Pergamon

Bioorganic & Medicinal Chemistry 9 (2001) 2445–2459

BIOORGANIC & MEDICINAL CHEMISTRY

Recognition and Cleavage at the DNA Major Groove

Edward B. Skibo,* Chenggou Xing and Thomas Groy

Department of Chemistry and Biochemistry, Arizona State University, Tempe, AZ 85287-1604, USA

Dedicated to Professor Peter Dervan for his Contributions to the Field of Bioorganic Chemistry

Received 20 March 2001; accepted 5 June 2001

Abstract—DNA recognition agents based on the indole-based aziridinyl eneimine and the cyclopent[*b*]indole methide species were designed and evaluated. The recognition process involved either selective alkylation or intercalating interactions in the major groove. DNA cleavage resulted from phosphate backbone alkylation (hydrolytic cleavage) and N(7) -alkylation (piperidine cleavage). The formation and fate of the eneimine was studied using enriched ^{13}C NMR spectra and X-ray crystallography. The aziridinyl eneimine specifically alkylates the N(7) position of DNA resulting in direction of the aziridinyl alkylating center to either the 3′- or 5′-phosphate of the alkylated base. The eneimine species forms dimers and trimers that appear to recognize DNA at up to three base pairs. The cyclopent[*b*]indole quinone methide recognizes the 3′-GT-5′ sequence and alkylates the guanine N(7) and the thymine 6-carbonyl oxygen causing the hydrolytic removal of these bases. In summary, new classes of DNA recognition agents are described and the utility of ^{13}C-enrichment and ^{13}C NMR to study DNA alkylation reactions is illustrated.

Introduction

The recognition of DNA sequences and subsequent cleavage by chemical nucleases have been the subject of intense interest. The pyrrolopeptide-based recognition agents of the DNA minor groove were pioneered in Professor Dervan's laboratory.[1–3] Attachment of the radical cleaving moiety, MPE·Fe (II), to the pyrrolopeptide-based agents led to the design of sequence specific DNA cleaving agents. [4–6] Largely due to Professor Dervan's work, the field of DNA minor groove recognition has grown with many investigators and with many variations of the pyrrolopeptide structure.[7–12] The DNA major groove has also been utilized for the design of sequence specific recognition agents. Usually DNA strands or synthetic analogues thereof recognize the major groove by Hoogsteen base pairing to afford a triple helix structure.[13–18] Because DNA recognition at the minor and major groove has value in cancer chemotherapy and molecular biology, efforts in this area will no doubt continue into the new millennium.[16]

This laboratory has been involved in the design of major groove recognition agents that cleave DNA at the phosphate backbone.[19–21] Currently, we are developing new reactive species capable of recognizing and cleaving at specific sites in the DNA major groove. This report describes the chemistry and DNA alkylating/cleaving properties of the aziridinyl eneimine and of the cyclopent[*b*]indole methide species shown in Chart 1. Either species has the capacity to trap two nucleophiles resulting in a crosslinking reaction.

These reactive species were considered to be responsible for the cytotoxic and antitumor properties of indole and cyclopent[*b*]indole based aziridinyl quinones.[22,23] This report confirms the formation and build-up of these reactive species in solution as well as the cleavage mechanism. Novel aspects of this report include:

- Rapid determination of hydrolytic DNA cleavage chemistry using the ^{13}C-labeling developed in this laboratory.[24]
- Crosslinking of the G-base by N(7) and 3′-phosphate alkylation.
- Recognition of 3′-GT-5′ and 3′-GGA-5′ with small molecules.

Eventually, more functionalized versions of these indole and cyclopent[*b*]indole-based aziridinyl quinones will permit the recognition and cleavage of specific sequences at the major groove of DNA.

*Corresponding author. Tel.: +1-480-965-3581; fax: +1-480-965-2747; e-mail: eskibo@asu.edu

0968-0896/01/$ - see front matter
PII: S0968-0896(01)00219-X

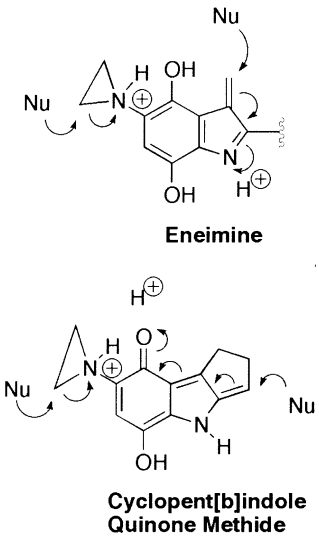


Chart 1.

Results and Discussion

The alkylating systems investigated

The indole (**1**) and cyclopent[*b*]indole (**2**) systems shown in Chart 2 were previously prepared and evaluated as antitumor agents possessing DNA alkylating capability.[22] The cytotoxicity of both systems prompted the detailed study of their DNA recognition and alkylation properties described in this report. Both systems require two-electron reduction to the hydroquinone in order to activate the alkylating centers. Hydroquinone formation activates the aziridinyl group as an alkylating center by permitting nitrogen protonation at neutrality.[20,21,25] Reduction also activates the acetate as a leaving group resulting in eneimine or quinone methide formation. The quinone methide species is an alkylating agent known to trap nucloephiles on DNA.[26–32] The eneimine species is related to the mitosene iminium ion known to alkylate DNA nucleophiles.[24,33,34] The metabolism of 3-methylindole in fact involves the formation of the eneimine species (also called the methylene imine) that traps nucleophiles.[35,36] Thus **1a** is potentially a triple alkylating agent capable of forming a quinone methide, an eneimine, and a protonated aziridine while **1b** and **1c** each have one or two alkylating center removed. Similarily, **2a** and **2b** possess either two or one alkylating centers respectively. The parallel study of these compounds will provide insights into how these multiple alkylating agents interact with DNA.

The ^{13}C-labeled analogues shown in the inset of Chart 2 were also prepared in order to assess the hydrolytic and DNA alkylation chemistry of **1**. The placement of ^{13}C at reacting centers permits the rapid assessment of the fate of highly reactive species.[24] The number of enriched ^{13}C resonances and their chemical shifts provide insights into the number and type of compounds formed in alkylation reactions. Incorporation of the ^{13}C-label at the 2α-position to afford **2α-^{13}C-1a** involved the procedure outlined in Scheme 1. Procedures for the conversion of 5-methoxyindole to **6** are provided in the Experimental Section. The conversion of **6** to **7** and the eventual preparation of **2α-^{13}C-1a** were carried out employing procedures previously reported for the unlabeled analogues.[22] The incorporation of the ^{13}C-label at the 3α-position was carried out by Vilsmeier formylation of **8** with ^{13}C-DMF to afford **9.** Multistep conversion of **9** to **3α-^{13}C-1a–d** was carried out employing reported procedures (Scheme 2).[22]

Eneimine formation and fate

The 3α-^{13}C labeled derivatives of **1c** and **1d** were used to verify that the eneimine can build-up in aqueous media and undergo complex reactions. Both of these compounds do not possess the aziridinyl group, so that only eneimine formation and fate can be studied.

Shown in Figure 1 is the 100 MHz ^{13}C NMR spectrum (obtained on a 400 MHz instrument) of two-electron reduced **3α-^{13}C-1c** obtained after 4 h of acquisitions. The presence of the methylene center of the eneimine is apparent at 98 ppm, along with starting material at 58 ppm and a reaction product at 18 ppm. The reaction product was isolated from a preparative reaction described below.

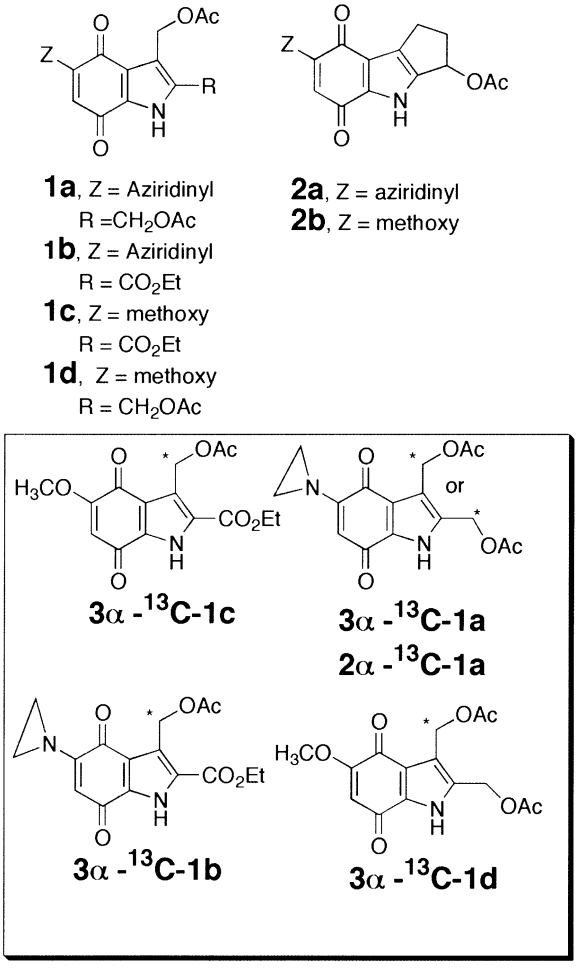
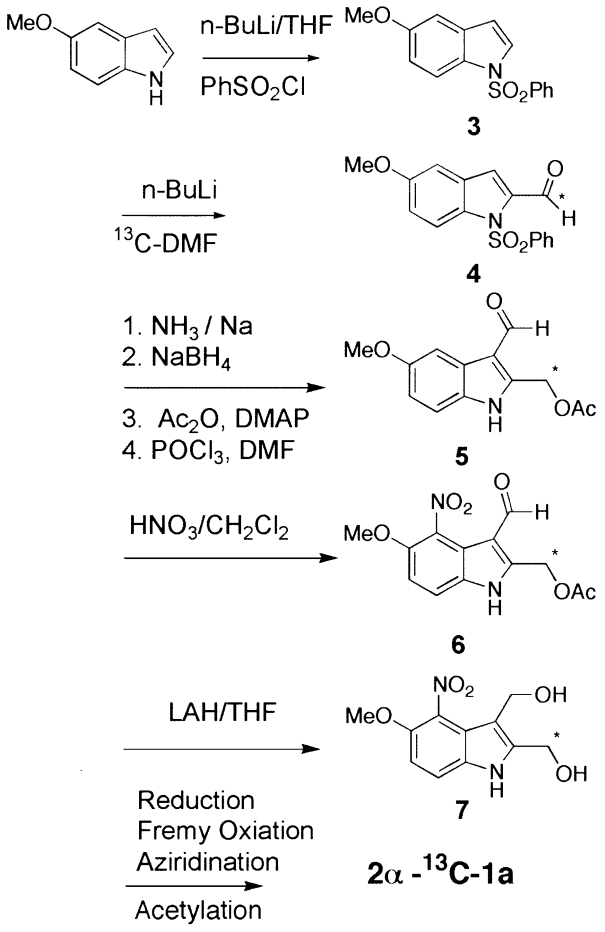


Chart 2.

Reduction of **3α-^{13}C-1c** and incubation in methanol afforded the complex mixture of products whose structures are shown in Scheme 3. An internal redox reaction (transfer of H_2 from the hydroquinone to the eneimine) results in formation of **10** and its transesterification product **11**. Product **10** is the same as that observed in the NMR tube reaction described above. The eneimime also traps methanol solvent to afford **12** as its hydroquinone, that then affords the corresponding quinone upon aerobic workup. However, most of the eneimine traps hydroquinone nucleophiles to form dimers **13–15** through bond formation at C-5 or C-7.

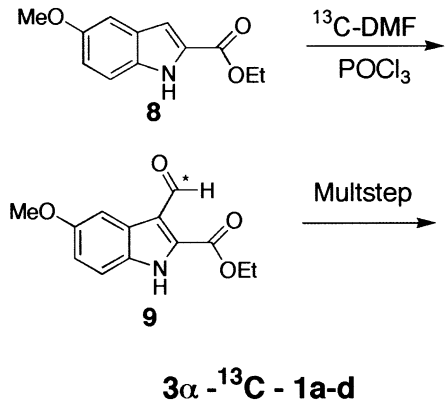


Scheme 1.

Reduction of **3α-^{13}C-1d** and incubation in methanol likewise afforded the complex mixture of products whose structures are shown in Scheme 4. The products arising from the internal redox reaction, **16** and **17**, are minor components of the mixture. The hydroquinone is the major nucleophile in solution resulting in the formation of dimers (**19–20** and **22–24**) and trimer **21** in addition to methanol trapping product **18**.

The structure of the dimers and trimer **21** in Schemes 3 and 4 were derived from ^{1}H NMR, ^{13}C NMR, COSY, HMQC, and HMBC. Furthermore the structure of trimer **21** was confirmed by X-ray crystallography. The incorporation of ^{13}C into the 3α position proved valuable in structural determinations, as illustrated by Figures 2–4. The ^{13}C NMR of dimers **23** and **20** shown in Figures 2 and 3 reveal the presence of the acetoxymethyl (57 ppm) as well as the linking methylene (38 ppm). The trimeric structure of **21** was readily assessed from the ^{13}C NMR shown in Figure 4 that shows two linking methylenes as well as the terminal acetoxymethyl. Methylene links resulting from direct carbonyl addition possess ^{13}C chemical shifts in the 35–38 ppm range. Previous studies of polymeric species derived from iminium ions showed that methylene links resulting from a Michael addition, such as those in **19** and **20**, possess ^{13}C chemical shifts in the 25–30 ppm range.[24]

The dimer and trimer structures shown in Schemes 3 and 4 can be readily determined spectroscopically, in spite of their complexity, as summarized:



Scheme 2.

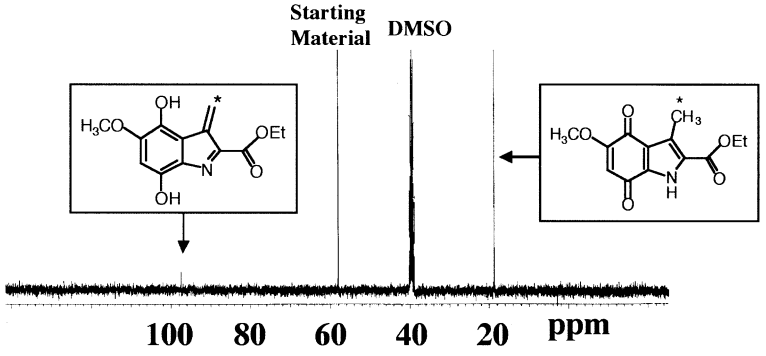
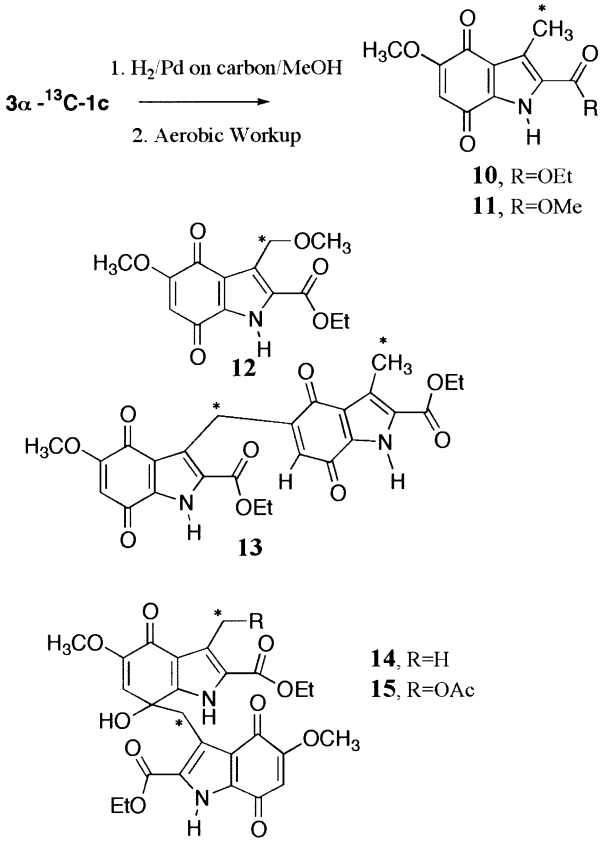


Figure 1. The ^{13}C NMR spectrum of eneimine formation from two-electron reduced **3α -^{13}C-1c**. The reaction contained 2 mg of starting material, 0.5 mL of DMSO-d_6, and 0.5 mL pD = 7.2 0.1 M phosphate D_2O buffer.

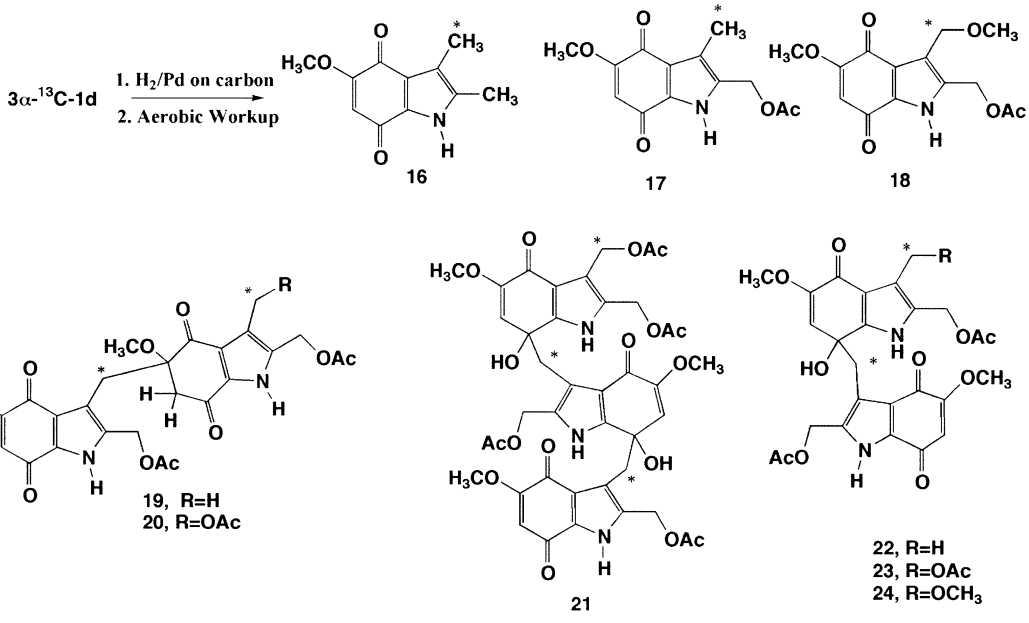
- Either the direct carbonyl or the Michael addition processes afford a stereocenter adjacent to the ^{13}C-methylene center. As a result, the H NMR spectrum of this methylene center shows two doublet of doublets. The methylene protons are chemically nonequivalent due to the adjacent stereocenter resulting in gem splitting (doublet of doublets). The ^{13}C carbon splits the doublets of doublets into two sets with a J value typically of 130 Hz.
- The ^{13}C-enriched HMQC spectrum of the dimer or trimer product readily identifies the H NMR chemical shifts of the ^{13}C-methylene center. Inspection of the COSY spectrum will show two adjacent 'boxes' corresponding to the doublets of doublets. A COSY spectrum of a mitosene dimer showing this feature was published recently.[24]
- The ^{13}C- enriched HMBC spectrum shows two-bond coupling between the ^{13}C-methylene center and a methylene center adjacent to a carbonyl indicating a Michael addition product.
- The indole NH protons are readily apparent between 9 and 10 ppm in the H NMR spectrum indicating that alkylation of this center had not occurred as well as the number of indole units (two for dimer and three for trimer).



Scheme 3.

The chemistry of the eneimine and the parent indole system can be readily explained by referring to the electron density of the indole ring based on molecular orbital calculations,[37] inset of Scheme 5. These chemical explanations are outlined below in conjunction with Schemes 5 and 6

- Electron density for the indole system is higher at the 3-position than at the 2-position resulting almost exclusive eneimine formation. The formation of the quinone methide species by acetate elimination from the 2-position contributes only if eneimine formation is not possible.
- Excess electron density resides on positions 5 and 7 of the indole ring resulting in electrophilic attack by the eneimine at these positions. The mechanism for the formation of the 5-linked dimers **13**, **19**, **20** is illustrated in Scheme 6. Trapping of the eneimine electrophile by the 3- and 8-positions was not observed in this study. A previous study reported



Scheme 4.

trapping of the iminium ion by the 3-position followed by loss of formaldehyde.[24] Trapping by the 8-position has never been observed probably because the product can never aromatize.

- The high specificity of the eneimine for nucleophilic sites on the indole system suggests that this electrophilic species will selectively trap nucleophiles on DNA.

Cyclopent[*b*]indole quinone methide

The formation of this reactive species has not been documented as well as the eneimine because ^{13}C incorporation at the electrophilic center has yet to be accomplished. Reduction of **2a** in anaerobic buffer produced the dimeric species **25a,b** suggesting that the quinone methide species can form in solution (see mechanism in the inset of Scheme 7). Modeling studies indeed suggested that the quinone methide would be stable.[23] The cyclopent[*b*]indole quinone methide proved to be a highly selective alkylating and cleaving agent of DNA.

Eneimine reaction with DNA

The most reactive nucleophilic center of DNA is the guanine N(7) position,[38–41] and the eneimine electrophile should selectively alkylate DNA at this position. In order to determine the DNA alkylation site, hexamer DNA with the sequence ATGCAT was treated with two-electron reduced **3α-^{13}C-1b**. The DNA hexamer was first purified by 20% PAGE as described in the Experimental and then reductively alkylated. The reduction was carried out catalytically and the DNA was recovered from precipitation. The ^{13}C NMR spectrum of this DNA was compared with that of the native hexamer. The alkylated DNA spectrum shown in Figure 5 has a strong peak at 42 ppm indicating that the ^{13}C-methylene is bound to a nitrogen. This result is consistent

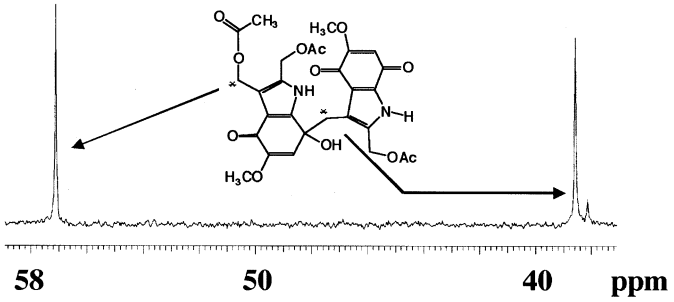


Figure 2. The ^{13}C NMR spectrum of dimer **23** from the hydrolysis of **3α-^{13}C-1c**.

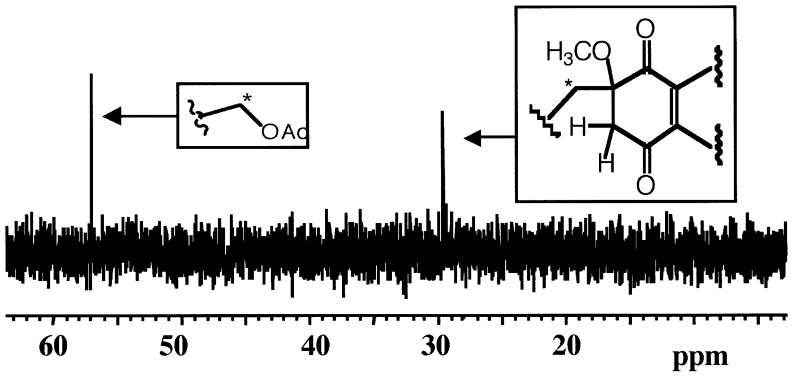


Figure 3. The ^{13}C NMR spectrum of dimer **20** from the hydrolysis of reduced **3α-^{13}C-1c**.

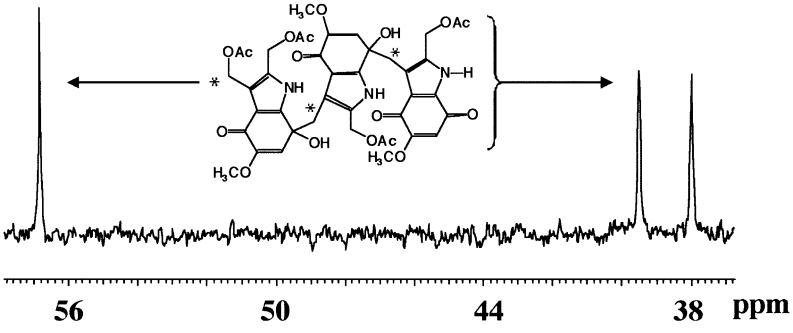


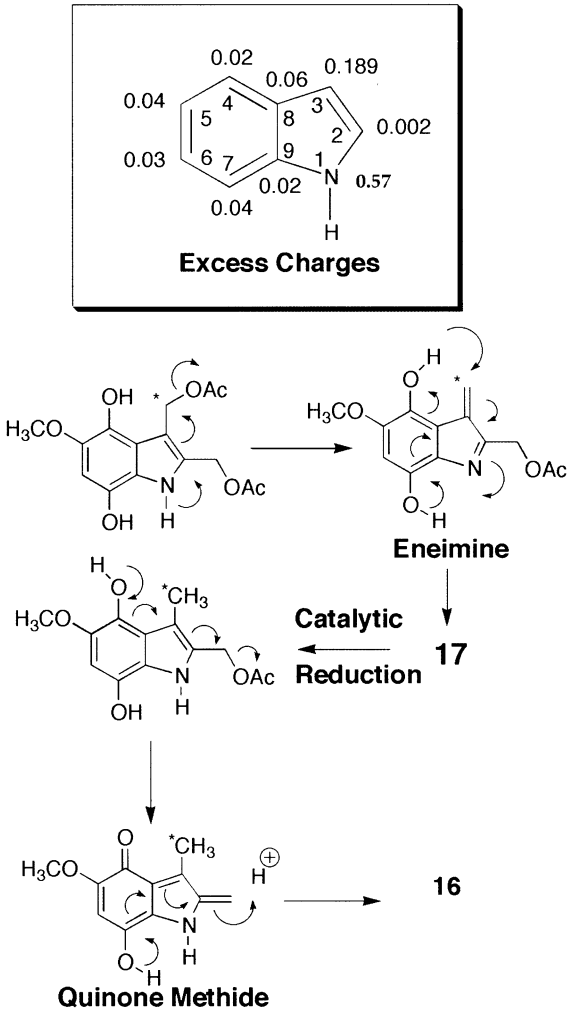
Figure 4. The ^{13}C NMR spectrum of trimer **21** from the hydrolysis of reduced **3α-^{13}C-1c**.

with alkylation of adenine or guanine N(7) as well the 2-amino of guanine. PAGE studies described below indicated that the N(7) positions are the alkylation sites.

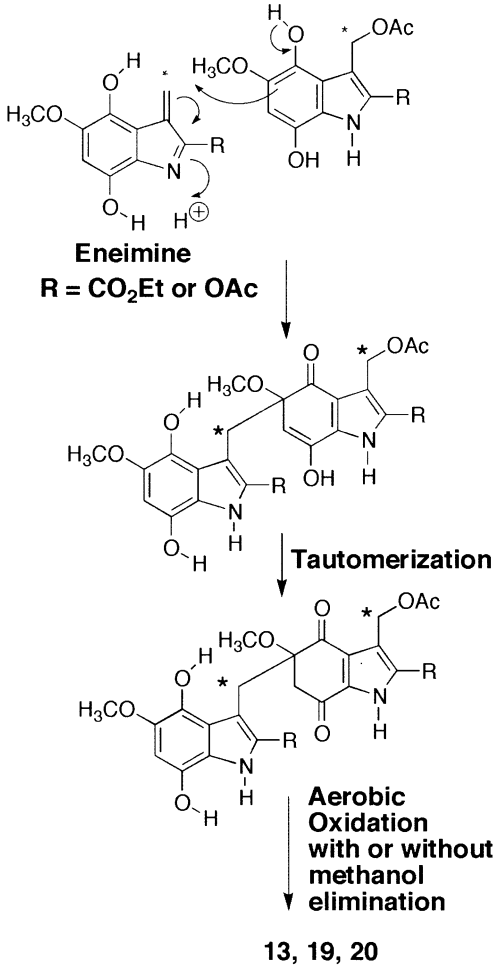
The hexamer was also treated with reduced **3α-^{13}C-1a** and **2α-^{13}C-1a** and ^{13}C NMR spectra of the alkylation products were obtained and compared with the ^{13}C natural abundance spectrum of the hexamer (Fig. 6). The series of spectra in this figure illustrate the limitation of using enriched ^{13}C NMR to study DNA alkylation reactions. The formation of multiple alkylation products as well as low yields could dilute the incorporated ^{13}C labels to the point that they are as intense as the natural abundance ^{13}C spectrum. The solution is to compare the natural abundance and enriched spectra and note differences. Comparison of the natural abundance ^{13}C NMR spectrum (part A) with the ^{13}C spectrum of 2α-labeled alkylating agent **2α-^{13}C-1a** (part B) reveals that the oxygen of acetate was still attached at the 2α-^{13}C center. In contrast, the 3α-labeled alkylating agent **3α-^{13}C-1a** afforded a ^{13}C spectrum showing that nitrogen alkylation had occurred at this position. These results are consistent with elimination of acetate from the 3α-position rather than from the 2α- position as observed in the hydrolysis study (Scheme 6).

The PAGE gel shown in Figure 7 reveals that the reductive cleavage of DNA by **1a** occurred at both the guanine (N-7) and the 3′ guanosine phosphate. The alkylation of either the adenine or guanine N(7) was determined by piperidine treatment according to Maxam and Gilbert.[42] The piperidine treatment also hydrolyzed the phosphotriester resulting from phosphate alkylation without detectable backbone cleavage. In contrast, the treatment of the alkylated DNA with basic loading buffer resulted in only phosphate backbone cleavage without detectable cleavage via N(7) adducts.[20]

The first two lanes of Figure 7 show the results of a limited alkylation reaction followed by piperidine and hydrolytic cleavage respectively. The ladder in the first lane shows the presence of specific guanine N(7) alkylation while the ladder in the second lane shows phosphate alkylation in the G+1 direction (of 5′-labeled DNA). The PAGE results and the ^{13}C NMR spectrum

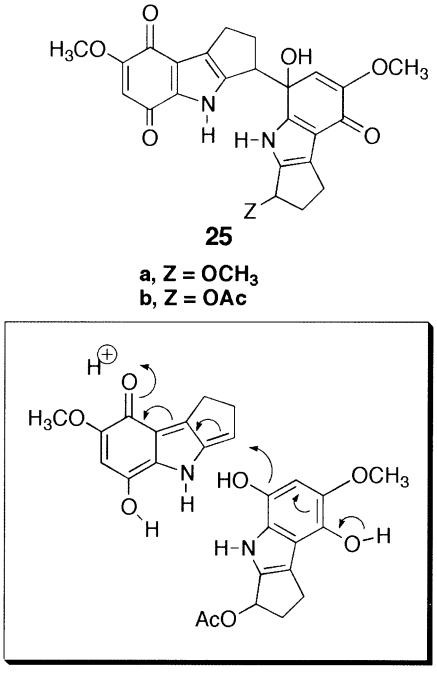


Scheme 5.



Scheme 6.

shown in Figure 5 support a process where the eneimine intermediate selectively reacts with guanine N(7) and the aziridinyl center is then directed to the 3′-guanosine phosphate)Scheme 8).



Scheme 7.

When borohydride reduction was employed, both adenine and guanine N(7) alkylation occurred (lanes 3 and 5) along with alkylation of the corresponding 3′-phophates (G +1 and A +1 cleavage ladders). Reductive activation with borohydride results in the rapid formation of the eneimine alkylating agent. In contrast, the slower catalytic reduction affords limiting amounts of eneimine and hence limited cleavage (lanes 1 and 2).

The PAGE gel in Figure 8 compares the reductive cleavage by **1a** with cleavage by **1b,d** and an analogue

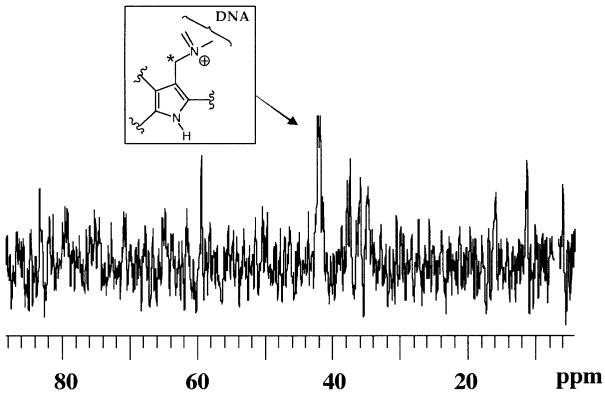


Figure 5. The ^{13}C NMR spectrum of alkylated hexamer 5′-ATGCAT-3′ obtained by treatment of the DNA with two-electron reduced **3α-^{13}C-1c.**.

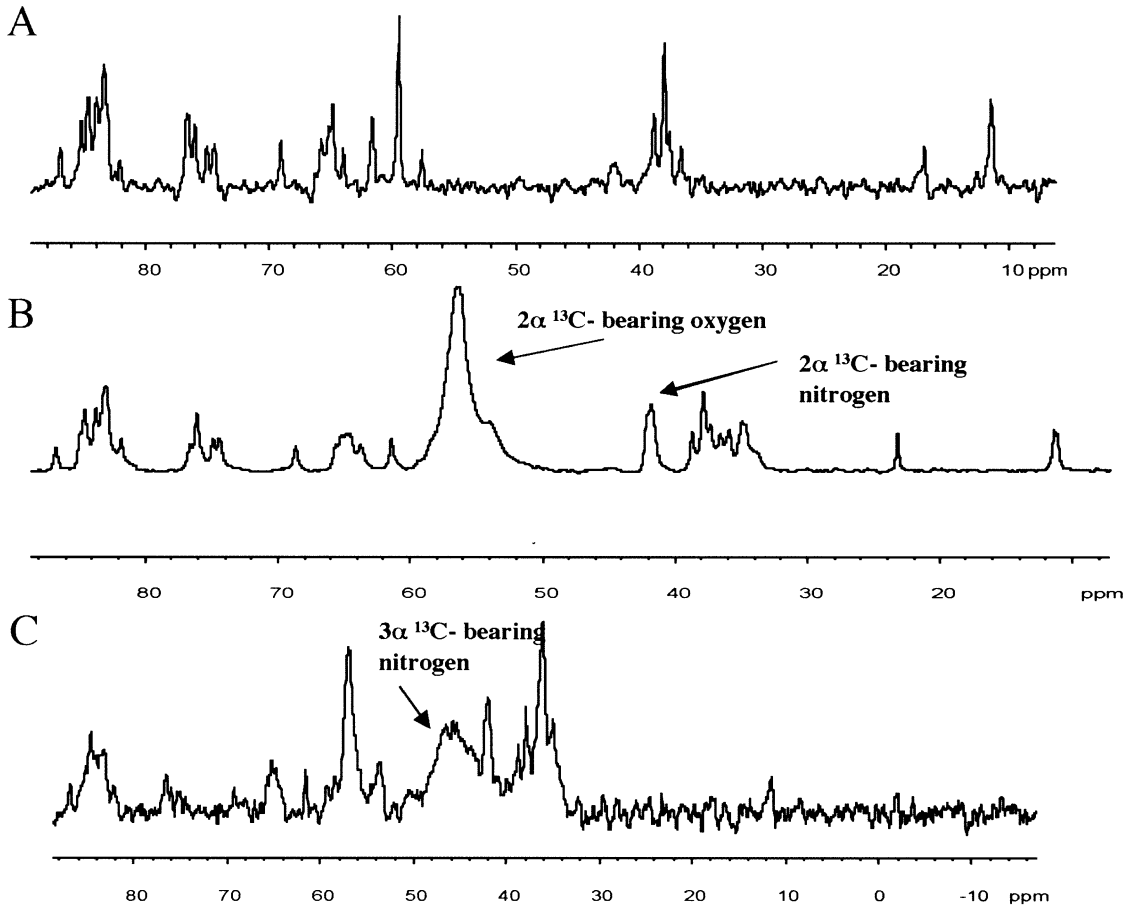


Figure 6. ^{13}C NMR Spectra of native hexamer 5′-ATGCAT-3′ (part A) and alkylation products obtained by treatment of the hexamer with two-electron reduced **2α-^{13}C-1a** (part B) and **3α-^{13}C-1a** (part C).

without a 3α leaving group. Lanes 1 and 2 show piperidine and hydrolytic cleavage, respectively, by reduced **1a**. These lanes were obtained for comparative purposes and possess the same cleavage patterns as those in Figure 7.

The change from the 2-acetoxymethyl (**1a**) to the 2-ethoxycarbonyl group (**1b**) does not influence eneimine trapping of the adenine and the guanine N(7) nucleophiles (compare lanes 1 and 3). However, the hydrolysis ladder in lane 4 does not show the G + 1 and the A + 1 ladder observed for **1a** (lane 2). In some sequences, the piperidine and hydrolysis cleavage patterns of lane 4 are identical while in other sequences there is no corresponding hydrolytic cleavage. These results suggest that attachment of the **1b** eneimine to the N(7) center permits access of the aziridinyl alkylating center to the 5′ phosphate or to neither phosphate. Note that DNA cleavage by N(7) alkylation and piperidine treatment affords the same products as 5′ phosphate hydrolysis.

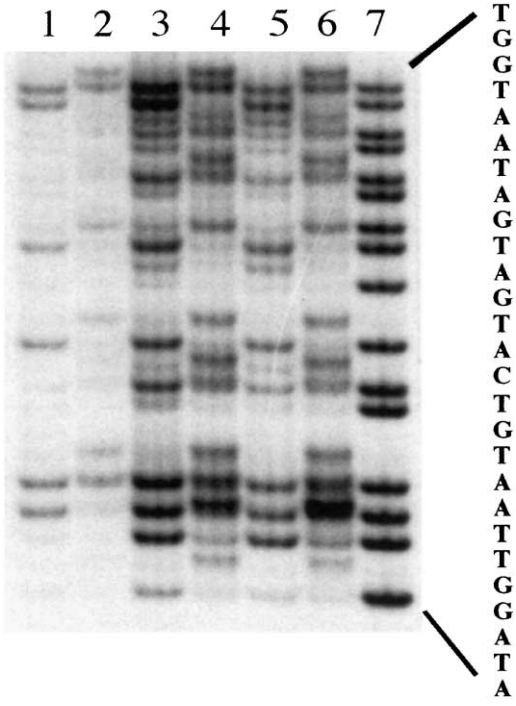
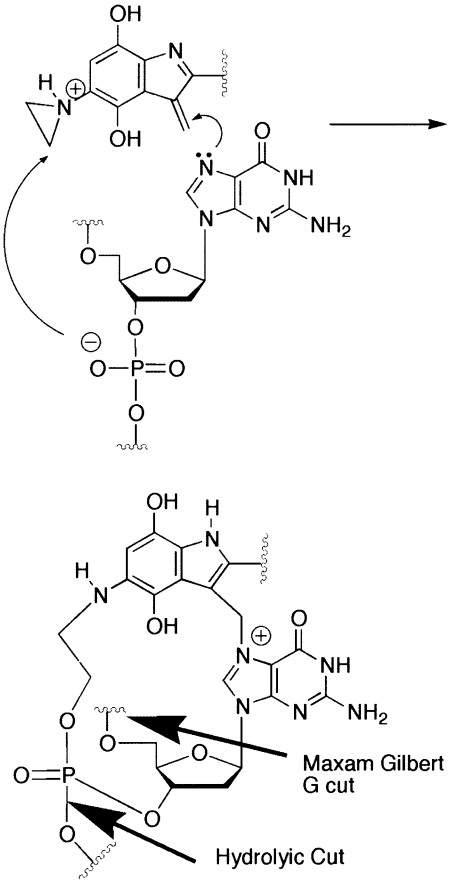


Figure 7. Autoradiogram of the 1a-mediated reductive cleavage of a 5′-^{32}P end-labeled 514 bps restriction fragment from pBR322 DNA (Ecor I/Rsa I).

	Compound	Conditions	Cleavage Pattern
Lane 1	**1a**	Catalytic reductive activation piperidine cleavage	G
Lane 2	**1a**	Catalytic reductive activation, hydrolytic cleavage	G + 1 (3′ direction)
Lane 3	**1a**	$NaBH_4$ reductive activation, piperidine cleavage	G & A
Lane 4	**1a**	$NaBH_4$ reductive activation, hydrolytic cleavage	G + 1 (3′ direction) A + 1 (3′ direction)
Lane 5	**1a**	$NaBH_4$ reductive activation, piperidine cleavage	G & A
Lane 6	**1a**	$NaBH_4$ reductive activation, hydrolytic cleavage	G + 1 (3′ direction) A + 1 (3′ direction)
Lane 7	DMS	Piperdine cleavage	G & A

the piperidine and hydrolysis cleavage ladders shown in lanes 5 and 6, respectively, were determined for an aziridinyl indoloquinone without a 3α leaving group. the weak hydrolysis ladder in lane 5 in fact indicates that phosphate alkylation is not an important process when the eneimine species cannot form. these results are consistent with the mechanism whereby the eneimine is the initial dna alkylating species followed by aziridinyl alkylation of the phosphate backbone.

The cleavage ladders shown in lanes 7 and 8 show the unusual multibase recognition capability of **1d**. The hydrolysis ladder in lane 8 shows that recognition occurs only at 3-GGA-5′ with phosphate alkylation and cleavage occurring largely at the center G base. The reason for multibase recognition by such a small molecule may be the presence of dimers and trimers. Recall that the hydrolysis of reduced **1d** afforded a variety of such species in good yields, **19**–**24** in Scheme 4. Any of these species could form a quinone methide or eneimine species capable of alkylating the phosphate backbone or N(7) centers. We are currently investigating if the trimer **21** and dimers could actually recognize and cleave DNA.



Scheme 8.

Cyclopent[*b*]indole quinone methide reaction with DNA

The DNA cleavage ladders obtained with reduced cyclopent[*b*]indoles **2a,b** and an isomer of **2a** are shown in Figure 9. These ladders clearly show that the aziridinyl group was not involved in the alkylation and cleavage process. Thus repositioning of the aziridinyl ring, or even its removal, resulted in the same cleavage pattern. Furthermore, the weak or nonexistent hydrolytic ladders in Figure 9 reveal the absence of phosphate alkylation. The reduced cyclopent[*b*]indole recognizes 3′-GT-5′ and cuts at both the G and T bases upon treatment with piperidine.

The mechanism of cyclopent[*b*]indole DNA recognition and cleavage is as yet unclear, but insights into these processes were arrived at from the PAGE results shown in Figure 9 and molecular modeling. A cyclopent[*b*]indole quinone methide would have to recognize

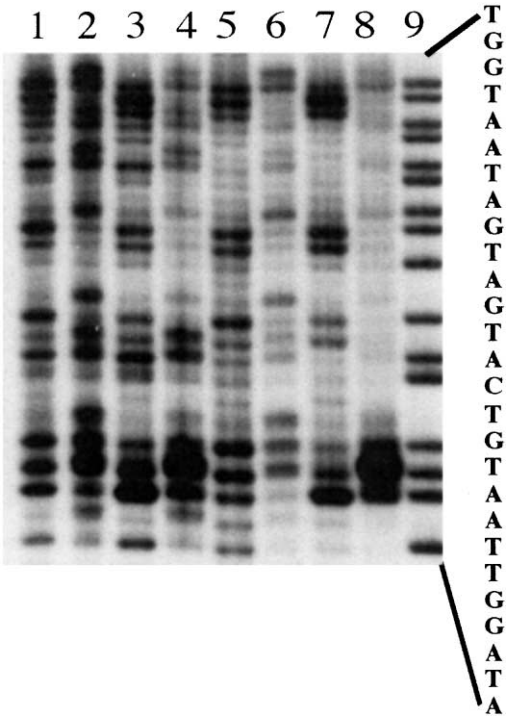
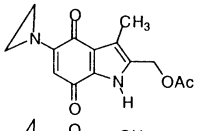
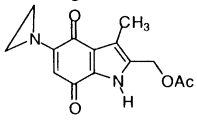


Figure 8. Autoradiogram of the reductive cleavage of a 5′-^{32}P end-labeled 514 bps restriction fragment from pBR322 DNA (Ecor I/Rsa I) by 1a,b,d and the reported 3-methyl analogue.[22]

	Compound	Conditions	Cleavage Pattern
Lane 1	**1a**	$NaBH_4$ reductive activation, piperidine cleavage	G & A
Lane 2	**1a**	$NaBH_4$ reductive activation, hydrolytic cleavage	G + 1 (3′ direction) A + 1 3′ direction)
Lane 3	**1b**	$NaBH_4$ reductive activation, piperidine cleavage	G & A Preference for A
Lane 4	**1b**	$NaBH_4$ reductive activation, hydrolytic cleavage	Some G & A
Lane 5		$NaBH_4$ reductive activation, piperidine cleavage	Some G & A
Lane 6		$NaBH_4$ reductive activation, hydrolytic cleavage	Weak Cleavage
Lane 7	**1d**	$NaBH_4$ reductive activation, piperidine cleavage	G T (5′′ direction)
Lane 8	**1d**	$NaBH_4$ reductive activation, hydrolytic cleavage	3′-GGA-5′ at the center G
Lane 9	**DMS**	Piperdine cleavage	G & A

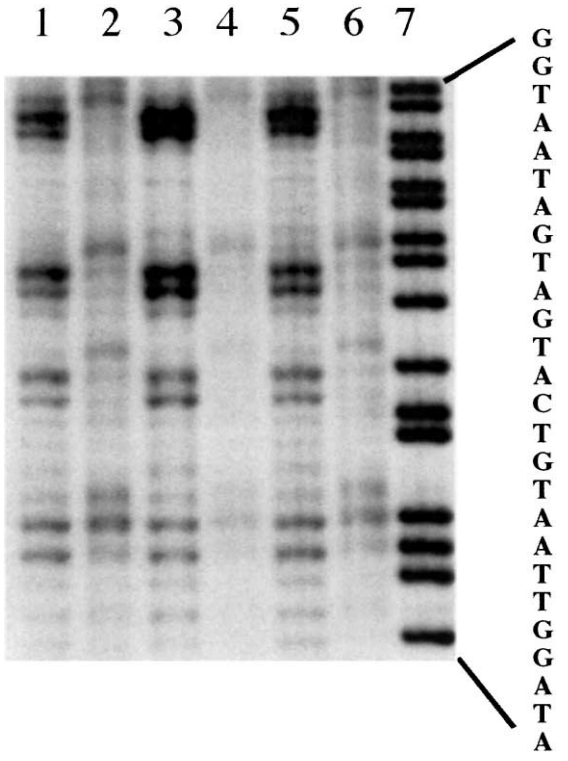
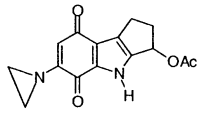
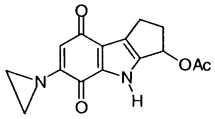


Figure 9. Autoradiogram of the reductive cleavage of a 5′-^{32}P end - labeled 514 bps restriction fragment from pBR322 DNA (Ecor I/ Rsa I) by **2a**,**b** and the reported 6- aziridinyl isomer of **2a**.[22]

	Compound	Conditions	Cleavage Pattern
Lane 1	**2a**	$NaBH_4$ reductive activation, piperidine cleavage	G & T of 3′-GT-5′
Lane 2	**2a**	$NaBH_4$ reductive activation, hydrolytic cleavage	Weak Cleavage at G
Lane 3	**2b**	$NaBH_4$ reductive activation, piperidine cleavage	G & T of 3′-GT-5′
Lane 4	**2b**	$NaBH_4$ reductive activation, hydrolytic cleavage	No Significant Cleavage
Lane 5		$NaBH_4$ reductive activation, piperidine cleavage	G & T of 3′-GT-5′
Lane 6		$NaBH_4$ reductive activation, hydrolytic cleavage	No Significant Cleavage
Lane 7	DMS	Piperdine cleavage	G & A

the 3′-GT-5′ sequence by interacting with the G C and AT base pairs in the major groove so as to place the alkylating center near the guanine N(7) and the adjacent to the oxygen near the thymine 6-carbonyl. Alkylation of either of these centers would result in hydrolytic removal of either base followed by piperidine-mediated strand cleavage. When the methide center was docked near the guanine N(7) in the 3′-GT-5′ sequence and the resulting complex minimized, the methide had intercalated between the base pairs without any distinctive hydrogen bonding interactons. In addition, the model indicated that placement of the methide near the guanine N(7) will permit reaction with the thymine oxygen.

Conclusions

DNA recognition agents based on the aziridinyl eneimine and the cyclopent[*b*]indole methide species were designed and evaluated utilizing ^{13}C NMR and PAGE. The recognition process involved either selective alkylation or intercalating interactions in the major groove. DNA cleavage resulted from phosphate backbone alkylation (hydrolytic cleavage) and N(7)-alkylation (piperidine cleavage). The following conclusions are made.

- The aziridinyl eneimine, resulting from elimination of acetate from the indole 3α position, specifically alkylates the N(7) position of DNA resulting in direction of the aziridinyl alkylating center to either the 3′- or 5′- phosphate of the alkylated base. One reported aziridinyl eneimine was found to be specific for guanine N(7) and this base's 3′-phosphate.
- The nonaziridinated indole eneimine was studied in order to gain insights into eneimine chemistry in the absence of other reacting centers. Spectroscopic evidence was obtained for the buildup of the eneimine in solution. Both dimers and trimers are formed by electrophilic attack of the eneimine at either the 5- or 7- positions of the indole ring. NMR and X-ray crystallography were employed to characterize these novel products. The eneimine trimers and dimers appear to be capable of multibase recognition of DNA.
- Generation of an indole-base quinone methide by acetate elimination from the 2α-position did not occur either in hydrolysis reactions or in DNA alkylation reactions. Since the indole 3-position is more electron rich than the 2-position, eneimine formation rather than quinone methide formation is the predominate reaction. This conclusion is relevant to the mechanism of indole -based antitumor agents that rely on leaving group elimination from the 2-position, such as WV-15[43] and EO9.[44,45] Our results suggest that the postulated quinone methides or extended quinone methides resulting from leaving group elimination may not be important intermediates.
- The cyclopent[*b*]indole quinone methide forms in solution and can trap itself (dimerize) and trap DNA nucleophiles [gunaine N(7) and the thymine 6-carbonyl oxygen]. The aziridinyl group is not required for alkylation and cleavage of DNA; these processes are mediated entirely by the quinone methide species.

The finding enumerated above are starting points for further studies that will lead to new classes of DNA cleaving and recognition agents.

Experimental

The ^{13}C NMR spectra was taken on Varian Inova 400 either in Nolorac dual PFG for hexamer or Varian indirect PFG for hydrolytic products at 100 MHz.

Acetylation procedure for 1 and 2

The reported compounds in Chart 2 (**1a**, **1b**, and **2a**) were prepared as previously described.[22] Compounds **1c**, **1d**, and **2b** were prepared by acetylation of the previously reported alcohol derivatives.[22]

To a solution of 0.2 mmol of the alcohol in 5 mL of CH_2Cl_2 and 1 mL of acetone containing 50 mg of DMAP was added 100 λ acetic anhydride. The solution was stirred at room temperature for 5 min and directly flash chromatographed using ethyl acetate as the eluent. The eluted solution containing the product was washed with $NaHCO_3$, dried over Na_2SO_4, and vacuum dried. The product was recrystallized from ethyl acetate and hexane to afford a near quantitative yield of the acetylated product. Physical properties of these acetylated products are provided below for **1c**, **1d**, and **2b.**

Ethyl 3-acetoxymethyl-5-methoxyindol-4,7-dione-2-carboxylate (1c). Mp 194–196 °C; TLC (dichloromethane/MeOH 98:2), R_f 0.55; ^1H NMR ($CDCl_3$) δ 9.86 (1H, bs, indole proton), 5.79 (1H, s, 6-proton), 5.84 and 5.28 (2H, d, $J = 168.3$ Hz, 3-methylene protons), 4.41 (2H, q, $J = 7.2$ Hz, 2-methylene of ethyl), 3.87 (3H, s, 5-methoxy), 2.06 (3H, s, 3-acetoxylmethyl), 1.39 (3H, t, $J = 7.2$ Hz, 2-methyl of ethyl); ^{13}C NMR ($CDCl_3$) δ 55.48 ppm; IR (KBr pellet) 3491, 3354, 3302, 3211, 3041, 2943, 2862, 1675, 1558, 1431, 1384, 1200, 1173, 1128, 919 cm^{-1}; MS (EI) 322 (M^+), 307 (M^+-CH_3), 280, 268, 251, 234, 207. Anal. calcd: C, 56.21; H, 4.69; N, 4.34. Found: C, 56.03; H, 4.79; N, 4.21.

2,3-Diacetoxymethyl-5-methoxyindol-4,7-dione (1d). Mp 179–181 °C. TLC (dichloromethane/MeOH 98:2), R_f 0.50. ^1H NMR ($CDCl_3$) δ 9.74 (1H, bs, indole proton), 5.72 (1H, s, 6-proton), 5.79 and 5.24 (2H, d, $J = 165.6$ Hz, 3-methylene protons), 5.20 (2H, s, 2-methylene protons), 3.84 (3H, s, 5-methoxy), 2.11 and 2.04 (6H, 2s, 2,3-acetoxymethyls). ^{13}C NMR ($CDCl_3$) δ 56.47 ppm. IR (KBr pellet) 3443, 3371, 3342, 3260, 2993, 2918, 2827, 1674, 1634, 1569, 1404, 1374, 1241, 1139, 919, 827 cm^{-1}. MS (EI) 322 (M^+), 280 (M^+ −acetyl), 269, 251, 231, 216, 207, 179. Anal. calcd: C, 56.21; H, 4.69; N, 4.34. Found: C, 56.18; H, 4.72; N, 4.33.

3-Acetoxy-7-methoxy-1,2,3,4-tetrahydrocyclopent[*b*]indole-5,8-dione (2b). Mp 173–175 °C. TLC: (dichloromethane/MeOH 95:5), R_f 0.55. ^{1}H NMR ($CDCl_3$) δ 10.03 (1H, bs, indole proton), 5.68 (1H, s, 6-protons), 5.64 (1H, m, 3-hydroxymethylene proton), 3.82 (3H, s, 7-methoxy), 2.96, 2.73 and 2.25 (4H, m, methylenes of cyclopentyl), 2.04 (3H, s, 3-acetyl methyl). IR (KBr pellet) 3452, 3206, 3035, 2949, 2892, 1701, 1623, 1577, 1502, 1433, 1322, 1219, 1167, 1024, 955, 834 cm^{-1}. MS (EI) 275 (M^+), 247 (M^+-CO), 232, 214, 202, 187, 160, 131. Anal. calcd: C, 61.08; H, 4.76; N, 5.09. Found: C, 60.94; H, 4.81; N, 5.00.

5-Methoxy-1-phenylsulfonylindole (3). To 2.5 g (17 mmol) of 5-methoxyindole in 13 mL of dry THF, cooled by dry ice/acetone bath over a nitrogen atmosphere, was added 12 mL of 1.6 M *n*-butyllithium in hexane over 10 min. The reaction solution was stirred for 1.2 h at 0 °C and then chilled in a dry ice acetone bath. To this chilled reaction was added 2.6 mL of $PhSO_2Cl$ over 15 min and the resulting mixture stirred for 12 h at room temperature. The reaction mixture was then mixed with 50 mL of 3% $NaHCO_3$ and stirred for 30 min. The solution was then extracted 4× with 100 mL portions of CH_2Cl_2. The extracts were dried over Na_2SO_4 and vacuum dried to a yellow oil. The oil was then mixed with 4 mL of THF followed by addition of 10 mL of hexane and kept at −10 °C for 5h. The crystallized product was then filtered and collected: 3.99g (82%) yield; mp 54–56 °C; TLC (CH_2Cl_2) R_f 0.75; IR (KBr pellet) 3342, 3211, 3017, 2939, 2902, 2878, 1622, 1547, 1301, 1187, 1125, 1003, 842 cm^{-1}; ^{1}H NMR ($CDCl_3$) δ 7.86, 7.52, 7.45, 6.93, and 6.59 (10H, m, aromatic protons), 3.80 (3H, s, 5-methoxy); MS (EI mode) *m*/*z* 287 (M^+), 272 (M^+-CH_3), 147. Anal. calcd ($C_{15}H_{13}NO_3S$) : C, 62.71; H, 4.56; N, 4.88. Found: C, 62.43; H, 4.71; N, 4.80.

2-Formyl-5-methoxy-1-phenylsulphonylindole (4). To 1.44 g of **3** in 10 mL of THF, cooled at −78 °C under a nitrogen atmosphere, was added 4 mL of 1.6 M *n*-butyllithium in hexane over 10 min. The reaction solution was stirred for 25 min followed by addition of 0.7 mL ^{13}C-DMF and stirring at 0 °C for 30 min. The completed reaction was mixed with 100 mL of saturated NH_4Cl solution and extracted 5×100 mL portion of CH_2Cl_2. The extracts were dried over Na_2SO_4 and dried to brown solid, which was further purifed by flash chromatography employing dichloromethane as the eluant. Yield: 1.21 g (76%); mp 123–124 °C; TLC (dichloromethane), R_f 0.44; ^{1}H NMR ($CDCl_3$) δ 10.82 and 10.19 (1H, d, J = 189 Hz, 2-formyl protons), 8.13, 7.73, 7.53, 7.39, 7.13, and 6.99 (9H, m, aromatic protons), 3.82 (3H, s, 5-methoxy); ^{13}C NMR ($CDCl_3$) δ 182.7; IR (KBr pellet) 3354, 3326, 2917, 2889, 1679, 1532, 1514, 1347, 1236, 1218, 1036, 945, 834 cm^{-1}; MS (EI) 316 (M^+), 301 (M^+-CH_3), 287 ($M^+-{}^{13}CO$), 274, 257, 172, 146. Anal. calcd ($C_{16}H_{13}N\ O_4S$): C, 61.07; H, 4.14; N, 4.43. Found: C, 60.94; H, 4.20; N, 4.39.

2-Acetoxymethyl-3-for mLy-5-methoxyindole (5). To a solution of 200 mg Na metal in 10 mL of NH_3, cooled at −78 °C, was added 630 mg of **4** dissolved in 10 mL of dry THF. The resulting solution was stirred for 15 min and then the reaction allowed to warm to room temperature after the addition of 0.5 g NH_4Cl. After the liquid NH_3 was gone, the solution was mixed with 20 mL of methanol containing 300 mg of $NaBH_4$. The solution was stirred for 10 min, diluted with 30 mL of water, and then extracted 4× with 100 mL portions of methylene chloride. The extracts were dried over Na_2SO_4 and evaporated to a residue. The residue was dissolved in 10 mL of methylene chloride containing 100 mg of dimethylaminopyridine. To this solution was added 300 λ acetic anhydride. The solution was stirred for 10 min and directly purified by a flash chromatography using methylene chloride as the eluent. The product was disolved in 10 mL of methylene chloride and this solution was added to a solution of 200 λ $POCl_3$ and 600 λ DMF in 10 mL of methylene chloride at 0 °C. The solution was stirred overnight and neutralized with saturated $NaHCO_3$ followed by extraction 6× with 50 mL portions of methylene chloride. The extract was dried over Na_2SO_4 and concentrated to a residue, which was further purifed by flash chromatography using methylene chloride as the eluent: Yield 123 mg (25%); Mp 189–192 °C; TLC (dichloromethane, methanol 98:2), R_f 0.68; ^{1}H NMR ($CDCl_3$) δ 10.27 (1H, s, 3-formyl proton), 9.00 (1H, bs, indole proton), 7.74 (1H, d, J = 2.7 Hz, 4-proton), 7.29 (1H, d, J = 8.7 Hz, 7-proton), 6.95 (1H, dd, J = 2.7, 8.7 Hz, 6-proton), 5.79 and 5.29 (2H, d, J = 150.9 Hz, 2-acetoxylmethyl), 3.89 (3H, s, 5-methoxy), 2.16 (3H, s, 2-acetoxylmethyl). ^{13}C NMR ($CDCl_3$) δ 56.10 ppm; IR (KBr pellet) 3494, 3337, 3169, 3088, 2962, 2812, 1665, 1616, 1537, 1518, 1198, 1174, 1102, 1031 cm^{-1}; MS (EI) 248 (M^+), 206 (M^+-COCH_2), 188, 172, 159, 145, 116. Anal. calcd ($C_{13}H_{13}NO_4$): C, 63.30; H, 5.28; N, 5.64. Found: C, 62.99; H, 5.39; N, 5.52.

2-Acetoxymethyl-3-formyl-5-methoxy-4-nitroindole (6). To a solution of 80 mg of **5** in 10 mL CH_2Cl_2 was added 50 λ of 70% HNO_3. The reaction was stirred for 0.5 h and made basic with $NaHCO_3$ saturated solution. The basic solution was then extracted 4× with 40 mL portions of methylene chloride. The extracts were dried over Na_2SO_4 and dried, which was further purifed by flash chromatography. Yield: quantitative. Mp: 201–203 °C. TLC (dichloromethane/MeOH 98:2), R_f 0.58. ^{1}H NMR ($CDCl_3$) δ 10.04 (1H, s, 3-formyl proton), 9.29 (1H, bs, indole proton), 7.51 and 7.09 (2H, 2d, J = 8.7 Hz, 6, 7-protons), 5.88 and 5.38 (2H, d, J = 152.4 Hz, 2-acetoxylmethyl), 3.96 (3H, s, 5-methoxy), 2.19 (3H, s, 2-acetoxylmethyl). ^{13}C NMR ($CDCl_3$) δ 57.53 ppm. IR (KBr pellet) 3467, 3323, 3198, 3001, 2983, 2822, 1701, 1643, 1502, 1471, 1304, 1229, 1137, 1039, 927, 864 cm^{-1}. MS (EI) 293 (M^+), 251 (M^+-COCH_2), 233, 215, 201, 185, 172, 156, 146, 116. Anal. calcd ($C_{13}H_{12}N_2O_6$) C, 53.58; H, 4.12; N, 9.55. Found: C, 52.99; H, 4.28; N, 9.32.

^{13}C NMR chemical shifts (δ) of labeled products. 2**α-^{13}C-1a**, 56.28 ppm; **3α-^{13}C-1a**, 55.98 ppm; **3α-^{13}C-1b**, 55.5 ppm; **3α-^{13}C-1c**, 55.5 ppm; **3α-^{13}C-1d**, 56.5 ppm.

^{13}C NMR study for the eneimine formation. A solution of 2 mg of **3α-^{13}C-1c** in 0.5 mL DMSO-d_6, containing 2 mg of 5% Pd on carbon, was then mixed with 0.5 mL of 0.1M pD = 7.2 phosphate D_2O buffer. The solution was degassed with argon for 5 min followed by purging with H_2 until the quinone color of the solution disappeared. The solution was then degassed again with argon for 5 min, and the catalyst was removed by centrifuge in a nitrogen box. The supernatant was then transferred into a 3 mm NMR tube for the ^{13}C NMR spectrum obtained on a 100 MHz instrument after 4 h of scanning.

Hydrolysis of 3α-^{13}C-1c. To a solution of 80 mg SM in 50 mL MeOH was added 60 mg 5% Pd on carbon. The solution was degassed with argon for 5 min, followed by passing H_2 through for 3 min. The solution was then degassed with argon for another 5 min and opened to air. The catalyst was filtered off through Celite and the filtrate was dried. The solid residue was purified by preparative TLC plate employing different eluents. The following products were isolated and characterized.

Ethyl 5-methoxy-3-methylindole-4,7-dione-2-carboxylate (10). Yield 2.36 mg (3.59%); TLC (CH_2Cl_2), R_f 0.62; ^{1}H NMR ($CDCl_3$) δ 9.61 (1H, bs, indole nitrogen proton), 5.79 (1H, s, 6-proton), 4.39 (2H, q, J = 7.2 Hz, 2-ethyl methylene), 3.86 (3H, s, 5-methoxy methyl), 2.63 (3H, d, J = 129.6 Hz, 3-methyl), 1.40 (3H, t, J = 7.2 Hz, 2-ethyl methyl); ^{13}C NMR ($CDCl_3$) δ 10.99; MS (EI) 264 (M^+), 245, 234, 217, 202.

Methyl 5-methoxy-3-methylindole-4,7-dione-2-carboxylate (11). Yield 0.31 mg (0.52%); TLC (CH_2Cl_2), R_f 0.61; ^{1}H NMR ($CDCl_3$) δ 9.61 (1H, bs, indole nitrogen proton), 5.79 (1H, s, 6-proton), 3.86 and 3.76 (6H, 2s, 2,5-methyls), 2.61 (3H, d, J = 129.6 Hz, 3-methyl); ^{13}C NMR ($CDCl_3$) δ 11.21; MS (EI) 250 (M^+), 234, 217, 202.

Ethyl 5-Methoxy-3-methoxymethyl-lindole-4,7-dione-2-carboxylate (12). Yield 1.73 mg (2.4%); TLC (CH_2Cl_2), R_f 0.56; ^{1}H NMR ($CDCl_3$) δ 9.61 (1H, bs, indole nitrogen proton), 5.78 (1H, s, 6-proton), 4.76 (2H, d, J = 132.3 Hz, 3-methylene), 4.37 (2H, q, J = 7.5 Hz, 2-ethyl methylene), 3.86 (3H, s, 5-methoxy methyl), 3.82 (3H, d, J = 6.6 Hz, 3-methoxy methyl), 1.40 (3H, t, J = 7.5 Hz, 2-ethyl methyl); ^{13}C NMR ($CDCl_3$) δ 64.73; MS (EI) 294 (M^+), 262, 250, 233, 202.

Ethyl 3-(ethyl 3′-methyl-indol-4′,7′-dione-2′-carboxylate-5′-yl) methyl-5-methoxy-indol-4,7-dione-2-carboxylate (13). Yield 2.37 mg (4.00%); TLC (CH_2Cl_2/acetone 98:2), R_f 0.73; ^{1}H NMR ($CDCl_3$) δ 9.83 and 9.58 (2H, 2bs, indole nitrogen protons), 5.96 (1H, t, J = 1.2 Hz, 6′-proton), 5.83 (1H, s, 6-proton), 4.38 (4H, m, methylenes of ethyls), 4.35 (2H, doublet of triplet, J_1 = 1.2 Hz, J_2 = 132.8 Hz, 3-methylene), 3.84 (3H, s, 5-methoxy methyl), 2.65 (3H, d, J = 129.6 Hz, 3′-methyl), 1.40 and 1.31 (6H, m, methyls of ethyls); ^{13}C NMR ($CDCl_3$) δ 24.58 and 10.90; MS (EI) 496 (M^+).

Ethyl 3-(ethyl 5′-Methoxy-3′-methyl-indole-4′-one-7′-hydroxy-7′-ly) methyl-5-methoxy-indol-4,7-dione-2-carboxylate (14). Yield 5.28 mg (7.73%); TLC (CH_2Cl_2/acetone 95:5), R_f 0.70; ^{1}H NMR ($CDCl_3$) δ 9.99 and 9.53 (2H, 2bs, indole nitrogen protons), 5.84 (1H, s, 6′-proton), 5.53 (1H, s, 6-proton), 4.28 and 4.16 (4H, m, methylenes of ethyls), 3.87 (6H, s, 5 and 5′-methoxy methyls), 3.74 (2H, dd, J_1 = 12.8 Hz, J_2 = 133.6 Hz, 3′-methylene), 1.69 (3H, d, J = 124 Hz, 3′-methyl), 1.33 and 1.21 (6H, m, methyls of ethyls); ^{13}C NMR ($CDCl_3$) δ 38.38 and 10.56; MS (EI) 528 (M^+). Anal. calcd ($C_{26}H_{26}N_2O_{10}$) : C, 59.09; H, 5.31; N, 5.33. Found: C, 58.68; H, 5.46; N, 5.12.

Ethyl 3-(ethyl 5′-methoxy-3′-acetoxymethyl-indole-4′-one-7′-hydroxy-7′-ly) methyl-5-methoxy-indol-4,7-dione-2-carboxylate (15). Yield 7.03 mg (10.0%); TLC (CH_2Cl_2/acetone 95:5), R_f 0.68; ^{1}H NMR ($CDCl_3$) δ 9.88 and 9.74 (2H, 2bs, indole nitrogen protons), 5.85 (1H, s, 6-proton), 5.67 (1H, s, 6′-proton), 4.72 (2H, dd, J_1 = 6.8 Hz, J_2 = 144.8 Hz, 3-methylene), 4.30 and 4.14 (4H, m, methylenes of ethyls), 3.87 (6H, s, 5 and 5′-methoxy methyls), 3.85 (2H, dd, J_1 = 12.8 Hz, J_2 = 164.8 Hz, 3′-methylene), 1.92 (3H, s, 3-acetate methyl), 1.32 and 1.22 (6H, m, methyls of ethyls); ^{13}C NMR ($CDCl_3$) δ 56.64 and 38.53; MS (EI) 584 (M^+). Anal. calcd ($C_{28}H_{28}N_2O_{12}$): C, 57.53; H, 4.83; N, 4.79. Found: C, 57.17; H, 5.02; N, 4.68.

To a solution of 80 mg SM in 50 mL MeOH was added 60 mg 5% Pd on carbon. The solution was degassed with argon for 5 min, followed by passing H_2 through for 3 min. The solution was then degassed with argon for another 5 min and opened to air. The catalyst was filtered off through Celite and the filtrate was dried. The solid residue was purified by preparative TLC plate employing different eluents. The following products were isolated and characterized (reaction in MeOH with $NaBH_4$ as the reductant affords the same hydrolysis products).

2,3-Dimethyl-5-methoxy-indol-4,7-dione (16). Yield 1.16 mg (2.27%); TLC (CH_2Cl_2/acetone 98:2), R_f 0.70; ^{1}H NMR ($CDCl_3$) δ 9.37 (1H, bs, indole nitrogen proton), 5.61 (1H, s, 6-proton), 3.81 (3H, s, 5-methoxy methyl), 2.24 (3H, d, J = 128.1 Hz, 3-methyl), 2.23 (3H, s, 2-methyl); ^{13}C NMR ($CDCl_3$) δ 9.69; MS (EI) 206 (M^+), 191.

2-Acetoxy-5-methoxy-3-methylindol-4,7-dione (17). Yield 1.10 mg (1.68%); TLC (CH_2Cl_2/acetone 98:2), R_f 0.68; ^{1}H NMR ($CDCl_3$) δ 9.42 (1H, bs, indole nitrogen proton), 5.69 (1H, s, 6-proton), 5.05 (2H, s, 2-methyl), 3.82 (3H, s, 5-methoxy methyl), 2.35 (3H, d, J = 128.7 Hz, 3-methyl), 2.09 (3H, s, 2-acetate methyl); ^{13}C NMR ($CDCl_3$) δ 9.58; MS (EI) 264 (M^+), 249, 209, 202, 149.

2-Acetoxy-5-methoxy-3-methoxymethylindol-4,7-dione (18). Yield 3.16 mg (4.33%); TLC (CH_2Cl_2/acetone 98:2), R_f 0.65; ^{1}H NMR ($CDCl_3$) δ 9.63 (1H, bs, indole nitrogen proton), 5.71 (1H, s, 6-proton), 5.20 (2H, s, 2-methylene), 4.71 (2H, d, J = 143.7 Hz, 3-methylene), 3.83 (3H, s, 5-methoxy methyl), 3.42 (3H, d, J = 5.7 Hz, 3-methoxy methyl), 2.02 (3H, s, 2-acetate methyl); ^{13}C NMR ($CDCl_3$) δ 64.57; MS (EI) 294 (M^+), 279, 261, 250, 233, 218, 204, 190. Anal. calcd ($C_{14}H_{15}NO_6$) : C, 57.13; H, 5.14; N, 4.76. Found: C, 56.97; H, 5.32; N, 4.73.

2-Acetoxymethyl-3-(2′-acetoxymethyl-5′-methoxy-3′-methyl-5′,6′-dihydroindol-4′,7′-dione-5′-yl) methyl-5-methoxyindol-4,7-dione (19). Yield 0.75 mg (1.15%); TLC (CH_2Cl_2/acetone 95:5), R_f 0.56; ^1H NMR ($CDCl_3$) δ 9.54, 9.19 (2H, 2bs, indole nitrogen protons), 5.76 (1H, s, 6-proton), 4.94 (2H, m, 2-methylene), 4.52 (2H, m, 2′-methylene), 3.87 (3H, s, 5-methoxy methyl), 3.85 and 3.02 (2H, m, 3-methylene), 3.63 (3H, s, 5′-methoxy methyl), 3.62 and 2.64 (2H, m, 6′-protons), 2.06 and 2.02 (6H, 2s, acetate methyls), 1.78 (3H, d, $J=128.1$ Hz, 3′-methyl); ^{13}C NMR ($CDCl_3$) δ 29.36, 8.03; MS (EI) 528 (M^+).

2-Acetoxymethyl-3-(2′,3′-diacetoxymethyl-5′-methoxy-5′,6′-dihydroindol-4′,7′-dione-5′-yl) methyl-5-methoxy-indol-4,7-dione (20). Yield 1.03 mg (1.42%); TLC (CH_2Cl_2/acetone 95:5), R_f 0.54; ^1H NMR ($CDCl_3$) δ 9.49, 9.31 (2H, 2bs, indole nitrogen protons), 5.72 (1H, s, 6-proton), 5.01 (2H, m, 2-methylene), 4.70 (2H, m, 2′-methylene), 4.28 (2H, m, 3′-methylene), 3.84 (3H, s, 5-methoxy methyl), 3.85 and 3.08 (2H, m, 3-methylene), 3.63 (3H, s, 5′-methoxy methyl), 3.61 and 2.72 (2H, m, 6′-protons), 2.07, 2.04 and 1.97 (9H, 3s, acetate methyls); ^{13}C NMR ($CDCl_3$) δ , 56.94, 29.36; MS (EI) 586 (M^+).

2-Acetoxymethyl-3-(2′-acetoxymethyl-3′-(2″,3″-diacetoxymethyl-5″-methoxy-indole-4″-one-7″-hydroxy-7″-ly-) methyl-5′-methoxy-indole-4′-one-7′-hydroxy-7′-ly-) methyl-5-methoxy-indol-4,7-dione (21). Yield 4.28 mg (6.27%); TLC (CH_2Cl_2/acetone 90:10), R_f 0.32; ^1H NMR ($CDCl_3$) δ 10.37, 9.96 and 9.75 (3H, 3bs, indole nitrogen protons), 5.77 (1H, s, 6″-proton), 5.55 and 5.34 (2H, 2s, 6, 6′-protons), 5.60 and 5.40 (2H, 2bs, 7, 7′-hydroxy protons), 5.06 (2H, m, 2′-methylene), 4.54 (2H, d, $J=93.6$ Hz, 3-methylene), 4.41 (2H, m, 2″-methylene), 3.89, 3.84 and 3.65 (9H, 3s, 5, 5′ and 5″-methoxy methyls), 4.12, 3.92, 3.74, 3.00 and 2.59 (4H, m, 3′and 3″-methylenes), 2.06, 2.01, 1.97 and 1.93 (12H, 4s, acetate methyls); ^{13}C NMR ($CDCl_3$) δ 56.84, 39.52 and 37.99; MS (EI) 828 (M^+). Anal. calcd ($C_{39}H_{41}N_3O_{17}$): C, 56.52; H, 5.34; N, 5.10. Found: C, 56.31; H, 5.52; N, 5.02.

2-Acetoxymethyl-3-(2′-acetoxymethyl-5′-methoxy-3′-methyl-indole-4′-one-7′-hydroxy-7′-ly) methyl-5-methoxy-indol-4,7-dione (22). Yield 0.78 mg (1.20%); TLC (CH_2Cl_2/acetone 90:10), R_f 0.53; ^1H NMR ($CDCl_3$) δ 9.57, 9.22 (2H, 2bs, indole nitrogen protons), 5.74 (1H, s, 6′-proton), 5.42 (1H, s, 6-proton), 4.99 (2H, m, 2′-methylene), 4.36 (2H, 2m, 2-methylene), 3.87 and 3.83 (6H, 2s, 5, 5′-methoxy methyls), 3.82, 3.62, 3.18 and 2.75 (2H, m, 3′-methylene), 2.07 and 2.03 (6H, 2s, acetate methyls), 1.72 (3H, d, $J=126$ Hz, 3-methy); ^{13}C NMR ($CDCl_3$) δ 38.46, 8.87; MS (EI) 528 (M^+).

2-Acetoxymethyl-3-(2′-acetoxymethyl-5′-methoxy-3′-methoxymethyl-indole-4′-one-7′-hydroxy-7′-ly-) methyl-5-methoxy-indol-4,7-dione (23). Yield 8.84 mg (12.2%); TLC (CH_2Cl_2/acetone 90:10), R_f 0.52; ^1H NMR ($CDCl_3$) δ 9.59 and 9.48 (2H, 2bs, indole nitrogen protons), 5.73 (1H, s, 6′-proton), 5.43 (1H, s, 6-proton), 5.10 (2H, m, 2′-methylene), 4.66 (2H, d, $J=147.0$ Hz, 3-methylene), 4.56 (2H, m, 2-methylene), 3.85 and 3.81 (6H, 2s, 5 and 5′-methoxy methyls), 3.98, 3.54, 3.29 and 2.89 (2H, m, 3′-methylene), 2.07, 2.04 and 1.99 (9H, 3s, acetate methyls); ^{13}C NMR ($CDCl_3$) δ 57.02 and 38.58; MS (EI) 586 (M^+). Anal. calcd ($C_{28}H_{28}N_2O_{12}$) : C, 57.33; H, 4.81; N, 4.80. Found: C, 57.18; H, 4.87; N, 4.69.

2-Acetoxymethyl-3-(2′, 3′-Diacetoxymethyl-5′-methoxy-indole-4′-one-7′-hydroxy-7′-ly-) methyl-5-methoxy-indol-4,7-dione (24). Yield 0.65 mg (1.00%); TLC (CH_2Cl_2/acetone 90:10), R_f 0.51; ^1H NMR ($CDCl_3$) δ 9.78 and 9.60 (2H, 2bs, indole nitrogen protons), 5.70 (1H, s, 6′-proton), 5.41 (1H, s, 6-proton), 5.09 (2H, m, 2′-methylene), 4.58 (2H, m, 2-methylene), 4.46 (2H, d, $J=92.4$ Hz, 3-methylene), 3.84 and 3.79 (6H, 2s, 5 and 5′-methoxy methyls), 3.89, 3.54, and 3.00 (2H, m, 3′-methylene), 3.23 (3H, d, $J=5.7$ Hz, 3-methoxy methyl), 2.07 and 2.04 (6H, 2s, acetate methyls); ^{13}C NMR ($CDCl_3$) δ 65.05 and 38.15; MS (EI) 558 (M^+).

5′-Hydroxy-7,3′,7′-trimethoxy-1,2,3,4,2′,3′,4′,5′-octahydro-1′*H*-[3,5′]bi[cyclopent[*b*] indolyl]-5,8,8′-trione (25a). TLC (CH_2Cl_2) R_f 0.55; ^1H NMR ($CDCl_3$) δ 9.47 and 9.04 (2H, 2bs, 4 and 4′ protons), 5.65 and 5.46 (2H, 2s, 6 and 6′ protons), 4.77 (1H, m, 3′-proton), 3.87 (1H, m, 3-proton), 3.85, 3.48, and 3.38 (9H, 3s, 7, 3′, and 7′ methoxys), 2.89, 2.75, 2.47, 2.28, 2.09, 1.88, and 0.87 (8H, mm, 1, 2, 1′, and 2′ protons).

3′-Acetoxy-5′-hydroxy-7,7′-dimethoxy-1,2,3,4,2′,3′,4′,5′-octahydro-1′*H*-[3,5′]bi[cyclopent[*b*] indolyl]-5,8,8′-trione (25b). TLC (CH_2Cl_2) R_f 0.52; ^1H NMR ($CDCl_3$) δ 9.66 and 9.40 (2H, 2bs, 4 and 4′ protons), 5.61 and 5.40 (2H, 2s, 6 and 6′ protons), 4.60 (1H, m, 3′-proton), 3.84 (1H, m, 3-proton), 3.85 and 3.34 (6H, 2s, 7 and 7′ methoxys), 3.59 (1H, bs, 5′-hydroxy proton), 2.61, 2.45, 2.01, 1.73, 1.10, and 0.87 (8H, mm, 1, 2, 1′, and 2′ protons).

Preparation of the ^{13}C-labeled indoloquinone–DNA adduct

The DNA hexamer $d(ATCGAT)_2$ was prepared by the phosphorimidate method and purified by 20% preparative polyacrylamide gel. The DNA adducts were prepared by mixing 8.0 mg (2.0 μmol of strand) of $d(ATCGAT)_2$ in 0.05 M of Tris buffer (pH 7.4) with 1.0 mg (3.0 μmol) of the quinone in 0.25 mL of DMSO. To the mixture was added 0.2 mg of 5% Pd on carbon. The resulting mixture was degassed under argon for 30 min and followed by purging with H_2 at 1 atm for 20 min. The mixture was then purged with argon for 10 min and incubated at 30 °C for 2 h. The reaction was opened to the air and the catalyst was centrifuged off. To the supernatant was added 0.45 mL of 7.5 M ammonium acetate, 10 mL of cold ethanol and left in a −20 °C freezer for 24 h. The mixture was centrifuged at 12,000*g* for 20 min and the supernatant was removed. The DNA pellet was redissolved in water, ethanol precipitated again and dried. The pellet was then dissolved in D_2O and lyophilized twice before diluting in 0.7 mL 100% D_2O.

Materials for sequencing studies

Quinone stock solutions were made up in DMSO at 10, 1, and 0.1 mM. Electrophoresis-grade acryalmide and bis (acrylamide) were purchased from Sigma, ultrapure urea and agarose from GIBCO-BRL Life Technologies (Grand Island, NY, USA), and piperidine from Sigma (St Louis, MI, USA). The pBR322 plasmid DNA, restriction enzymes *Eco*RI and *Rsa*I, T4 polynucleotide kinase (PNK), and bacterial alkaline phosphatase (BAP) were purchased from Gibco BRL Life Technologies. The 5′-[γ-^{32}P]ATP were purchased from New England Nuclear Research Products (Wilmington, DE, USA). For sequencing labeled DNA, ^{32}P-end labeled 514 base pair restriction fragments were prepared by first digesting supercoiled pBR322 plasmid DNA with EcorR I restriction endonuclease. 5′-end labeling was achieved by the treatment of the EcorR I pre-cleaved DNA with alkaline phosphatase, [γ-^{32}P]ATP and T4 polynucleotide kinase. Following the end-labeling, the DNA was further digested with RsaI to yield the 514 base pair fragment that was purified by 6% preparative non-denaturing gel electrophoresis with TBE buffer (TBE electrophoresis buffer is 90mM Tris, 90mM boric acid, and 2mM EDTA at pH 8.3).

Cleavage of ^{32}P-end labeled DNA. Indoloquinone-induced cleavage reactions were carried out in two different reaction conditions. For catalytic reduction, the reaction was carried out in 200 μL total volumes containing calf thymus DNA (500 μM nucleotide concentration) 2×10^4 cpm ^{32}P-end labeled restriction fragment and 100 μM of drug in DMSO in 25 mM phosphate buffer (pH 7.4) with 0.05 mg 5% Pd on carbon as catalyst. Reaction mixtures were degassed under argon for 5 min before being initiated by the bubbling of H_2 at 1 atm for 30 s. The reaction was then degassed under argon for 5 min and incubated at 37 °C for 30 min. For sodium boromhydride reduction, the reaction was carried out in 200 μL total volumes containing calf thymus DNA (500 μM nucleotide concentration) 2×10^4 cpm ^{32}P-end labeled restriction fragment and 100 μM of drug in DMSO in 25 mM phosphate buffer (pH 7.4). Reaction mixtures were degassed under argon for 5 min before being initiated by the addition of 100 mM sodium boromhydride in water. The reaction incubated at 37×C for 30 min.Under both reaction condition, the reactions were quenched by ethanol precipitation in the presence of 0.3M sodium acetate. The DNA pellet was washed with 70% ethanol, dried, and resuspended in 3 μL of 80% formamide loading dye. Piperidine treatment was also carried out by dissolving the DNA pellet in 20 μL of fresh 0.2 M piperidine, heated at 90 °C for 10 min and lyophilized. The samples, along with the Maxam-Gilbert A + G sequencing reaction mixtures, were then loaded onto 12% polyacrylamide/7.5 M urea denaturing gels and electrophoresed at 1200 V for 5 h in the TBE buffer described above. Autoradiography of the gel was conducted at −70 °C using Kodak X-omat film.

Modeling into DNA

INSIGHT II from Molecular Simulations, Inc. (San Diego) was used for modeling studies and consistent valence force field (CVFF) was utilized for all minimization protocols.

The cyclopent[*b*]indole quinone methide and the B-DNA sequence were constructed utilizing fragments available in fragment libraries of BUILDER module and minimized utilizing consistent valence force field (500–1000 iterations). The quinone methide reaction center was contrained to within a bond length of either the guanine N(7) or thymine O(6) nucleophiles. The best-fit structure was achieved by using the DOCKING module of INSIGHT II employing the consistent valence forcefield (CVFF). The orientation with the lowest energy (van der Waals and electrostatic) was chosen as the optimal DNA-bound structure.

Acknowledgements

The authors gratefully acknowledge support from the National Institutes of Health, National Science Foundation, and the Arizona Disease Control Research Commission.

References and Notes

1. Wemmer, D. E.; Dervan, P. B. *Curr. Opin. Struct. Biol.* **1997**, *7*, 355.
2. White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B. *Nature* **1998**, *391*, 468.
3. Dervan, P. B.; Burli, R. W. *Curr. Opin. Chem. Biol.* **1999**, *3*, 688.
4. Herzberg, R. P.; Dervan, P. B. *J. Am. Chem. Soc.* **1982**, *104*, 313.
5. Van Dyke, M. W.; Herzberg, R. P.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 5470.
6. Taylor, J. S.; Schultz, P. G.; Dervan, P. B. *Tetrahedron* **1984**, *40*, 457.
7. Iida, H.; Jia, G. F.; Lown, J. W. *Curr. Opin. Biotechnol.* **1999**, *10*, 29.
8. Fox, K. R.; Yan, Y. F.; Gong, S. *Anti-Cancer Drug Des.* **1999**, *14*, 219.
9. Cozzi, P. *Farmaco* **2000**, *55*, 168.
10. Denny, W. A. *Expert Opin. Ther. Patents* **2000**, *10*, 459.
11. Hamy, F.; Albrecht, G.; Florsheimer, A.; Bailly, C. *Biochem. Biophys. Res. Commun.* **2000**, *270*, 393.
12. Guelev, V. M.; Harting, M. T.; Lokey, R. S.; Iverson, B. L. *Chem. Biol.* **2000**, *7*, 1.
13. Nielsen, P. E. Antisense properties of peptide nucleic acid. In *Antisense Technology, Pt A*; Phillips, M. I., Ed.; Academic: San Diego, 2000; p 156.
14. Nielsen, P. E. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 167.
15. Majumdar, A.; Khorlin, A.; Dyatkina, N.; Lin, F. L. M.; Powell, J.; Liu, J.; Feiz, Z. Z.; Khripine, Y.; Watanabe, K. A.; George, J.; Glazer, P. M.; Seidman, M. M. *Nat. Genet.* **1998**, *20*, 212.
16. Thurston, D. E. *Brit. J. Cancer* **1999**, *80*, 65.
17. Ren, J. S.; Chaires, J. B. *Biochemistry* **1999**, *38*, 16067.
18. Crooke, S. T. *Bba Gene Struct. Express* **1999**, *1489*, 31.
19. Huang, X.; Suleman, A.; Skibo, E. B. *Bioorg. Chem.,* **2000**, *28*, 324.
20. Schulz, W. G.; Nieman, R. A.; Skibo, E. B. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 11854.

21. Skibo, E. B.; Schulz, W. G. *J. Med. Chem.* **1993**, *36*, 3050.
22. Xing, C.; Skibo, E. B.; Dorr, R. T. *J. Med. Chem.*, submitted for publication.
23. Xing, C. G.; Wu, P.; Skibo, E. B.; Dorr, R. T. *J. Med. Chem.* **2000**, *43*, 457.
24. Ouyang, A.; Skibo, E. B. *Biochemistry* **2000**, *39*, 5817.
25. Gutierrez, P. L. *Front Biosci.* **2000**, *5*, D629.
26. Kang, H. M.; Rokita, S. E. *Nucleic Acids Res.* **1996**, *24*, 3896.
27. Zeng, Q. P.; Rokita, S. E. *J. Org. Chem.* **1996**, *61*, 9080.
28. Rokita, S. E.; Yang, J. H.; Pande, P.; Greenberg, W. A. *J. Org. Chem.* **1997**, *62*, 3010.
29. Lemus, R. L.; Skibo, E. B. *J. Org. Chem.* **1988**, *53*, 6099.
30. Skibo, E. B. *J. Org. Chem.* **1992**, *57*, 5874.
31. Skibo, E. B. *J. Org. Chem.* **1986**, *51*, 522.
32. Zhou, Q. B.; Turnbull, K. D. *J. Org. Chem.* **1999**, *64*, 2847.
33. Franck, R. W.; Tomasz, M. The Chemistry of Mitomycins. In *The Chemistry of Antitumor Agents*; Wilman, D. E., Ed.; Blackie & Sons; Glasgow, UK, 1990; p 379.
34. Boruah, R. C.; Skibo, E. B. *J. Org. Chem.* **1995**, *60*, 2232.
35. Skiles, G. L.; Yost, G. S. *Chem. Res. Toxicol.* **1996**, *9*, 291.
36. Ruangyuttikarn, W.; Skiles, G. L.; Yost, G. S. *Chem. Res. Toxicol.* **1992**, *5*, 713.
37. Coulson, C. A.; Streitwieser, A. *Dictionary of π-Electron Calculations*; W. H Freeman: San Francisco, 1965.
38. Kim, H. S.; LeBreton, P. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 3725.
39. Kim, H. S.; Yu, M.; Jaing, Q.; LeBreton, P. R. *J. Am. Chem. Soc.* **1993**, *115*, 6169.
40. Broch, H.; Hamza, A.; Vasilescu, D. *J. Biomol. Struct. Dyn.* **1996**, *13*, 903.
41. Hamza, A.; Broch, H.; Vasilescu, D. *J. Biomol. Struct. Dyn.* **1996**, *13*, 915.
42. Maxam, A. M.; Gilbert, W. *Methods Enzymol.* **1980**, *65*, 499.
43. Maliepaard, M.; deMol, N. J.; Tomasz, M.; Gargiulo, D.; Janssen, L. H. M.; vanDuynhoven, J. P. M.; vanVelzen, E. J. J.; Verboom, W.; Reinhoudt, D. N. *Biochemistry* **1997**, *36*, 9211.
44. Maliepaard, M.; Wolfs, A.; Groot, S. E.; de Mol, N. J.; Janssen, L. H. M. *Br. J. Cancer* **1995**, *71*, 836.
45. Workman, P.; Binger, M.; Kooistra, K. L. *Int. J. Radiat. Oncol. Biol. Phys* **1992**, *22*, 713.