **18**, were approximately 10 times less effective than the natural enantiomers.

Agent	k_{rel}
(+)-DSA-indole$_2$	3000
(+)-indole$_2$-DSA-Ac (**23**)	3
(+)-indole$_2$-DSA-BOC (**18**)	2
(+)-indole$_2$-DSA-H (**24**)	1

[a] w836 DNA

(+)-DSA-indole$_2$ >> (+)-**23** > (+)-**18** > (+)-**24**

Figure 7. Relative rates of DNA Alkylation.[a]

DNA Alkylation Properties of the Sandwiched Analogs of Duocarmycin SA. Consistent with the noncovalent binding model, both enantiomers of CDPI-DSA-CDPI (**30**) alkylated the same sites, and their selectivity proved distinct from either enantiomer of the extended or reversed agents DSA-CDPI$_2$ or CDPI$_2$-DSA (**20**). The consensus alkylation sequence for (+)- and *ent*-(−)-**30** is summarized in Table 1, and a statistical treatment of the available alkylation sites is provided in the Supporting Information. All alkylation sites were found to be adenine, and essentially all adenine N3 alkylation sites were flanked by a 5′ and 3′ A or T base. The preference for this 3-base sequence follows the order: 5′-A<u>A</u>A > 5′-A<u>A</u>T ≥ 5′-T<u>A</u>A > 5′-T<u>A</u>T (Supporting Information). In addition, there was a very strong preference for both of the second 5′ and 3′ bases to be A or T. Exceptions typically involved one but not

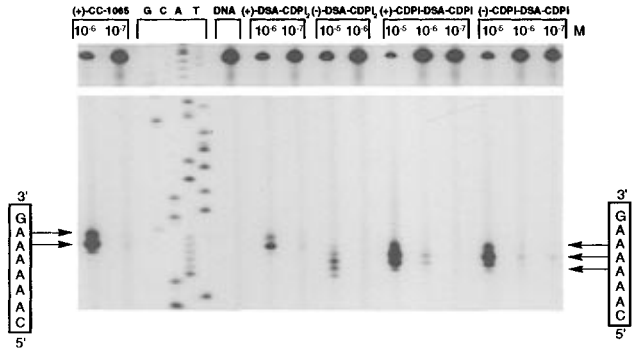


Figure 8. Thermally-induced strand cleavage of w836 DNA (146 bp, nucleotide 5189-91) after DNA−agent incubation at 25 °C (24 h), removal of unbound agent by EtOH precipitation and 30 min thermolysis (100 °C), followed by denaturing 8% PAGE and autoradiography. Lanes 1 and 2, (+)-CC-1065 (1×10^{-6} and 1×10^{-7} M); lanes 3−6, Sanger G, C, A and T sequencing standards; lane 7, control DNA; lanes 8 and 9, (+)-DSA-$CDPI_2$ (1×10^{-6} and 1×10^{-7} M); lanes 10 and 11, *ent*-(−)-DSA-$CDPI_2$ (1×10^{-5} and 1×10^{-6} M); lanes 12−14, (+)-CDPI-DSA-CDPI (1×10^{-5} to 1×10^{-7} M); lanes 15−17, (−)-CDPI-DSA-CDPI (1×10^{-5} to 1×10^{-7} M).

both of these locations, and the preference was strongest on the 5′ side of the alkylation site *e.g.* 5′-AAAAG > 5′-CAAAA) (Figure 8). Thus, alkylation was observed at adenines central to a 5 base-pair AT-rich sequence (*e.g.*, 5′-AAAAA). This is illustrated nicely in Figure 8 with w836 DNA where (+)- and *ent*-(−)-**30** exhibited identical alkylation profiles within the stretch of 6 adenines and alkylated those central to the sequence rather than the 3′ or 5′ terminal adenines characteristic of (+)- or *ent*-(−)-DSA-$CDPI_2$. The comparisons highlighted in Figure 8 are actually misleading in that the distinctions between **30**, and the extended or reversed analogs of duocarmycin SA are more pronounced than this might suggest. Very few of the alkylation sites overlap with those of either enantiomer of the extended or reversed analogs, and those that do are typically found in a long stretch of A's containing multiple alkylation sites for all agents.

Models of the DNA Alkylation Reactions. Both enantiomers of *N*-BOC-DSA exhibit an identical 2 base-pair alkylation selectivity (5′-AA > 5′-TA) with a requirement for a single 5′ A or T base adjacent to the adenine N3 alkylation site. For the natural enantiomer, this involves 3′-adenine N3 alkylation with agent binding in the 3′ → 5′ direction relative to the alkylated strand across the adjacent 5′ base (Figure 9). For the unnatural enantiomer, this similarly involves adenine N3 alkylation but with a reversed 5′ → 3′ binding orientation. As a consequence of the diastereomeric relationship of the adducts and in spite of the reversed 5′ → 3′ binding orientation, *ent*-(−)-*N*-BOC-DSA covers the same adjacent 5′ base as (+)-*N*-BOC-DSA. Thus, both enantiomers occupy the same binding site surrounding the alkylated adenine.

Models of alkylation at the high-affinity w794 DNA site by the natural enantiomer of the extended agent (+)-DSA-$CDPI_2$ and the unnatural enantiomer of the reversed agent *ent*-(−)-$CDPI_2$-DSA are illustrated in Figure 10 and highlight the origin of the switch in the inherent enantiomeric alkylation selectivity with the reversed analogs. The natural enantiomer, (+)-DSA-$CDPI_2$, alkylates the 3′-adenine and extends in the typical 3′ → 5′ direction over the adjacent 4 5′ bases (5′-AATTA). The unnatural enantiomer of the reversed analog, *ent*-(−)-$CDPI_2$-DSA, also alkylates the 3′-adenine albeit with the reversed orientation of the alkylation subunit but with the DNA-binding subunits extending in the atypical 3′ → 5′ direction over the same adjacent 4 5′ bases (5′-AATTA). Thus, the natural enantiomer of DSA-$CDPI_2$ and the unnatural enantiomer of the reversed analog $CDPI_2$-DSA bind and cover the exact same AT-rich 5 base-pair region surrounding the adenine N3 alkylation site.

Similar models of the unnatural enantiomer of *ent*-(−)-DSA-$CDPI_2$ and the natural enantiomer of the reversed agent (+)-$CDPI_2$-DSA are illustrated in Figure 11. The unnatural enantiomer of the typical agent alkylates adenine N3 and binds across a 5 base region in the 5′ → 3′ direction in the minor groove which is opposite that of the typical natural enantiomers. Because of the diastereomeric nature of the adducts and analogous to *ent*-(−)-*N*-BOC-DSA (*cf.*, Figure 9), the binding site starts at the 5′ site adjacent to the alkylation site and extends in the 5′ → 3′ direction covering the adenine N3 alkylation site and the following 3 3′ bases (5′-AATTT). Consequently, the 5 base-pair AT-rich binding site surrounding the alkylation site for *ent*-(−)-DSA-$CDPI_2$ is analogous to that of the natural enantiomer except that it extends in the reverse direction in the minor groove relative to the alkylated adenine and is offset by

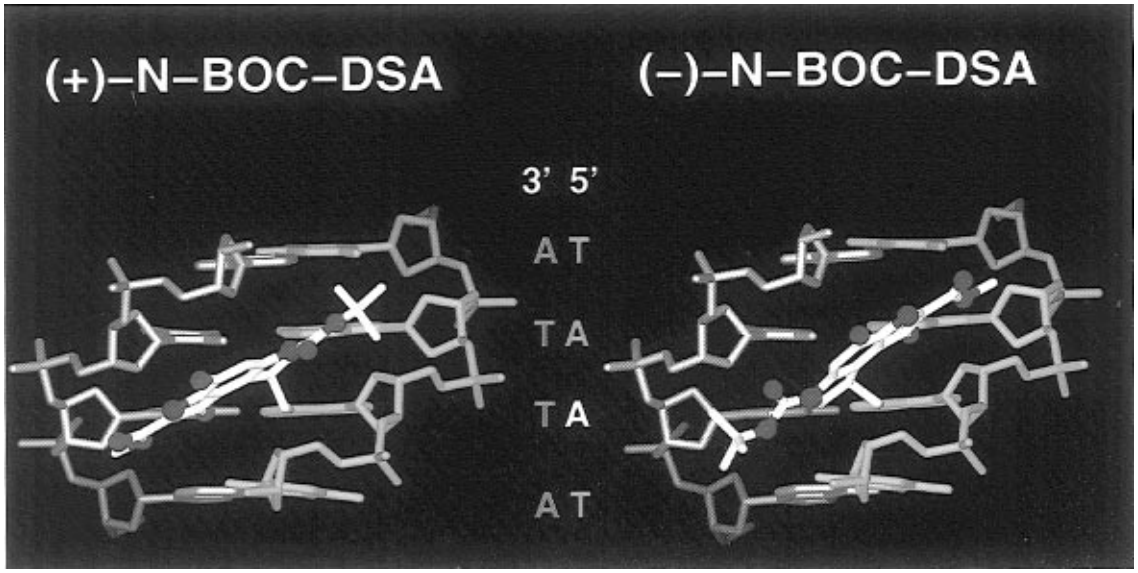


Figure 9. Stick models illustrating the alkylation of the same site within w794 DNA by (+)- and *ent*-(−)-*N*-BOC-DSA (**4**). The binding of the natural enantiomer extends in the 3′ → 5′ direction from the adenine N3 alkylation site across 5′-AA while that of the unnatural enantiomer binds in the reverse 5′ → 3′ direction but across the same 5′-AA site.

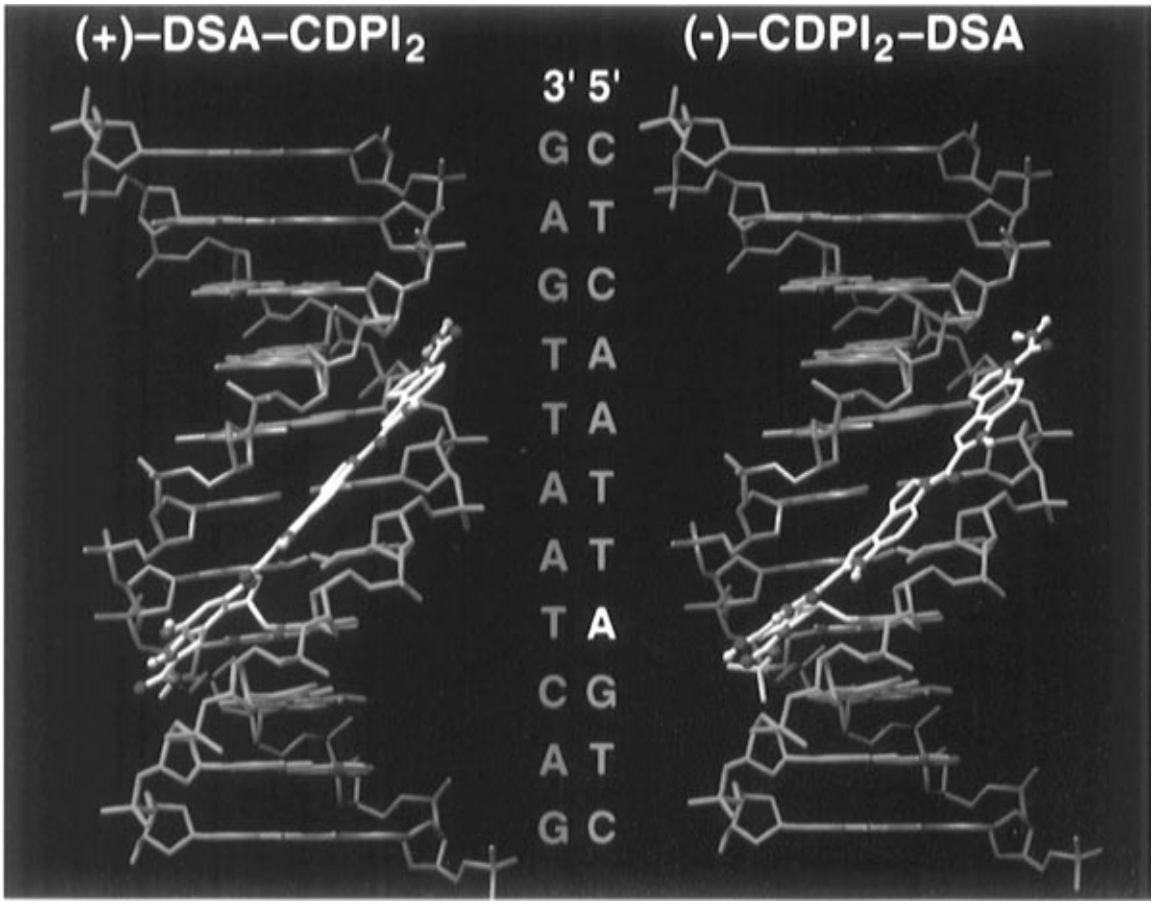


Figure 10. Stick models of (+)-DSA-$CDPI_2$ (left) and *ent*-(−)-$CDPI_2$-DSA (right, **20**) alkylation at the high-affinity site in w794 DNA, 5′-CTCAATTAGTC. The natural enantiomer of the typical agent extends in the 3′ → 5′ direction across the adenine N3 alkylation site across the 5-base AT-rich site 5′-AATTA. The unnatural enantiomer of the reversed analog binds in the same but atypical 3′ → 5′ direction across the identical 5-base AT-rich site 5′-AATTA.

1 base-pair. The natural enantiomer of the reversed analog, (+)-$CDPI_2$-DSA, covers the same 5 base-pair AT-rich region surrounding the alkylated adenine but with the reversed orientation of the alkylation subunit and with the DNA binding subunits extending in the atypical 5′ → 3′ direction.

Figure 12 illustrates the alkylation of the two enantiomers of CDPI-DSA-CDPI (**30**). The natural enantiomer binds in the minor groove with the alkylation subunit extending in the 3′ → 5′ direction with the binding subunits covering 2 base-pairs on both the 5′ and 3′ sides of the alkylation site. The binding extends slightly farther to the 5′ side (2.5 base-pairs) than the 3′ side (2 base-pairs), and this accounts nicely for the stronger and slightly extended AT preference on the 5′ side of the alkylation site. The unnatural enantiomer exhibits identical characteristics except that the alkylation subunit binds in the minor groove with the alkylation subunit extending in the reverse 5′ → 3′ direction relative to the alkylation strand covering the same 5-base AT-rich site surrounding the central alkylated adenine. This unusual feature of the two enantiomers binding and alkylating the same 5 base-pair AT-rich sites is analogous to (+)- and *ent*-(−)-*N*-BOC-DSA, except that the latter smaller agents only cover 2 base-pairs (*cf.*, Figure 9).

Sequence Preferences. Each of the three classes of agents has been shown to exhibit an adenine N3 alkylation selectivity that contains a 3-base AT sequence including and surrounding the alkylated adenine (Table 2). Although it is tempting to assign a special significance to the sequence preferences, the results are most consistent with a purely statistical preference. The most frequently alkylated sequence for each class is 5′-AAA, and the extent of alkylation diminishes as the A content gets smaller. In contrast to 5′-AAA, the mixed sequences contain competitive alkylation sites on the complementary unlabeled strand which diminish the apparent alkylation efficiency on the labeled strand. It is likely that a majority of the apparent selectivity is simply a statistical preference exaggerated by competitive unlabeled strand alkylation rather than unique characteristics embodied in the individual sequences. The exception to this generalization is the unusually effective alkylation of 5′-TTA by (+)-DSA-$CDPI_2$ representative of the typical natural enantiomers including (+)-CC-1065 and (+)-duocarmycin SA and the under represented alkylation of 5′-ATA by the unnatural enantiomers of the reversed analogs including *ent*-(−)-$CDPI_2$-DSA. The significance of this is not yet known.

Rates and Efficiencies of DNA Alkylation. In the course of examining the reversed and sandwiched agents, we observed substantial differences in the rates and efficiencies of DNA alkylation. First and foremost, we observed extraordinarily slow rates for the reversed agents at sites that are alkylated rapidly by the typical agents and we observed extraordinarily fast rates by the sandwiched agents at new sites not previously observed. This suggests that the characteristics responsible for the effective alkylation are not uniquely imbedded in the DNA sites or associated with the alkylated adenine but with structural features of the agents and intimately related to the source of catalysis for the reaction.

Consequently, we quantitated these differences by establishing the relative rate constants (k_{rel}, 5×10^{-6} M agent, 25 °C) for alkylation of the w836 high-affinity alkylation sites within the 6-base A sequence (5′-AAAAAA; *cf.*, Figure 8) that both

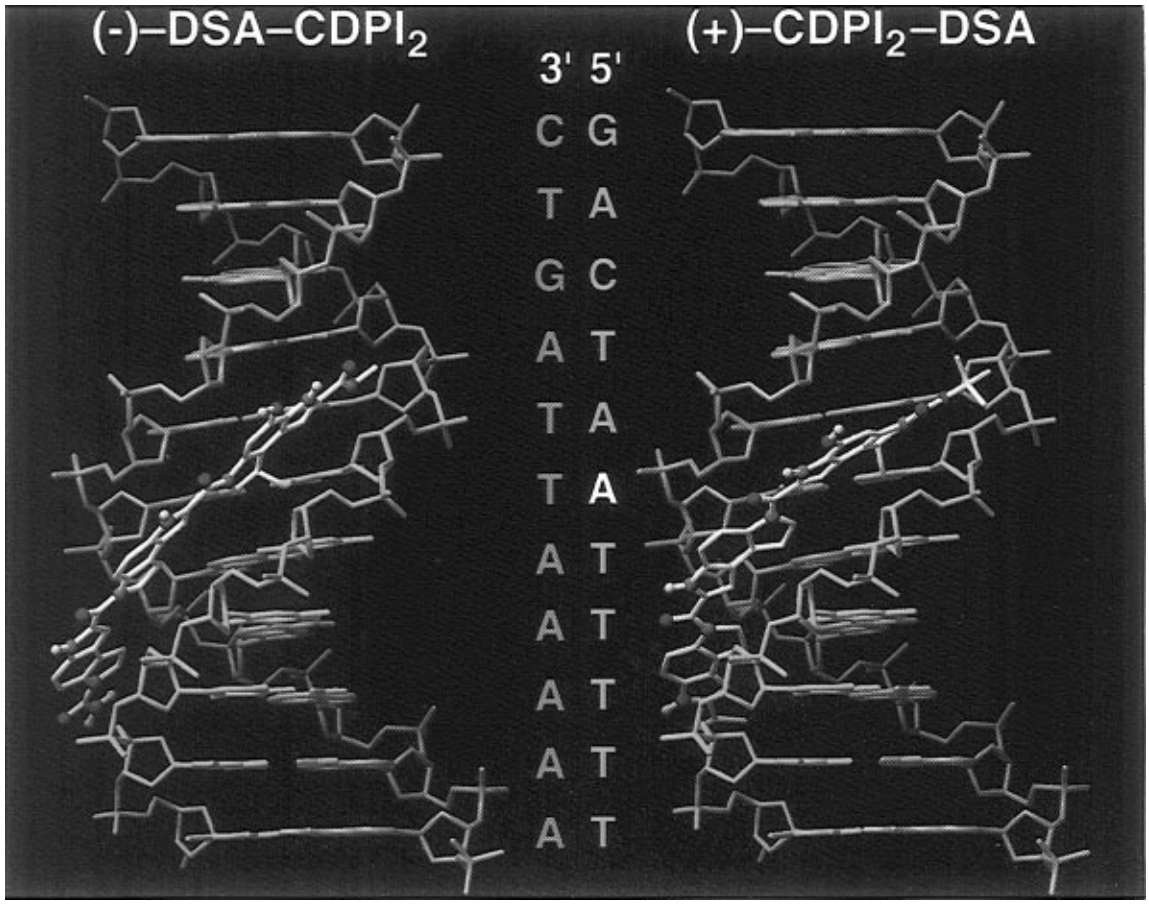


Figure 11. Stick models of *ent*-(−)-DSA-$CDPI_2$ (left) and (+)-$CDPI_2$-DSA (right, **20**) alkylation at the high-affinity site in w794 DNA, 5′-GACTAATTTTT-3′. The unnatural enantiomer of the typical agent binds in the 5′ → 3′ direction across the 5-base AT-rich site 5′-AATTT. The natural enantiomer of the reversed analog binds in the same but atypical 5′ → 3′ direction across the identical 5-base AT-rich site 5′-AATTT.

Table 2. Sequence Preferences

(+)- and *ent*-(−)-*N*-BOC-DSA	(+)-DSA-$CDPI_2$ and *ent*-(−)-$CDPI_2$-DSA	*ent*-(−)-DSA-$CDPI_2$ and (+)-$CDPI_2$-DSA	(+)-CDPI-DSA-CDPI and *ent*-(−)-CDPI-DSA-CDPI
5′-AA (75%)[a]	5′-AAA (62%, 74%)	5′-AAA (72%, 67%)	5′-AAA (59%, 59%)
5′-TA (40%)	5′-TTA (53%, 21%)	5′-AAT (40%, 28%)	5′-AAT (43%, 38%)
	5′-TAA (22%, 39%)	5′-TAA (39%, 33%)	5′-TAA (28%, 28%)
	5′-ATA (22%, 06%)	5′-TAT (13%, 13%)	5′-TAT (00%, 00%)

[a] Percent (%) frequency of alkylation, *e.g.*, 75% of all available 5′-AA sites were alkylated by (+)- and *ent*-(−)-*N*-BOC-DSA.

enantiomers of all three agent classes effectively alkylate. The results are summarized in Figure 13. Illustrating the differences more dramatically, Supporting Information Figure 3 records the extent of alkylation at selected time points taken from the assay. After just 1−5 min, the natural enantiomer of the typical extended agent, (+)-DSA-$CDPI_2$ (+E), and both enantiomers of the sandwiched analog **30**, CDPI-DSA-CDPI (+ and −S), exhibit extensive alkylation. In contrast, the unnatural enantiomer of the typical extended agent, *ent*-(−)-DSA-$CDPI_2$ (−E), requires 6 h to approach the same level of DNA alkylation and both enantiomers of the reversed analog, $CDPI_2$-DSA (+ and −R), require 72 h to reach a detectable and diminished level of DNA alkylation. Although the reversed natural enantiomer was faster than the unnatural enantiomer, neither comes even close to the rates exhibited by (+)-DSA-$CDPI_2$ or (+)- and *ent*-(−)-CDPI-DSA-CDPI (**30**) and more closely approximate the rates of DNA alkylation observed with *N*-BOC-DSA (**4**) which lacks the DNA-binding subunits altogether.

Additional Structural Requirement for Catalysis of the DNA Alkylation Reaction and Source of Catalysis. The DNA alkylation rate of the extended as well as the sandwiched agents is exceptionally fast and typical of this class of agents while that of all reversed agents is exceptionally slow, proceeding at rates similar to those of *N*-BOC-DSA (**4**). Although there are many potential explanations for this, it is not due to differences in the noncovalent binding affinity of the agent. In addition, the DNA alkylation selectivity of the reversed versus extended analogs simply reversed with these agents, and no new sites were detected. Thus, it was the rates but not the sites that were altered.

Complementing these observations, the rapid rate of DNA alkylation by the sandwiched analogs was observed at a new set of alkylation sites independent of the absolute configuration of the agent indicating that the source of catalysis was not uniquely imbedded in the original DNA alkylation sites. Rather, the distinguishing feature between the extended or sandwiched analogs and the reversed analogs is the presence of the right-hand heteroaryl N^2 amide. Thus, a rigid extended N^2 amide substituent is required for rapid and effective alkylation of duplex DNA. With the sandwiched analogs, this effect is independent of the sites of DNA alkylation and the enantiomeric configuration of the alkylation subunit. We suggest that upon binding to DNA with the adoption of a helical-bound conformation, the inherent twist of the alkylation subunit N^2 amide in

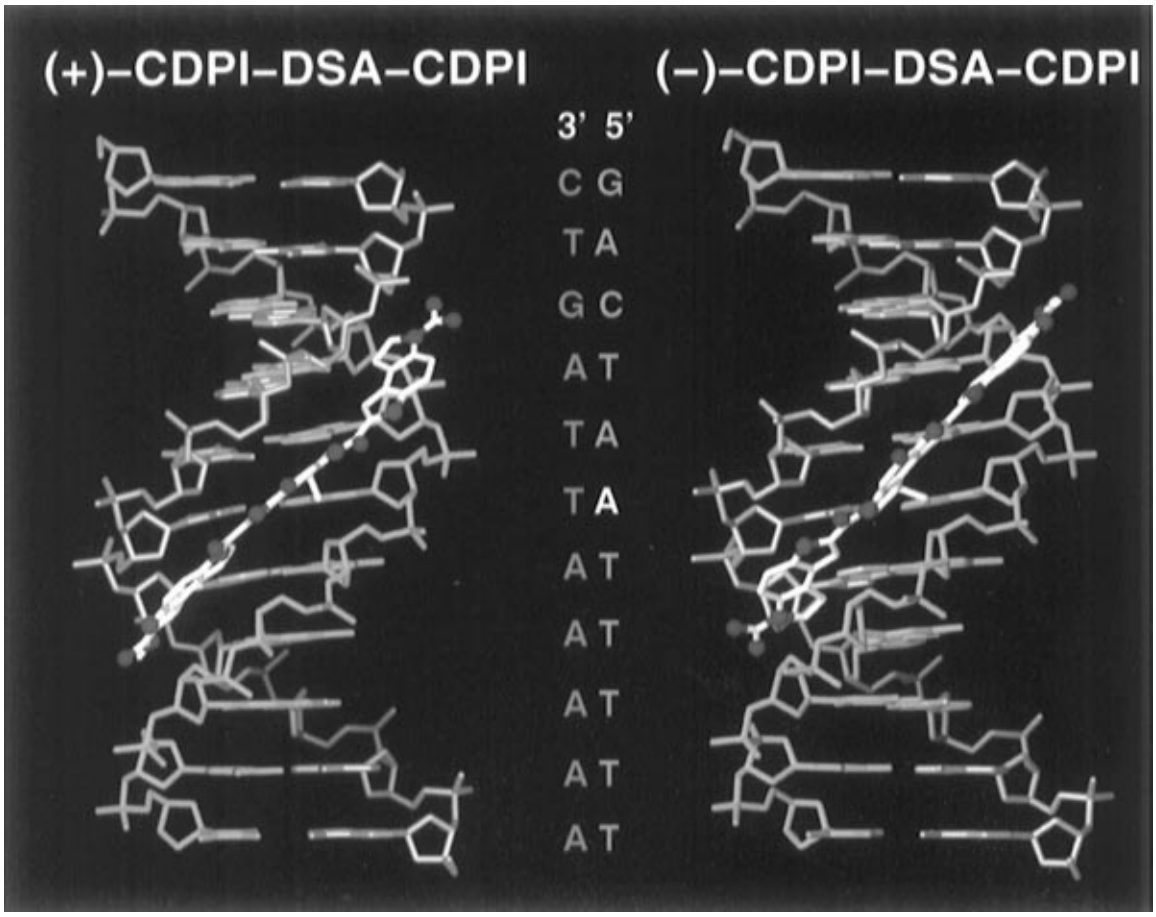


Figure 12. Comparison models of (+)- and *ent*-(−)-CDPI-DSA-CDPI (**30**) alkylation and binding at the high-affinity site in w794 DNA, 5′-GACTAATTTTT. Both enantiomers bind and alkylate the 5′-TAATT site but with reversed binding orientations.

Agent	k_{rel}
(+)-CDPI-DSA-CDPI (**30**)	16500
(+)-DSA-$CDPI_2$	13000
ent-(−)-CDPI-DSA-CDPI (**30**)	10000
ent-(−)-DSA-$CDPI_2$	100
(+)-$CDPI_2$-DSA (**20**)	4
ent-(−)-$CDPI_2$-DSA (**20**)	1

Natural enantiomers:
CDPI-DSA-CDPI = DSA-$CDPI_2$ >> $CDPI_2$-DSA (10^3-10^4x)
Unnatural enantiomers:
CDPI-DSA-CDPI > DSA-$CDPI_2$ (100x) > $CDPI_2$-DSA (10^4x)
Extended (typical) analogs:
(+)-DSA-$CDPI_2$ > *ent*-(−)-DSA-$CDPI_2$ (100x)
Sandwiched analogs:
(+)-CDPI-DSA-CDPI ≥ *ent*-(−)-CDPI-DSA-CDPI (1-2x)
Reversed analogs:
(+)-$CDPI_2$-DSA ≥ *ent*-(−)-$CDPI_2$-DSA (4x)

Figure 13. Relative rates of DNA alkylation.

the reversed analogs is not altered and, thus, not activated for nucleophilic addition. Consequently, they undergo DNA alkylation at rates comparable to those of the simple derivatives such as **4**. The simple derivatives including **4** and the reversed agents **17**−**20** both require extended reaction times (2−7 days, 25−37 °C) for substantial or complete alkylation with **17**−**20** being only marginally faster and ultimately 10−100 times more efficient. These small differences can be attributed to the more effective noncovalent minor groove binding properties of **17**−**20**. The larger 10^3−10^4 rate differences between **1**−**3** and **17**−**20** may be attributed to the DNA binding-induced conformational change uniquely imposed on **1**−**3** and related agents. In the absence of the extended right-hand subunit, DNA minor groove binding no longer requires an induced twist in the N^2 amide linkage depriving the agent of this additional activation toward DNA alkylation.

Cytotoxic Activity. The in vitro cytotoxic activity of the agents is summarized in Table 3. Several important trends were observed which parallel the observations made in the DNA alkylation studies. Foremost, the natural enantiomers of the reversed analogs were 100−1000 times less potent than the corresponding extended analogs while the two sandwiched analogs **30** and **31** were essentially equipotent with the typical agents. The reversed analogs proved essentially indistinguishable (**17**, **18**, and **23**) or less than 10 times more potent (**19** and **20**) than the simple derivatives, *N*-BOC-DSA (**4**) and *N*-Ac-DSA, that lack DNA binding subunits altogether. This is analogous to the observations made in the DNA alkylation studies where **17**−**20** and **23**−**24** were found to alkylate DNA with a rate and efficiency similar to those of **4** rather than **1**−**3**. Consistent with the trends observed in the DNA alkylation studies, the unnatural enantiomers of the agents generally were approximately 10 times less potent than the natural enantiomers. The exceptions to this generalization are DSA-$CDPI_2$, $CDPI_2$-DSA, and the sandwiched analogs where the two enantiomers proved nearly equipotent.

Like the natural enantiomers, the unnatural enantiomers of the reversed agents were also much less potent than the corresponding unnatural enantiomers of the extended analogs although the magnitude of the differences was somewhat smaller (20−100 versus 100−1000 times). However, they proved to be even more comparable in potency to the unnatural enantiomers of *N*-BOC-DSA (**4**) and *N*-Ac-DSA lacking the attached

Table 3. In Vitro Cytotoxic Activity, L1210 (pM)

reversed agents	IC_{50}	sandwiched agents	IC_{50}	extended agents	IC_{50}
		Natural Enantiomers			
17, (+)-indole$_1$-DSA	6000			(+)-DSA-indole$_1$	65
18, (+)-indole$_2$-DSA	1000			(+)-DSA-indole$_2$	3
19, (+)-CDPI$_1$-DSA	400			(+)-DSA-CDPI$_1$	4
20, (+)-CDPI$_2$-DSA	500	**30**, (+)-CDPI-DSA-CDPI	5	(+)-DSA-CDPI$_2$	4
23, (+)-indole$_2$-DSA-$COCH_3$	2000	**31**, (+)-CDPI-DSA-TMI	10	(+)-*N*-Ac-DSA	1000
24, (+)-indole$_2$-DSA-NH	20000			(+)-*N*-BOC-DSA	6000
		Unnatural Enantiomers			
17, (−)-indole$_1$-DSA	50000			(−)-DSA-indole$_1$	1700
18, (−)-indole$_2$-DSA	10000			(−)-DSA-indole$_2$	150
19, (−)-CDPI$_1$-DSA	3000			(−)-DSA-CDPI$_1$	130
20, (−)-CDPI$_2$-DSA	500	**30**, (−)-CDPI-DSA-CDPI	6	(−)-DSA-CDPI$_2$	20
23, (−)-indole$_2$-DSA-$COCH_3$	30000	**31**, (−)-CDPI-DSA-TMI	20	(−)-*N*-Ac-DSA	45000
24, (−)-indole$_2$-DSA-NH	30000			(−)-*N*-BOC-DSA	60000

DNA-binding subunits. The nature of the terminal N^2 acyl substituent did not alter these observations in a substantial manner. The corresponding natural and unnatural enantiomers of **18** and **23** proved essentially indistinguishable while **24** was noticeably less potent. This latter effect could be anticipated given the remarkable stability of the alkylation subunits lacking a N^2 acyl substituent and the resulting less effective DNA alkylation. Thus, consistent with the trends observed in the rates and efficiencies of DNA alkylation and independent of the nature of the simple N^2 substituent, the reversed analogs proved comparable in cytotoxic potency to the simple derivatives **4** and *N*-Ac-DSA lacking the DNA-binding subunits rather than **1**−**3** and the related extended analogs.

The potent cytotoxic activity of the sandwiched analogs is observed with a class of agents that exhibit a significantly different selectivity of DNA alkylation than preceding agents and one which is the same for both the natural and unnatural enantiomers. Since both enantiomers are essentially equipotent and exhibit the same DNA alkylation selectivity, this could permit the use of the more readily available racemic agents in the development of potential clinical candidates with confidence that both enantiomers contribute productively and equivalently to the expression of the resulting properties.

The cytotoxic activities of the more potent, extended derivatives plateau at 3−4 pM and closely follow established trends relating chemical stability and cytotoxic potency.[2] Since the duocarmycin SA alkylation subunit is among the most stable examined to date, the cytotoxic activity of such analogs is among the most potent yet described.

The seco precursor agents (*e.g.* **13**−**16**, **22**, **28** and **29**) exhibited cytotoxic activity that was not distinguishable from the corresponding cyclopropane containing agents **17**−**20**, **23**, **30** and **31**, respectively, indicating that their spirocyclization to the biologically potent agents under the conditions of assay is not limiting.

Conclusions. A number of key features contribute to the sequence-selective DNA alkylation by members of this class of agents. The reaction constitutes nucleophilic addition of the most sterically accessible of the two most nucleophilic sites in the minor groove (adenine N3 versus guanine N3).[2,11−13] The clean regioselectivity for exclusive addition to the least substituted cyclopropane carbon represents the stereoelectronically-preferred site of attack[1−4,20,21] which is further reinforced by the destabilizing torsional and steric interactions[1,2,20] that would accompany addition to the more substituted carbon with ring expansion and is characteristic of S_N2 addition of a hindered nucleophile. Consistent with the noncovalent binding model, the length-dependent AT-rich alkylation selectivity is derived from the preferential noncovalent binding selectivity of the agents in the deeper, narrower AT-rich minor groove.[14−16] The reverse and offset AT-rich alkylation selectivity of the enantiomers and the switch in the inherent enantiomeric alkylation selectivity of the reversed analogs establish that both enantiomers are subject to the same polynucleotide recognition features.[11−13] Finally, a DNA binding-induced conformational change in the agent induces a twist in the linking N^2 amide resulting in loss of the alkylation subunit vinylogous amide stabilization catalyzing the DNA alkylation reaction.[1,19] Since the binding-induced conformational change is greatest within the narrower, deeper AT-rich minor groove, this leads to selective catalysis within the preferred binding sites. An important ramification of the binding-induced substrate ground state destabilization is that it further serves to stabilize the inherently reversible DNA alkylation reaction.[1,2,11−13,22] As such, our original characterization of the DNA alkylation reaction as "accessible hydrophobic binding-driven-bonding" may prove more accurate today than when originally detailed.[2,23]

This alternative source of catalysis for the DNA alkylation reaction explains a number of observations. It is consistent with the lack of a substantial pH dependence on the rate of reaction[1] and explains the growing number of instances where the rates of DNA alkylation were not found to follow the relative rates of acid-catalyzed reactivity.[19] It explains the extraordinarily slow DNA alkylation rates of the reversed analogs and suggests that the small differences between the reversed analogs and simple derivatives such as **4** constitute those attributable to the enhanced minor groove binding. The remaining larger differences between the reversed analogs and the typical agents, and thus the bulk of the distinctions between **1**−**3** and **4**−**6** constitute the catalysis derived from the DNA binding-induced conformational change in the agent. It is consistent with the well-established view that the unusual stability of the agents is derived from the vinylogous amide conjugation and that its disruption should lead to significant increases in reactivity.[19] Moreover, it offers new insights into the reversibility of the DNA alkylation reaction and suggests that both the rate of retroalkylation and the reaction equilibrium is shifted to favor adduct formation through a binding-induced destabilization of the substrate ground state. It explains the important effects of selected substituents

(20) Boger, D. L.; McKie, J. A.; Nishi, T.; Ogiku, T. *J. Am. Chem. Soc.* **1997**, *119*, 311.

(21) 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.

(22) Warpehoski, M. A.; Harper, D. H.; Mitchell, M. A.; Munroe, J. J. *Biochemistry* **1992**, *31*, 2502. Lee, C.-S.; Gibson, N. W. *Biochemistry* **1993**, *32*, 9108. Asai, A.; Nagamura, S.; Saito, H.; Takahasi, I.; Nakano, H. *Nucleic Acids Res.* **1994**, *22*, 88.

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

that are located in the minor groove of the bound agents (*e.g.*, duocarmycin SA C6-CO_2Me and C5′-OMe)[1] and even offers a new insight into the origin of the distinctions between the enantiomers of the typical agents such as **1**−**3**.[19] This shape-selective catalysis coincides within the preferred noncovalent binding sites. That the sequence selectivity is controlled by the noncovalent binding selectivity is defined by the fact that the identical selectivities are observed with agents not subject to this source of catalysis[11,12,16] including the reversed analogs albeit with alkylation at much slower rates.

Experimental Section

2-((*tert*-Butyloxy)carbonyl)-4-oxo-1,2,8,8a-tetrahydrocyclopropa-[*c*]pyrrolo[3,2-*e*]indol-6-carboxylic Acid (7). A solution of **4** (56 mg, 163 μmol) in THF/CH_3OH/H_2O (3:2:1, 1.6 mL) was treated with 0.17 mL of aqueous 1 N LiOH (1.05 equiv), and the reaction mixture was warmed at 60 °C under Ar for 1 h. An additional 0.041 mL of aqueous 1 N LiOH (0.25 equiv) was added, and the solution was warmed at 60 °C for an additional 30 min. The reaction mixture was allowed to cool to 25 °C, and the solvent was removed under a stream of N_2. H_2O (2 mL) was added, and the aqueous layer was extracted with EtOAc (2 × 2 mL). EtOAc (2 mL) was added to the aqueous layer, followed by aqueous 1 N $KHSO_4$ (0.21 mL, 1 equiv). The mixture was diluted with H_2O (35 mL) and extracted with EtOAc (3 × 35 mL). The combined organic extract was washed with H_2O (35 mL) and dried (Na_2SO_4). Concentration under reduced pressure provided **7** (50.0 mg, 93%) as a yellow film: ^{1}H NMR (acetone-d_6, 400 MHz) δ 11.15 (s, 1H, NH), 6.77 (s, 1H, C3-H), 6.75 (d, 1H, J = 2.0 Hz, C7-H), 4.03−4.01 (m, 2H, C1-H_2), 2.90−2.86 (m, 1H, C8a-H), 1.70 (dd, 1H, J = 7.7, 4.4 Hz, C8-H), 1.52 (s, 9H, $C(CH_3)_3$), 1.40 (t, 1H, J = 4.4 Hz, C8-H); IR (film) ν_{max} 3204, 2963, 1721, 1667, 1597, 1393, 1277, 1258, 1155, 1137, 798 cm^{-1}; FAB HRMS (NBA/CsI) *m/z* 331.1290 (M^+ + H, $C_{17}H_{18}N_2O_5$ requires 331.1294).

(+)-**7**: $[\alpha]^{23}_D$ +190° (*c* 0.2, CH_3OH).

ent-(−)-**7**: $[\alpha]^{23}_D$ −193° (*c* 0.2, CH_3OH).

General Procedure for the Preparation of 13−16. 7-[Methyl 1,2-Dihydro(3*H*-pyrrolo[3,2-*e*]indol-3-yl)carbonyl-7-carboxylate]-3-((*tert*-butyloxy)carbonyl)-1-(chloromethyl)-5-hydroxy-1,2-dihydro-3*H*-pyrrolo[3,2-*e*]indole (15). A solution of **7** (3.1 mg, 9.4 μmol) in EtOAc (0.70 mL) under Ar was treated with 4 N HCl/EtOAc (12 μL, 47 μmol, 5 equiv) over 10 s. The yellow reaction mixture was stirred at 25 °C for 30 min and was concentrated under a stream of N_2. The resultant yellow residue containing **8** was dissolved in DMF (0.17 mL) and treated sequentially with **11** (1.7 mg, 7.9 μmol, 0.9 equiv) and EDCI (3.4 mg, 17.6 μmol, 2.0 equiv). The reaction mixture was stirred for 2 h at 25 °C before the solvent was removed under reduced pressure, and the residual solid was slurried in 0.5 mL H_2O. The solid was collected by centrifugation and washed with 1% aqueous HCl (1 × 0.5 mL) and H_2O (1 × 0.5 mL). Drying the solid *in vacuo* afforded **15** (4.3 mg, 96%) as a tan solid: mp 188−190 °C (dec.); ^{1}H NMR (acetone-d_6, 400 MHz) δ 10.99 (s, 1H, NH), 10.40 (s, 1H, NH), 8.81 (br s, 1H, OH), 8.41 (m, 1H, C4′-H), 7.59 (br s, 1H, C4-H), 7.41 (d, 1H, J = 8.9 Hz, C5′-H), 7.22 (d, 1H, J = 1.5 Hz), 7.12 (s, 1H), 4.80−4.62 (m, 2H), 4.20−4.10 (m, 2H), 4.04 (dd, 1H, J = 11.8, 4.0 Hz), 4.04−3.92 (m, 1H, C1-H), 3.89 (s, 3H, OCH_3), 3.73 (t, 1H, J = 9.3 Hz, C2′-H), 3.49 (t, 2H, J = 8.3 Hz, C1′-H_2), 1.55 (s, 9H, $C(CH_3)_3$); IR (neat) ν_{max} 3327, 2975, 2933, 1690, 1669, 1607, 1524, 1436, 1389, 1369, 1327, 1255, 1213, 1141 cm^{-1}; FAB HRMS (NBA/CsI) *m/z* 697.0809 (M^+ + Cs, $C_{29}H_{29}N_4O_6Cl$ requires 697.0830).

(1*S*)-**15**: $[\alpha]^{22}_D$ −23° (*c* 0.2, THF).

ent-(1*R*)-**15**: $[\alpha]^{23}_D$ +27° (*c* 0.2, THF).

General Procedure for the Preparation of 17−20, Method A. 6-[7-[Methyl 1,2-Dihydro(3*H*-pyrrolo[3,2-*e*]indol-3-yl)carbonyl-7-carboxylate]-1,2-dihydro(3*H*-pyrrolo[3,2-*e*]indol-3-yl)carbonyl]-2-((*tert*-butyloxy)carbonyl)-1,2,8,8a-tetrahydrocyclopropa[*c*]pyrrolo-[3,2-*e*]indol-4-one (20, CDPI$_2$-DSA). A portion of NaH (1.1 mg, 60%, 26.7 μmol, 10 equiv) in DMF at 0 °C under Ar was treated with a solution of **16** (2.0 mg, 2.67 μmol, 1.0 equiv) in DMF (0.25 mL), and the reaction mixture was stirred for 45 min at 0 °C. The reaction mixture was directly subjected to flash chromatography (0.5 × 5 cm SiO_2, 15% DMF/toluene) to afford **20** (1.8 mg, 95%) as a pale yellow solid: mp 213−215 °C (dec.); ^{1}H NMR (DMF-d_7, 400 MHz) δ 11.99 (s, 1H, NH), 11.79 (s, 1H, NH), 11.73 (d, 1H, J = 1.6 Hz, NH), 8.42−8.20 (m, 2H), 7.46 (app t, 2H, J = 9.2 Hz), 7.20 (d, 1H, J = 1.6 Hz), 7.17 (d, 1H, J = 1.4 Hz), 6.81 (s, 1H), 6.68 (s, 1H), 4.75 (t, 2H, J = 8.4 Hz), 4.55 (app q, 2H, J = 7.9 Hz), 4.07 (dd, 1H, J = 11.2, 4.9 Hz), 4.04−3.99 (m, 1H), 3.93 (s, 3H), 3.57−3.45 (m, 4H, partially obscured by H_2O), 3.00 (dt, 1H, J = 7.8, 4.3 Hz, C8a-H), 1.82 (dd, 1H, J = 7.7, 3.7 Hz, C8-H), 1.53 (s, 9H, $C(CH_3)_3$), 1.49 (t, 1H, J = 4.4 Hz, C8-H); IR (neat) ν_{max} 3324, 2958, 2931, 2871, 1709, 1703, 1620, 1582, 1530, 1511, 1503, 1434, 1378, 1348, 1256, 1211, 1163, 1143, 1022 cm^{-1}; FAB HRMS (NBA) *m/z* 713.2757 (M^+ + H, $C_{40}H_{36}N_6O_7$ requires 713.2724).

(+)-CDPI$_2$-DSA (**20**): $[\alpha]^{23}_D$ +56° (*c* 0.1, DMF).

ent-(−)-CDPI$_2$-DSA (**20**): $[\alpha]^{23}_D$ −60° (*c* 0.1, DMF).

Acknowledgment. We gratefully acknowledge the financial support of the National Institutes of Health (CA55276), the Skaggs Institute for Chemical Biology, and the award of predoctoral (D.S.J., ACS Organic Division sponsored by Zeneca Pharmaceutical, 1993−94; Eli Lilly, 1995−96) and postdoctoral fellowships (D.L.H., NIH CA09303; B.B., Deutsch Fellowship Bo-1309/1).

Supporting Information Available: Experimental procedures and characterization of **10**, **13**, **14**, **16**−**19**, **22**−**24**, and **28**−**31**, three detailed and four summary tables of the statistical treatment of the alkylation sites for (+)-**20**, *ent*-(−)-**20**, and (+)/(−)-**30**, gel figures of (+)-**17**−**20** in w794 DNA (Figure 1), (+)- and *ent*-(−)-**17**−**20** in w836 DNA (Figure 2), and a gel of representative time points in the relative rate study summarized in Figure 13 (Figure 3) (22 pages). See any current masthead page for ordering and Internet access instructions.

JA9637210