Groove- and Sequence-Selective Alkylation of DNA by Sulfonate Esters Tethered to Lexitropsins†

Yi Zhang, Fa-Xian Chen, Pratibha Mehta, and Barry Gold*

Eppley Institute for Research in Cancer and Allied Diseases and Department of Pharmaceutical Sciences, University of Nebraska Medical Center, Omaha, Nebraska 68198-6805

Received December 3, 1992; Revised Manuscript Received March 18, 1993

ABSTRACT: A series of sulfonate esters that are attached to a noncationic minor-groove-binding *N*-methylpyrrole dipeptide (Lex) related to netropsin have been synthesized. The compounds prepared differ in two respects: (1) the length [$(CH_2)_2$ vs $(CH_2)_8$] of the tether between the DNA affinity binding portion of the molecule and the sulfonate ester and (2) whether a methyl group [$MeOSO_2(CH_2)_n$–Lex] or the dipeptide including the aliphatic tether [$MeSO_2O(CH_2)_n$–Lex] is covalently transferred to the DNA. The DNA-cleavage patterns of these bimolecular alkylating compounds have been mapped in ^{32}P-end-labeled restriction fragments using neutral thermal hydrolysis and alkali treatment to expose single-strand breaks at bases with thermally labile modifications. In contrast to the alkylation of DNA by simple alkyl alkanesulfonate esters, that predominantly yield major-groove alkylation at *N*7-guanine, the modification of DNA by $MeOSO_2(CH_2)_n$–Lex and $MeSO_2O(CH_2)_n$–Lex occurs primarily at *N*3-adenine residues associated with previously footprinted Lex DNA affinity binding regions. The ratio for the formation of *N*3-methyladenine (minor groove) to *N*7-methylguanine (major groove) in calf thymus DNA is 1:7 for dimethyl sulfate, while only the former adenine product is observed with $MeOSO_2(CH_2)_n$–Lex, indicating the change in groove specificity. DNA cleavage by $MeOSO_2(CH_2)_n$–Lex and $MeSO_2O(CH_2)_n$–Lex is efficiently inhibited by the coaddition of distamycin; however, only the DNA damage generated by the latter is blocked by NaCl. As expected, increasing the length of the $(CH_2)_n$ tether from $n = 2$ to $n = 8$ moves the alkylation site by 1–2 base pairs further from the affinity binding domain. Finally, a comparison of the methylation patterns of $MeOSO_2(CH_2)_n$–Lex as a function of tether length provides an insight into Lex sequence and orientational preferences.

The cytotoxicity of many antibiotics, including some of clinical value, is a direct consequence either of their affinity binding to DNA or of a subsequent event that is initiated by complexation with DNA (Sarma & Sarma, 1988). The minor groove of DNA frequently serves as a receptor for these compounds, and binding often involves A/T-rich regions. The antibiotics distamycin and netropsin (Figure 1), based on *N*-methylpyrrolecarboxamide subunits and commonly referred to as lexitropsins (Lex),[1] are classic examples of such minor-groove binders, and their interactions with DNA have been probed by NMR (Patel, 1982; Patel & Shapiro, 1968; Lee et al., 1988; Pelton & Wemmer, 1989; Pelton & Wemmer, 1990; Sarma et al., 1990), crystallography (Kopka et al., 1985; Coll et al., 1987; Coll et al., 1989), visible, UV, or CD spectroscopy (Luck et al., 1974; Zimmer et al., 1975), thermodynamic studies (Marky & Breslauer, 1987; Dabrowiak et al., 1990) and chemical methods (Shultz et al., 1982; Taylor et al., 1984; Shultz et al., 1984; Portugal & Waring, 1987; Baker et al., 1989; Church et al., 1990; Churchill et al., 1990). On the basis of these studies it is generally agreed that VDW contacts, H-bonding, and electrostatics stabilize the DNA–drug complex and play a role in sequence-selective DNA recognition. The specific details of binding include (a) the formation of H-bonds between the peptide amide N-H's and the N3-A and O^2-T atoms that line the floor of the minor groove, (b) H-bonds between the ionized terminal amidine or guanidinium groups and the same DNA H-bond acceptor atoms, (c) VDW contacts between the pyrrole C3-H's and the C2-H of A, and (d) VDW contacts between the aliphatic terminus and the C2-H of A (Figure 1).

While the above information regarding Lex binding to DNA provides a basis to rationalize experimentally verified binding domains, it is not possible to predict a priori which A/T-rich regions will be preferred binding sites within large DNA fragments. In fact, it is not uncommon to observe a G residue contained in a strong binding site (Shultz et al., 1982; Taylor et al., 1984; Shultz & Dervan, 1984; Youngquist & Dervan, 1985; Churchill et al., 1990). It is assumed that VDW interactions are pivotal in the sequence recognition, and the precise sequence-dependent conformational requirements for this type of binding are not readily predicted or modeled.

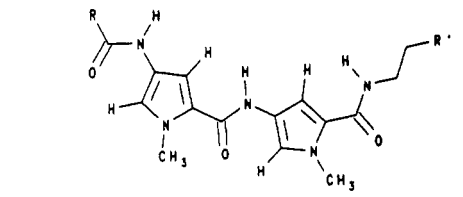
Because of an interest in regulating groove- and sequence-selective DNA adduct formation by small alkylating agents, we have prepared minor-groove binders related to netropsin with alkyl sulfonate ester functionalities. In designing these molecules several important changes have been made with respect to the natural product. The amidine and guanidinium groups on the C- and N-termini of netropsin have been replaced, respectively, with an *N*-propyl amide and a sulfonate ester ($MeOSO_2$– or $MeSO_2O$–) attached with a $(CH_2)_n$ chain ($n = 2$ or 8) (Figure 1). This deletion of charged terminal groups removes the electrostatic interactions between the

† This work was supported by USPHS Research Grant CA29088 and Core Grant CA36727 awarded by the National Cancer Institute and by American Cancer Society Center Grant ACS SIG-16.

* Author to whom correspondence should be addressed.

[1] Abbreviations: bp, base pair; $BrCH_2CO$-Dis, *N*-(bromoacetyl)distamycin; DMF, *N*,*N*-dimethylformamide; DMS, dimethyl sulfate; EDTA, ethylenediaminetetraacetic acid; FAB-MS, fast-atom bombardment mass spectrometry; HPLC, high-performance liquid chromatography; Lex, lexitropsin (information-reading peptide); MMS, methyl methanesulfonate; MNU, *N*-methyl-*N*-nitrosourea; *N*3-MeA, *N*3-methyladenine; *N*7-MeG, *N*7-methylguanine; MPE, methidiumpropyl–EDTA; O^6-MeG, O^6-methylguanine; THF, tetrahydrofuran; TLC, thin-layer chromatography; Tris, tris(hydroxymethyl)aminomethane; VDW, van der Waal.

0006-2960/93/0432-7954$04.00/0



netropsin: R = $CH_2NHC(=NH)NH_2$; R' = $C(=NH)NH_2$

$MeOSO_2(CH_2)_n$-lex: R = $CH_3OSO_2(CH_2)_n$; R' = H (n = 2 or 8)

$MeSO_2O(CH_2)_n$-lex: R = $CH_3SO_2O(CH_2)_n$; R' = H (n = 2 or 8)

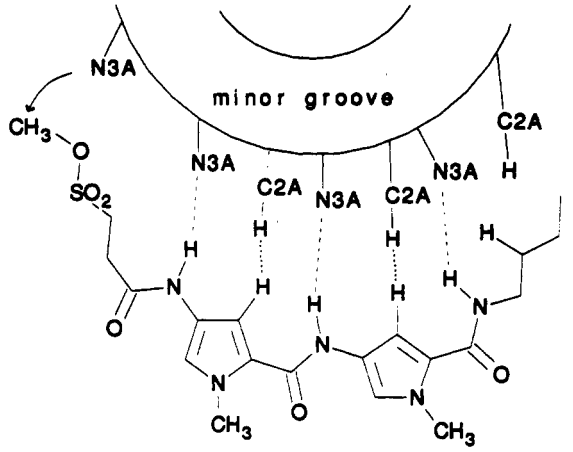


FIGURE 1: Structure of compounds and model for equilibrium binding of Lex to DNA.

peptide and DNA and was intended to reduce the number of factors involved in the equilibrium binding process, and possibly increase the degree of cellular uptake. The alkylating property of the molecules, as mapped in DNA restriction fragments, serves to report where the peptide binds, since the mechanism of alkylation involves a concerted process (Lawley, 1984). Finally, the employment of hydrocarbon linkers of different lengths provides confirmation of the relationship between binding and bonding sites as well as orientational preferences within binding sites.

MATERIALS AND METHODS

^1H and ^{13}C NMR were recorded on a Varian XL-300 spectrometer in DMSO-d_6, and 2D-NMR techniques (COSY and phase-sensitive DQCOSY) were used in structural assignments. Mass spectral data were collected on an AEI MS-9 (Eppley Institute, University of Nebraska Medical Center) or a Kratos MS-50 (Midwest Center for Mass Spectrometry, University of Nebraska—Lincoln) spectrometer utilizing fast-atom bombardment ionization techniques. Analytical and preparative silica TLC employed 0.25- and 2.0-mm Merck Kieselgel glass plates. All flash column chromatography was performed with 40 μm silica gel.

Synthesis (Figure 2). (*A*) [*1-Methyl-4-[1-methyl-4-(2-propenamido)pyrrole-2-carboxamido]pyrrole-2-carboxamido]propane* (**1a**). [1-Methyl-4-(1-methyl-4-nitropyrrole-2-carboxamido)pyrrole-2-carboxamido]propane (333 mg, 1 mmol) (Church et al., 1990), 10% Pd/C (600 mg), and cyclohexene (15 mL) in 95% EtOH (50 mL) were refluxed for 4–6 h. The mixture was filtered and the catalyst thoroughly washed with MeOH. The combined filtrates were concentrated in vacuo, and the residue containing [1-methyl-4-(1-methyl-4-aminopyrrole-2-carboxamido)pyrrole-2-carboxamido]propane was dissolved in acetone (25 mL) containing diisopropylethylamine (600 μL, 3 mmol). The flask was cooled to −20 °C, and acryloyl chloride (90 μL, 1.2 mmol) was slowly added. The reaction was stirred at −10 °C for 20 h and then concentrated in vacuo. Upon addition of H_2O to the residue, a solid formed which was collected, dried, and purified by silica column chromatography (EtOAc): yield 221 mg (62%); ^1H NMR δ 10.11 (s, 1 H, CONH), 9.89 (s, 1 H, CONH), 8.00 (t, 1 H, CONH), 7.27 (d, 1 H, pyrrole CH), 7.18 (d, 1 H, pyrrole CH), 6.92 (d, 1 H, pyrrole CH), 6.85 (d, 1 H, pyrrole CH), 6.36 (q, 1 H, CH_2=C*H*CO), 6.19 (q, 1 H, *E*-HC*H*=CHCO), 5.66 (q, 1 H, *Z*-HC*H*=CHCO), 3.84 (s, 3 H, NCH_3), 3.79 (s, 3 H, NCH_3), 3.12 (m, 2 H, $CH_2CH_2CH_3$), 1.48 (m, 2 H, CH_2CH_3), 0.86 (t, 3 H, CH_2CH_3); ^{13}C NMR δ 161.09 (CONH), 160.65 (CONH), 157.74 (CONH), 130.96 (CH_2=CHCO–), 124.95 (CH_2=CHCO–), 122.52, 121.39, 121.13, 117.85, 117.18, 103.56, 103.43, 35.57, 35.32, 22.01, 10.86.

(*B*) [*1-Methyl-4-[1-methyl-4-(3-sulfopropanamido)pyrrole-2-carboxamido]pyrrole-2-carboxamido]propane Ammonium Salt* (**1b**). The olefin **1a** prepared above (178 mg, 500 μmol), 45% NH_4HSO_3 (600 μL, 3 mmol), and 30% H_2O_2 (100 μL, 1 mmol) were refluxed in H_2O (15 mL) for 3 h. The solution was concentrated in vacuo and the resulting residue thoroughly triturated with anhydrous EtOH. The EtOH was concentrated in vacuo to yield 200 mg (91%) of product: ^1H NMR δ 9.97 (s, 1 H, CONH), 9.85 (s, 1 H, CONH), 8.00 (t, 1 H, CONH), 7.17 (d, 1 H, pyrrole CH), 7.15 (d, 1 H, pyrrole CH), 7.11 (br s, 4 H, NH_4), 6.85 (s, 2 H, pyrrole CH's), 3.82 (s, 3 H, NCH_3), 3.79 (s, 3 H, NCH_3), 3.10 (m, 2 H, $NHCH_2CH_2CH_3$), 2.69 (t, 2 H, SCH_2), 2.54 (t, 2 H, SCH_2CH_2), 1.49 (m, 2 H, CH_2CH_3), 0.87 (t, 3 H, CH_2CH_3); ^{13}C NMR δ 167.94 (CONH), 160.67 (CONH), 157.82 (CONH), 122.52, 122.13, 121.53, 121.47, 117.53, 117.12, 103.55, 103.36, 47.02 (SCH_2), 35.47, 35.29, 31.97, 22.01, 10.87; IR (KBr disk) 3189, 2359, 2246, 1633, 1619, 1543, 1174, 1043 cm^{-1}; MS (FAB, Xe, 9 keV) *m/z* 440 (12, M + 1 − NH_3).

(*C*) [*1-Methyl-4-[1-methyl-4-(3-sulfopropanamido)pyrrole-2-carboxamido]pyrrole-2-carboxamido]propane* (**1c**). The ammonium sulfonate salt **1b** (50 mg, 110 μmol) was dissolved in H_2O and the pH adjusted to 2.0 with 1 N HCl. After 5–10 min, the solution was concentrated in vacuo and the remaining solid was repeatedly triturated with MeOH. The combined MeOH extracts were concentrated to afford the corresponding sulfonic acid in quantitative yield: ^1H NMR δ 10.02 (s, 1 H, CONH), 9.89 (s, 1 H, CONH), 8.02 (t, 1 H, CONH), 7.18 (s, 1 H, pyrrole CH), 7.16 (s, 1 H, pyrrole CH), 6.87 (s, 1 H, pyrrole CH), 6.85 (s, 1 H, aryl CH), 3.82 (s, 3 H, NCH_3), 3.79 (s, 3 H, NCH_3), 3.11 (m, 2 H, $NHCH_2CH_2CH_3$), 2.70 (t, 2 H, SCH_2), 2.56 (t, 3 H, SCH_2CH_2), 1.48 (m, 2 H, CH_2CH_3), 0.87 (t, 3 H, CH_2CH_3); ^{13}C NMR δ 167.72 (CONH), 160.68 (CONH), 157.82 (CONH), 122.52, 122.14, 121.46, 117.54, 117.14, 103.56, 103.37, 46.97 (SCH_2), 35.48, 35.30, 31.69, 22.00, 10.86; IR (KBr disk) 3133, 1653, 1560, 1401, 1220, 1189, 1173, 1046 cm^{-1}.

(*D*) [*1-Methyl-4-[1-methyl-4-(3-(methoxysulfonyl)propanamido)pyrrole-2-carboxamido]pyrrole-2-carboxamido]propane* ($MeOSO_2(CH_2)_7$–*Lex*). The sulfonic acid (50 mg, 118 μmol) was suspended in anhydrous THF (10 mL), cooled in an ice–H_2O bath, and repeatedly treated with ethereal CH_2N_2 until all the starting material had dissolved in the THF. The solvent was removed in vacuo and the residue purified by preparative TLC (silica, EtOAc): yield 29 mg

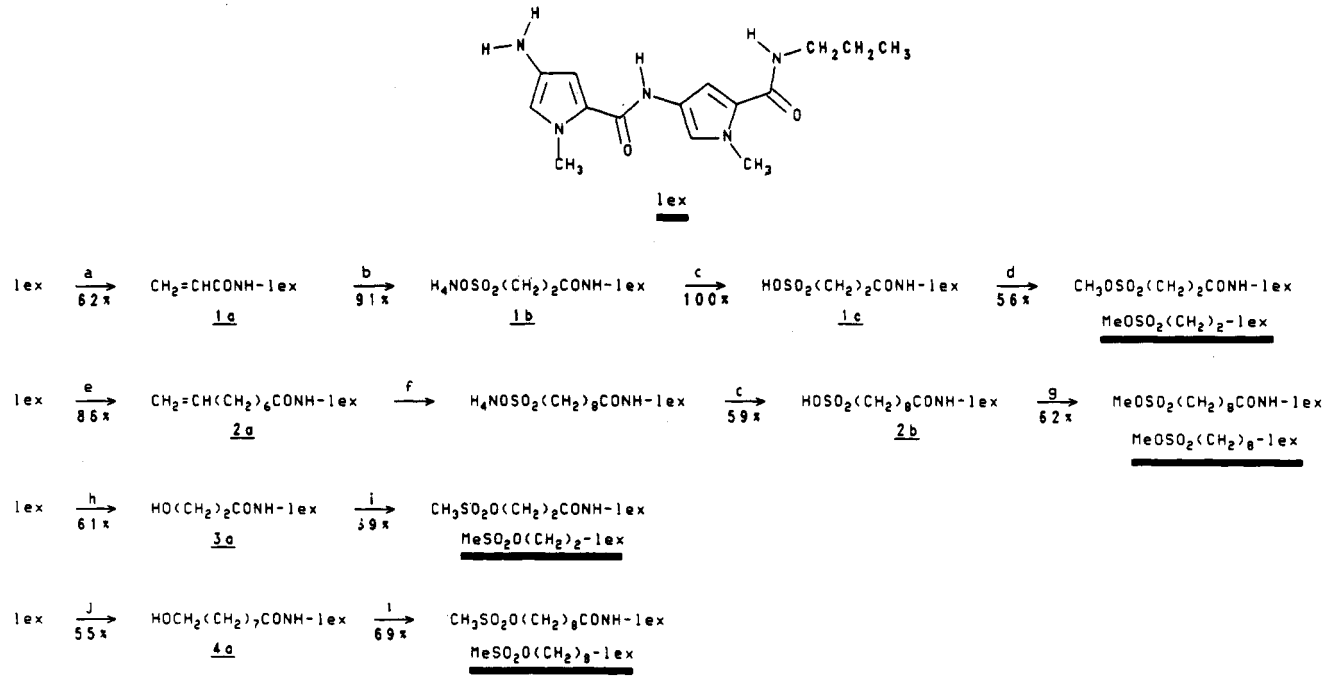


FIGURE 2: Synthesis of $MeOSO_2(CH_2)_n$–Lex and $MeSO_2O(CH_2)_n$–Lex compounds: a, CH_2CH= COCl; b, $PhCO_3$-*t*-Bu–NH_4HSO_3; c, pH 2.0; d, CH_2N_2–MeOH; e, CH_2=$CH(CH_2)_6COCl$; f, H_2O_2–NH_4HSO_3; g, *p*-NO_2-PhNHN=NCH3; h, β-propiolactone; i, $MeSO_2Cl$; j, $HOCH_2(CH_2)_7CO_2H$–dipyridyl sulfide–Ph_3P.

(56%): UV (MeOH) 234 and 294 nm; ^{1}H NMR δ 10.07 (s, 1 H, CONH), 9.87 (s, 1 H, CONH), 8.00 (t, 1 H, CONH), 7.17 (s, 2 H, pyrrole CH's), 6.87 (s, 1 H, pyrrole CH), 6.84 (s, 1 H, pyrrole CH), 3.86 (s, 3 H, OCH_3), 3.83 (s, 3 H, NCH_3), 3.79 (s, 3 H, NCH_3), 3.62 (t, 2 H, SCH_2), 3.11 (m, 2 H, $CH_2CH_2CH_3$), 2.74 (t, 2 H, SCH_2CH_2), 1.48 (m, 2 H, CH_2CH_3), 0.86 (t, 3 H, CH_2CH_3); ^{13}C NMR δ 167.96 (CONH), 160.66 (CONH), 157.81 (CONH), 122.52, 121.53, 121.47, 121.07, 117.58, 117.17, 103.55, 103.35, 47.01 (CH_3OS), 43.52 (OSO_2CH_2), 35.46, 35.30, 31.99, 22.00, 10.86; MS (FAB, Xe, 9 keV) *m/z* 454 (20, M + 1), 395 (23, M – $NHCH_2CH_2CH_3$), 273 (73, M – 180), 358 (37); IR (KBr disk) 3384, 2961, 2362, 1637, 1582, 1540, 1437, 1362, 1166 cm^{-1}.

(*E*) [*1-Methyl-4-[1-methyl-4-(8-nonenamido)pyrrole-2-carboxamido]pyrrole-2-carboxamido]propane* (**2a**). To Mg turnings (0.6 g, 25 mmol) suspended in Et_2O (10 mL) containing a trace amount of I_2 was added 6-bromohexene (200 μL). When the reaction was initiated, additional 6-bromohexene (3.2 mL, 25 mmol) in 35 mL of Et_2O was introduced. After all the Mg had dissolved, the solution was refluxed for 15 min and then cooled to –50 °C. CuCl (5 mg) was added followed by the dropwise addition of ethyl acrylate (83 mg, 8.3 mmol) in 20 mL of Et_2O over a 3-h period with vigorous stirring. At each 15-min interval during the addition, another 5 mg of CuCl was added while maintaining the system under N_2 atmosphere. The last portion of CuCl was added just after completion of the addition of ethyl acrylate. A total of 65 mg (610 μmol) of CuCl was used. The cooling bath was then removed and the reaction mixture stirred at room temperature for 1 h. The dark solution was poured onto a mixture of crushed ice and concentrated HCl with vigorous stirring. The H_2O layer was extracted with Et_2O several times, and the combined extracts were dried ($MgSO_4$). Upon concentration, the ester was obtained as an oil (1.14 g). The crude ester was dissolved in a minimum amount of MeOH, and then KOH (4 g) in 15 mL of H_2O was added. The resulting solution was stirred at room temperature for 24 h and then poured onto a mixture of crushed ice and concentrated HCl with vigorous stirring. The solution was extracted with ether several times, and the combined ether layers were dried ($MgSO_4$). After evaporation of solvent, the residue was treated with $SOCl_2$ (15 mL) and the resulting solution refluxed for 1 h. The reaction solution was concentrated in vacuo, and the 8-nonenoyl chloride product was used in the next reaction step without further purification. Nitrodipetide (333 mg, 1 mmol) was reduced to the amine (as described above) and then dissolved in MeCN (25 mL) containing diisopropylethylamine (600 μL, 3 mmol) and cooled in an ice–H_2O bath. The 8-nonenoyl chloride (1.2 equiv) in MeCN (10 mL) was added dropwise. The resulting solution was stirred for 1 h and then concentrated in vacuo. The crude product was purified by flash chromatography (silica, EtOAc) to yield 378 mg (86%) of product: ^{1}H NMR δ 9.82 (s, 1 H, CONH), 9.75 (s, 1 H, CONH), 7.99 (t, 1 H, $CONHCH_2$), 7.15 (s, 1 H, pyrrole CH), 7.17 (s, 1 H, pyrrole CH), 6.85 (s, 2 H, pyrrole CH's), 5.75 (m, 1 H, CH_2dCH), 4.95 (m, 2 H, CH_2dCH), 3.83 (s, 3 H, NCH_3), 3.79 (s, 3 H, NCH_3), 3.10 (m, 2 H, $CH_2CH_2CH_3$), 2.20 (t, 2 H, CH_2CO), 1.98 (m, 2 H, CH_2dCHCH_2), 1.55 (m, 2 H, CH_2CH_2CO), 1.48 (m, 2 H, CH_2CH_3), 1.28 (m, 6 H, $(CH_2)_3$), 0.86 (t, 3 H, CH_2CH_3); ^{13}C NMR δ 168.80 (CONH), 160.53 (CONH), 157.68 (CONH), 138.07 (CH_2=CH), 122.40, 122.03, 121.37, 121.32, 117.37, 117.00, 113.94 (CH_2dCH), 103.41, 103.23, 35.34, 35.18, 34.93, 32.44, 27.83, 27.60, 27.49, 24.63, 21.88, 10.72; MS (FAB, Xe, 9 keV) *m/z* 442 (25, M + 1); IR (KBr disk) 3239, 2929, 2854, 1633, 1544, 1539, 1465, 1434 cm^{-1}.

(*F*) [*1-Methyl-4-[1-methyl-4-(9-sulfononanamido)pyrrole-2-carboxamido]pyrrole-2-carboxamido]propane* (**2b**). Olefin **2a** (144 mg, 326 μmol), *tert*-butyl peroxybenzoate (3.16 g, 16.3 μmol), and 45% NH_4HSO_3 (220 μL, 1 mmol) were dissolved in MeOH containing sufficient H_2O to make a clear solution. The solution was stirred at 60 °C for 4 h, the solvent removed in vacuo, and the residue extracted with EtOH. The ammonium salt product was purified by flash chromatography (silica, MeOH–EtOAc, 3:7) and then dissolved in H_2O, and the pH was adjusted to 2.0 with 1 N HCl. The resulting solution was stirred at room temperature for 15 min and the

solvent removed in vacuo to afford 100 mg (59%) of product: ^{1}H NMR δ 9.86 (s, 1 H, CONH), 9.77 (s, 1 H, CONH), 8.00 (t, 1 H, CON*H*CH_2), 7.17 (s, 1 H, pyrrole CH), 7.15 (s, 1 H, pyrrole CH), 6.84 (s, 2 H, pyrrole CH's), 3.81 (s, 3 H, NCH_3), 3.78 (s, 3 H, NCH_3), 3.11 (m, 2 H, $CH_2CH_2CH_3$), 2.38 (t, 2 H, SO_3CH_2), 2.21 (t, 2 H, CH_2CO), 1.25–1.56 (br m, 14 H, $(CH_2)_6$ and CH_2CH_3), 0.86 (t, 3 H, CH_2CH_3); MS (FAB, Xe, 9 keV) m/z 524 (1, M + 1); IR (KBr disk) 3408, 2930, 2855, 1653, 1559, 1437, 1209 cm^{-1}.

(*G*) [*1-Methyl-4-*[*1-methyl-4-*(*9-*(*methoxysulfonyl*)*nonanamido*)*pyrrole-2-carboxamido*]*pyrrole-2-carboxamido*]*propane* (*MeOSO*$_2$(*CH*$_2$)$_8$–*Lex*). Sulfonic acid **2b** (30 mg, 57 μmol) in THF (10 mL) was cooled in an ice–H_2O bath, and methyl-*p*-tolyltriazine (10 mg, 67 μmol) in THF was added dropwise. The resulting solution was stirred at room temperature for 4 h, and the solvent was then removed in vacuo. The product was purified by preparative TLC (silica, MeOH–CH_2Cl_2, 1:9, R_f 0.44) to yield 18.9 mg (62%): ^{1}H NMR δ 9.84 (s, 1 H, CONH), 9.76 (s, 1 H, CONH), 7.99 (t, 1 H, CONH), 7.15 (m, 2 H, pyrrole CH's), 6.84 (m, 2 H, pyrrole CH's), 3.83 (s, 3 H, CH_3OSO_2), 3.81 (s, 3 H, NCH_3), 3.79 (s, 3 H, NCH_3), 3.31 (t, 2 H, SO_2CH_2), 3.11 (m, 2 H, $CONHCH_2$), 2.22 (t, 2 H, CH_2CO), 1.47–1.69 (br m, 14 H, CH_2CH_3 and $(CH_2)_6$), 0.86 (t, 3 H, CH_2CH_3); ^{13}C NMR δ 168.81 (CONH), 160.53 (CONH), 157.69 (CONH), 122.42, 122.05, 121.38, 121.32, 117.37, 117.01, 103.43, 103.25, 55.58 (CH_3SO_2), 47.17 (OSO_2CH_2), 35.32, 35.16, 34.93, 27.88, 27.84, 27.64, 26.60, 24.63, 22.39, 21.87, 10.71; MS (FAB, Xe, 9 keV) m/z 538 (3, M + 1); IR (KBr disk) 3422, 2934, 1653, 1582, 1541, 1384, 1194, 1044 cm^{-1}.

(*H*) [*1-Methyl-4-*[*1-methyl-4-*(*3-hydroxypropanamido*)*pyrrole-2-carboxamido*]*pyrrole-2-carboxamido*]*propane* (**3a**). The nitrodipeptide (70 mg, 210 μmol) was reduced (as described above) and the resulting amine dissolved in MeCN (5 mL). β-Propiolactone (50 μL, 800 μmol) in MeCN (2 mL) was added in three equal portions over 6 h under Ar atmosphere and the resulting solution stirred at room temperature for 20 h. Upon removal of the solvent in vacuo, the pure product (48.3 mg, 61%) was obtained by TLC (silica, MeOH–EtOAc, 1:9, R_f 0.36): ^{1}H NMR δ 9.84 (s, 1 H, CONH), 9.81 (s, 1 H, CONH), 7.99 (t, 1 H, CON*H*CH_2), 7.17 (s, 2 H, pyrrole CH's), 6.86 (s, 1 H, pyrrole CH), 6.84 (s, 1 H, pyrrole CH), 4.65 (t, 1 H, O*H*), 3.82 (s, 3 H, NCH_3), 3.79 (s, 3 H, NCH_3), 3.67 (m, 2 H, $HOCH_2$), 3.12 (m, 2 H, $CH_2CH_2CH_3$), 2.39 (t, 2 H, CH_2CO), 1.49 (m, 2 H, CH_2CH_3), 0.87 (t, 3 H, CH_2CH_3); MS (FAB, Xe, 9 keV) m/z 376 (10, M + 1).

(*I*) [*1-Methyl-4-*[*1-methyl-4-*(*3-*((*methylsulfonyl*)*oxy*)*propanamido*)*pyrrole-2-carboxamido*]*pyrrole-2-carboxamido*]*propane* (*MeSO*$_2$*O*(*CH*$_2$)$_2$–*Lex*). Compound **3a** (10 mg, 27 μmol) and diisopropylethylamine (54 μmol) in MeCN (5 mL) were cooled in an ice bath, and methanesulfonyl chloride (15 μL, 270 μmol) in MeCN (0.5 mL) was slowly added. The resulting solution was stirred at room temperature for 10 h, and the solvent was then removed in vacuo. The crude product was purified by preparative TLC (silica, MeOH–EtOAc, 1:9, R_f 0.70) to afford 8.3 mg (69%) of product: ^{1}H NMR δ 10.07 (s, 1 H, CONH), 9.88 (s, 1 H, CONH), 8.01 (t, 1 H, CON*H*CH_2), 7.18 (s, 2 H, pyrrole CH's), 6.89 (s, 1 H, pyrrole CH), 6.85 (s, 1 H, pyrrole CH), 4.47 (t, 2 H, SO_2OCH_2), 3.79 (s, 3 H, NCH_3), 3.83 (s, 3 H, NCH_3), 3.14 (s, 3 H, CH_3SO_2), 3.12 (m, 2 H, $CH_2CH_2CH_3$), 2.72 (t, 2 H, CH_2CO), 1.48 (m, 2 H, CH_2CH_3), 0.87 (t, 3 H, CH_2CH_3); ^{13}C NMR δ 167.72 (CONH), 160.68 (CONH), 157.82 (CONH), 122.58, 121.39, 121.13, 117.85, 117.45, 103.56, 103.43, 66.25, 35.87, 35.34, 35.18, 34.20, 21.88, 10.73; MS (FAB, Xe, 9 keV) m/z 454 (3, M + 1); IR (KBr disk) 3431, 2940, 1633, 1583, 1437, 1347, 1200, 1058 cm^{-1}.

(*J*) [*1-Methyl-4-*[*1-methyl-4-*(*9-hydroxynonanamido*)*pyrrole-2-carboxamido*]*pyrrole-2-carboxamido*]*propane* (**4a**). Azelaic acid monomethyl ester (2.02 g, 10 mmol) in anhydrous THF (20 mL) was cooled in an ice–H_2O bath, and 20 mmol of $LiBEt_3H$ (1 M in THF) was slowly added under N_2 atmosphere. The resulting solution was stirred at 0 °C for 1 h and the reaction mixture carefully quenched with H_2O and neutralized with 1 N HCl. The product was extracted into EtOAc, and the organic layer was washed several times with H_2O and dried ($MgSO_4$). The EtOAc was removed in vacuo, and the resulting residue containing the desired 9-hydroxynonanoic acid was used without any further purification. Nitrodipeptide (333 mg, 1 mmol) was reduced (as described above) and the resulting amine dissolved in dry oxygen-free xylene (15 mL). In a separate flask, 9-hydroxynonanoic acid (174 mg, 1 mmol), 2,2′-dipyridyl disulfide (330 mg, 1.5 mmol), and triphenylphosphine (394 mg, 1.5 mmol) were dissolved in dry oxygen-free xylene (25 mL) under N_2 and stirred at ambient temperature for 5 h. At this time the aminopeptide in xylene was added and the resulting solution was stirred at room temperature for 24 h. The product (254 mg, 55%) was obtained by flash chromatography (silica, MeOH–CH_2Cl_2, 1:9): ^{1}H NMR δ 9.83 (s, 1 H, CH_2CONH), 9.75 (s, 1 H, CONH), 7.99 (t, 1 H, CON*H*CH_2), 7.16 (s, 1 H, pyrrole CH), 7.14 (s, 1 H, pyrrole CH), 6.83 (s, 2 H, pyrrole CH's), 4.32 (t, 1 H, $HOCH_2$), 3.80 (s, 3 H, NCH_3), 3.77 (s, 3 H, NCH_3), 3.35 (m, 2 H, $HOCH_2$), 3.11 (m, 2 H, $CH_2CH_2CH_3$), 2.21 (t, 2 H, CH_2CH_2CO), 1.54 (m, 2 H, CH_2CH_2CO), 1.48 (m, 2 H, CH_2CH_3), 1.39 (m, 2 H, $HOCH_2CH_2$), 1.25 (m, 8 H, $(CH_2)_4$), 0.85 (t, 3 H, CH_2CH_3); ^{13}C NMR δ 168.84 (CONH), 160.53 (CONH), 157.68 (CONH), 122.38, 122.01, 121.36, 121.30, 117.38, 117.00, 103.41, 103.23, 60.01 ($HOCH_2$), 35.34, 35.18, 34.95, 31.82, 28.17, 27.97, 24.79, 24.69, 21.88, 10.73; MS (FAB, Xe, 9 keV) m/z 460 (12, M + 1).

(*K*) [*1-Methyl-4-*[*1-methyl-4-*(*9-*((*methylsulfonyl*)*oxy*)*nonanamido*)*pyrrole-2-carboxamido*]*pyrrole-2-carboxamido*]*propane* (*MeSO*$_2$*O*(*CH*$_2$)$_8$–*Lex*). Alcohol **4a** (35 mg, 76 μmol), diisopropylethylamine (39.8 μL, 230 μmol), and MeCN (10 mL) were cooled in an ice bath, and methanesulfonyl chloride (17.7 μL, 230 μmol) in a minimum volume of MeCN was slowly added. The resulting solution was stirred at room temperature for 24 h and the solvent removed in vacuo. The crude product was purified by preparative TLC (silica, MeOH–CH_2Cl_2, 1:9, R_f 0.67) to yield 25 mg (69%) of product: ^{1}H NMR δ 9.85 (s, 1 H, CH_2CONH), 9.75 (s, 1 H, CONH), 8.00 (t, 1 H, CON*H*CH_2), 7.16 (s, 2 H, pyrrole CH's), 6.85 (s, 2 H, pyrrole CH's), 4.18 (t, 2 H, SO_2OCH_2), 3.81 (s, 3 H, NCH_3), 3.79 (s, 3 H, NCH_3), 3.16 (s, 3 H, CH_3SO_2), 3.11 (m, 2 H, $CH_2CH_2CH_3$), 2.22 (t, 2 H, CH_2CO), 1.65 (m, 2 H, $SO_2OCH_2CH_2$), 1.57 (m, 2 H, CH_2CH_2CO), 1.49 (m, 2 H, CH_2CH_3), 1.29 (m, 8 H, $(CH_2)_4$), 0.86 (t, 3 H, CH_2CH_3); ^{13}C NMR δ 168.79 (CONH), 160.52 (CONH), 157.68 (CONH), 122.40, 122.03, 121.36, 121.31, 117.37, 117.01, 103.40, 103.22, 69.74 (SO_2OCH_2), 35.87 (CH_3SO_2), 35.34, 35.18, 34.92, 27.96, 27.90, 27.80, 27.66, 24.64, 24.19, 21.88, 10.73; MS (FAB, Xe, 9 keV) m/z 538 (10, M + 1); IR (KBr disk) 3426, 2929, 1654, 1584, 1437, 1346, 1171, 1023 cm^{-1}.

Preparation of ^{32}*P-End-Labeled DNA Restriction Fragments.* 5′- and 3′-^{32}P-End-labeled DNA restriction fragments (Figure 3 shows resolved regions of restriction fragments) were prepared from a 3220-bp DNA clone containing the

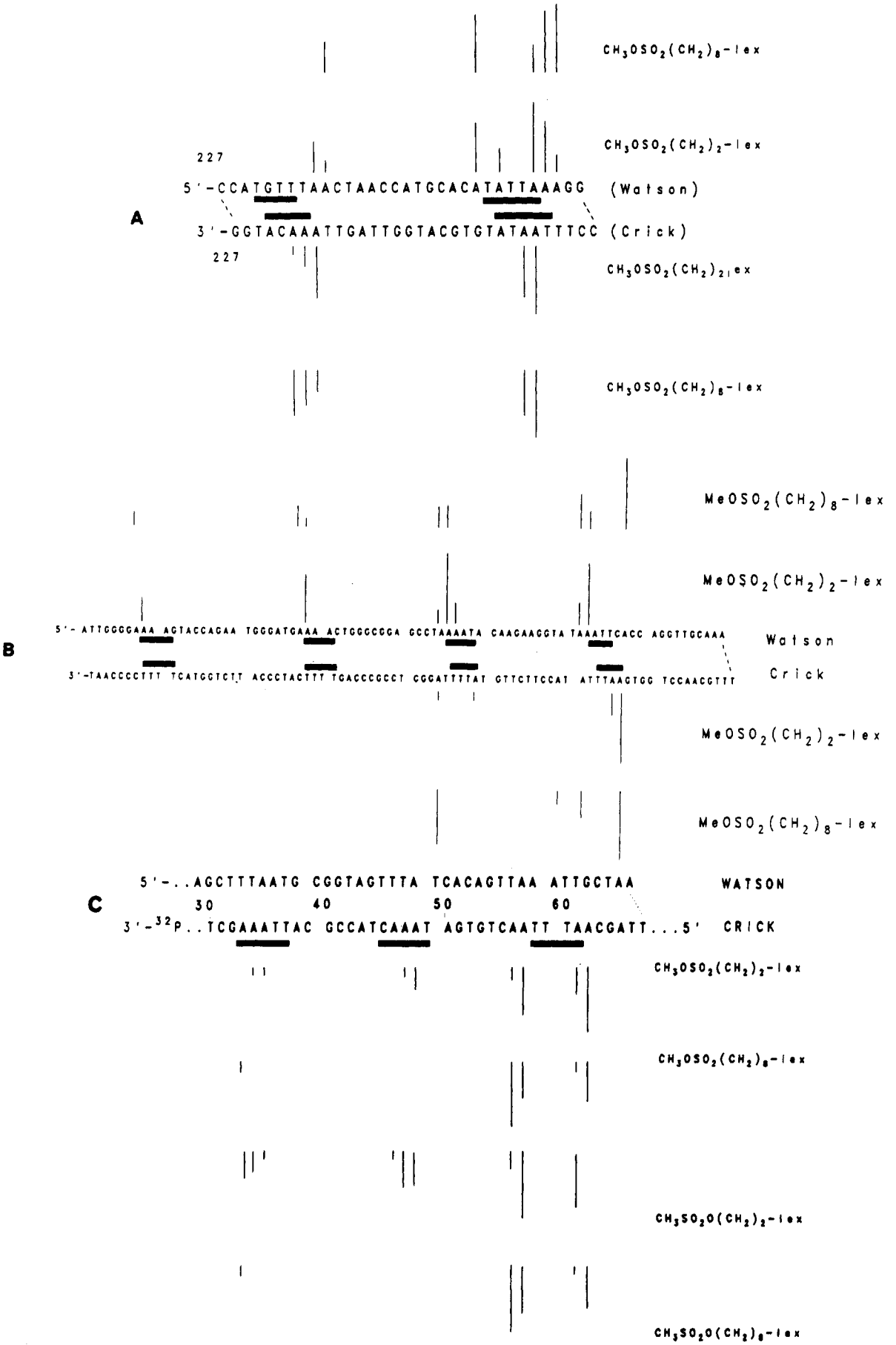


FIGURE 3: Resolved sequence of 85 (A), 576 (B), and 167 (C) bp fragments with Lex-binding sites (underlined). Location and normalized intensities for bands (based on strongest cleavage band for each strand with each compound) for the sulfonate esters are indicated by vertical lines. The cleavage analysis is derived, in part, from Figures 4–7 (gel data for the 3′-labeled 85-bp and 5′-labeled 576-bp fragments are not shown).

promotor region of the coat protein gene of the canine parvovirus (Rhode, 1985) using standard procedures (Maxam & Gilbert, 1980) as previously described (Wurdeman & et al., 1989). A 3′-end-labeled 167-bp fragment was prepared from pBR322 plasmid by sequential EcoRI endonuclease digestion, treatment of the linearized DNA with [α-^{32}P]dATP in the presence of the Klenow fragment of polymerase I, and restriction with RsaI endonuclease (Figure 3). The end-labeled fragments were purified by electrophoresis on 5% polyacrylamide gels prior to use.

Reactions of $MeOSO_2(CH_2)_n$–Lex and $MeSO_2O(CH_2)_n$–Lex with ^{32}P-End-Labeled DNA Fragments. The restriction fragment (80 000–100 000 cpm) and sonicated calf thymus DNA (final concentration 83 μM, phosphate) were dissolved in 10 mM sodium cacodylate buffer (pH 7.8) containing, when specified, the desired concentration of inorganic salt or cationic DNA affinity binder. This DNA solution was incubated at 37 °C for a specified time with freshly prepared anhydrous MeOH solutions of sulfonate esters. The reactions were terminated by cooling and precipitation of the DNA with NaOAc and EtOH. The DNA was washed with cold 70% EtOH and dried in vacuo.

Generation and Analysis of DNA Strand Breaks. Strand breaks in the reacted DNA were generated by either neutral thermal hydrolysis (Lawley & Brookes, 1973) or Maxam–Gilbert G-specific reaction (Maxam & Gilbert, 1980) as previously described (Church et al., 1990).

Analysis of DNA Adducts from $MeOSO_2(CH_2)_2$–Lex. Calf thymus DNA (100 μM nucleotide), that had been dialyzed against 10 mM Tris–HCl buffer (pH 7.8) overnight, was reacted with $MeOSO_2(CH_2)_2$–Lex (100 μM) at room temperature for 24 h in 10 mM Tris–HCl buffer (pH 7.8) in a total volume of 1 mL. The DNA was precipitated from the reaction solution with 3 M NaOAc and EtOH, and the precipitated DNA was washed with EtOH. The DNA was redissolved in 10 mM Tris buffer (pH 7.0) to a total volume of 500 μL and then heated at 90 °C for 30 min to release 3- and 7-alkylpurines. The partially apurinic DNA was precipitated at 0 °C by adding 0.1 volume of 0.1 N HCl (Beranek et al., 1980) and the supernatant analyzed by HPLC (column, Partisil-10 SCX strong cation exchange; eluant, 175 mM ammonium formate buffer, pH 3.0; flow rate, 1 mL/min; detection, UV at 270 nm to monitor *N*7-MeG and *N*3-MeA). Authentic standards were used to confirm the identification of adducts and to calculate response factors.

RESULTS

Synthesis (Figure 2). The preparation of the peptide portion of the Lex molecules is based on the synthesis of distamycin (Lown & Krowicki, 1985) and has been previously described in detail (Church et al., 1990). The elaboration of the $MeSO_2O$ functionality on the N-terminus of $MeSO_2O(CH_2)_n$–Lex employed treatment of the primary amine with β-lactone (n = 2) or 9-hydroxynonanoic acid and dipyridyl sulfide–triphenylphosphine (n = 8) (Mukaiyama et al., 1970; Corey et al., 1976). The resulting alcohols were treated with methanesulfonyl chloride to afford the desired products in ~70% yield after purification. The synthesis of the $MeOSO_2(CH_2)_n$–Lex compounds entailed the initial attachment of a methylene tether with a terminal olefin onto the amino terminus of the dipeptide. The olefin was then converted into the ammonium sulfonate salt by treatment with ammonium bisulfite and oxidation with H_2O_2 or peracid (Gilbert, 1965). After conversion of the ammonium salt into the sulfonic acid, the *O*-methyl sulfonate ester was prepared in ~60% yield by treatment with diazomethane or *N*-methyl-*N*′-(4-nitrophenyl)triazene in MeOH, the latter being the reagent of choice because of convenience.

DNA Alkylation. The general feature of the DNA-cleavage pattern produced by $MeOSO_2(CH_2)_n$–Lex and $MeSO_2O(CH_2)_n$–Lex is the dominance of A-cleavage sites (Figures 4–6). This contrasts with the pattern generated by DMS (Figures 4–6, Maxam–Gilbert G-lane) and MMS (Figure 5, lane x; Figure 6, lane p) that is essentially restricted to G's. Furthermore, the bands induced by $MeOSO_2(CH_2)_n$–Lex and $MeSO_2O(CH_2)_n$–Lex are restricted to A's located in previously footprinted Lex-binding sites in the parvovirus (Church et al., 1990) and pBR322 (Schultz & Dervan, 1984) fragments (Figure 3). $MeOSO_2(CH_2)_n$–Lex and $MeSO_2O(CH_2)_n$–Lex share another common feature: distamycin inhibits their facility to cleave DNA (Figure 4, lanes l, o, y; Figure 5, lanes g, l, q, v; Figure 6, lanes e, h, k, n). In contrast, distamycin has virtually no effect on DMS or MMS (Figure 6, lane q; Wurdeman & Gold, 1988). The modifications of DNA by $MeOSO_2(CH_2)_n$–Lex and $MeSO_2O(CH_2)_n$–Lex differ in two important aspects. Firstly, the rate of DNA modification by $MeOSO_2(CH_2)_n$–Lex is faster than that for $MeSO_2O(CH_2)_n$–Lex. The bands for $MeOSO_2(CH_2)_n$–Lex reach a maximum intensity near 6 h (Figure 4), while the cleavage by $MeSO_2O(CH_2)_n$–Lex continues to intensify up through 24 h (data not shown). Despite differences in rate, the same band patterns are observed at low concentrations or with short (15 min) incubation periods (data not shown) for each compound. The second difference relates to the effect that the coaddition of 200 mM NaCl has on DNA adduct formation. DNA methylation by $MeOSO_2(CH_2)_n$–Lex is completely refractory to NaCl, $MgCl_2$, and spermine, while alkylation by $MeSO_2O(CH_2)_n$–Lex is substantially inhibited (Figure 4, lane p; Figure 5, lane h; Figure 6, lanes f, i; data not shown).

Cleavage Patterns. The A-cleavage sites for $MeOSO_2(CH_2)_n$–Lex and their relative intensities (based on densitometry) in the three restriction fragments are shown in Figure 3. It is stressed that the sequencing gel data are normalized in Figure 3 for each compound on each strand; therefore, absolute comparisons between strands or compounds are not intended. The most intense cleavage bands in the 85-bp fragment at the 2-h time point for $MeOSO_2(CH_2)_2$–Lex are at A_{219}, A_{205}, A_{200}, and A_{199} (Watson strand, Figure 4) and at $A_{221-220}$ and $A_{202-201}$ (Crick strand, data not shown). For $MeOSO_2(CH_2)_8$–Lex, the bands at A_{218}, A_{205}, A_{199}, and A_{198} (Watson strand) and $A_{222-221}$ and A_{201} (Crick strand) are the most intense. The nonmethylating analogues $MeSO_2O(CH_2)_2$–Lex and $MeSO_2O(CH_2)_8$–Lex cleave selectively in the Watson strand at A_{199} and A_{198}, and A_{198}, respectively. The corresponding cleavage sites in the Crick strand for the two compounds are at $A_{221-220}$, $A_{202-201}$, and A_{201}, respectively (data not shown). Other A's become prominent cleavage sites at longer time periods, but as noted above, the relative cleavage pattern for each compound seems to be qualitatively insensitive to increasing drug concentration within the range studied.

In the 576-bp fragment the sequence-selective methylation of DNA by $MeOSO_2(CH_2)_2$–Lex shows a clear preference for $A_{267-268}$, A_{288}, A_{305}, and $A_{321-322}$ (Watson strand, data not shown) and A_{303}, A_{308}, and $A_{324-325}$ (Crick strand, Figure 5), which are all associated with distamycin binding sites. The $MeOSO_2(CH_2)_8$–Lex analogue shows an A-cleavage pattern restricted to Lex affinity binding sites, but with a 1–2-bp displacement relative to the C_2 compound. It is of interest that the increase in the length of the tether in $MeOSO_2(CH_2)_8$–Lex allows strong methylation of A_{285}, which is separated from a Lex binding site by a G–C base pair (Figure 5, lane

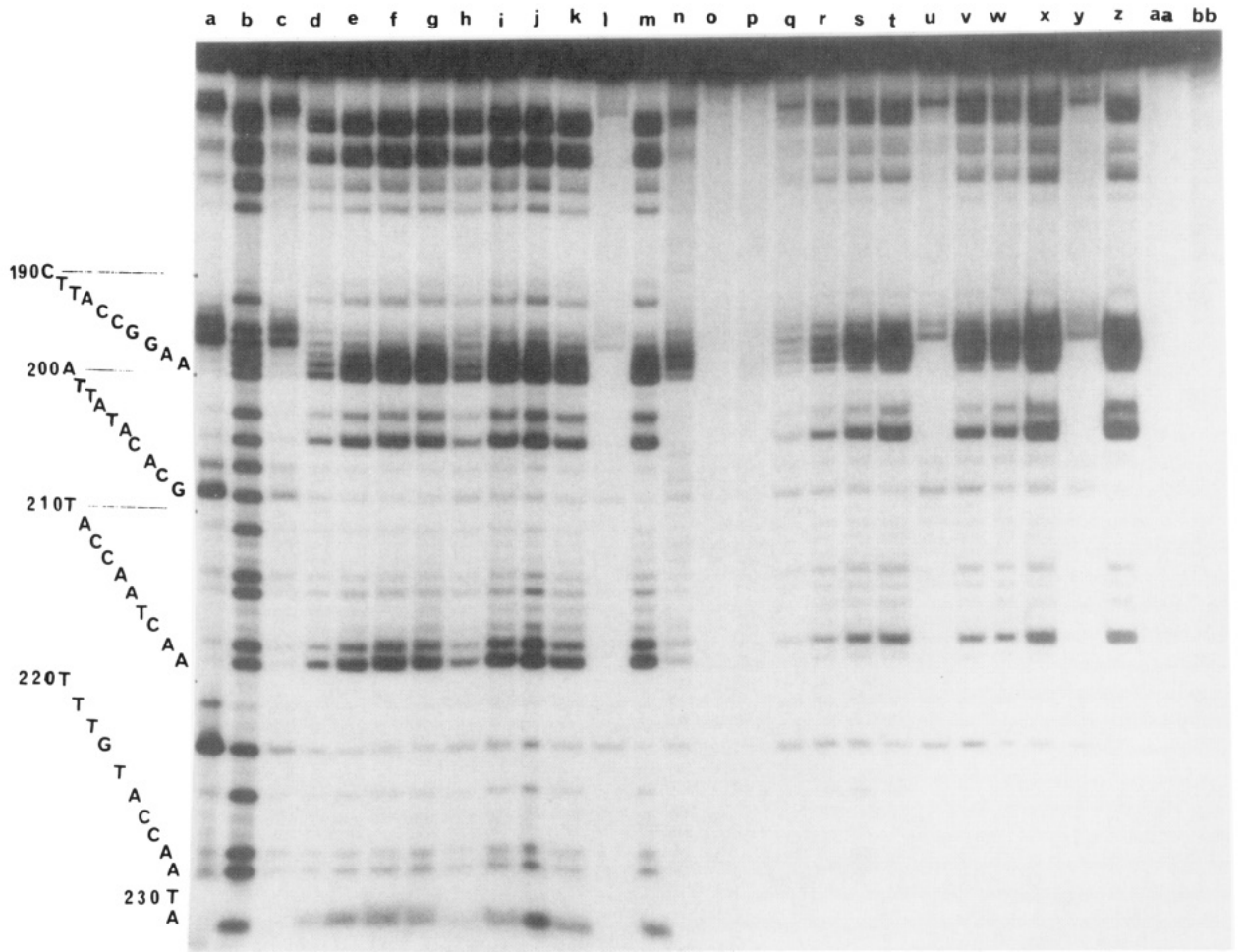


FIGURE 4: 5′-Labeled 85-bp fragment. Time course and dose response for methylation of DNA by $MeOSO_2(CH_2)_n$–Lex and alkylation of DNA by $MeSO_2O(CH_2)_2$–Lex: lane a, G; lane b, G + A; lane c, 500 μM MNU (2 h); lanes d–g, 250 μM $MeOSO_2(CH_2)_2$–Lex (with incubation times of 2, 6, 12, and 24 h, respectively); lanes h–m, 500 μM $MeOSO_2(CH_2)_2$–Lex (with incubation times of 2, 6, 12, 24, 24, and 24 h, respectively); lanes n–p, 500 μM $MeSO_2O(CH_2)_2$–Lex (with 24-h incubation time); lanes q–t, 250 μM $MeOSO_2(CH_2)_8$–Lex (with incubation times of 2, 6, 12, and 24 h, respectively); lanes u–z, 500 μM $MeOSO_2(CH_2)_8$–Lex (with incubation times of 2, 6, 12, 24, 24, and 24 h, respectively); lanes l, o, and y, 100 μM distamycin; lanes m, p, and z, 200 mM NaCl; lane aa, control (2-h incubation); lane bb, control (24-h incubation).

s). Apparently the N^2-amino of C–G_{286} does not interfere with the extension of the aliphatic tether in the minor groove. The prominent A bands generated by $MeSO_2O(CH_2)_2$–Lex are at $A_{267-268}$, A_{288}, A_{305}, A_{311}, and A_{322} (Watson strand, data not shown) and at A_{324} (Crick strand, Figure 5). For $MeSO_2O(CH_2)_8$–Lex, A_{268}, A_{288}, A_{305}, and A_{322} are the dominant cleavage sites.

The reaction of the pBR322 3′-end-labeled 167-bp fragment with $MeOSO_2(CH_2)_n$–Lex and $MeSO_2O(CH_2)_2$–Lex yields bands associated with previously MPE–Fe(II)-footprinted distamycin-binding sites (Schultz et al., 1982; Taylor et al., 1984). $MeOSO_2(CH_2)_2$–Lex targets A_{57} and A_{62} while $MeOSO_2(CH_2)_8$–Lex reacts at A_{56} and A_{59}. Weak bands are seen at A_{33}, A_{34}, A_{47}, A_{48}, A_{57}, and A_{61} in the three different Lex recognition domains with $MeSO_2(CH_2)_2$–Lex. The C_8-linked analogue affords three moderate cleavage sites at A_{56}, A_{57}, and A_{62}. As seen in the other fragments, the addition of distamycin and NaCl inhibits alkylation.

DNA Adducts from $MeOSO_2(CH_2)_2$–Lex. The HPLC analysis of the reaction products of calf thymus DNA and $MeOSO_2(CH_2)_2$–Lex is limited to thermally labile adducts, and only the *N*3-MeA adduct was detected: no *N*7-MeG was observed even upon analysis of large amounts of supernatant. The yield of *N*3-MeA product is presented as the ratio of *N*3-MeA to A to correct for DNA recovery. The *N*3-MeA/A ratio for $MeOSO_2(CH_2)_2$–Lex is 0.0056 ± 0.0007.

DISCUSSION

Alkylation Pattern. The methylation of DNA by either DMS or MMS affords as the dominant product (75–80%) *N*7-MeG along with lesser amounts of *N*3-MeA (10–15%), *N*7-methyladenine (2%), and *N*3-methylguanine (1%) (Beranek et al., 1980; Den Engelse et al., 1986; Beranek, 1990). These alkylating agents show little sequence selectivity in their modification of the dominant nucleophilic site (N7–G), and the yield of DNA lesions is not sensitive to the addition of inorganic salts, e.g. Mg^{2+} and Na^+, or DNA affinity binding cations, including distamycin, spermine, and ethidium bromide (Wurdeman & Gold, 1988). $CH_3SO_2OCH_2CH_3$, which can be considered as a model for $MeSO_2O(CH_2)_n$–Lex, produces a similar alkylation profile with several notable differences: the *N*7-ethylguanine and *N*3-ethyladenine adducts are reduced to ~60% and 4% of the total adducts, respectively. There is also a concomitant increase in alkylation of the phosphate backbone and O^6–G to ~12% and 2% of the total products, respectively (Beranek et al., 1980; Den Engelse et al., 1986).

$MeOSO_2(CH_2)_n$–Lex affords methylation patterns that are clearly different from what might be expected from a sulfonate ester in that the predominant cleavages occur at A's, and specifically at A's that are associated with distamycin and netropsin equilibrium binding sites. This result was not completely unanticipated, since Baker & Dervan had dem-

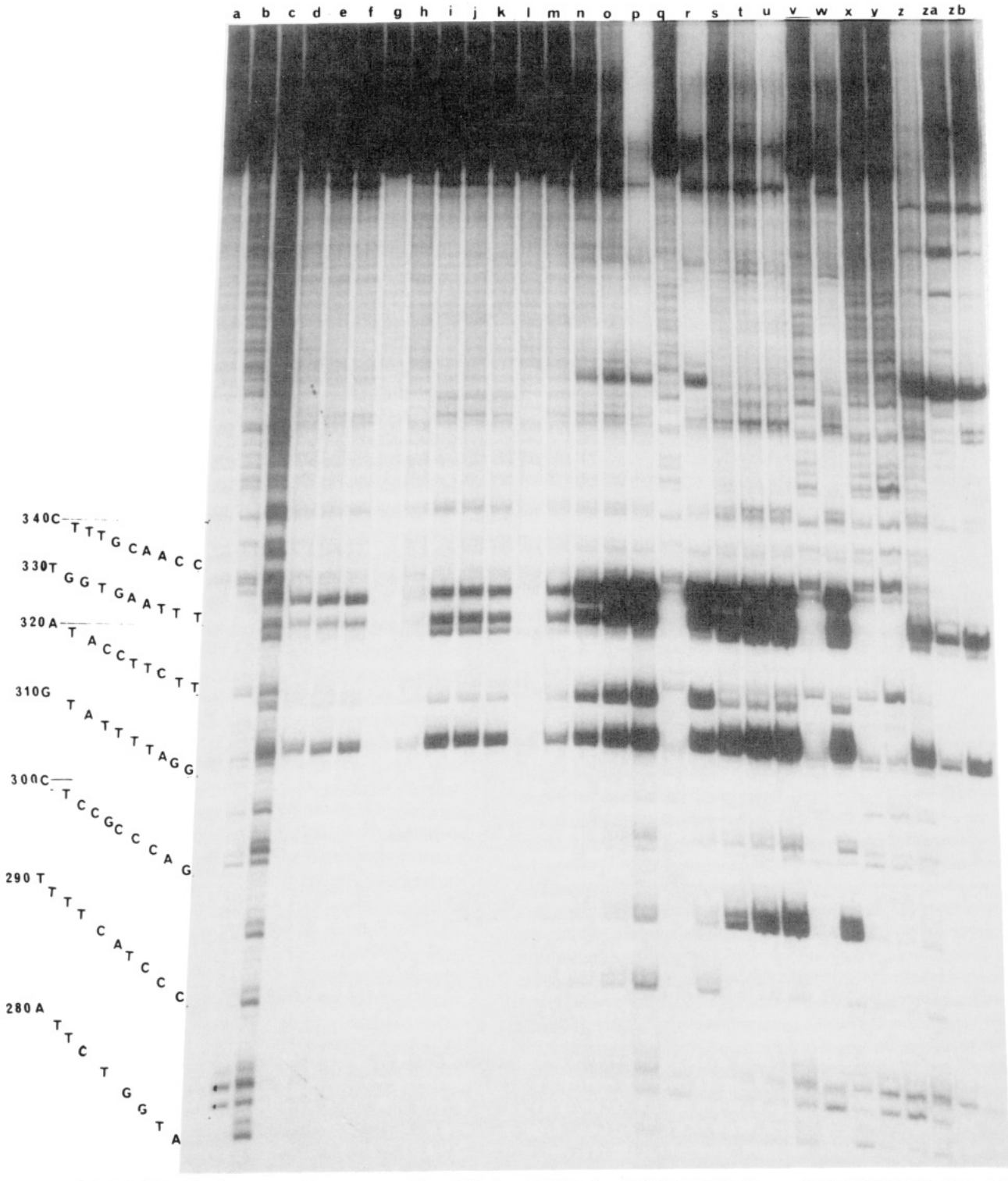


FIGURE 5: 3′-Labeled 576-bp fragment. Dose response for alkylation of DNA by $MeOSO_2(CH_2)_n$–Lex and $MeSO_2O(CH_2)_n$–Lex, and alkylation of DNA by CC-1065 (all incubations were run for 24 h except lanes z–zb which were for 1 h): lane a, control; lane b, G; lane c, G + A; lanes d–h, $MeSO_2O(CH_2)_2$–Lex (250, 500, 1000, 1000, and 1000 μM, respectively); lanes i–m, $MeSO_2O(CH_2)_8$–Lex (250, 500, 1000, 1000, and 1000 μM, respectively); lanes n–r, $MeOSO_2(CH_2)_2$–Lex (50, 100, 250, 250, and 250 μM, respectively); lanes s–w, $MeOSO_2(CH_2)_8$–Lex (50, 100, 250, 250, and 250 μM, respectively); lanes x and y, 500 μM $MeOSO_2OMe$; lanes z–zb, 0.5 μM CC-1065; lanes g, l, q, v, y, and za, 100 μM distamycin; lanes h, m, r, w, and zb, 200 mM NaCl.

onstrated that a haloketone-functionalized lexitropsin afforded alkylation in the minor groove of DNA at N3-A (Baker & Dervan, 1985, 1989). Similarly, it was reported that FCE 24517, an aniline mustard appended to the N-terminus of distamycin, reacts at A (presumably N3-A) and not at N7-G as do all other simple mustards (Broggini et al., 1991). In addition, we have previously synthesized *N*-alkyl-*N*-nitrosoureas (alkyl = methyl, ethyl, and 2-chloroethyl) linked to lexitropsins in order to prepare minor-groove alkylating agents (Church et al., 1990). This approach gave mixed results with only the (chloroethyl)nitrosoureas reacting in the minor groove as envisioned; the *N*-methyl and -ethyl analogues gave no minor groove adducts.

The dramatic decrease in DNA methylation by the coaddition of distamycin suggests that the antibiotic and $MeOSO_2(CH_2)_n$–Lex compete for the same minor-groove binding sequences. The structure of the A adduct that is responsible for the abasic sites after neutral thermal hydrolysis

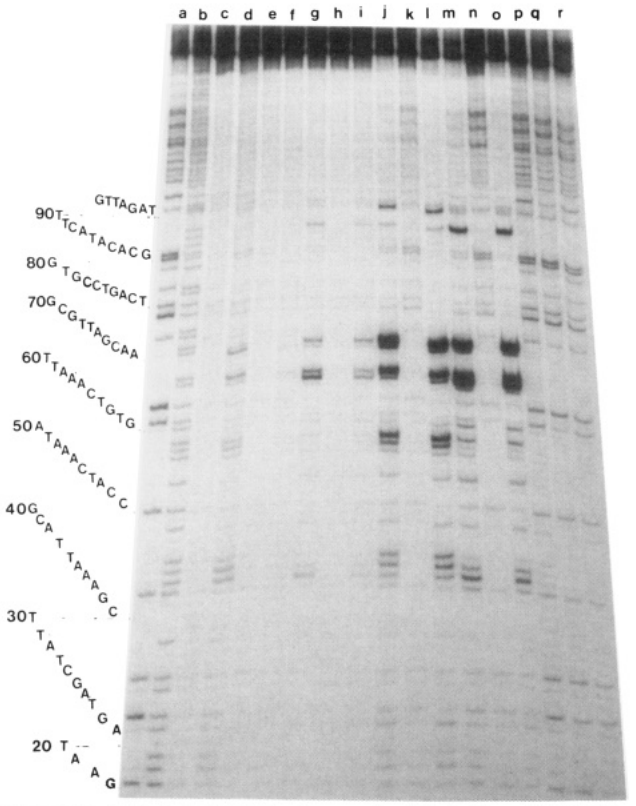


FIGURE 6: 3′-Labeled 167-bp fragment. Alkylation of DNA by $MeOSO_2(CH_2)_n$–Lex and $MeSO_2O(CH_2)_n$–Lex, and effect of distamycin and salt: lane a, G; lane b, G + A; lane c, control; lanes d–f, 1000 μM $MeSO_2O(CH_2)_2$–Lex; lanes g–i, 1000 μM $MeSO_2O(CH_2)_8$–Lex; lanes j–l, 500 μM $MeOSO_2(CH_2)_2$–Lex; lanes m–o, 500 μM $MeOSO_2(CH_2)_8$–Lex; lanes p–r, 1000 μM $MeOSO_2Me$; lanes e, h, k, n, and q, 100 μM distamycin; lanes f, i, l, o, and r, 200 mM NaCl.

was confirmed by reacting $MeOSO_2(CH_2)_2$–Lex with calf thymus DNA. $MeOSO_2(CH_2)_2$–Lex afforded only *N*3-MeA. The failure to observe the *N*7-MeG lesion is consistent with the sequencing gel data and contrasts with the normal predominance of the *N*7-MeG adduct: the *N*7-MeG/*N*3-MeA ratio is 7:1 for DMS (Beranek et al., 1980). It is assumed that the N3-A position is also the site for electrophilic attack by $MeSO_2O(CH_2)_n$–Lex. The formation of heat-labile *N*3-alkylA residues from $MeSO_2O(CH_2)_2$–Lex is in keeping with the observation that the A bands are much weaker when Maxam–Gilbert G-lane chemistry (Maxam & Gilbert, 1977, 1980) is used to generate strand breaks (data not shown). Additional evidence for the focus of alkylating events in the minor groove is the observation that G residues immediately adjacent to strong A-cleavage sites are not alkylation targets for $MeOSO_2(CH_2)_n$–Lex and $MeSO_2O(CH_2)_n$–Lex. A detailed and elegant study has been reported for the reaction of DNA with $BrCH_2CO$–Dis, an N-terminal alkylating pyrrolecarboxamide tripeptide with a dimethylamino group on the carboxy terminus (Baker & Dervan, 1985, 1989). The reactions of $BrCH_2CO$–Dis with restriction fragments were carried out in 10 mM sodium phosphate (pH 7.0), and DNA alkylation is very slow (similar to what is observed for $MeSO_2O(CH_2)_n$–Lex) and very selective for a small subset of A residues associated with distamycin affinity binding sites. The structure of the *N*3-(CH_2CO–Dis)–A adduct was confirmed by NMR (Baker & Dervan, 1989).

Binding Preferences. Efforts to footprint the equilibrium binding of $MeSO_2O(CH_2)_n$–Lex in the DNA restriction fragments with MPE–Fe(II) (Van Dyke et al., 1982; Hertzberg & Dervan, 1984) were unsuccessful. $MeSO_2O(CH_2)_n$–Lex, the more stable of the two Lex analogues ($t_{1/2}$ of ~35 h under reaction conditions), was used to avoid any potential confusion about performing experiments in the presence of the sulfonic acid decomposition product derived from the hydrolysis of $MeOSO_2(CH_2)_n$–Lex. The challenge in obtaining footprinting evidence for the binding of a neutral ligand using the competitive inhibition approach was anticipated, since previous efforts to map the association of the nitrosourea $Cl(CH_2)_2$-$N(NdO)CONH(CH_2)_2$–Lex with DNA had also failed (Church et al., 1990). However, the substitution on the carboxyl terminus of a $-NHCH_2CH_2N(CH_3)_2$ group that is ionized at near-physiological pH in place of $-NHCH_2CH_2CH_3$ resulted in a footprint of the nitrosourea that overlaps with distamycin-binding regions (Church et al., 1990). The observation of sequence-selective alkylation by $MeSO_2O$-$(CH_2)_n$–Lex, which is a weak binder, is not unique, since des-ABC, an unreactive analogue of the efficient minor-groove DNA-alkylating agent CC-1065, also cannot be footprinted by conventional DNA protection methods (Hurley et al., 1988). The cleavage pattern induced by CC-1065 is shown in Figure 5 (lanes z–zb). CC-1065 is a much more powerful alkylating agent than the sulfonate esters, since intense cleavage bands are evident with CC-1065 at concentrations as low as 0.5 μM using a 1-h incubation time.

While footprinting data are useful, the methylation patterns for $MeOSO_2(CH_2)_n$–Lex and their shift as a function of tether length in some ways provide more valuable information on the relative preference of the Lex compounds for binding sites. In fact, the effect of the tether (C_2 vs C_8) to displace the cleavage pattern by 1–2 bp's is strong evidence that equilibrium binding precedes and determines the site of DNA methylation. The same process is proposed for $MeSO_2O(CH_2)_n$–Lex, where the length of the tether also changes the cleavage sites. We have previously noted that the methylation of N3-A by MNU via $CH_3N_2^+$ (Church et al., 1990) and DMS (unpublished data) shows virtually no sequence selectivity, so the alkylation patterns of $MeOSO_2(CH_2)_n$–Lex, do not reflect sequence-dependent differences in the nucleophilic reactivity or accessibility of N3-A.

In order to use the DNA methylation pattern as a marker of peptide binding, the distance between the CH_3 group in $MeOSO_2(CH_2)_n$–Lex that is transferred to DNA and the amide N-*H* on the N-terminus of the peptide (Figure 1) needs to be known in a drug–DNA complex. The calculation of this distance, a subject of a previous investigation (Goodsell & Dickerson, 1986), translates to 1 and 2 bp's, respectively, for the $(CH_2)_2$ and $(CH_2)_8$ linkers (see Figure 7 for model), assuming that the peptide is symmetrically bound in the minor groove of B-DNA with a rise per bp of 3.38 Å, a helical twist angle of 36°, and a mean distance between the ligand and DNA axis of 5 Å. The other feature of the alkylation patterns is the 1–2-bp stagger between the Watson and Crick strands. This is because the individual DNA-reading atoms on the Lex peptide are not associated with Watson–Crick base pairs (Figures 3 and 7) (Kopka et al., 1985; Coll et al., 1987; Pelton & Wemmer, 1988; Lee et al., 1988; Coll et al., 1989; Sarma et al., 1990). On the basis of extrapolation from NMR, crystal structures, and molecular modeling (Kopka et al., 1985; Coll et al., 1987; Pelton & Wemmer, 1988; Lee et al., 1988; Coll et al., 1989; Sarma et al., 1990; Boehncke et al., 1991), a 3 bp/strand recognition site for the Lex peptide, involving two amido H-bonds and three VDW contacts, is suggested (Figure 1). A VDW contact between the C-terminal propyl group and A-C2-H that is stabilized by an electrostatic interaction between the charged C-terminus and minor groove atoms is commonly proposed for the binding of netropsin and distamycin

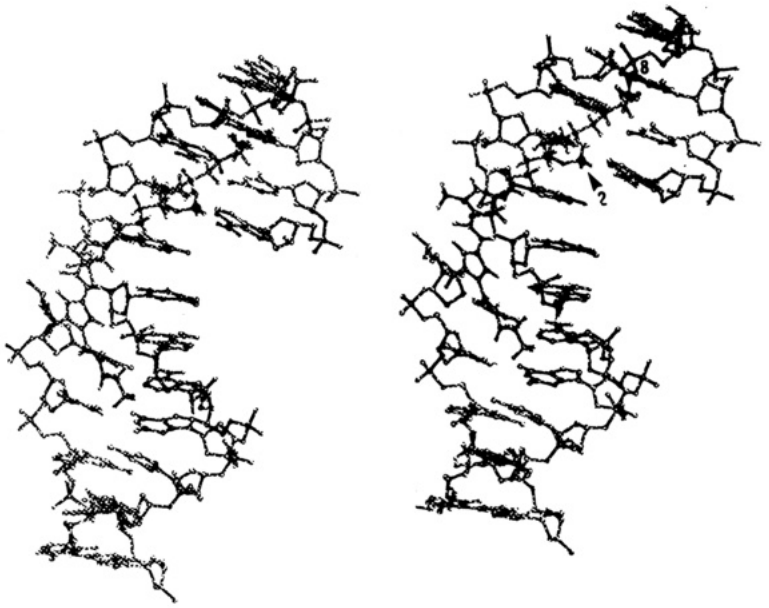


FIGURE 7: Relaxed stereoscopic view of $MeOSO_2(CH_2)_n$–Lex modeled in the minor groove of 5′-d(CGCGAATTTACG)·3′-d(GCGCTTAAATCG). The CH_3 group that is delivered to DNA is darkened and is denoted by 2→ and 8→ for $MeOSO_2(CH_2)_2$–Lex and $MeOSO_2(CH_2)_8$–Lex compounds, respectively. Part of one strand of the DNA has been deleted to facilitate viewing of the ligands.

(Lee et al., 1988). In the absence of the charged terminus, this VDW contact is not envisioned for $MeOSO_2(CH_2)_n$–Lex or $MeSO_2O(CH_2)_n$–Lex, so the interaction of ligand and DNA truncates with the H-bond made by the C-terminus amide N-H. The proposed equilibrium binding arrangement translates to the following rules to interpret the methylation patterns: (i) $MeOSO_2(CH_2)_2$–Lex binds to three bp's and predominantly methylates the fourth (adjacent) base if it is an A and (ii) $MeOSO_2(CH_2)_8$–Lex binds to the same three bp's and predominantly methylates the fifth or sixth base if it is an A (Figure 7). Because of the flexibility of the aliphatic tether, there is probably a limited distribution of methylation sites derived from an individual binding complex rather than there being a unique methylation site for each individual bound Lex. Notwithstanding, this does not affect the interpretation of the data.

As noted above, $MeOSO_2(CH_2)_8$–Lex can methylate the fifth base, assuming it is an A, even if a G intervenes (Figure 5, lane s), meaning that the N^2-amino group of G which protrudes into the minor groove and blocks Lex peptide binding does not significantly affect the ability of the aliphatic tether to deliver the sulfonate ester to N3-A. Obviously, minor groove atoms in d–s DNA that react poorly with sulfonate esters, e.g. O^2-T, N3-G, N^2-G, and O^2-C, will not be efficiently methylated, and methylated bases that are not heat labile, e.g. N^2-MeG, will not afford abasic sites that can be converted into single-strand breaks. Accordingly, no measurable bands are observed at T's, G's, or C's near binding sites. It should be noted that O^2- and N3-alkylC and O^2-alkylT depyrimidinate at 100 °C at pH 7.0 with $t_{1/2}$ of <5 and 160 min, respectively (Singer et al., 1978), so the failure to observe bands is not due to the inability to convert these adducts into abasic sites.

Nucleophilic atoms on bases that cannot readily adopt the required transition-state geometry to facilitate the alkyl group transfer will also not be readily modified. For $MeSO_2O(CH_2)_n$–Lex this requirement appears to be important, and it is proposed that the reduced rate of alkylation and the smaller number of cleavage sites for $MeSO_2O(CH_2)_n$–Lex arise from steric interactions between N3-A and the hindered sulfonate ester. A similar argument has been made for the highly selective and lethargic modification of DNA by $BrCH_2CO$–Dis (Baker & Dervan, 1989). A related issue is the inhibition of DNA cleavage seen only with $MeSO_2O(CH_2)_n$–Lex in the presence of 200 mM NaCl, 10 mM $MgCl_2$, or 100 μM spermine. This probably reflects the stiffening of the DNA resulting in reduced conformational flexibility and accessibility of the nucleophilic N3-A site. Even the longer more flexible aliphatic tether in $MeSO_2O(CH_2)_8$–Lex does not overcome the inhibition by salt. Cations could also exert a modulating effect on the affinity binding of the peptide ligand. However, the failure to observe any consequence of added cations on $MeOSO_2(CH_2)_n$–Lex-mediated DNA methylation, either quantitative or qualitative, makes it unlikely that a decrease in the affinity binding of the neutral Lex is involved. We are not aware of any other S_N2 alkylating ligand showing a similar salt effect, although it is possible that the mechanism for the alkylation of DNA by $MeSO_2O(CH_2)_n$–Lex involves a transition state with significant positive charge buildup. In contrast to $MeSO_2O(CH_2)_n$–Lex, the intensity of the cleavage pattern of CC-1065 does not change with the coaddition of NaCl or $MgCl_2$ (data not shown).

On the basis of the above arguments and the observed methylation patterns, the Lex-binding sites are depicted in Figure 3 and involve essentially the same bp's as footprinted for a charged Lex compound (Church et al.,1990). Within the binding domains there are favored 3-bp recognition sequences and orientational preferences (see below). This is illustrated for $MeOSO_2(CH_2)_2$–Lex in the 85-bp fragment at the $T_{204(Watson)}$ATTA region (Figure 3A). The relative intensities of the bands at $A_{200-198(Watson)}$ are approximately 4:3:1, and this implies a binding preference order of 5′-ATT > TTA > TAA. This interpretation is confirmed by the methylation pattern generated by $MeOSO_2(CH_2)_8$–Lex in the same region, which shows that the strongest methylation site has moved toward the 3′ end by 1–2 bp's. The potential to have intense methylation at A_{200} and A_{198} with the C_8 compounds confirms that the methylation pattern from the C_2 Lex's results from binding preferences and not differences in N3-A nucleophilicity. A 5′-ATT sequence is also associated with the most intense methylation bands in the Crick strand of the 576-bp fragment at A_{325} with $MeOSO_2(CH_2)_2$–Lex and in the Watson strand at A_{327} for $MeOSO_2(CH_2)_8$–Lex.

To put this sequence recognition ranking in perspective, the difference in relative intensities between the A's in the 5′-ATATTAAA (Watson) sequence in the 85-bp fragment ranges ~4-fold, a value similar to binding preferences observed for distamycin using NMR (Pelton & Wemmer, 1990b) or a DNA-footprinting approach (Schultz et al., 1982; Taylor et al., 1984).

It is particularly informative to compare the binding preference of the neutral dipeptide Lex's to that of the monocationic tripeptides. DNA alkylation by $BrCH_2CO$–Dis and oxidative cleavage by an Fe(II)–EDTA–Dis has been studied in the same 167-bp restriction fragment. The preferred binding site of the two-ring Lex compounds is at T_{58}–A_{62} (Crick strand). This same preference was observed for Fe(II)–EDTA–Dis (Shultz & Dervan, 1984). In addition, both compounds share the same orientational preference with the N-terminus pointing toward the 5′-end of the Crick strand. $BrCH_2CO$–Dis, which we incorrectly assumed would behave similarly to $MeSO_2O(CH_2)_2$–Lex, alkylates almost exclusively at A_{48} (Baker & Dervan, 1985; 1989). A_{48} is not a strong alkylation site for any of the sulfonate ester Lex compounds. Calculations indicate that in the minor groove the deepest negative electrostatic well exists at A–T-rich regions (Pullman, 1988). Accordingly, it has been argued that the binding of mono- and dicationic Lex compounds to A–T-rich regions has a significant electrostatic component (Lee et al., 1988; Marky & Breslauer, 1987). The current work cannot address the question of A–T sequence-specific electrostatic interactions between the neutral Lex peptide and DNA; however, the similar sequence and orientational preferences of monocationic tripeptide Dis and neutral dipeptide Lex strongly argue against the role of electrostatics in sequence or orientational selectivity.

Orientational Preferences. It was first noted by Dervan and co-workers that a monocationic distamycin analogue derivatized with a DNA cleaver can adopt both possible orientations within an individual equilibrium binding region. In general the N-terminus of the peptide points toward the 3′-end of the T-rich strand (Schultz et al., 1982; Taylor et al., 1984). NMR and crystallographic studies, although limited in number, tend to confirm this orientational bias (Coll et al., 1987; Pelton & Wemmer, 1990b). This preference was also seen in the interaction of $BrCH_2CO$–Dis with the 167-bp pBR322 fragment (Baker & Dervan, 1989). $MeOSO_2(CH_2)_2$–Lex at some binding sites shows a similar orientational preference, while at other sites none is obvious. The orientational preference of the Lex molecule used in this study is most easily seen in the Watson strand at the A_4 stretches starting at 5′-A_{267} (data not shown), -A_{287}, and -A_{304} in the 576-bp fragment (Figure 3B). In all three sets the favored orientation has the methylating end (N-terminus) of the molecule toward the 3′-end of the T-rich strand and the most reactive A is the one second from the 5′-end (Figure 3). In the binding sites flanked by G–C bp's (beginning at 5′-A_{267} and 5′-A_{287}), the $MeOSO_2(CH_2)_2$–Lex must associate with the 3′-GC bp in order to bind in the site and still methylate the A that it does. These results indicate that the Lex prefers not to bind across the central two A's, since this would result in preferential methylation at the 5′-terminal A in the runs. The observation of an orientational preference at a GAAAAG run implies the sequence is not symmetrical. This is in accord with NMR data that demonstrate that A_4 runs are bent and that there is a B–B′ conformational transition between the middle A's with associated changes in propeller twisting (Coll et al., 1987; Sarma et al., 1990; Chuprina et al., 1991). That the neutral Lex compounds show the same interactions with these sequences as the ionized distamycin tripeptide indicates that the origin of the orientational preference can be narrowed down to the dipeptide structure. The crystal structure data on distamycin disclose several differences on the interactions of the two ends of the antibiotic with DNA. On the carboxyl terminus, the amide N-H (H-bond donor) and the C3-H (VDW contact) of the associated pyrrole point toward each other and sterically interact. The result is a significant torsional angle between these two H's. On the amino terminus there is no similar steric effect, with the N-H and C3-H being virtually parallel and nearly coplanar to each other. Exactly how the different ends of the peptide interact with the asymmetrical minor-groove recognition sequence remains to be determined.

CONCLUSIONS

In an attempt to enhance our control over DNA alkylation, we have prepared alkyl sulfonate esters that are tethered to a DNA affinity binding peptide. These are bimolecular alkylating agents; therefore, no diffusible species is involved, and the alkyl group is only delivered to DNA sites it can "touch". Sequencing gel studies show that high groove and sequence specificity, dictated by the binding of the peptide, are obtained with both the $MeSO_2O(CH_2)_n$–Lex and $MeOSO_2(CH_2)_n$–Lex compounds. The adduct data show that the predominant point of electrophilic attack occurs at N3-A. To our knowledge, $MeOSO_2(CH_2)_n$–Lex are the first compounds that predominantly methylate this position in duplex DNA. The ability to alter the "normal" alkylation patterns will be useful in understanding the contributing role of different DNA adducts in toxicity and mutagenicity, in the design of novel chemotherapeutic agents, and in unraveling the activity of different DNA-repair enzymes.

ACKNOWLEDGMENT

We are grateful to the late Dr. Phillip Issenberg and Mr. David Babcook for the mass spectroscopy data, to Dr. Solon Rhode for helpful discussions and originally providing the parvovirus plasmid, to Dr. Peter Dervan for providing the MPE reagent, and to The Upjohn Company for a sample of CC-1065.

REFERENCES

Baker, B. F., & Dervan, P. B. (1985) *J. Am. Chem. Soc. 107*, 8266–8268.

Baker, B. F., & Dervan, P. B. (1989) *J. Am. Chem. Soc. 111*, 2700–2712.

Beranek, D. T., Weis, C. C., & Swenson, D. H. (1980) *Carcinogenesis 1*, 595–606.

Boehncke, K., Nonella, M., Schulten, K., & Wang, A. H.-J. (1991) *Biochemistry 30*, 5465–5475.

Broggini, M., Erba, E., Ponti, M., Ballinari, D., Geroni, C., Spreafico, F., & D'Incalci, M. (1991) *Cancer Res. 51*, 199–204.

Chuprina, V. P., Fedoroff, O. Y., & Reid, B. R. (1991) *Biochemistry 30*, 561–568.

Church, K. M., Wurdeman, R. L., Zhang, Y., Chen, F.-X., & Gold, B. (1990) *Biochemistry 29*, 6827–6838.

Churchill, M. E. A., Hayes, J. J., & Tullius, T. D. (1990) *Biochemistry 29*, 6043–6050.

Coll, M., Frederick, C. A., Wang, A. H.-J., & Rich, A. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8385–8389.

Coll, M., Aymami, J., van der Marel, G. A., van Boom, J. H., Rich, A., & Wang, A. H.-J. (1989) *Biochemistry 28*, 310–320.

Corey, E. J., Brunelle, D. J., & Stork, P. J. (1976) *Tetrahedron Lett. 38*, 3405–3408.
Dabrowiak, J. C., Goodisman, J., & Kissinger, K. (1990) *Biochemistry 29*, 6139–6145.
Friedmann, T., & Brown, D. M. (1978) *Nucleic Acids Res. 5*, 615–622.
Gilbert, E. E. (1965) in *Sulfonation and Related Reactions*, Wiley-Interscience, New York.
Goodsell, D., & Dickerson, R. E. (1986) *J. Med. Chem. 29*, 727–733.
Hartley, J. A., Gibson, Kohn, K. W., & Mattes, W. B. (1986) *Cancer Res. 46*, 1943–1947.
Hertzberg, R. P., & Dervan, P. B. (1984) *Biochemistry 23*, 3934–3945.
Hurley, L. H., Lee, C.-S., McGovren, J. P., Warpehoski, M. A., Mitchell, M. A., Kelly, R. C., & Aristoff, P. A. (1988) *Biochemistry 27*, 3886–3892.
Jensen, D. E., & Reed, D. J. (1978) *Biochemistry 17*, 5098–5107.
Johnston, B. H., & Rich, A. (1985) *Cell 42*, 713–724.
Kopka, M. L., Yoon, C., Goodsell, D., Pjura, P., & Dickerson, R. E. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 1376–1380.
Lawley, P.D. (1984) in *Chemical Carcinogens* (Searle, C. E., Ed.) 2nd ed., pp 325–484, ACS Monograph 182, American Chemical Society, Washington, DC.
Lee, M., Krowicki, K., Hartley, J. A., Pon, R. T., & Lown, J. W. (1988) *J. Am. Chem.* Soc. *110*, 3641–3649.
Lück, G., Triebel, H., Waring, M., & Zimmer, C. (1974) *Nucleic Acids Res. 1*, 503–530.
Lück, G., Zimmer, C., Reinert, K.-E., & Arcamone, F. (1977) *Nucleic Acids Res. 4*, 2655–2670.
Marky, L. A., & Breslauer, K. J. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 4359–4363.
Maxam, A. M., & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 560–564.
Maxam, A. M., & Gilbert, W. (1980) *Methods Enzymol. 65*, 499–560.
Mukaiyama, T., Matsueda, R., & Suzuki, M. (1970) *Tetrahedron Lett. 22*, 1901–1904.
Patel, D. J. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 6424–6428.
Patel, D. J., & Shapiro, L. (1986) *J. Biol. Chem. 261*, 1230–1240.
Pelton, J. G., & Wemmer, D. E. (1988) *Biochemistry 27*, 8088–8096.
Pelton, J. G., & Wemmer, D. E. (1990a) *J. Am. Chem. Soc. 112*, 1393–1399.
Pelton, J. G., & Wemmer, D. E. (1990b) *J. Biomol. Struct. Dyn. 8*, 81–97.
Portugal, J., & Waring, M. J. (1987) *FEBS Lett. 225*, 195–200.
Pullman, B. (1988) in *Structure & Expression Volume 2: DNA & Its Drug Complexes* (Sarma, M. H., & Sarma, R. H., Eds.), pp 237–249, Adenine Press, Schenectady, NY.
Rajalakshmi, S., Rao, P. M., & Sarma, D. S. R. (1978) *Biochemistry 17*, 4515–4518.
Sarma, M. H., & Sarma, R. H. (1988) *Structure & Expression Volume 2: DNA & Its Drug Complexes*, Adenine Press, Schenectady, NY.
Sarma, M. H., Gupta, G., Garcia, A. E., Umemoto, K., & Sarma, R. H. (1990) *Biochemistry 29*, 4723–4734.
Shultz, P. G., & Dervan, P. B. (1984) *J. Biomol. Struct. Dyn. 1*, 1133–1147.
Shultz, P. G., Taylor, J. S., & Dervan, P. B. (1982) *J. Am. Chem. Soc. 104*, 6861–6863.
Singer, B., Kröger, M., & Carrano, M. (1978) *Biochemistry 17*, 1246–1250.
Taylor, J. S., Schultz, P. G., & Dervan, P. B. (1984) *Tetrahedron 40*, 457–465.
Van Dyke, M. W., Hertzberg, R. P., & Dervan, P. B. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 5470–5474.
Wurdeman, R. L., & Gold, B. (1988) *Chem. Res. Toxicol. 1*, 146–147.
Wurdeman, R. L., Church, K. M., & Gold, B. (1989) *J. Am. Chem. Soc. 111*, 6408–6412.
Youngquist, R. S., & Dervan, P. B. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 2565–2568.
Zimmer, C. (1975) *Prog. Nucleic Acids Res. Mol. Biol. 15*, 285–318.