Pergamon

Bioorganic & Medicinal Chemistry, Vol. 3, No. 6, pp. 611–621, 1995

0968-0896/95 $9.50 + .00

0968-0896(95)00048-8

CC-1065 CBI Analogs: an Example of Enhancement of DNA Alkylation Efficiency Through Introduction of Stabilizing Electrostatic Interactions

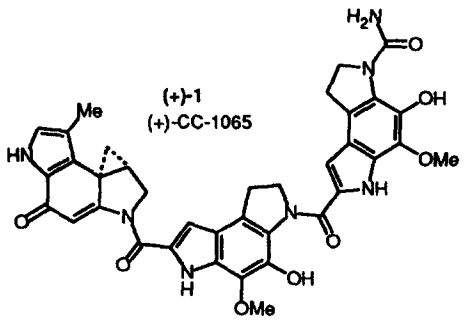
Dale L. Boger,* Weiya Yun, Nianhe Han and Douglas S. Johnson

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

Abstract—The three trimethylammonium salts **3–5** proved to be 100 times more efficient at alkylating DNA than **2** and exhibited DNA alkylation efficiencies identical to that of (+)-CC-1065 (**1**). In addition, the agents **3** and **4** exhibited DNA alkylation selectivities identical to that of **2**. This may be attributed to spatially well-defined stabilizing electrostatic interactions between the positively charged trimethylammonium salt lying on the peripheral face of the agents and the bracketing, negatively charged phosphates located in the DNA backbone that enhance the DNA noncovalent binding affinity without affecting DNA binding or alkylation selectivity. The agent **5** exhibited an altered and more discriminating AT-rich adenine N3 alkylation selectivity than **2–4** that may be attributed to the groove placement of the large trimethylammonium salt.

Introduction

(+)-CC-1065 (**1**), disclosed in 1978,[1] and the duocarmycins, disclosed in 1988 and 1990,[2–4] represent the initial members of a class of exceptionally potent antitumor antibiotics that derive their biological effects through the reversible, sequence selective alkylation of DNA.[5–13] Subsequent to their disclosure, extensive efforts have been devoted to establish their DNA alkylation selectivity and its structural origin,[5–13] to establish the link between DNA alkylation and the ensuing biological properties,[14] and to define the fundamental principles underlying the relationships between structure, chemical reactivity, and biological properties.[15–24]



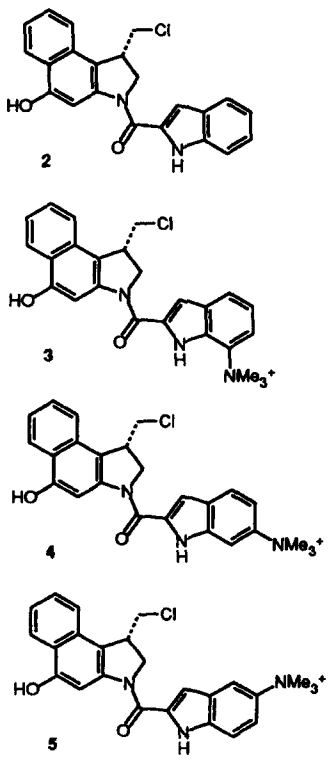
Early in the efforts, we demonstrated that the noncovalent binding affinity of CC-1065 for the DNA minor groove was derived nearly exclusively from stabilizing van der Waals contacts and hydrophobic binding and that removal of the peripheral methoxy and hydroxy substituents had no effect on its noncovalent AT-rich binding selectivity and binding affinity.[25] This dependence on hydrophobic binding stabilization leads to preferential binding in the narrower, deeper AT-rich regions of the minor groove where the stabilizing van der Waals contacts are maximal ($\Delta G°$ = 9.5–11.5 kcal mol^{-1}).[25] Moreover, we have advanced a model where this preferential AT-rich five base-pair noncovalent binding selectivity, coupled with the steric accessibility to the alkylation site accomplished only with the deep groove penetration at an AT-rich site, are the primary determinants controlling the sequence selectivity of the DNA alkylation reaction.[6,9,11,13,19,20] Not only does this model explain the behavior of both the natural and unnatural enantiomers of CC-1065[9] and the duocarmycins[6] but it also accommodates the behavior and DNA alkylation selectivity of the simplified structural analogs that have been disclosed to date.[6,9,10,19–22] Although these proposals have not received universal acceptance and alternative explanations have been advanced,[8,12,26] a past[27] and more recent definitive examination[28] proved consistent with this model and inconsistent with alternative models that have been advanced.[8,12,26]

In addition, we have examined the DNA binding affinity and selectivity impact of electronegative and electropositive substituents placed on the peripheral face of the agents. In these studies, we found an apparent destabilizing contribution to the DNA binding affinity that results from the introduction of a strong electronegative substituent on the outer face of the agents and found a substantial enhancement of noncovalent minor groove binding affinity that results from introduction of an electropositive substituent.[29] This may be attributed to a spatially well-defined destabilizing or stabilizing electrostatic interaction with the DNA phosphate backbone and was found to have little impact on the intrinsic preferential AT-rich binding selectivity of the agents.[29] Herein, we report the extension of these studies[29,30] to the examination of **3–5** versus **2**.

Results and Discussion

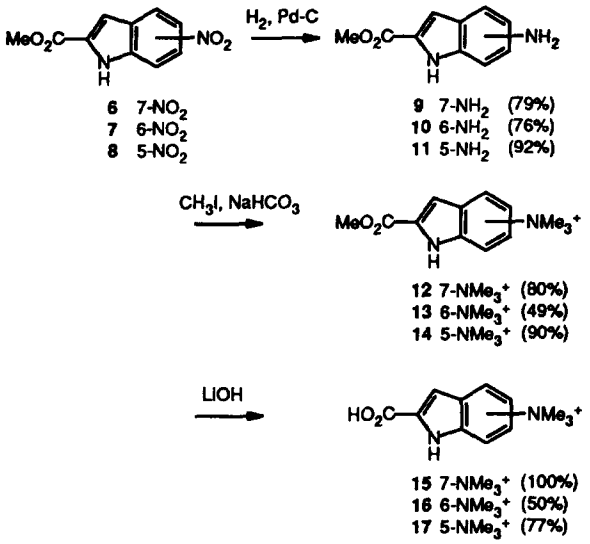
Preceding studies of the CBI-based analogs of CC-1065 have shown that they are chemically more stable (4 ×), biologically more potent (4 ×) and synthetically more accessible than the corresponding agents incorporating the natural CPI alkylation subunit of CC-1065.[20] Moreover, the CBI-based analogs alkylate DNA with an unaltered sequence selectivity at an enhanced rate and with a slightly greater efficiency than the corresponding CPI analog.[20] Consequently, we have conducted the studies detailed herein with this synthetic CBI class of agents. In addition, the *seco* agents including **2–5** which readily close to the corresponding cyclopropane containing agents have been shown in past studies to exhibit biological properties and DNA alkylation efficiencies and selectivities identical to those of the agent containing the preformed cyclopropane.[20] Consequently, for the studies detailed herein, the *seco* agents **2–5** were used in the comparative examinations.

In the preceding studies, shortened or aborted CBI analogs of CC-1065 including *N*-BOC-CBI and (+)-CBI-indole$_1$ (**2**) were found to be less effective DNA alkylating agents and biologically less potent derivatives than the natural products or its close analogs.[20] This may be attributed to their diminished DNA alkylation capabilities resulting from a kinetically slower rate of alkylation with formation of a thermodynamically less stable, more reversible adduct.[6] The net result of the removal and substantial reduction of the size of the DNA binding subunits is an intrinsically less efficient alkylation of duplex DNA. Herein, we report that introduction of a trimethylammonium substituent on to the peripheral face of the agent **2**, which provides a spatially well-defined stabilizing electrostatic interaction, substantially enhances the DNA alkylation efficiency of such derivatives and provides agents that behave comparably to CC-1065 and its more advanced analogs.



Synthesis of 2–5

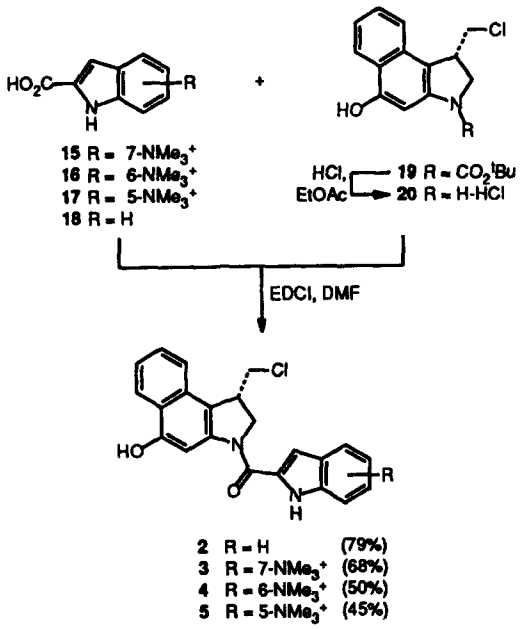
Catalytic hydrogenation of the readily available methyl 7-, 6-, and 5-nitroindole-2-carboxylates **6–8**[31] (1 atm H_2, 0.1 wt equiv. 10% Pd/C, EtOAc, 25 °C, 4–5 h) provided corresponding amines[31] **9** (79%), **10** (76%), and **11** (92%), respectively (Scheme 1). Exhaustive methylation of **9–11** (10 equiv. CH_3I, 5 equiv. $NaHCO_3$, DMF, 25 °C, 4–7 h) cleanly provided the trimethylammonium iodides **12–14** (80–90%) which were conveniently purified by simple recrystallization (50–90% recovery) from CH_3CN. Subsequent hydrolysis of the methyl esters **12–14** (3 equiv. LiOH, THF:CH_3OH:H_2O 3:1:1, 25 °C, 6 h) followed by coupling of the resultant carboxylic acids **15–17**, which were purified by simple recrystallization from CH_3CN, with the indoline hydrochloride **20** (2 equiv. EDCI, DMF, 25 °C, 12 h) deliberately conducted in the absence of added base provided the *seco*-CBI-indole-NMe_3^+ salts **3–5** in excellent conversions (68–45%, Scheme 2). Under these conditions no competitive closure of **20** or **3–5** to the corresponding cyclopropane was detected. Notably, **3–5** were readily purified by SiO_2 chromatography (BuOH:H_2O:EtOAc:HOAc 5:5:5:3 eluent) to provide pure samples of the final agents and no evidence of contaminant cyclopropane ring closure was observed upon reaction or purification. The water solubility of the agents **3–5** and their diminished solubility in conventional nonprotic solvents coupled with the rapid hydrolysis of the cyclopropane containing CBI agents generally observed under basic, protic conditions discouraged any attempt at ring closure of **3–5**.



Scheme 1.

Similarly, *seco*-CBI-indole$_1$ (**2**) was prepared by acid-catalyzed deprotection of **19** followed by coupling of the crude hydrochloride salt **20** with indole-2-carboxylic

acid (**18**) to provide **2** in excellent overall yield (79%, Scheme 2).



Scheme 2.

DNA alkylation properties of 2–5

The DNA alkylation properties of the agents were examined within two 150 base-pair segments of duplex DNA for which comparative results are available for related agents. Two clones of phage M13mp10 were selected for study and contain SV40 nucleosomal DNA inserts: w794 (nucleotide No. 5238-138) and its complement w836 (nucleotide No. 5189-91).[32] The alkylation site identification and the assessment of the relative selectivity among the available sites were obtained by thermally-induced strand cleavage of the singly 5' end-labeled duplex DNA after exposure to the agents. Following treatment of the end-labeled duplex DNA with a range of agent concentrations (24 h), the unbound agent was removed by EtOH precipitation of the DNA. Redissolution of the DNA in aqueous buffer, thermolysis (100 °C, 30 min) to induce strand cleavage at the sites of DNA alkylation, denaturing high resolution polyacrylamide gel electrophoresis (PAGE) adjacent to Sanger dideoxynucleotide sequencing standards,[33] and autoradiography led to identification of the DNA cleavage and alkylation sites. It has been established in our prior studies that rapid thermolysis at 100 °C at neutral to low pH (pH 6–7) and at low salt concentration minimize a potentially competitive reversible DNA alkylation reaction and optimize thermal depurination leading to stoichiometric strand cleavage at the sites of DNA alkylation.[6] Thermal depurination reactions conducted at lower temperatures (50–90 °C), at higher pH (8.0), and to a lesser extent at high ionic strength favor reversal of the DNA alkylation reaction and result in less efficient strand cleavage. This may be attributed to base-catalyzed phenol deprotonation required for retroalkylation observed at the higher pH or maintenance of duplex (favors retroalkylation) versus denatured (favors depurination) DNA observed at the lower reaction temperatures or at the higher ionic strength. The full details of this procedure have been disclosed and discussed elsewhere.[10] The DNA alkylation reaction selectivities observed under the incubation conditions of 4 °C (24 h) for the natural enantiomers of **1–5** have proven identical to the alkylation selectivities observed with shorter or extended reaction periods or when the reactions were conducted at higher temperatures (37 or 25 °C, 0.5–7 days).

The introduction of the trimethylammonium group onto the outer face of the indole resulted in a substantial increase in DNA alkylation efficiency (*ca* 100 ×) and this is nicely illustrated in the comparisons provided in Figure 1. Each of the three agents **3–5** were found to alkylate DNA with nearly identical efficiencies as (+)-CC-1065 (**1**) and (+)-duocarmycin SA (10^{-7} M alkylation) and substantially more efficiently than (+)-CBI-indole$_1$ (**2**, 10^{-5} M alkylation). This increase in DNA alkylation efficiency of **3** and **4** versus **2** (100 ×) may be attributed to the strategic placement of the trimethylammonium group on the outer face of DNA–agent complex between the walls of minor groove and proximal to the phosphates in the DNA backbone, Figures 2 and 3. This proximal placement of the positively charged trimethylammonium group next to the backbone negatively charged phosphates provides additional and substantial stabilizing electrostatic binding interactions enhancing the relative efficiencies of DNA alkylation.[29] Importantly, this increased alkylation efficiency provided by the enhanced binding affinity derived from the stabilizing electrostatic interactions is of such a magnitude that it overrides the effect of the increased water solubility of the agents that would be expected to diminish their inherent hydrophobic binding affinity. Notably, this occurs without altering the expected DNA alkylation selectivity for **3** and **4** which proved identical to (+)-CBI-indole$_1$ (**2**). Like the distinctions previously noted between (+)-duocarmycin SA and (+)-CC-1065 (**1**) where a 3.5 versus five base-pair AT-rich adenine N3 alkylation selectivity was observed,[6] the agents **2–4** exhibited a slightly smaller and expected three base-pair AT-rich adenine N3 alkylation selectivity (i.e. 5'-AAA and 5'-TTA). This is nicely illustrated in Figure 1 where **2–4** more prominently alkylate the 5'-ACTAA site than (+)-duocarmycin SA and (+)-CC-1065 (**1**). This diminished alkylation of the 5'-ACTAA site by (+)-CC-1065 (**1**) and (+)-duocarmycin SA may be attributed to their increasingly strong preference but not absolute requirement that the fourth base be A or T in order to accommodate groove binding of the agent within the appropriately sized AT- rich binding site (5 and 3.5 bp, respectively) surrounding the alkylation site.

Interestingly, the agent **5** exhibited an altered and more discriminating alkylation selectivity than **2–4** that may be attributed to the placement of the quaternary trimethylammonium salt within the minor groove in contrast to its placement on the peripheral face as with **3** and **4**. Initially we had expected that this non-ideal

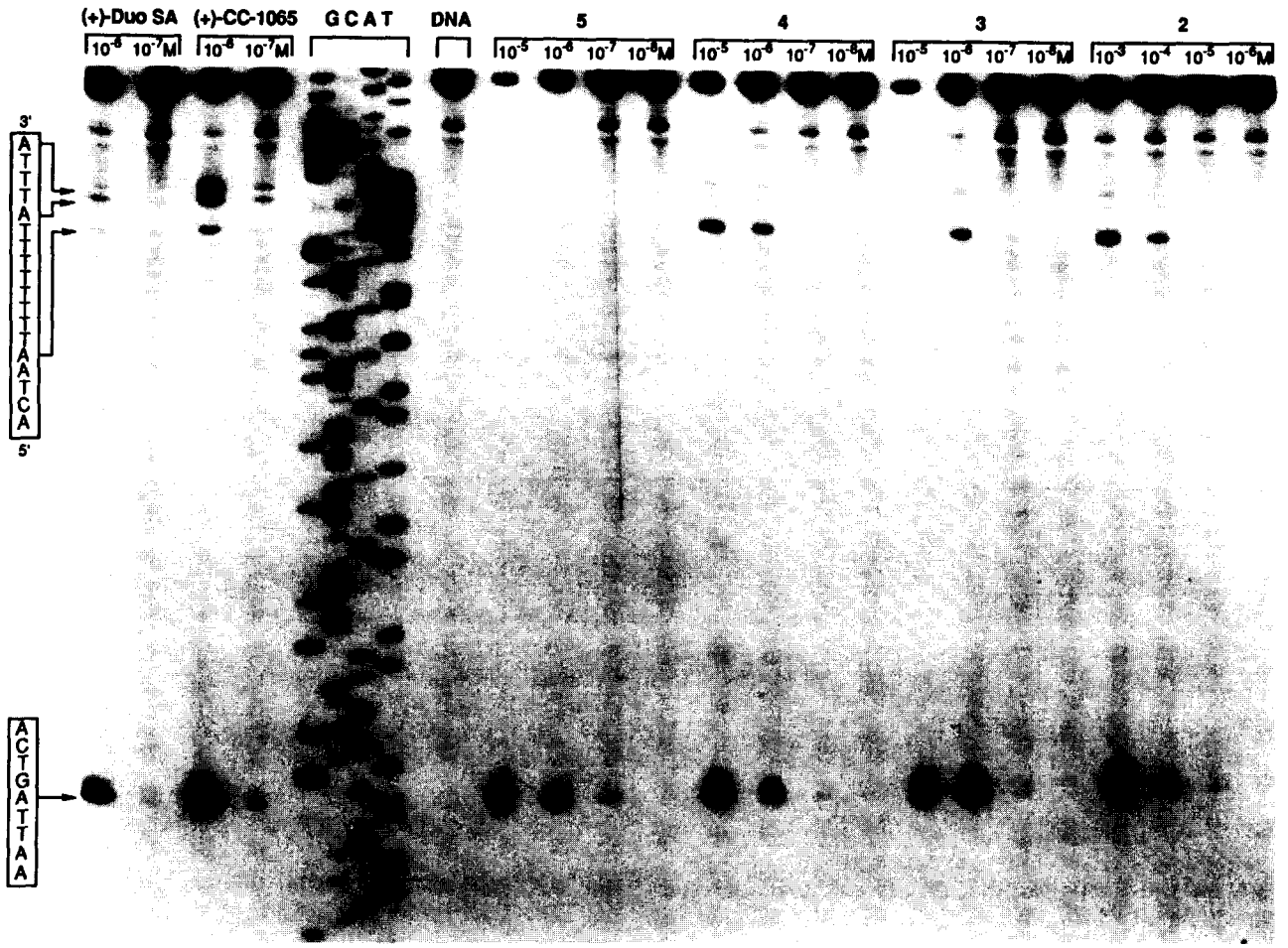


Figure 1. Thermally-induced strand cleavage of w794 DNA (144 bp, nucleotide No. 138-5238) after DNA–agent incubation at 4 °C (24 h), removal of unbound agent by EtOH precipitation and 30 min thermolysis (100 °C), followed by denaturing 8% PAGE and autoradiography. Lanes 1–2, (+)-duocarmycin SA (1×10^{-6} and 1×10^{-7} M); lanes 3–4, (+)-CC-1065 (1×10^{-6} and 1×10^{-7} M); lanes 5–8, Sanger G, C, A, and T sequencing standards; lane 9, control DNA; lanes 10–13, **5** (1×10^{-5} M to 1×10^{-8} M); lanes 14–17, **4** (1×10^{-5} to 1×10^{-8} M); lanes 18–21, **3** (1×10^{-5} to 1×10^{-8} M); lanes 22–25, **2** (1×10^{-3} to 1×10^{-6} M).

placement of the quaternary trimethylammonium salt might alter the DNA alkylation selectivity and perhaps diminish the DNA alkylation efficiency. This latter concern proved unfounded and **5**, like **3** and **4**, proved to be 100 × more efficient than **2** and equivalent to (+)-CC-1065. However, the incorporation of the trimethylammonium salt at the 5-position of the indole did alter the DNA alkylation selectivity. This is nicely illustrated in Figure 1 where **5**, like **1–4**, alkylates the high affinity 5'-AATTA site but fails to alkylate the lower affinity 5'-ACTAA site. Presumably, this may be attributed to the placement of large quaternary trimethylammonium salt in the minor groove proximal to the third and fourth 5' bases from the alkylation site. For the minor 5'-ACTAA site, this places the large quaternary trimethylammonium salt adjacent to the C2-amine of guanine at the fourth CG base-pair of this alkylation sequence and results in a destabilizing steric interaction, Figures 4 and 5. Presumably, this destabilizing interaction is of such a significant magnitude that it precludes alkylation at this site. This is most clearly seen in the space-filling model of Figure 4 and the expanded stick model of Figure 5 where one methyl group of the trimethylammonium salt sterically eclipses the guanine C2-amine precluding full penetration of the agent into the minor groove at this site.

Conclusions

Thus, the introduction of the trimethylammonium substituents into **3–5** substantially increased their DNA alkylation efficiencies relative to **2** (100 ×) and provided agents comparable to (+)-CC-1065. For **3** and **4**, this is the result of spatially well-defined stabilizing electrostatic interactions between the positively charged trimethylammonium salt lying on the peripheral face of the agents and the negatively charged phosphates located in the DNA backbone that enhance noncovalent binding affinity without altering DNA binding or alkylation selectivity. In contrast, the agent **5** exhibited an altered and more discriminating AT-rich adenine N3 alkylation selectivity that we attribute to the groove placement of the large trimethylammonium substituent. Further studies on these and related water soluble analogs of (+)-CC-1065 are in progress and will be reported in due course.

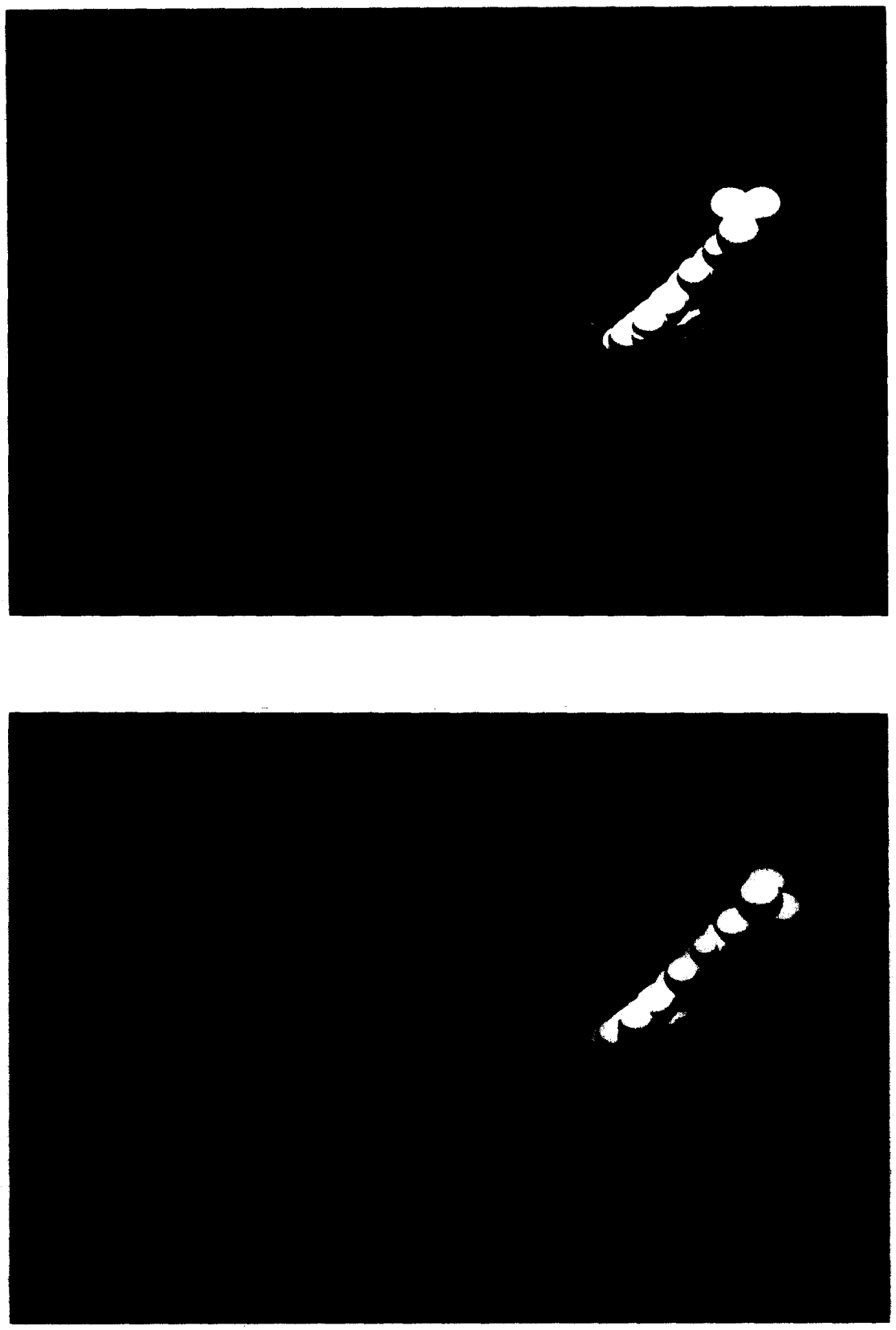


Figure 2. Stick and space filling models of the **3** (top) and **4** (bottom) alkylation of the 5'-GACTAATTTTT site of w794 DNA with agent binding across the 3 base AT-rich 5'-TAA site.

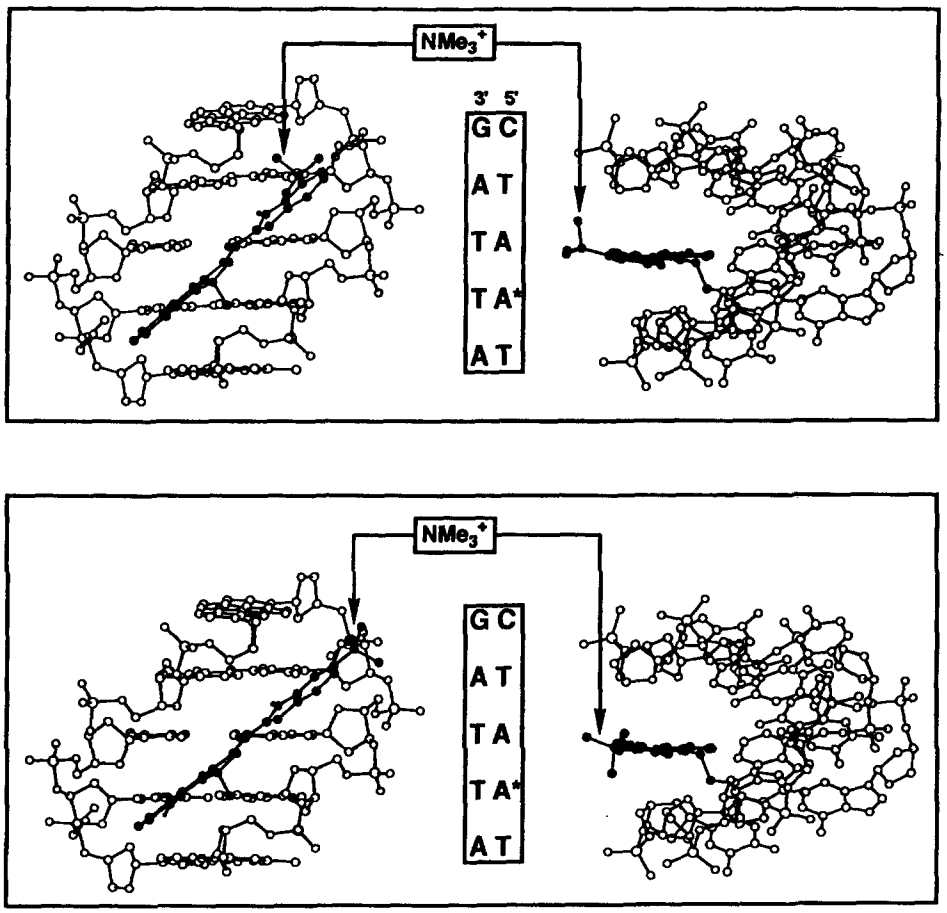


Figure 3. Expanded and rotated views of the alkylation of **3** (top) and **4** (bottom) at the 5'-CTAAT site of w794 DNA illustrating the placement of the trimethylammonium salt.

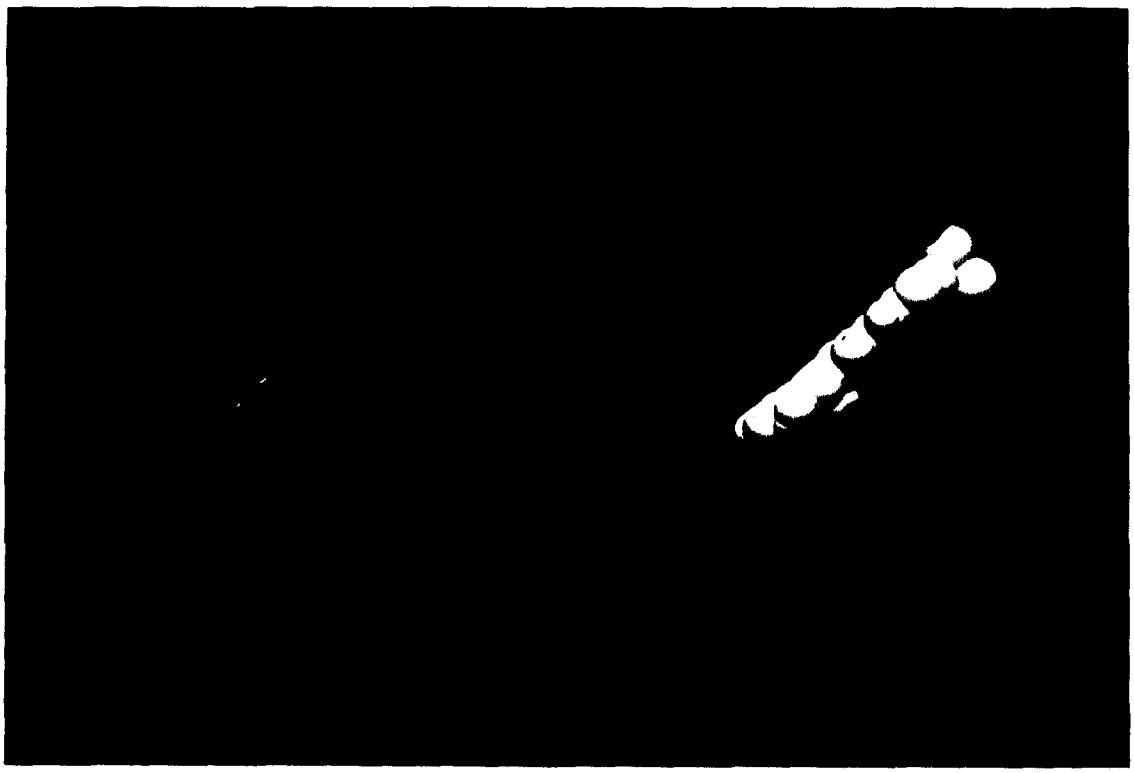


Figure 4. Stick and space filling models of the hypothetical alkylation of the 5'-GACTAATTTTT site in w794 DNA by **5**.

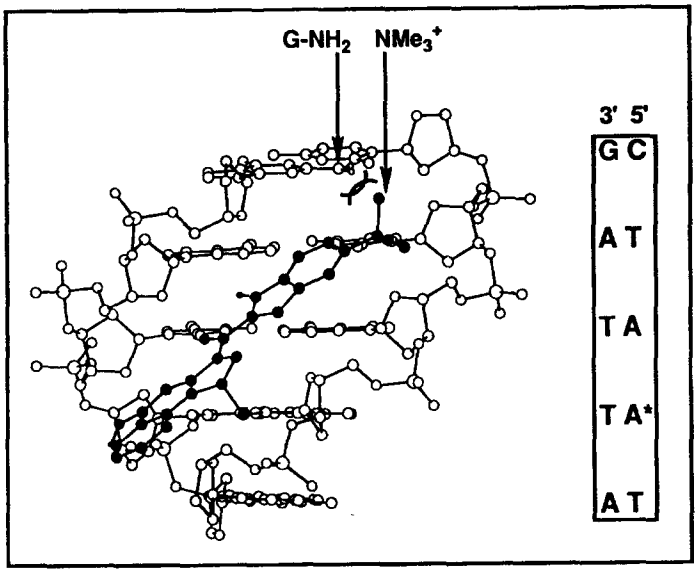


Figure 5. Expanded stick model of the hypothetical alkylation of the 5'-GACTAATTTTT site in w794 DNA by **5** illustrating the groove location of the quaternary trimethylammonium salt extending up to the fourth GC base-pair producing a destabilizing steric interaction with the G C2-amine precluding deep groove penetration and alkylation at this site.

Experimental

3-[(Indol-2'-yl)carbonyl]-1-(chloromethyl)-5-hydroxy-1,2-dihydro-3H-benz[e]indole (seco-*CBI-indole*$_1$, *2*)

A sample of **19** (4.0 mg, 0.012 mmol) was treated with anhydrous 4 M HCl–EtOAc (1 mL) at 25 °C for 30 min. The solvent was removed *in vacuo* to afford crude unstable **20** (quantitative). A mixture of **20**, [3-(dimethylamino)propyl]ethylcarbodiimide hydrochloride (EDCI, 5.8 mg, 0.030 mmol, 2.5 equiv.), and indole-2-carboxylic acid (**18**, 2.9 mg, 0.018 mmol, 1.5 equiv.) in 0.2 mL of DMF was stirred at 25 °C under Ar for 16 h. The mixture was diluted with 0.3 mL of H_2O and extracted with EtOAc (0.4 mL × 4). The combined organic layer was concentrated. Chromatography (SiO_2, 40% EtOAc–hexane) afforded **2** (3.4 mg, 4.3 mg theoretical, 79%) as a pale yellow solid: ^{1}H NMR (THF-d_8, 400 MHz) δ 11.04 (*br s*, 1H, NH), 9.31 (*s*, 1H, OH), 8.21 (*d*, 1H, *J* = 8.3 Hz, C6-H), 8.02 (*br s*, 1H, C4-H), 7.78 (*d*, 1H, *J* = 8.3 Hz, C9-H), 7.67 (*d*, 1H, *J* = 7.9 Hz, C4'-H), 7.48 (*dd*, 1H, C8-H partially obscured by overlapping C7'-H), 7.47 (*d*, 1H, *J* = 8.3 Hz, C7'- H), 7.30 (*dd*, 1H, *J* = 8.0, 8.3 Hz, C7-H), 7.22 (*dd*, 1H, *J* = 7.1, 8.3 Hz, C6'-H), 7.17 (*s*, 1H, C3'-H), 7.06 (*dd*, 1H, *J* = 7.1, 7.9 Hz, C5'-H), 4.78 (*m*, 2H, C2-H_2), 4.17 (*m*, 1H, C1-H), 4.00 (*dd*, 1H, *J* = 3.2, 11.1 Hz, C<u>H</u>HCl), 3.61 (*m*, 1H, CH<u>H</u>Cl); IR (film) ν_{max} 3427, 3225, 3056, 2965, 2865, 1608, 1578, 1512, 1417, 1394, 1363, 1338, 1316, 1252, 1140, 1058, 850, 804, 743 cm^{-1}; FAB-HRMS (NBA) m/z 377.1065 (M^+ + H, $C_{22}H_{17}ClN_2O_2$ requires 377.1057). Natural (1*S*)-**2**: $[\alpha]_D^{23}$ +8.8 (*c* 0.17, THF).

General method for the preparation of trimethylammonium substituted indole-2-carboxylate methyl esters: methyl 5-(trimethylammonio)indole-2-carboxylate iodide (14)

Compound **11**[31] (76 mg, 0.4 mmol) was dissolved in DMF (3 mL) and treated with $NaHCO_3$ (168 mg, 2.0 mmol, 5 equiv.) and CH_3I (568 mg, 248 μL, 4.0 mmol, 10 equiv.). The reaction mixture was stirred at 24 °C under N_2 for 4 h before the solvent was removed *in vacuo*. The dry residue was slurried in H_2O and precipitate was collected by centrifugation. Recrystallization from CH_3CN afforded **14** (129 mg, 144 mg theoretical, 90%) as a pale yellow solid: mp 228 °C (dec); ^{1}H NMR (DMSO-d_6, 400 MHz) δ 12.37 (*br s*, 1H, NH), 8.26 (*d*, 1H, *J* = 2.6 Hz, C4-H), 7.89 (*dd*, 1H, *J* = 2.6, 9.3 Hz, C6-H), 7.64 (*d*, 1H, *J* = 9.3 Hz, C7-H), 7.29 (*s*, 1H, C3-H), 3.86 (*s*, 3H, CO_2CH_3), 3.65 (*s*, 9H, $N(CH_3)_3$); ^{13}C NMR (DMSO-d_6, 100 MHz) δ 161.4 (C), 140.6 (C), 136.4 (C), 129.7 (C), 125.9 (C), 116.9 (CH), 114.1 (CH), 113.9 (CH), 108.9 (CH), 56.9 (three CH_3), 52.2 (CH_3); IR (film) ν_{max} 3446, 1708, 1537, 1437, 1339, 1259, 1205, 995, 937, 770, 742 cm^{-1}; FABHRMS (NBA) m/z 233.1290 (M^+ – I, $C_{13}H_{17}IN_2O_2$ requires 233.1290). Anal. calcd for $C_{13}H_{17}IN_2O_2$: C, 43.35; H, 4.76; N, 7.78. Found: C, 42.99; H, 4.62; N, 7.51.

Methyl 7-(trimethylammonio)indole-2-carboxylate iodide (12)

Mp 151.5 °C (dec., pale green fine needles); ^{1}H NMR (CD_3CN, 400 MHz) δ 10.27 (*br s*, 1H, NH), 8.07 (*d*, 1H, *J* = 8.0 Hz, C4-H), 7.81 (*d*, 1H, *J* = 8.0 Hz, C6-H), 7.53 (*s*, 1H, C3-H), 7.41 (*t*, 1H, *J* = 8.0 Hz, C5-H), 4.05 (*s*, 3H, CO_2CH_3), 3.89 (*s*, 9H, $N(CH_3)_3$); ^{13}C NMR (CD_3OD, 100 MHz) δ 162.8 (C), 133.5 (C), 133.2 (C), 131.5 (C), 128.0 (C), 127.1 (CH), 121.7 (CH), 117.8 (CH), 111.2 (CH), 56.7 (three CH_3), 52.8 (CH_3); IR (film) ν_{max} 3188, 1717, 1614, 1467, 1438, 1306, 1254, 1204, 1149, 944, 833, 731 cm^{-1}; FABHRMS (NBA) m/z 233.1297 (M^+ – I, $C_{13}H_{17}IN_2O_2$ requires 233.1290). Anal. calcd for $C_{13}H_{17}IN_2O_2$: C, 43.35; H, 4.76; N, 7.78. Found: C, 43.37; H, 4.73; N, 7.78.

Methyl 6-(trimethylammonio)indole-2-carboxylate iodide (13)

Mp 209 °C (dec., colorless crystals); ^{1}H NMR (DMSO-d_6, 400 MHz) δ 12.48 (*br s*, 1H, NH), 7.91 (*d*, 1H, *J* = 9.0 Hz, C4-H), 7.80 (*d*, 1H, *J* = 2.0 Hz, C7-H), 7.75 (*dd*, 1H, *J* = 2.1, 9.1 Hz, C5-H), 7.25 (*s*, 1H, C3-H), 3.86 (*s*, 3H, CO_2CH_3), 3.68 (*s*, 9H, $N(CH_3)_3$); ^{13}C NMR (DMSO-d_6, 100 MHz) δ 153.5 (C), 124.2 (C), 123.9 (C), 122.2 (C), 118.7 (C), 117.8 (CH), 112.4 (CH), 108.5 (CH), 101.9 (CH), 47.4 (three CH_3), 43.5 (CH_3); IR (film) ν_{max} 3409, 1716, 1605, 1564, 1489, 1433, 1325, 1226, 1005, 942 cm^{-1}; FABHRMS (NBA) m/z 233.1290 (M^+ – I, $C_{13}H_{17}IN_2O_2$ requires 233.1290). Anal. calcd for $C_{13}H_{17}IN_2O_2$: C, 43.35; H, 4.76; N, 7.78. Found: C, 43.36; H, 4.72; N, 7.81.

General method for the preparation of trimethylammonium substituted indole-2-carboxylic acids: 5-(trimethylammonio)indole-2-carboxylic acid (17)

A solution of **14** (100 mg, 0.28 mmol) in THF:CH_3OH:H_2O (3:1:1, 2.6 mL) was treated with LiOH–H_2O (35 mg, 0.83 mmol, 3 equiv.), and the reaction mixture was stirred at 24 °C for 6 h. The solvent was removed and the dry residue was mixed with H_2O (10 mL) and saturated aqueous NaCl (5 mL). The solution was acidified to pH 1 with the addition of 1 N aqueous HCl and extracted with CH_3CN (10 mL each) until no UV active material was detected in aqueous solution. The extracts were combined, dried (Na_2SO_4) and concentrated. Recrystallization from CH_3CN afforded **17** (73.8 mg, 96.2 mg theoretical, 77%) as a pale yellow solid: ^{1}H NMR (DMSO-d_6, 400 MHz) δ 13.31 (*br s*, 1H, CO_2H), 12.19 (*br s*, 1H, NH), 8.23 (*d*, 1H, *J* = 1.8 Hz, C4-H), 7.88 (*d*, 1H, *J* = 9.2 Hz, C6-H or C7-H), 7.59 (*d*, 1H, *J* = 9.2 Hz, C6-H or C7-H), 7.19 (*d*, 1H, *J* = 1.1 Hz, C3-H), 3.65 (*s*, 9H, $N(CH_3)_3$); ^{13}C NMR (DMSO-d_6, 100 MHz) δ 162.4 (C), 140.5 (C), 136.3 (C), 131.3 (C), 126.1 (C), 116.5 (CH), 133.8 (two CH), 108.2 (CH), 56.8 (three CH_3); IR (film) ν_{max} 3342, 3016, 1697, 1538, 1469, 1419, 1339, 1226, 1194, 938, 852, 772 cm^{-1}; FABHRMS (NBA) m/z 219.1143 (M^+ – Cl, $C_{12}H_{15}ClN_2O_2$ requires 219.1134.)

7-(Trimethylammonio)indole-2-carboxylic acid (15)

Mp > 198 °C (dec., white solid); ^{1}H NMR (DMSO-d_6, 400 MHz) δ 13.41 (*br s*, 1H, CO_2H), 12.23 (*br s*, 1H, NH), 7.94 (*d*, 1H, *J* = 7.9 Hz, C4-H), 7.74 (*d*, 1H, *J* = 8.0 Hz, C6-H), 7.38 (*s*, 1H, C3-H), 7.26 (*t*, 1H, *J* = 8.0 Hz, C5-H), 3.79 (*s*, 9H, $N(CH_3)_3$); ^{13}C NMR (DMSO-d_6, 100 MHz) δ 162.2 (C), 132.2 (C), 131.5 (C), 131.3 (C), 126.6 (C), 125.3 (CH), 120.1 (CH), 116.9 (CH), 109.5 (CH), 55.5 (three CH_3); IR (film) ν_{max} 3327, 1694, 1477, 1444, 1416, 1388, 1328, 1173, 1141, 940, 730 cm^{-1}; FABHRMS (NBA) m/z 219.1141 (M^+ – Cl, $C_{12}H_{15}ClN_2O_2$ requires 219.1134).

6-(Trimethylammonio)indole-2-carboxylic acid (16)

Mp > 195 °C (dec., off-white needles); ^{1}H NMR (DMSO-d_6, 400 MHz) δ 13.12 (*br s*, 1H, CO_2H), 12.33 (*br s*, 1H, NH), 7.88 (*d*, 1H, *J* = 9.0 Hz, C4-H), 7.81 (*d*, 1H, *J* = 2.2 Hz, C7-H), 7.73 (*d*, 1H, *J* = 2.2, 9.0 Hz, C5-H), 7.17 (*d*, 1H, *J* = 1.7 Hz, C3-H), 3.66 (*s*, 9H, $N(CH_3)_3$); ^{13}C NMR (DMSO-d_6, 100 MHz) δ 162.3 (C), 143.7 (C), 135.7 (C), 131.6 (C), 126.9 (C), 123.6 (CH), 112.5 (CH), 107.0 (CH), 104.5 (CH), 56.5 (three CH_3); IR (film) ν_{max} 3260, 1689, 1530, 1328, 1222, 1131, 835, 778 cm^{-1}; FABHRMS (NBA) m/z 219.1142 (M^+ – Cl, $C_{12}H_{15}ClN_2O_2$ requires 219.1134).

General method for the coupling of seco-N-*BOC-CBI (19) with 15–17: 3-[7'-((trimethylammonio)indol-2'-yl)carbonyl]-1-(chloromethyl)-5-hydroxy-1,2-dihydro-3*H*-benz[e]indole (3)*

Phenol **19**[20] (9.0 mg, 0.027 mmol) was treated with anhydrous 3 M HCl–EtOAc (2 mL) at 24 °C for 30 min. The solvent was removed *in vacuo* to afford crude unstable **20** (quantitative). A mixture of **20**, [3-(dimethylamino)propyl]ethylcarbodiimide hydrochloride (EDCI, 10.3 mg, 0.054 mmol, 2.0 equiv.), and **15** (9.3 mg, 0.027 mmol, 1.0 equiv.) in 0.5 mL of DMF was stirred at 24 °C under Ar for 12 h. The solvent was removed *in vacuo* and the dry residue was mixed with H_2O (3 mL) and saturated aqueous NaCl (2 mL). The mixture was extracted with CH_3CN (5 mL × 3). The organic layer was dried (Na_2SO_4) and concentrated. Chromatography (SiO_2, *n*-BuOH:H_2O:EtOAc:HOAc, 5:5:5:3) afforded **3** (10.3 mg, 15.2 mg theoretical, 68%) as a pale yellow solid: mp > 152 °C (dec.); ^{1}H NMR (CD_3OD, 400 MHz) δ 8.22 (*d*, 1H, *J* = 8.3 Hz, C6-H), 8.02 (*d*, 1H, *J* = 7.9 Hz, C4'-H), 7.97 (*br s*, 1H, C4-H), 7.81 (*d*, 1H, *J* = 8.6 Hz, C6'-H or C9-H), 7.79 (*d*, 1H, *J* = 8.4 Hz, C6'-H or C9-H), 7.55 (*t*, 1H, *J* = 8.2 Hz, C8-H), 7.43 (*s*, 1H, C3'-H), 7.39 (*t*, 1H, *J* = 8.2 Hz, C7-H), 7.35 (*t*, 1H, *J* = 8.0 Hz, C5'-H), 4.73–4.77 (*m*, 1H, C2-H), 4.65 (*dd*, *J* = 1.7, 11.0 Hz, C2-H), 4.17–4.21 (*m*, 1H, C1-H), 4.00 (*dd*, 1H, *J* = 3.1, 11.2 Hz, C<u>H</u>HCl), 3.90 (*s*, 9H, $N(CH_3)_3$), 3.69 (apparent *t*, 1H, *J* = 10.6 Hz, CH<u>H</u>Cl); IR (film) ν_{max} 3354, 1624, 1584, 1466, 1414, 1326, 1259 cm^{-1}; FABHRMS (NBA) m/z 434.1648 (M^+ – Cl, $C_{25}H_{25}Cl_2N_3O_2$ requires 434.1635). Natural (1*S*)-**3**: $[\alpha]_D^{23}$ –9.9 (*c* 0.10, CH_3OH).

*3-[6'-((Trimethylammonio)indol-2'-yl)carbonyl]-1-(chloromethyl)-5-hydroxy-1,2-dihydro-3*H*-benz[e]indole (4)*

^{1}H NMR (CD_3OD, 400 MHz) δ 8.21 (*d*, 1H, *J* = 8.3 Hz, C6-H), 7.99 (*d*, 1H, *J* = 8.8 Hz, C4'-H), 7.87 (*br s*, 1H, C4-H), 7.80 (*d*, 1H, *J* = 8.3 Hz, C9-H), 7.65 (*dd*, 1H, *J* = 2.3, 9.3 Hz, C5'-H), 7.64 (*s*, 1H, C7'-H), 7.54 (*t*, 1H, *J* = 8.2 Hz, C8-H), 7.36-7.45 (*m*, 1H, C7- H), 7.30 (*s*, 1H, C3'-H), 4.75–4.82 (*m*, 1H, C2-H), 4.70 (*dd*, 1H, *J* = 1.8, 10.9 Hz, C2-H), 4.19–4.23 (*m*, 1H, C1-H), 4.00 (*dd*, 1H, *J* = 3.2, 11.2 Hz, C<u>H</u>HCl), 3.75 (*s*, 9H, $N(CH_3)_3$), 3.69 (*dd*, 1H, *J* = 3.0, 11.2 Hz, CH<u>H</u>Cl); IR (film) ν_{max} 3373, 1625, 1577, 1558, 1519, 1409, 1342, 1256 cm^{-1}; FABHRMS (NBA) m/z 434.1722 (M^+ – Cl, $C_{25}H_{25}Cl_2N_3O_2$ requires 434.1714). Natural (1*S*)-**4**: $[\alpha]_D^{23}$ +53 (*c* 0.04, CH_3OH).

3-[5'-((Trimethylammonio)indol-2'-yl)carbonyl]-1-(chloromethyl)-5-hydroxy-1,2-dihydro-3H-benz[e]indole (5)

^{1}H NMR (CD_3OD, 400 MHz) δ 8.27 (*d*, 1H, *J* = 2.6 Hz, C4'-H), 8.21 (*d*, 1H, *J* = 8.3 Hz, C6-H), 7.83 (*br s*, 1H, C4-H), 7.81 (*dd*, 1H, *J* = 2.8, 9.3 Hz, C6'-H), 7.80 (*d*, 1H, *J* = 8.3 Hz, C9-H), 7.73 (*d*, 1H, *J* = 9.2 Hz, C7'-H), 7.53 (*t*, 1H, *J* = 8.2 Hz, C8-H), 7.37 (*t*, 1H, *J* = 8.4 Hz, C7-H), 7.33 (*s*, 1H, C3'-H), 4.68–4.76 (*m*, 2H, partially obscured by H_2O, C2-H_2), 4.17–4.21 (*m*, 1H, C1-H), 3.99 (*dd*, 1H, *J* = 3.2, 11.2 Hz, CHHCl), 3.73 (*s*, 9H, $N(CH_3)_3$), 3.65 (*dd*, 1H, *J* = 8.8, 11.2 Hz, CHHCl); IR (film) ν_{max} 3374, 1557, 1416, 1342, 1265, 1232, 758 cm^{-1}; FABHRMS (NBA) m/z 434.1619 (M^+ – Cl, $C_{25}H_{25}Cl_2N_3O_2$ requires 434.1635). Natural (1*S*)-**5**: $[\alpha]_D^{23}$ +64 (*c* 0.10, CH_3OH).

DNA alkylation studies of 2–5: selectivity and efficiency

Eppendorf tubes containing singly ^{32}P 5'-end-labeled double-stranded DNA[10] (9 μL) in TE buffer (10 mM Tris, 1 mM EDTA, pH 7.5) were treated with the agents **1–5** in DMSO (1 μL, at the specified concentrations). The solutions were mixed by vortexing and brief centrifugation and subsequently incubated at 4 °C for 24 h. The covalently modified DNA was separated from unbound agent by EtOH precipitation of the DNA. The EtOH precipitations were carried out by adding tRNA as a carrier (1 μL, 10 μg μL^{-1}), a buffer solution containing salt (0.1 volume, 3 M NaOAc in TE) and –20 °C EtOH (2.5 volumes). The solutions were mixed and chilled at –78 °C in a Revco freezer for 1 h or longer. The DNA was reduced to a pellet by centrifugation at 4 °C for 15 min, washed with –20 °C 70% EtOH (in TE containing 0.2 M NaCl) and recentrifuged briefly. The pellets were dried in a Savant Speed Vac concentrator and resuspended in TE buffer (10 μL). The solutions of alkylated DNA were warmed at 100 °C for 30 min to induce cleavage at the adenine N3 alkylation sites. After brief centrifugation, formamide dye solution (5 μL) was added. Prior to electrophoresis, the samples were denatured by warming at 100 °C for 5 min, placed in an ice bath, centrifuged briefly, and the supernatant (2.8 μL) was loaded on to a gel. Sanger dideoxynucleotide sequencing reactions were run as standards adjacent to the agent treated DNA reaction samples. Polyacrylamide gel electrophoresis (PAGE) was run on an 8% sequencing gel under denaturing conditions (19:1 acrylamide:*N,N'*-methylenebisacrylamide, 8 M urea) in TBE buffer (100 mM Tris, 100 mM boric acid, 0.2 mM Na_2EDTA). PAGE was pre-run for 30 min with formamide dye solution prior to loading the samples. Autoradiography of dried gels was carried out at –78 °C using Kodak X-Omat AR film and a Picker Spectra™ intensifying screen.

Acknowledgments

We gratefully acknowledge the financial support of the National Institutes of Health (CA 41986) and the award of an ACS Organic Division fellowship sponsored by Zeneca Pharmaceuticals (D.S.J., 1993–1994).

References and Notes

1. Hanka, L. J.; Dietz, A.; Gerpheide, S. A.; Kuentzel, S. L.; Martin, D. G. *J. Antibiot.* **1978**, *31*, 1211; Chidester, C. G.; Krueger, W. C.; Mizsak, S. A.; Duchamp, D. J.; Martin, D. G. *J. Am. Chem. Soc.* **1981**, *103*, 7629.

2. Takahashi, I.; Takahashi, K.; Ichimura, M.; Morimoto, M.; Asano, K.; Kawamoto, I.; Tomita, F.; Nakano, H. *J. Antibiot.* **1988**, *41*, 1915; Yasuzawa, T.; Iida, T.; Muroi, K.; Ichimura, M.; Takahashi, K.; Sano, H. *Chem. Pharm. Bull.* **1988**, *36*, 3728; Ichimura, M.; Muroi, K.; Asano, K.; Kawamoto, I.; Tomita, F.; Morimoto, M.; Nakano, H. *J. Antibiot.* **1988**, *41*, 1285.

3. 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.

4. 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.

5. 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; Boger, D. L.; Yun, W.; Terashima, S.; Fukuda, Y.; Nakatani, K.; Kitos, P. A.; Jin, Q. *Bioorg. Med. Chem. Lett.* **1992**, *2*, 759.

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

7. 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; Asai, A.; Nagamura, S.; Saito, H. *J. Am. Chem. Soc.* **1994**, *116*, 4171.

8. Hurley, L. H.; Reynolds, V. L.; Swenson, D. H.; Petzold, G. L.; Scahill, T. A. *Science* **1984**, *226*, 843; Reynolds, V. L.; Molineaux, I. J.; Kaplan, D. J.; Swenson, D. H.; Hurley, L. H. *Biochemistry* **1985**, *24*, 6228; Hurley, L. H.; Lee, C.-S.; McGovren, J. P.; Warpehoski, M. A.; Mitchell, M. A.; Kelly, R. C.; Aristoff, P. A. *Biochemistry* **1988**, *27*, 3886; 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. Boger, D. L.; Johnson, D. S.; Yun, W.; Tarby, C. M. *Bioorg. Med. Chem.* **1994**, *2*, 115; 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.

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

11. Boger, D. L. *Acc. Chem. Res.* **1995**, *28*, 20; Boger, D. L. *Chemtracts: Org. Chem.* **1991**, *4*, 329; Boger, D. L. *Proceed. R. A. Welch Found. Conf. on Chem. Res., XXXV. Chem. at the Frontiers of Medicine* **1991**, *35*, 137; Boger, D. L. In: *Advances in Heterocyclic Natural Products Synthesis*, Vol. 2, pp. 1–188, Pearson, W. H., Ed.; JAI Press; Greenwich, CT, 1992; Boger, D. L. *Pure Appl. Chem.* **1993**, *65*, 1123; Boger, D. L. *Pure Appl. Chem.* **1994**, *66*, 837.

12. Warpehoski, M. A. In: *Advances in DNA Sequence Specific Agents*, Vol. 1, p. 217, Hurley, L. H., Ed.; JAI Press; Greenwich, CT, 1992; Warpehoski, M. A.; Hurley, L. H. *Chem. Res. Toxicol.* **1988**, *1*, 315; 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, MI, 1993; Hurley, L. H.; Needham-VanDevanter, D. R. *Acc. Chem. Res.* **1986**, *19*, 230.

13. Coleman, R. S.; Boger, D. L. In: *Studies in Nat. Prod. Chem.*, Vol. 3, p. 301 Atta-ur-Rahman, Ed.; Elsevier; Amsterdam, 1989; 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.; Johnson, D. S.; Wrasidlo, W. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 631.

15. Wierenga, W. *J. Am. Chem. Soc.* **1981**, *103*, 5621; Magnus, P.; Gallagher, T.; Schultz, J.; Or, Y.-S.; Ananthanarayan, T. P. *J. Am. Chem. Soc.* **1987**, *109*, 2706; Kraus, G. A.; Yue, S.; Sy, J. *J. Org. Chem.* **1985**, *50*, 283; Boger, D. L.; Coleman, R. S. *J. Am. Chem. Soc.* **1988**, *110*, 1321, 4796; Boger, D. L.; Coleman, R. S. *J. Org. Chem.* **1988**, *53*, 695; Bolton, R. E.; Moody, C. J.; Pass, M.; Rees, C. W.; Tojo, G. *J. Chem. Soc., Perkin Trans. 1* **1988**, 2491; Sundberg, R. J.; Baxter, E. W.; Pitts, W. J.; Ahmed-Schofield, R.; Nishiguchi, T. *J. Org. Chem.* **1988**, *53*, 5097; Sundberg, R. J.; Pitts, W. J. *J. Org. Chem.* **1991**, *56*, 3048; Martin, V. P. *Helv. Chim. Acta* **1989**, *72*, 1554; Toyota, M.; Fukumoto, K. *J. Chem. Soc., Perkin Trans. 1* **1992**, 547; Tietze, L. F.; Grote, T. *J. Org. Chem.* **1994**, *59*, 192.

16. 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; Muratake, H.; Abe, I.; Natsume, M. *Tetrahedron Lett.* **1994**, *35*, 2573.

17. Fukuda, Y.; Itoh, Y.; Nakatani, K.; Terashima, S. *Tetrahedron* **1994**, *50*, 2793; Fukuda, Y.; Nakatani, K.; Terashima, S. *Tetrahedron* **1994**, *50*, 2809; Fukuda, Y.; Nakatani, K.; Terashima, S. *Bioorg. Med. Chem. Lett.* **1992**, *2*, 755; Fukuda, Y.; Nakatani, K.; Ito, Y.; Terashima, S. *Tetrahedron Lett.* **1990**, *31*, 6699.

18. 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.

19. CI-based analogs: Boger, D. L.; Zarrinmayeh, H.; Munk, S. A.; Kitos, P. A.; Suntornwat, O. *Proc. Natl Acad. Sci. USA* **1991**, *88*, 1431; 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; Sundberg, R. J.; Baxter, E. W. *Tetrahedron Lett.* **1986**, *27*, 2687; Wang, Y.; Lown, J. W. *Heterocycles* **1993**, *36*, 1399; Wang, Y.; Gupta, R.; Huang, L.; Lown, J. W. *J. Med. Chem.* **1993**, *36*, 4172; Tietze, L. F.; Grote, T. *Chem. Ber.* **1993**, *126*, 2733; see also Refs 5 and 10.

20. CBI-based analogs: Boger, D. L.; Munk, S. A. *J. Am. Chem. Soc.* **1992**, *114*, 5487; Boger, D. L.; Yun, W. *J. Am. Chem. Soc.* **1994**, *116*, 7996; Boger, D. L.; Munk, S. A.; Ishizaki, T. *J. Am. Chem. Soc.* **1991**, *113*, 2779; Boger, D. L.; Yun, W. *J. Am. Chem. Soc.* **1994**, *116*, 5523; 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.; Sakya, S. M.; Munk, S. A.; Kitos, P. A.; Jin, Q.; Besterman, J. M. *Bioorg. Med. Chem. Lett.* **1991**, *1*, 115; Boger, D. L.; Yun, W.; Teegarden, B. R. *J. Org. Chem.* **1992**, *57*, 2873; Drost, K. J.; Cava, M. P. *J. Org. Chem.* **1991**, *56*, 2240; 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.

21. 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. *Bioorg. Med. Chem.* **1993**, *1*, 27.

22. CBQ-based analogs: Boger, D. L.; Mesini, P.; Tarby, C. M. *J. Am. Chem. Soc.* **1994**, *116*, 6461; Boger, D. L.; Mesini, P. *J. Am. Chem. Soc.* **1994**, *116*, 11335.

23. CFI-based analogs: Mohamadi, F.; Spees, M. M.; Staten, G. S.; Marder, P.; Kipka, J. K.; Johnson, D. A.; Boger, D. L.; Zarrinmayeh, H. *J. Med. Chem.* **1994**, *37*, 232.

24. A *p*-quinonemethide analog: Boger, D. L.; Nishi, T.; Teegarden, B. R. *J. Org. Chem.* **1994**, *59*, 4943.

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

26. Warpehoski, M. A.; McGovren, J. P.; Mitchell, M. A. In: *Molecular Basis of Specificity in Nucleic Acid–Drug Interactions*, p. 531, Pullman, B.; Jortner, J., Eds; Kluwer; Netherlands, 1990.

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

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

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

30. Boger, D. L.; Ishizaki, T.; Zarrinmayeh, H.; Kitos, P. A.; Suntornwat, O. *Bioorg. Med. Chem. Lett.* **1991**, *1*, 55.

31. Condensation of 3-nitrobenzaldehyde with methyl 2-azidoacetate (8 equiv., 6 equiv. $NaOCH_3$, CH_3OH, –23 to 0 °C, 6 h, 88%) followed by thermolysis of the resulting methyl 2-azidocinnamate (xylene, reflux, 4.5 h, 81%) provided a readily separable mixture (4:1) of methyl 5- and 7-nitroindole-2-carboxylate. For methyl 7-nitroindole-2-carboxylate (**6**): mp 122–125 °C (CH_2Cl_2, light yellow fine needles); ^{1}H NMR ($CDCl_3$, 400 MHz) δ 10.37 (*br s*, 1H, NH), 8.31 (*d*, 1H, *J* = 8.0 Hz, C4-H), 8.06 (*d*, 1H, *J* = 8.0 Hz, C6-H), 7.36 (*d*, 1H, *J* = 2.4 Hz, C3-H), 7.28 (*t*, 1H, *J* = 8.0 Hz, C5-H), 3.99 (*s*, 3H, CO_2CH_3); IR (film) ν_{max} 3372, 1727, 1531, 1445, 1344, 1298, 1249, 1188, 1107, 830, 763 cm^{-1}; FABHRMS (NBA) m/z 221.0560 (M^+ + H, $C_{10}H_8N_2O_4$ requires 221.0562). For methyl 5-nitroindole-2-carboxylate (**8**): ^{1}H NMR (DMSO-d_6, 400 MHz) δ 12.65 (*br s*, 1H, NH), 8.73 (*d*, 1H, *J* = 2.3 Hz, C4-H), 8.14 (*dd*, 1H, *J* = 2.0, 8.0 Hz, C6-H), 7.60 (*d*, 1H, *J* = 8.0 Hz, C7-H), 7.45 (*d*, 1H, *J* = 0.7 Hz, C3-H), 3.90 (*s*, 3H, CO_2CH_3); IR (film) ν_{max} 3316, 1701, 1614, 1531, 1435, 1343, 1261, 1203, 992, 746 cm^{-1}. Similarly, condensation of 4-

nitrobenzaldehyde with methyl 2-azidoacetate (8 equiv., 6 equiv. $NaOCH_3$, CH_3OH, -23 to 0 °C, 7 h, 84%) followed by thermolysis (xylene, reflux, 4 h, 83%) provided methyl 6-nitroindole-2-carboxylate (**7**): ^1H NMR ($CDCl_3$, 400 MHz) δ 9.27 (*br s*, 1H, NH), 8.39 (*d*, 1H, *J* = 2.0 Hz, C7-H), 8.04 (*dd*, 1H, *J* = 2.0, 8.0 Hz, C5-H), 7.78 (*d*, 1H, *J* = 8.0 Hz, C4-H), 7.28 (*d*, 1H, *J* = 2.3 Hz, C3-H), 4.00 (*s*, 3H, CO_2CH_3). Catalytic hydrogenation of **6–8** (1 atm H_2, 0.1 wt equiv. 10% Pd/C, EtOAc, 25 °C, 4–5 h) provided the corresponding amines. For methyl 7-aminoindole-2-carboxylate (**9**): 79%; mp 184 °C (dec., pale green crystals); ^1H NMR ($CDCl_3$, 400 MHz) δ 9.47 (*br s*, 1H, NH), 7.21 (*s*, 1H, C3-H), 7.20 (*d*, 1H, *J* = 7.4 Hz, C6-H), 6.99 (*t*, 1H, *J* = 7.5 Hz, C5-H), 6.67 (*d*, 1H, *J* = 7.4 Hz, C4-H), 3.97 (*s*, 3H, CO_2CH_3), 2.30 (*br s*, 2H, NH_2); IR (film) ν_{max} 3205, 2815, 1693, 1547, 1437, 1345, 1247, 1211, 1112, 997, 827, 783, 734 cm^{-1}; FABHRMS (NBA) m/z 190.0747 ($M^+ + H$, $C_{10}H_{10}N_2O_2$ requires 190.0742). For methyl 6-aminoindole-2-carboxylate (**10**): 76%, ^1H NMR ($CDCl_3$, 400 MHz) δ 8.58 (*br s*, 1H, NH), 7.45 (*d*, 1H, *J* = 8.4 Hz, C4-H), 7.11 (*d*, 1H, *J* = 2.1 Hz, C3-H), 6.62 (*d*, 1H, *J* = 1.9 Hz, C7-H), 6.59 (*dd*, 1H, *J* = 1.9, 8.4 Hz, C5-H), 3.89 (*s*, 3H, CO_2CH_3), 3.79 (*br s*, 2H, NH_2); IR (film) ν_{max} 3351, 2922, 1694, 1629, 1528, 1440, 1271, 1206, 1130, 999, 834, 736, 668 cm^{-1}; FABHRMS (NBA) m/z 190.0740 ($M^+ + H$, $C_{10}H_{10}N_2O_2$ requires 190.0742). For methyl 5-aminoindole-2-carboxylate (**11**): 92%, mp 150–152 °C (CH_2Cl_2); ^1H NMR ($CDCl_3$, 400 MHz) δ 8.72 (*br s*, 1H, NH), 7.23 (*d*, 1H, *J* = 8.6 Hz, C7-H), 7.03 (*dd*, 1H, *J* = 1.0, 2.1 Hz, C3-H), 6.93 (*dd*, 1H, *J* = 1.0, 2.0 Hz, C4-H), 6.81 (*dd*, 1H, *J* = 2.0, 8.6 Hz, C6-H), 3.93 (*s*, 3H, CO_2CH_3), 3.57 (*br s*, 2H, NH_2); ^{13}C NMR ($CDCl_3$, 100 MHz) δ 160.0 (C), 150.3 (C), 145.6 (C), 143.0 (C), 127.7 (C), 117.7 (CH), 113.5 (CH), 112.6 (CH), 106.1 (CH), 52.2 (CH_3); IR (film) ν_{max} 3320, 1691, 1628, 1531, 1437, 1376, 1337, 1232, 1034, 997, 766 cm^{-1}; FABHRMS (NBA) m/z 190.0746 ($M^+ + H$, $C_{10}H_{10}N_2O_2$ requires 190.0742).

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

33. Sanger, F.; Nicklen, S.; Coulsen, A. R. *Proc. Natl Acad. Sci. U.S.A.* **1977**, *74*, 5463.

(*Received in U.S.A.* 14 *September* 1994; *accepted* 27 *October* 1994)