Nucleotide Preferences for DNA Interstrand Cross-Linking Induced by the Cyclopropylpyrroloindole Analogue U-77,779

Chong-Soon Lee and Neil W. Gibson*

School of Pharmacy and Comprehensive Cancer Center, University of Southern California, Los Angeles, California 90033

Received September 14, 1992; Revised Manuscript Received December 29, 1992

ABSTRACT: Using a 21-base-pair duplex oligonucleotide containing a centrally located defined cross-linkable site, we have separated by gel electrophoresis DNA interstrand cross-links (ISC) from monofunctionally alkylated DNA (MA) and investigated the sequence selectivity for DNA ISC induced by the CC-1065 analogue U-77,779 (U-77). Sequencing gel analysis shows that U-77 induces two distinct types of DNA ISC. The first distinct form of DNA ISC spans six nucleotides and links two adenine N3 positions within an A/T-rich sequence. The second distinct DNA ISC spans seven nucleotides, also linking two adenine N3 positions, with a preference for contiguous runs of adenines. Three major 6-nucleotide DNA ISC's were identified and found to occur within 5′-TAATTA-3′, 5′-TAAATA-3′, and 5′-TAAAAA-3′ sequences. The major 7-nucleotide DNA ISC was found to occur within 5′-TAAAAAA-3′ sequences. Within this sequence, the formation of the 7-nucleotide DNA ISC was preferred over the 6-nucleotide DNA ISC by a ratio of approximately 2:1. DNA ISC formation within adenine tracts eliminated the inherent DNA bending associated with such sequences. Further, chemical probing of each isolated DNA ISC with diethyl pyrocarbonate (A-specific) and potassium permanganate (T-specific) shows that the major DNA conformational changes, such as helical distortion, were localized within the cross-linked sequence. These results suggest that a significant degree of DNA distortion may occur as a consequence of interstrand cross-linking.

CC-1065 is an extremely potent antitumor antibiotic isolated from *Streptomyces zelensis* by The Upjohn Co. (Hanka et al., 1978; Martin et al., 1981). Structurally, the CC-1065 molecule consists of three repeating pyrroloindole subunits (Figure 1), one of which contains the DNA-reactive cyclopropylpyrroloindole (CPI)[1] function while the other two subunits mediate noncovalent binding interactions with DNA. CC-1065 alkylates the N3 position of adenine in the minor groove of double-helical DNA in a sequence-selective manner, and a common consensus sequence of 5′-(T/A)(T/A)A* (where the asterisk indicates the site of alkylation) was observed (Reynolds et al., 1985; Hurley et al., 1984, 1988, 1990; Warpehoski & Hurley, 1988; Weiland & Dooley, 1991; Boger et al., 1991a–c). The preferred site of alkylation by CC-1065 was found to occur within 5′-TTA* sequences.

Despite the highly potent activity of CC-1065 in vitro (Li et al., 1982; Bhuyan et al., 1982) and in vivo (Martin et al., 1981), clinical development was precluded by an unusual hepatotoxicity, which led to delayed death in mice at therapeutic doses (McGovren et al., 1984). The structural element of CC-1065 which produces delayed toxicity has been elucidated, but its biochemical mechanism is still unknown (Lin et al., 1991). CC-1065 analogues which maintain potent antitumor activity without exhibiting delayed death were subsequently synthesized. Adozelesin, a monofunctional DNA-alkylating agent, has recently completed phase I clinical trials (Fleming et al., 1992; Burris et al., 1992). Subsequently, a bifunctional DNA-reactive agent, U-77,779 (U-77), was synthesized by linking two DNA-reactive CPI functions with a rigid bis(indolecarboxylic acid) linker (Mitchell et al., 1991). Numerous human carcinoma cell lines were 2–30-fold more sensitive to U-77 than adozelesin (Lee & Gibson, 1991). U-77 also displayed curative antitumor efficacy in vivo against the L1210 mouse leukemia (Mitchell et al., 1991). Due to this promising antitumor efficacy and lack of delayed toxicity, U-77 is currently being developed for clinical trials in humans by the U.S. National Cancer Institute in conjunction with The Upjohn Co. (McGovren, 1992).

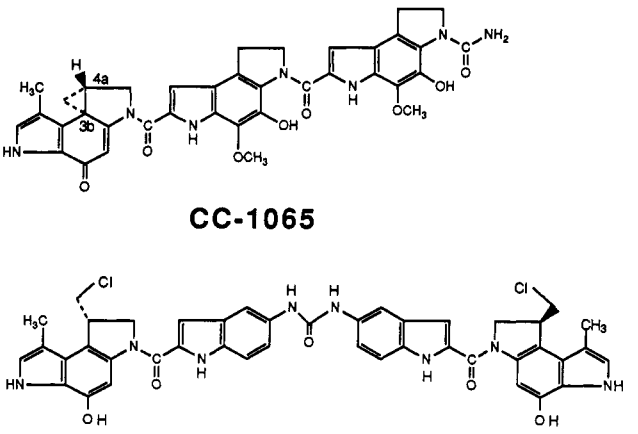


FIGURE 1: Structures of CC-1065 and U-77,779 (U-77).

In this study, we wished to determine the nucleotide preferences for U-77-induced DNA ISC and to examine the influence that the intervening sequence between the two cross-linked bases may have upon the efficiency of DNA ISC formation. We have employed a 21-bp duplex oligonucleotide containing centrally located A tracts or A/T-rich sequences. The results show that U-77 induces either a 6- or 7-nucleotide DNA ISC with the efficiency of formation being dependent upon the intervening sequence between the two cross-linked bases. U-77 was also found to induce monofunctional alkylation of DNA within potential sites of DNA ISC

* To whom correspondence should be addressed.

[1] Abbreviations: ISC, DNA interstrand cross-links; MA, monofunctional alkylations; A tracts, adenine tracts; CPI, cyclopropylpyrroloindole; U-77, U-77,779; DEPC, diethyl pyrocarbonate; DMS, dimethyl sulfate; EDTA, ethylenediaminetetraacetic acid.

0006-2960/93/0432-2592$04.00/0

NAME	SEQUENCES (5' to 3')
A: 6A	GGTCGTCGA**TAAAAA**TCGACG GCAGCT**ATTTTT**AGCTGCCCA
6T	GGTCGTCGA**TAAATA**TCGACG GCAGCT**ATTTAT**AGCTGCCCA
5A	GGTCGTCGA**TAAAA**TTCGACG GCAGCT**ATTTT**AAGCTGCCCA
B: 7A	GGTCGTCGA**TAAAAAA**CGACG GCAGCT**ATTTTTT**GCTGCCCA
7T	GGTCGTCGA**TAAAATA**CGACG GCAGCT**ATTTTAT**GCTGCCCA
6G	GGTCGTCGA**TAGACA**TCGACG GCAGCT**ATCTGT**AGCTGCCCA
C: 6I	GGTCGTCGA**TTAAAA**TCGACG GCAGCT**AATTTT**AGCTGCCCA
6II	GGTCGTCGA**TTTAAA**TCGACG GCAGCT**AAATTT**AGCTGCCCA
6III	GGTCGTCGA**TAATTA**TCGACG GCAGCT**ATTAAT**AGCTGCCCA
6IV	GGTCGACGA**TGATTA**TCGTCG GCTGCT**ACTAAT**AGCAGCCCA

FIGURE 2: Sequences of duplex oligodeoxynucleotides used in these studies. Potential cross-linking sites are shown in bold. Groups A and B were used to study the formation of a 6- and/or 7-nucleotide DNA ISC in bent A tracts. Group C were used to study the nucleotide preferences for formation of a 6-nucleotide DNA ISC formation in nonbent DNA.

formation. DNA conformational changes induced by U-77, such as helical distortion (which was only observed within the cross-linked sequences) and DNA bending, will be discussed.

MATERIALS AND METHODS

Chemicals and Reagents. U-77 (purity 98.5%) was generously supplied by The Upjohn Co. U-77 was dissolved in dimethyl sulfoxide, and the concentration was determined using the extinction coefficient E_{340} = 35 000. DEPC, $KMnO_4$, and DMS were purchased from Sigma. T4 polynucleotide kinase and *Bam*HI linker (12 bp) were obtained from Stratagene, T4 DNA ligase was from Bethesda Research Laboratories, and Spin X centrifuge filter units (0.22-mm nitrocellulose) were from Costar.

Preparation of Oligonucleotides. A series of oligonucleotides (Figure 2) were synthesized on an automated DNA synthesizer (Applied Biosystems 391) and deprotected with saturated ammonium hydroxide at 55 °C overnight. After ethanol precipitation, DNA oligonucleotides were purified on a 20% denaturing polyacrylamide gel.

Purification of DNA ISC and MA. Approximately 10 μg of each oligonucleotide was 5′-end-labeled with T4 polynucleotide kinase and [γ-^{32}P]ATP. After removal of unincorporated [γ-^{32}P]ATP by ethanol precipitation, an equal amount of complementary strand was added, and the mixture was heated to 65 °C and then cooled to room temperature overnight to form annealed duplex in 25 μL of 60 mM NaCl, 6 mM Tris, pH 8.0. Half of a microliter of 1 mM U-77 was added to annealed duplex and incubated at 37 °C for 20 h. DNA ISC and MA produced by U-77 were purified by running a 20% denaturing polyacrylamide gel (mono:bis acrylamide ratio = 29:1, 8 M urea) until the xylene cyanol marker had migrated 15–17 cm. After the gel was exposed to X-ray films, each adduct was excised, soaked at 4 °C overnight, filtered through Spin X centrifuge filter unit, and precipitated with ethanol and sodium acetate. The percentage of DNA ISC in each duplex oligonucleotide was determined by microdensitometry using an LKB ultra scan XL laser densitometer.

Determination of Alkylation Site. Aliquots of drug-induced DNA ISC and MA were heated in 40 μL of 10 mM phosphate buffer, pH 7.0, at 92 °C for 30 min. A subsequent β-elimination reaction, which hydrolyzes the phosphate backbone, was achieved by adding 4.4 μL of 10 M piperidine and heating the mixture at 92 °C for 30 min. After lyophilization, each sample was resuspended in tracking dye containing 80% formamide and 1 mM EDTA and then subjected to a 20% denaturing polyacrylamide gel electrophoresis in parallel with modified Maxam and Gilbert DNA sequencing reactions (Maxam & Gilbert, 1980).

Chemical Probing of DNA ISC and DNA Sequencing Reaction. Control DNA duplex and DNA ISC were purified on an 8% nondenaturing gel and a 20% denaturing gel, respectively, and then reacted with DEPC, $KMnO_4$, and DMS as described below.

A-specific reaction: 10 μL of DEPC was added to 50-μL reactions in 25 mM sodium cacodylate, pH 7.2, and the mixtures were incubated 37 °C for 20 min. Reaction mixes were vortexed at the start of the DEPC reaction and also after 7 min (Herr, 1985). DEPC reaction was terminated with ethanol precipitation.

T-specific reaction: 5 μL of $KMnO_4$ was added to 50-μL reactions in 25 mM sodium cacodylate, pH 7.2, and the mixtures were incubated at 37 °C for 20 min. $KMnO_4$ reaction was terminated with 3 μL of β-mercaptoethanol and ethanol precipitation (McCarthy & Rich, 1991).

G-specific reaction: 30 μL of 10% DMS in ethanol was added to 20-μL reactions in distilled water, and the mixtures were incubated at 37 °C for 20 min. The DMS reaction was terminated with ethanol precipitation.

Purine-specific reaction: The reaction conditions are identical to those of the G-specific reaction except that 30 μL of 88% formic acid was added instead of 10% DMS.

The chemically modified DNA duplexes were resuspended in 40 μL of 1 M piperidine, heated at 92 °C for 30 min, lyophilized overnight (ethanol precipitation is not recommended), and then run on a 20% denaturing gel.

Determination of DNA Bending. Oligonucleotide duplexes (approximately 10 μg) were resuspended in 25 μL of 70 mM Tris-HCl (pH 7.6), 10 mM $MgCl_2$, and 5 mM dithiothreitol, and 5′-end-labeled with 40 μCi of [γ-^{32}P]ATP and 7 units of T4 polynucleotide kinase at 37 °C for 30 min. After end-labeling, 1 μL of 0.1 M cold ATP and 5 units of T4 polynucleotide kinase were added to this mixture and the phosphorylation reaction was continued at 37 °C for an additional 30 min. Phosphorylated duplexes were heated to 65 °C and cooled slowly to room temperature overnight to form hybrids. Annealed duplex DNA was purified by running on an 8% nondenaturing acrylamide gel. DNA ISC was also purified by running on a 20% denaturing acrylamide gel after drug incubation as described above.

Purified nonalkylated duplex DNA and DNA ISC were self-ligated to produce multimers in 20 μL of ligation buffer

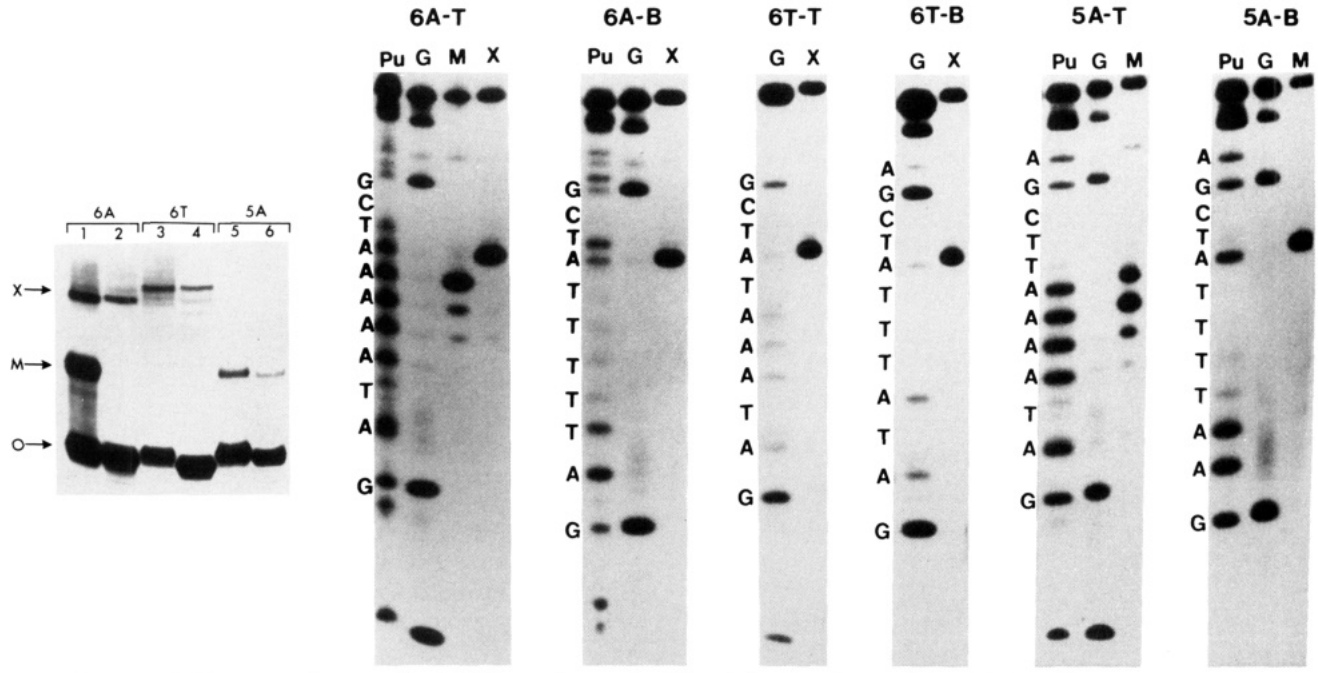


FIGURE 3: Panel A (left): Autoradiogram of a 20% denaturing acrylamide gel showing the two distinct DNA adducts (X and M) induced by U-77 in duplexes 6A (5′-TAAAAA), 6T (5′-TAAATA), and 5A (5′-TAAAAT). Odd numbers refer to the top strand labeled duplex, and even numbers refer to the bottom strand labeled duplex. O indicates original single-stranded DNA. Panel B (right): Determination of the alkylation site of DNA ISC (X) and DNA MA (M) shown in panel A on a 20% denaturing acrylamide gel. Abbreviations: T, the top strand; B, the bottom strand; Pu, purine-specific sequencing reaction: G, guanine-specific sequencing reaction.

with 1 μL (1 unit) of T4 DNA ligase at room temperature overnight. The ligation buffer contains 50 mM Tris-HCl (pH 7.6), 10 mM $MgCl_2$, 5% (w/v) poly(ethylene glycol) 8000, 1 mM ATP, and 1 mM dithiothreitol. The ligation reaction was terminated by quenching with 1/10 volume of 250 mM EDTA (pH 8.0). Ligated multimers were run on an 8% nondenaturing polyacrylamide gel (mono:bis acrylamide ratio = 29:1) until the bromophenol blue marker had migrated 23 cm at room temperature. R_L values, the ratio of apparent size to true size, were measured relative to *Bam*HI linkers (Koo et al., 1986). pBR322-*Hin*fI digests were also electrophoresed adjacent to multimers of *Bam*HI linkers (data not shown).

RESULTS

Identification of DNA ISC and MA Produced by U-77. In order to determine the nucleotide preferences for DNA ISC induced by U-77, we have analyzed numerous duplex oligonucleotides each with different centrally located sequences. Duplexes 6A (5′-TAAAAA), 6T (5′-TAAATA), and 5A (5′-TAAAAT) were treated with drug and were resolved on a denaturing 20% polyacrylamide gel. Panel A (left) in Figure 3 shows that when the top strand of duplex 6A was labeled, we observed two bands (bands X and M) with mobilities different from those of nonalkylated single-stranded oligonucleotide (band O, lane 1). In contrast, when the bottom strand of duplex 6A was labeled, only one band with a mobility different from control DNA was observed (lane 2). Band X which has a similar mobility regardless of whether the top or bottom strand had been end-labeled is taken to represent cross-linked DNA. Band M, however, which is only observed in duplex 6A, when the top strand is labeled, is taken to represent DNA MA. The slow electrophoretic mobility of DNA MA may be a result of the high molecular weight of U-77 and the formation of two extra positive charges on the exocyclic N6 nitrogen of covalently modified adenine upon cross-linking (Lin & Hurley, 1990).

In order to confirm that band X represents DNA ISC and that band M represents DNA MA, DNA sequencing analysis was performed upon the isolated reaction products (panel B (right), Figure 3). Heat treatment of the isolated bands at neutral pH followed by hot 1 M piperidine treatment produces a second β-elimination product which displays a mobility identical to that seen with the Maxam and Gilbert sequencing reactions (Reynolds et al., 1985). The major site of alkylation in band X (top strand labeled) was at the last adenine in a 5′-TAAAAA sequence (lane X in 6A-T). The major site of alkylation in band X (bottom strand labeled) was at the adenine in a 5′-TTTTTA sequence (lane X in 6A-B). When taken together, these data suggest that the cross-linking site observed in duplex 6A was within a 5′-TAAAAA*, where the asterisk indicates the site of alkylation on the top strand and T indicates the location of adenine alkylation on the complementary strand. Sequencing analysis of band M indicated that the major adenine involved in the formation of DNA MA within duplex 6A was the penultimate adenine in the 5′-TAAAA*A sequence (lane M in 6A-T).

In order to test whether this interpretation was valid, duplexes 6T (5′-TAAATA) and 5A (5′-TAAAAT) were synthesized and analyzed in a fashion identical to that described above. Duplex 6T contained a thymine substituted for an adenine at the major site of DNA MA. In contrast, duplex 5A contained a thymine substituted for an adenine at the base involved in the DNA ISC on the top strand. U-77 induced DNA ISC (band X) in duplex 6T (5′-TAAATA) but not in duplex 5A (lanes 3–6, panel A (left), Figure 3). U-77 induced DNA MA in duplex 5A (5′-TAAAAT) but not in duplex 6T (lanes 3–6, panel A (left), Figure 3). These results were confirmed by sequencing analysis (panel B (right), Figure 3). Thus, it can be concluded that band X is a DNA ISC and that band M is a DNA MA. In addition, U-77 induces a 6-nucleotide DNA ISC within 5′-TAAAAA and 5′-TAAATA sequences but does not induce a 5-nucleotide DNA ISC within a 5′-TAAAAT sequence.

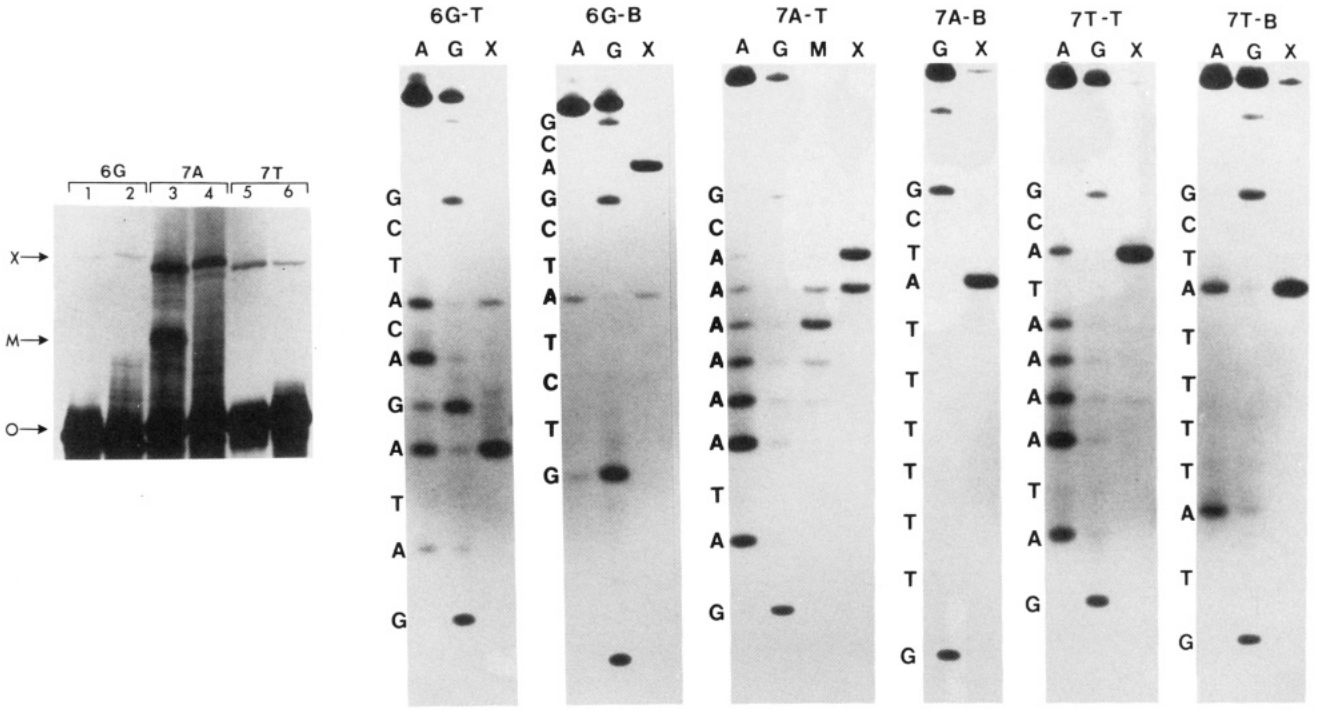


FIGURE 4: Panel A (left): Autoradiogram of a 20% denaturing acrylamide gel showing the two distinct DNA adducts (X and M) induced by U-77 in duplexes 6G (5′-TAGACA), 7A (5′-TAAAAAA), and 7T (5′-TAAAATA). Odd numbers refer to the top strand labeled duplex, and even numbers refer to the bottom strand labeled duplex. O indicates original single-stranded DNA. Panel B (right): Determination of the alkylation site of DNA ISC (X) and DNA MA (M) shown in panel A on a 20% denaturing acrylamide gel. Abbreviations: T, the top strand; B, the bottom strand; A, adenine-specific sequencing reaction; G, guanine-specific sequencing reaction.

Identification of a 7-Nucleotide DNA ISC in A_6 Tracts. Sequences rich in contiguous adenines induce DNA bending toward the minor groove of DNA (Zinkel & Crothers, 1987; Crothers et al., 1990) and are potential sites for U-77-induced DNA ISC. Therefore, we have determined whether bent DNA has any influence upon the distance which a DNA ISC can span as well as the efficiency of its formation. For this purpose, duplexes which contained centrally located 5′-TAAAAAA (duplex 7A) and 5′-TAAAATA sequences (duplex 7T) were analyzed as above. In these experiments, a 5′-TAGACA sequence (duplex 6G) was used as a control.

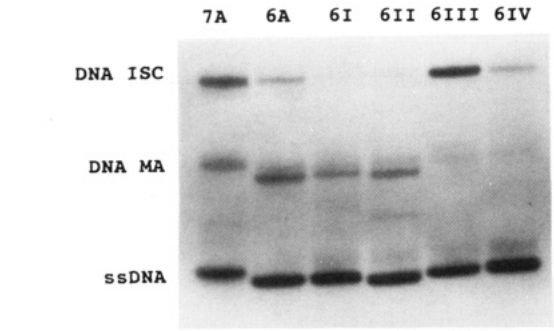
A significant quantity of DNA ISC was observed within 5′-TAAAAAA sequences (7A in panel A (left), Figure 4), while only a very small quantity of DNA ISC (10% of 5′-TA_6) was observed within a 5′-TAAAATA sequence (7T in panel A (left), Figure 4). The negligible amount of DNA ISC observed within the 5′-TAGACA sequence (6G in panel A (left), Figure 4) was expected as substitution of adenine with either guanine or cytosine in the cross-linked site is thought to inhibit the drug from binding within the minor groove of DNA.

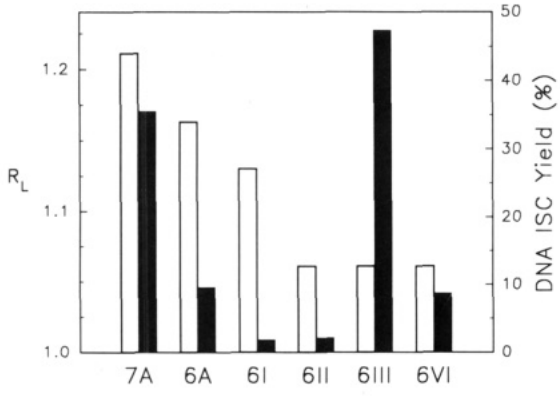
Subsequent sequencing gel analysis was performed to determine alkylation sites of DNA ISC and DNA MA in duplexes 7A, 7T, and 6G (panel B (right), Figure 4). When duplex 7A was labeled on the top strand, DNA cleavage occurred at two adenines, thus implicating two sites of alkylation within this DNA ISC band (lane X, 7A–T, Figure 4B). In contrast, when duplex 7A was labeled on the bottom strand, only one site of alkylation (at an adenine) was found to be associated with the DNA ISC band (lane X, 7A–B, panel B (right), Figure 4). This result suggests that U-77 induces two distinct DNA ISC, with similar mobility, within a 5′-TAAAAAA-3′ sequence. Desnitometric analysis of the sequencing gel suggests that U-77 induces the 7-nucleotide DNA ISC (5′-TAAAAAA*) to a 2-fold greater degree than the 6-nucleotide DNA ISC (5′-TAAAAA*A) when both potential DNA ISC sites are located within the same duplex DNA.

U-77 was also found to induce DNA ISC in duplex 7T, but the efficiency of formation of this 7-nucleotide DNA ISC, however, was much less than that observed with duplex 7A (panel A (left), Figure 4). Sequencing analysis identifies the cross-link as spanning 7 nucleotides within a 5′-TAAAATA sequence and linking the two adenines at the ends of this particular sequence (5′-TAAAATA*) (lane X in 7T–T and 7T–B, panel B (right)). It can be concluded that 7-nucleotide DNA ISC can be induced at either 5′-TAAAAAA or 5′-TAAAATA sequences.

Further, some minor DNA interstrand cross-linking was observed within duplex 6G, indeed cross-linking occurs within two sequences in this duplex. The major cross-link is formed outside the centrally located 5′-TAGACA-3′ sequence and may reflect the ability of U-77 to alkylate the 5′-ATA-3′ triplet within this duplex. An adenine is located 6 nucleotides distant on the opposite strand thus allowing DNA ISC to form within the 5′-TCGATA*-3′ sequence after the initial alkylation event (panel B (right), Figure 4). It is unlikely that the initial alkylation occurs within the 5′-CGA-3′ triplet as this is not a favored site for alkylation by the CPI class of molecules (Hurley et al., 1990; Boger et al., 1991c).

Comparison of DNA ISC Formation in Bent and Nonbent DNA. The high yield of DNA ISC observed in these bent DNA duplexes (duplexes 7A and 6A) suggested that bent DNA may be a preferred target for DNA ISC formation by U-77. In order to address this question, the efficiency of DNA ISC formation and the inherent DNA bending associated with each duplex DNA sequence were determined in duplexes 7A, 6A, 6I, 6II, 6III, and 6IV (Figure 5). Panel A (top left) shows the separation of nonmodified DNA from DNA MA and DNA ISC on a 20% denaturing gel. The degree of DNA bending associated with each duplex is visualized by the alteration in mobility of ligated multimers upon an 8%





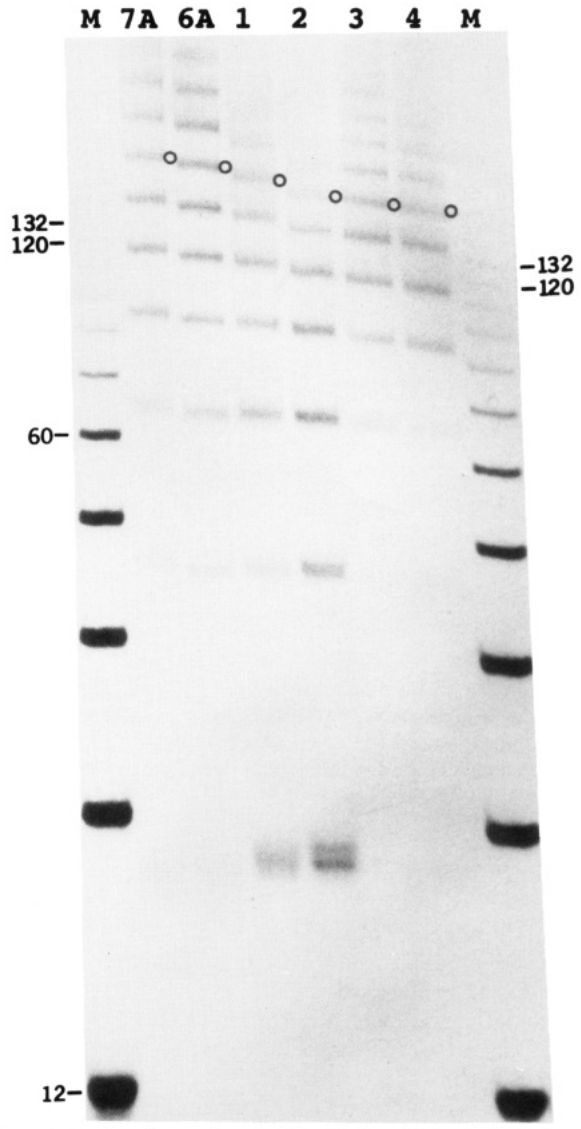


FIGURE 5: Panel A (top left) Autoradiogram of a 20% denaturing acrylamide gel showing formation of DNA ISC and MA in duplexes 7A (5′-TAAAAAA), 6A (5′-TAAAAA), 6I (5′-TTAAAA), 6II (5′-TTTAAA), 6III (5′-TAATTA), and 6IV (5′-TGATTA). The top strand labeled duplexes were purified on an 8% nondenaturing gel and then treated with drug. Panel B (right): Autoradiogram of a 8% nondenaturing gel showing electrophoretic mobility of the multimers of nonalkylated DNA duplexes. The position of the seven multimer sized ligation product is indicated by circles. Multimers of the *Bam*HI linkers (12 bp) were also electrophoresed (lane M) adjacent to a pBR322-*Hin*fI digest (data not shown) to confirm their normal gel mobility. Panel C (bottom left): Histogram of R_L values (open bars) of 7 multimers of nonalkylated DNA duplexes and yield (%) of DNA ISC formation (filled bars). R_L values were determined by comparison with the multimers of the *Bam*HI linker (Koo et al., 1986).

nondenaturing gel (panel B (right), Figure 5). The R_L value is the ratio of apparent size to true size for each multimer and provides a manner by which the degree of DNA bending can be quantitated. The histogram of R_L values (determined at the 7 multimer size) and the efficiency of DNA ISC formation for each duplex, as determined by microdensitometry, are shown in panel C (bottom left) in Figure 5. The degree of DNA bending associated with each duplex DNA decreased in the following order: 5′-TA_6 > 5′-TA_5 > 5′-T_2A_4 > 5′-T_3A_3 = 5′-TAATTA = 5′-TGATTA sequences. In contrast, the efficiency of U-77-induced DNA ISC formation decreased in the following order: 5′-TAATTA > 5′-TA_6 > 5′-TA_5 ≥ 5′-TGATTA > 5′-T_2A_4 = 5′-T_3A_3 sequences. These results suggest that there is no relationship between the efficiency of formation of the 6-nucleotide DNA ISC and the inherent DNA bending associated with each duplex DNA.

Sequencing gel analyses of U-77-induced DNA ISC and DNA MA within duplexes 6I, 6II, 6III, and 6IV are shown in Figure 6. Multiple monofunctional alkylation sites were observed within 5′-T_2A_4 (duplex 6I) and 5′-T_3A_3 sequences (duplex 6II), while a single alkylation site, which results in DNA ISC, was observed within 5′-TAATTA (duplex 6III) and 5′-TGATTA (duplex 6IV) sequences.

Table I shows a summary of the preferred sites for DNA alkylation and DNA ISC formation. The preferred sites for U-77-induced DNA ISC were found to be within 5′-TAATTA-3′ (6-nucleotide DNA ISC) and 5′-TAAAAAA-3′ (7-nucleotide DNA ISC) sequences. The low levels of DNA MA formation observed in duplexes 6T and 7T indicate that the preferred site for monofunctional alkylation within such sequences is at the adenine involved in DNA ISC formation in the top strand. In these duplexes, the initial alkylation event always leads to DNA ISC formation.

Elimination of Inherent DNA Bending Ability of A Tracts by U-77. The effect of DNA ISC formation upon the inherent DNA bending ability of short A_5 and A_6 tracts has been investigated. The DNA ISC formed within 5′-TAAAAA (duplex 6A), 5′-TAAAAAA (duplex 7A), and 5′-TAATTA

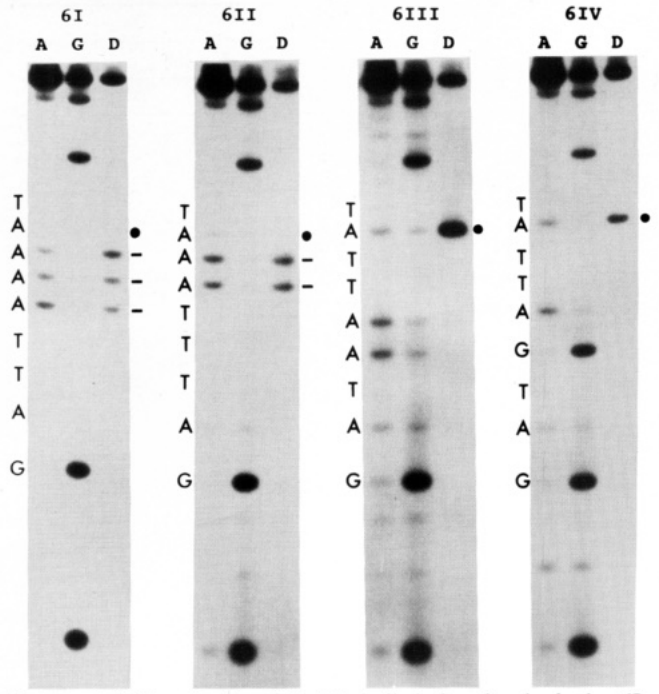


FIGURE 6: Determinationof the total alkylation sites in duplex 6I, 6II, 6III, and 6IV on a 20% denaturing acrylamide gel. The top strand labeled duplexes were purified on an 8% nondenaturing gel and then treated with drug. Filled circles represent the site of DNA ISC and dashed lines represent the site of DNA MA. Abbreviations: A, adenine-specific sequencing reaction; G, guanine-specific sequencing reaction; D, drug alkylation.

(duplex 6III) sequences were self-ligated to produce multimers, and their electrophoretic mobility was compared with that of nonalkylated duplex DNA. As shown in Figure 7, R_L values of the multimers of nonalkylated duplex DNA and of the respective DNA ISC were plotted against the actual length of multimers of the *Bam*HI linker (12 bp). In the case of small multimers (i.e., monomer, dimers, and trimers), the difference in R_L values between DNA ISC and control DNA is negligible. However, when the sizes of multimers are increased, the difference in R_L values between multimers of the DNA ISC and multimers of control DNA becomes apparent. The DNA bending ability associated with A_5 and A_6 tracts was eliminated as a consequence of DNA ISC formation by U-77 (Figure 7). U-77 induced DNA ISC formation in nonbent duplex 6III (5′-TAATTA) and did not alter the R_L value.

Chemical Probing of DNA ISC. Control or cross-linked DNA was reacted with DEPC (A-specific) or $KMnO_4$ (T-specific), and the susceptibility of DNA ISC to cleavage was determined by denaturing gel electrophoresis (penal A (top), Figure 8). Panel B (bottom) in Figure 8 shows a densitometric analysis of the cleavage patterns obtained after control DNA duplex (A and C) and DNA ISC (B and D) were reacted with DEPC (A and B) and $KMnO_4$ (C and D), respectively. The densitometric tracings obtained with the DNA ISC are distinct from those obtained with control DNA. The central adenine within the 5′-AAAAA sequence in DNA ISC was consistently hypersensitive to DEPC (B in panel B (bottom), Figure 8). The increased hypersensitivity of the 3′-A in cross-linked DNA presumably reflects the fact that this adenine is alkylated by U-77, a reaction which obviously does not occur in control DNA.

Hypersensitivity of the DNA ISC to $KMnO_4$ exposure occurred at the thymine immediately 5′ of the site of alkylation within the 5′-TTTTTA sequence (C and D, panel B (bottom), Figure 8). Further, a reduced sensitivity to $KMnO_4$ was observed at the thymine 3′ of the thymine opposite the adenine involved in the cross-link on the complementary strand. No hypersensitivity to $KMnO_4$ was observed at the other thymines within the cross-linked site (Figure 8). The thymine which is located four bases away from the DNA ISC did not exhibit any altered sensitivity. The DNA ISC was also reacted with dimethyl sulfate and this showed that none of the guanine residues were hyperreactive (data not shown). These results suggest that U-77 induces DNA structural changes which may be localized within the cross-linked sequence 5′-TAAAAA*.

DISCUSSION

In this study, we have found that U-77 induces two distinct DNA ISC, one which spans 6 nucleotides within 5′-TNNNNA-3′ sequences and one which spans 7 nucleotides within 5-TNNNNNA-3 sequences. No DNA ISC which spans 5 nucleotides was observed. The two strands of DNA are linked by alkylation of adenine at the N3 position, and thus the drug molecule is located within the minor groove of DNA. DNA ISC formation within our 6-base-pair target sequences was found to occur with either high (duplex 6III), intermediate (duplexes 6A, 6T, and 6IV), or low efficiency (duplexes 6I and 6II).

The most preferred DNA ISC was found to span 6 nucleotides and was located within 5′-TAATTA-3′ sequences. The ability of U-77 to induce a 6-nucleotide DNA ISC is consistent with molecular modeling studies, where the size of binding site was suggested to be six base pairs inclusive of the alkylated adenines (Mitchell et al., 1991), and with recent experimental data from Hurley and colleagues (Ding & Hurley, 1991). Further, CC-1065 preferentially alkylates the adenine N3 position within 5′-TTA* sequences, with 5′-AAA* and 5′-ATA* sequences being intermediate sites in terms of efficiency of alkylation (Hurley et al., 1990; Boger et al., 1991c). The 5′-TAATTA-3 sequence (duplex 6III) thus represents the joining of two adjacent 5′-TTA triplets, and thus U-77 alkylation and subsequent cross-linking follows a similar sequence selectivity of alkylation determined for CC-1065. That is, two preferred sites for alkylation (5′-TTA) are found on opposite strands of DNA with the two adenines being separated by 6 nucleotides.

Intermediate levels of DNA ISC were found to occur in duplexes 6A, 6T, and 6IV. In all cases, one 5′-TTA-3′ triplet, representing the preferred site for an initial monofunctional alkylation, was present. Formation of a DNA ISC is allowed by the presence (regardless of the intervening sequence) of an adenine located 6 nucleotides in the 3′ direction of the complementary strand. In all these sequences, U-77 induced approximately 5-fold less DNA ISC formation than in the 5′-TAATTA-3′ sequences.

The low level of DNA ISC formation observed in duplexes 6I and 6II can be rationalized in the following manner. First, they lack a 5′-TTA-3′ sequence (the preferred sequence for monofunctional alkylation). Next, an initial alkylation of the 3′-end adenine within an A/T-rich sequence, which has a 3′ T immediately adjacent, is unlikely to occur (Boger et al., 1991c). Alkylation therefore, at any of the other adenines within duplexes 6I and 6II would result in an adduct which cannot cross-link DNA. Indeed, we have observed that the preferred site for DNA MA within a 5′-TT(A/T)AAAT-3′ sequence is the penultimate adenine from the 3′ end of the run of adenines (Figure 6). Such an alkylation cannot lead to DNA ISC formation, and this explanation can also account

Table I: Analysis of Sequence Selectivity of DNA ISC by U-77[a]

name	cross-linked sequence[b]	two subsets of triplet	% yield MA	% yield ISC
	6 nucleotide DNA ISC			
6III	5'-TAATTA*	5'-TTA + 5'-TTA	1.4	47.3
6A	5'-TAAAAA*	5'-TTA + 5'-AAA	27.0	9.6
6T[c]	5'-TAAATA*	5'-TTA + 5'-ATA	1.3	10.7
6IV	5'-TGATTA*	5'-TTA + 5'-TCA	2.4	8.7
6I	5'-TTAAAA*	5'-TAA + 5'-AAA	14.9	1.8
6II	5'-TTTAAA*	5'-AAA + 5'-AAA	15.1	2.1
6G	5'-TAGACA*	5'-ACA + 5'-CTA	0.0	< 0.5[d]
	7 nucleotide DNA ISC			
7A	5'-TAAAAA*A*	5'-TTA + 5'-AAA	17.3	35.5
7T[e]	5'-TAAAATA*	5'-TTA + 5'-ATA	0.0	3.9
	5 nucleotide DNA ISC			
5A[c]	5'-TAAAATT	5'-TTA + 5'-ATT	9.6	0.0

[a] These data were obtained from Figure 5. [b] For simplicity, the sequences of the top strand are shown. The asterisk (*) indicates the cross-linked site on the top strand, and the underlined T indicates the cross-linked site on the bottom strand. The arrow indicates the monofunctional alkylation site. The bold arrow indicates the major monofunctional alkylation site. [c] Data obtained from Figure 3. [d] The percentage yield reflects a mixture of two cross-linked species occurring within a 5'-TCGATA*-3' (major site) and a 5'-TAGACA*-3' (minor site) sequence. [e] Data obtained from Figure 4.

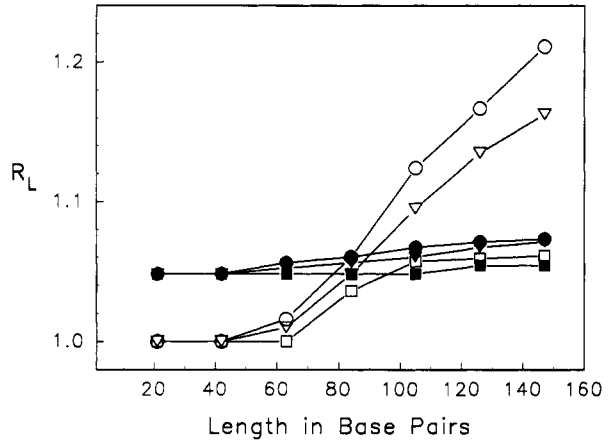


FIGURE 7: Plot of R_L values of multimers of DNA ISC and control DNA duplex against actual length of the multimers of the *Bam*HI linker. R_L values were determined by comparison with the multimers of the *Bam*HI linker (12 bp). Symbols: Open circles, control duplex 7A; closed circles, duplex 7A DNA ISC; open triangles, control duplex 6A; closed triangles, duplex 6A DNA ISC; open squares, control duplex 6III; closed squares, duplex 6III DNA ISC.

for the high formation of DNA MA observed in the top strand of duplex 6A.

It is important to note that a previous study failed to detect the presence of U-77-induced DNA MA (Ding & Hurley, 1991). The difference between the two studies is that we have used duplex oligonucleotides which are labeled on either the top or bottom strand of DNA. This has allowed us to differentiate between DNA ISC products and DNA MA products. We thus interpret the product designated band B in Ding and Hurley's work to be a DNA MA and not a DNA ISC (Ding & Hurley, 1991). Further, the lack of DNA bending associated with the 6-nucleotide 5'-TAATTA-3' sequence is in contrast to previous work (Ding & Hurley, 1991) but suggests that bent DNA is not required for the formation of a 6-nucleotide DNA ISC within this sequence.

The molecular mechanism of formation of the 7-nucleotide DNA ISC may be distinct from that controlling formation of a 6-nucleotide DNA ISC. Seven-nucleotide DNA ISC formation was observed preferentially within A_6 tracts. Such runs of contiguous adenines induce DNA bending by 19° toward the minor groove (Crothers et al., 1990). One altered base, replacement of an A with a T within this A_6 tract (duplex 7T), results in a dramatic reduction in U-77-induced DNA ISC formation, as well as a dramatic reduction in the inherent bending associated with such a sequence. This suggests that bent DNA might be a requirement for DNA ISC formation spanning 7 nucleotides. This can be rationalized by the fact that minor groove compression at the center of the A tracts (Zinkel & Crothers, 1987) could reduce the distance between the two adenine N3 positions required for DNA ISC formation. Consequently, the two DNA-reactive CPI functions might now be capable of spanning the distance between the two adenine N3 positions on opposite strands of DNA. In support of this, the yield of a 7-nucleotide DNA ISC observed in the non-adenine contiguous sequences, 5'-TAATTTA* and 5'-TAAGTTA*, is lower than that observed here (Ding & Hurley, 1991).

The formation of DNA ISC by U-77 in A tracts has been found to eliminate the intrinsic DNA bending associated with such sequences. This is in contrast to the observation that CC-1065-induced DNA MA at the 3'-end of A_5 tracts enhances their inherent DNA bending ability (Lin et al., 1991).

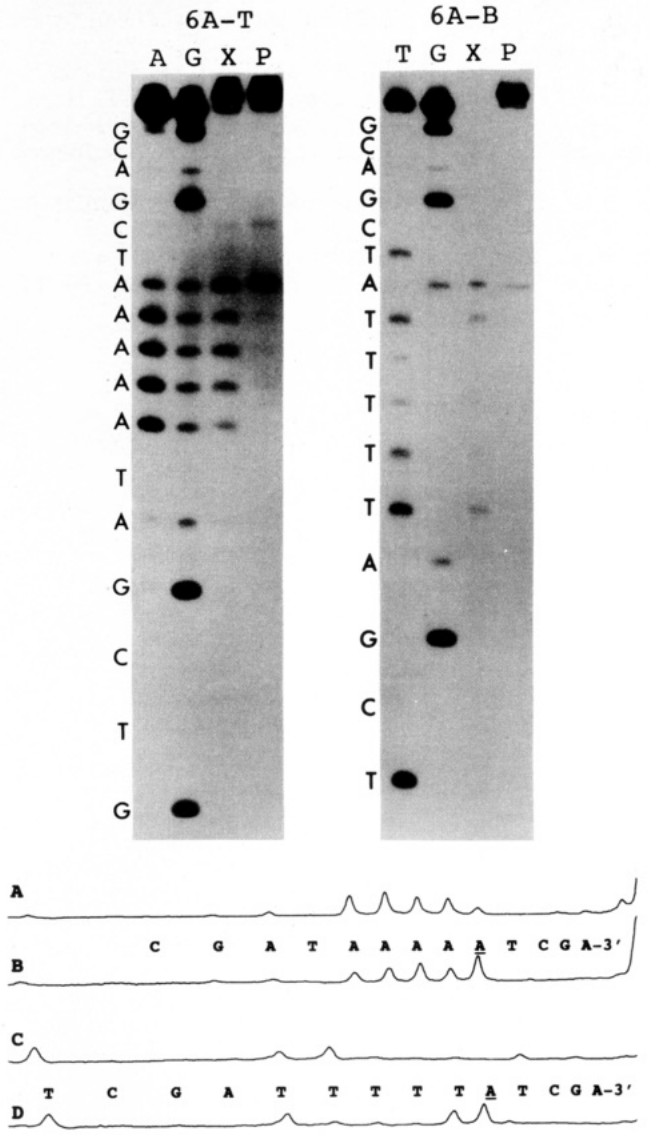


FIGURE 8: Panel A (top): Pattern of chemical reactivities of DNA ISC on a 20% denaturing acrylamide gel. Purified DNA ISC observed in the duplex 6A labeled at the top (T) or the bottom (B) strand was reacted with DEPC (T) or $KMnO_4$ (B), respectively. Abbreviations: A, adenine-specific sequencing reaction; G, guanine-specific sequencing reaction; X, DNA ISC; P, 1 M piperidine treatment only. Chemical reactivity of control DNA duplex is not shown. Panel B (bottom): Densitometric analysis of the chemical reactivity of control DNA duplex and DNA ISC. Abbreviations: A, DEPC reaction of control DNA duplex; B, DEPC reaction of DNA ISC; C, $KMnO_4$ reaction of control duplex DNA; D, $KMnO_4$ reaction of DNA ISC. Underlined A represents the cross-linking site. Nucleotide sequences are 5′ to 3′.

We propose that drug-induced DNA ISC formation causes a significant distortion of phosphate backbone and/or sugar moiety and that this may disrupt the overall structure of unusual A tracts. The biological consequences of intrinsic DNA bending (Crothers et al., 1990; Hagerman, 1990) is thought to be important as A tracts have been identified in various gene regulatory regions (Schroth et al., 1992; Zahn & Blatter, 1987; Bossi & Smith, 1984; Ross et al., 1982). Modification of such sequences, such as the elimination of DNA bending by U-77 DNA ISC formation, could affect DNA protein binding and ultimately the processes of gene regulation.

In summary, U-77 was found to induce two distinct forms of DNA ISC. One DNA ISC spans 6 nucleotides and mimics the nucleotide preferences for CC-1065-induced DNA MA, where the preferred target sequence contains two 5′-TTA-3′ sequences within the 5′-TAATTA-3′ sequence. The second DNA ISC spans 7 nucleotides and is found to prefer bent A tracts. The formation of this DNA ISC was accompanied by the pronounced structural changes within cross-linked sequences and elimination of intrinsic DNA bending in A tracts. Finally, it is clear that the rational design of drugs which can alkylate and cross-link DNA at specific nucleotide sequences can be achieved. This raises the possibility that molecules may be designed to mimic the sequence selectivity of DNA-binding proteins (Mitchell & Tijian, 1989). Whether such DNA-reactive agents could mimic the function of the protein products of native tumor suppressor genes (Madden et al., 1991; Robbins et al., 1990) and thus play a role in cancer chemotherapy is an intriguing speculation.

REFERENCES

Bhuyan, B. K., Newell, K. A., Crampton, S. L., & Von Hoff, D. D. (1982) *Cancer Res. 42*, 3532–3537.

Boger, D. L., Munk, S. A., & Zarrinmayeh, H. (1991a) *J. Am. Chem. Soc. 113*, 3980–3983.

Boger, D. L., Zarrinmayeh, H. J., Munk, S. A., Kitos, P. A., & Suntornwat, O. (1991b) *Proc. Natl. Acad. Sci. U.S.A. 88*, 1431–1435.

Boger, D. L., Munk, S. A., Zarrinmayeh, H., Ishizaki, T., Haught, J., & Bina, M. (1991c) *Tetrahedron 47*, 2661–2682.

Bossi, L., & Smith, D. M. (1984) *Cell (Cambridge, Mass.) 39*, 643–652.

Burris, H., Earhart, R., Kuhn, J., Shaffer, D., Smith, L., Weiss, G., Kasunic, D., Radbury, G., Campbell, L., & Von Hoff, D. (1992) *Proc. Am. Assoc. Cancer Res. 33*, 520.

Crothers, D. M., Haran, T. E., & Nadeau, J. G. (1990) *J. Biol. Chem. 265*, 7093–7096.

Ding, Z.-M., & Hurley, L. H. (1991) *Anti-Cancer Drug Des. 6*, 427–452.

Fleming, G. F., Ratain, M. J., O'Brien, S. M., Vogelzang, N. J., & Earhart, R. H. (1992) *Proc. Am. Assoc. Cancer Res. 33*, 265.

Hagerman, P. J. (1990) *Annu. Rev. Biochem. 59*, 755–781.

Hanka, L. J., Dietz, A., Gerpheide, S. A., Kuentzel, S. L., & Martin, D. G. (1978) *J. Antibiot. 31*, 1211–1217.

Haran, T. E., & Crothers, D. M. (1988) *Biochemistry 27*, 6967–6971.

Herr, W. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 8009–8013.

Hurley, L. H., Reynolds, V. L., Swenson, D. H., Petzold, G. L., & Scahill, T. A. (1984) *Science 226*, 843–844.

Hurley, L. H., Lee, C.-S., McGovren, J. P., Mitchell, M. A., Warpehoski, M. A., Kelly, R. C., & Aristoff, P. A. (1988) *Biochemistry 27*, 3886–3892.

Hurley, L. H., Warpehoski, M. A., Lee, C.-S., McGovren, J. P., Scahill, T. A., Kelly, R. C., Wicnienski, N. A., Gebhard, I., & Bradford, V. S. (1990) *J. Am. Chem. Soc. 112*, 4633–4649.

Johnson, N. P., & Rich, A. (1985) *Cell (Cambridge, Mass.) 42*, 713–724.

Koo, H.-S., Woo, H.-M., & Crothers, D. M. (1986) *Nature 320*, 501–506.

Lee, C.-S., & Gibson, N. W. (1991) *Cancer Res. 51*, 6586–6591.

Li, L. H., Swenson, D. H., Schpok, S., Kuentzel, S. L., Dayton, B. D., & Krueger, W. C. (1982) *Cancer Res. 42*, 999.

Lin, C. H., & Hurley, L. H. (1990) *Biochemistry 29*, 9503–9507.

Lin, C. H., Sun, D., & Hurley, L. H. (1991) *Chem. Res. Toxicol. 4*, 21–26.

Madden, S. L., Cook, D. M., Morris, J. F., Gashler, A., Sukhatme, V. P., & Rascher, F. J., III (1991) *Science 253*, 1550–1553.

Martin, D. G., Biles, C., Gerpheide, S. A., Hanka, L. J., Krueger, W. C., McGovren, J. P., Mizsak, S. A., Neil, G. L., Stewart, J. C., & Visser, J. (1981) *J. Antibiot. 34*, 1119–1125.

Maxam, A., & Gilbert, W. (1980) *Methods Enzymol. 65*, 499–560.

McCarthy, J. G., & Rich, A. (1991) *Nucleic Acid Res. 19*, 3421–3429.
McGovren, J. P. (1992) The Upjohn Co., private communication.
McGovren, J. P., Clarke, G. L. Pratt, E. A., & DeKoning, T. F. (1984) *J. Antibiot. 37*, 63–70.
Mitchell, M. A., Kelly, R. C., Winnienski, N. A., Hatzenbuhler, N. T., & Williams, M. G. (1991) *J. Am. Chem. Soc. 113*, 8994–8995.
Mitchell, P. J., & Tijian, R. (1989) *Science 245*, 371–378.
Reynolds, V. L., Molineux, I. J., Kaplan, D. J., Swenson, D. H., & Hurley, L. H. (1985) *Biochemistry 24*, 6228–6237.
Robbins, P. D., Horowitz, J. M., & Mulligan, R. C. (1990) *Nature 346*, 668–671.
Ross, W., & Landy, A. (1982) *J. Mol. Biol. 156*, 505–529.
Schroth, G. P., Siino, J. S., Cooney, C. A., Th'ng, J. P., Ho, S., & Bradbury, E. M. (1992) *J. Biol. Chem. 267*, 9958–9964.
Warpehoski, M. A., & Hurley, L. H. (1988) *Chem. Res. Toxicol. 1*, 315–333.
Weiland, K. L., & Dooley, T. P. (1991) *Biochemistry 30*, 7559–7565.
Zahn, K., & Blatter, F. R. (1987) *Science 236*, 416–422.
Zinkel, S. S., & Crothers, D. M. (1987) *Nature 328*, 178–181.